Chemical Geology 581 (2021) 120402

Contents lists available at ScienceDirect

Chemical Geology

journal homepage: www.elsevier.com/locate/chemgeo

Invited Research Article

Trend reversal of soil *n*-alkane Carbon Preference Index (CPI) along the precipitation gradient and its paleoclimatic implication

Yan Yan [a,b,c,*], Bingyan Zhao [d,e], Luhua Xie [a,b], Zhaoyu Zhu [a,b]

[a] Key Laboratory of Ocean and Marginal Sea Geology, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, China
[b] CAS Center for Excellence in Deep Earth Science, Guangzhou 510640, China
[c] Geology & Geophysics Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA
[d] University of Chinese Academy of Sciences, Beijing 100049, China
[e] State Key Laboratory of Organic Geochemistry, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, China

ARTICLE INFO

Editor: Dr. Karen Johannesson

Keywords:
n-alkane
CPI
Precipitation gradient
Paleoclimate reconstruction

ABSTRACT

Homologue distribution proxies of leaf wax *n*-alkanes (normal chain alkanes) have been widely used for chemotaxonomy and paleoclimate reconstruction, while modern calibrations of them reported contrasting conclusions in different climate settings, especially between the cold arid and other climate regions. To further explore *n*-alkane-abiotic factor correlation in cold arid regions, this study surveyed *n*-alkane patterns of soil and leaf wax of three plant species along confined temperature and precipitation gradients in the Junggar Basin of northwest China. Our results show that salinity has a possible strong constraint on the Average Chain Length (ACL) of soil *n*-alkanes in high salinity settings, while precipitation or water stress is the leading constraint on the Carbon Preference Index (CPI) of those in low salinity settings. Precipitation also shows strong control over CPI or ACL of leaf wax *n*-alkanes of different plant species. Our results confirmed that *n*-alkane patterns are sensitive to, and constrained by abiotic factors independently on levels of plant species and averaged biome. A compilation of soil CPI from transect studies across the hot-humid to cold-arid regions of China further revealed precipitation being the first-order constraint of soil CPI, and a CPI trend reversal pattern along the precipitation gradient, where CPI increases from arid to semiarid climate and then decreases toward humid climate. Further analysis of existing CPI records showed that this CPI trend reversal may help differentiate three climate states qualitatively, i.e. arid, semiarid-semihumid and humid climates, which could be a valuable complement to existing proxies for paleoclimate reconstruction.

1. Introduction

Leaf wax is predominantly composed of long-chain aliphatic compounds, *n*-alkanes (normal chain alkanes) being one major group of them, which forms the outermost layer of leaf cuticle to protect plants from extreme climatic and environmental conditions that can impose negative influence on plant growth, i.e. abiotic stresses (Eglinton and Hamilton, 1967; Eglinton and Logan, 1991; Edwards et al., 1996). It had been observed that the *n*-alkane chain-length distribution patterns and abundances can vary dramatically across different plant types and species, and in response to various abiotic stresses within the same plant species and types, such as radiation, heat, cold, drought, and extreme pH and salinity (Post-Beittenmiller, 1996; Jetter et al., 2006; Shepherd and Wynne Griffiths, 2006; Bush and McInerney, 2013; and the references therein). Because of their relative recalcitrance and omnipresence in sediments (Eglinton and Logan, 1991; Peters et al., 2005), diverse *n*-alkane based proxies had been constructed and calibrated in modern settings as chemotaxonomy and direct abiotic-factor proxies, and then applied to geologic deposits to infer past climatic and environmental changes (Castañeda and Schouten, 2011; Sachse et al., 2012; Diefendorf and Freimuth, 2017), including the Dominant Chain Length (DCL), Average Chain Length (ACL), Carbon Preference Index (CPI) and various ratios of different homologue combinations.

Based on limited data, early studies concluded that plant *n*-alkanes show systematically strong odd-over-even carbon number predominance, and different *n*-alkanes are attributed to dominant contributions from diverse plant types, e.g. C_{23}, C_{25}, C_{27}, C_{29} and C_{31} *n*-alkanes were deemed dominant in Sphagnum moss, aquatic plants, trees, shrubs and

* Corresponding author at: 511 Kehua Street, Guangzhou 510640, China.
E-mail address: yanyan@gig.ac.cn (Y. Yan).

https://doi.org/10.1016/j.chemgeo.2021.120402
Received 27 November 2020; Received in revised form 15 June 2021; Accepted 16 June 2021
Available online 19 June 2021

grasses, respectively; therefore, variations of the DCL, ACL and ratios of various homologue combinations of *n*-alkanes in geologic records were used to infer past vegetation and thus environmental and climatic changes, while CPI measuring the strength of odd-over-even predominance was used as a petroleum source-rock maturity proxy, or to distinguish terrestrial from bacterial *n*-alkanes (Bray and Evans, 1961; Eglinton and Hamilton, 1967; Corrigan et al., 1973; Poynter et al., 1989; Meyers and Ishiwatari, 1993; Ficken et al., 1998; Nott et al., 2000; Meyers, 2003). Recent studies having included more specimens from different plant types and species, however, showed that although dramatic variations can exist among some coarse divisions of plants, DCL can shift and overlap notably within and across plant species or types (Bush and McInerney, 2013), rendering most vegetation chemotaxonomy attempts using DCL, ACL and various ratios of *n*-alkane combinations problematic. In contrast, since a primary function of leaf wax is to protect plants from excessive water loss through cuticle (Kerstiens, 1996; Riederer and Schreiber, 2001), direct linkages between *n*-alkane proxies and abiotic factors were continuously explored and calibrated with a special emphasis on aridity related factors (e.g. Hoffmann et al., 2013; Bush and McInerney, 2015; Lu et al., 2020).

A few transect studies of leaf wax and/or soil *n*-alkanes along various climate gradients covering arid regions were conducted in recent years. Sachse et al. (2006) found increasing chain length of broadleaf tree leaf wax *n*-alkanes toward warmer and dryer regions along a European vegetation gradient. Similar increasing chain length or ACL with increasing aridity had also been noted along a C_3-C_4 grass (Rommerskirchen et al., 2006), a rainforest-savannah-C_4 grass (Vogts et al., 2009) and a succulent-fynbos (Carr et al., 2014) vegetation gradient in southern Africa, and a Mean Annual air Temperature (MAT = 3–23 °C, with limited Mean Annual Precipitation, MAP = 600–800 mm) gradient in America (Bush and McInerney, 2015). This seemingly consistent relationship was, however, at odds between *n*-alkane ACL of *Acacia* and *Eucalyptus* along a precipitation gradient (MAP = 400–1600 mm) in Australia (Hoffmann et al., 2013), and absent along a precipitation gradient (MAP = 150–2000 mm) in China (Luo et al., 2012). Instead, Luo et al. (2012) found that CPI negatively correlates with relative humidity, precipitation and temperature. Similar negative CPI-MAP correlation can also be observed from results along a humid-semiarid climate transect (Rao et al., 2009; MAP = 380–1900 mm) and an arid-semihumid transect (Lu et al., 2020; MAP = 150–410 mm). A preliminary study in arid Chinese deserts and gobi deserts (MAP = 30–310 mm), however, revealed possible positive correlations of ACL and CPI with precipitation, relative humidity, temperature, etc. (Yan, 2017), which is similar to the CPI trend of surface sediments of multiple lakes along a 40–470 mm MAP gradient on the Tibetan Plateau (Tian et al., 2017), but in clear contrast with all other previous observations.

Several possibilities can account for the contrasting observations. One is the different sample types that they used, i.e. leaf (Rommerskirchen et al., 2006; Vogts et al., 2009; Hoffmann et al., 2013; Carr et al., 2014) versus soil (Rao et al., 2009; Luo et al., 2012; Yan, 2017; Lu et al., 2020), lake sediment (Tian et al., 2017), or leaf plus soil samples (Bush and McInerney, 2015), which revealed *n*-alkane-abiotic factor correlations on different levels from individual plant species or types to averaged biome that is represented by soil and lake sediment samples. Another important cause is the large scatter of vegetation and climate gradients that these studies covered, which may have revealed incomplete or regionalized patterns that are not suitable for direct comparison. For example, the arid regions investigated by other studies are mostly hot to mild arid zones (temperature of coldest month, $T_c > 10$ °C), while Yan (2017) and Tian et al. (2017) mainly focused on cold arid regions ($T_c < 0$ °C). Luo et al. (2012) and Lu et al. (2020) do have several samples in cold arid regions; their results, however, are heavily weighted by data from the humid-semiarid regions. Additionally, most vegetation and climate gradients surveyed so far are not confined to single-factor gradients (e.g. Luo et al., 2012; Hoffmann et al., 2013; Bush and McInerney, 2015; Yan, 2017), whose results are thus open to interpretation multiplicity. Recognizing the above-mentioned diverse correlations being true in their respective climate settings, we hypothesize that *n*-alkane distribution patterns of plant species and averaged biomes are sensitive to abiotic factors, but they "respond" to these factors differently in the cold arid climate.

To test this hypothesis, we surveyed *n*-alkane patterns of coupled soil and leaves of three dominant plant species along more confined temperature and precipitation gradients within a typical cold arid zone, the Junggar Basin in Xinjiang Province of China. Our aims are: 1) to test if *n*-alkane patterns are constrained by abiotic factors on levels of individual plant species and averaged biome; 2) to assess the validity of the known *n*-alkane proxy-abiotic factor correlations in the cold arid climate; and 3) to assess the applicability of *n*-alkane proxies in paleoclimate reconstruction, CPI in particular in this study, by analysing existing section records.

2. Material and methodology

2.1. Sampling

Junggar Basin is in northwest China. The central flatland of the basin is largely covered by fixed to semi-fixed sand dunes, belonging to the Gurbantunggut desert, which has an average elevation of around 400 m, a typical MAT of 3–7.5 °C and a MAP of 150–250 mm. Long-lasting snow cover in winter makes this basin a typical cold arid zone (Thomas, 2011). Based on Chinese climatological MAT and MAP contour maps (Xu and Zhang, 2017), we conducted a sampling campaign in 2018 along a MAP-confined MAT gradient (2–7.5 °C, MAP = 134–148.1 mm), and a MAT-confined MAP gradient (134.5–280.9 mm, MAT = 4.8–8.1 °C). Additional samples along an altitude gradient (~100 m apart in between) and from typical salt-crusted soil were also collected to assess possible joint influence of MAT and MAP, which generally covary along altitude gradients, and the influence of salinity, respectively. Sampling sites were distributed along the gradients at intervals of around 50–100 km apart, with additional ones more scattered (Fig. 1, Supplementary Table S1). Leaf (including assimilative shoot) samples of plant species dominant at individual sites were collected from more than five individual plants into paper bags and air-dried. Note that by "dominant" we mean it as more commonly occurring plants by visual assessment in the field. Samples of the top 5 cm of soil were collected from under more than five plants using a spatula, wrapped in ashed aluminium foil and stored in plastic bags. After collection, all samples were shipped back immediately and lyophilized before further processing. Aiming at intra-species comparison, only results of three most commonly occurring plant species were analyzed and presented here, including *Ephedra distachya*, *Haloxylon ammodendron*, and *Artemisia songarica*. Only samples deep in desert and gobi desert areas, where human influence is minimized, were analyzed and reported, including 20 soil, 8 *Artemisia*, 5 *Ephedra*, and 14 *Haloxylon* samples (please refer to Table S1 for their distribution along the gradients, and Appendix Fig. A1 for pictures of a typical sampling site and three plant species).

2.2. Sample processing and measurements

Soil water content (SWC) was calculated from weights of the samples before and after lyophilization following the equation:

$$SWC = (W_f - W_d)/W_d * 100\%$$

where W_f is the weight of fresh samples, W_d is the weight of dried samples.

Soil salinity and pH were measured from soil solution extracts (1:2.5 g/ml for pH and 1:1 g/ml for salinity) using a HACH HQ40D Portable Multi Meter. For Total Organic Carbon content (TOC) measurements, 1 g of samples was ground, passed through a 100 mesh sieve, reacted with 2 N hydrogen chloride (HCl) overnight, washed using deionised water,

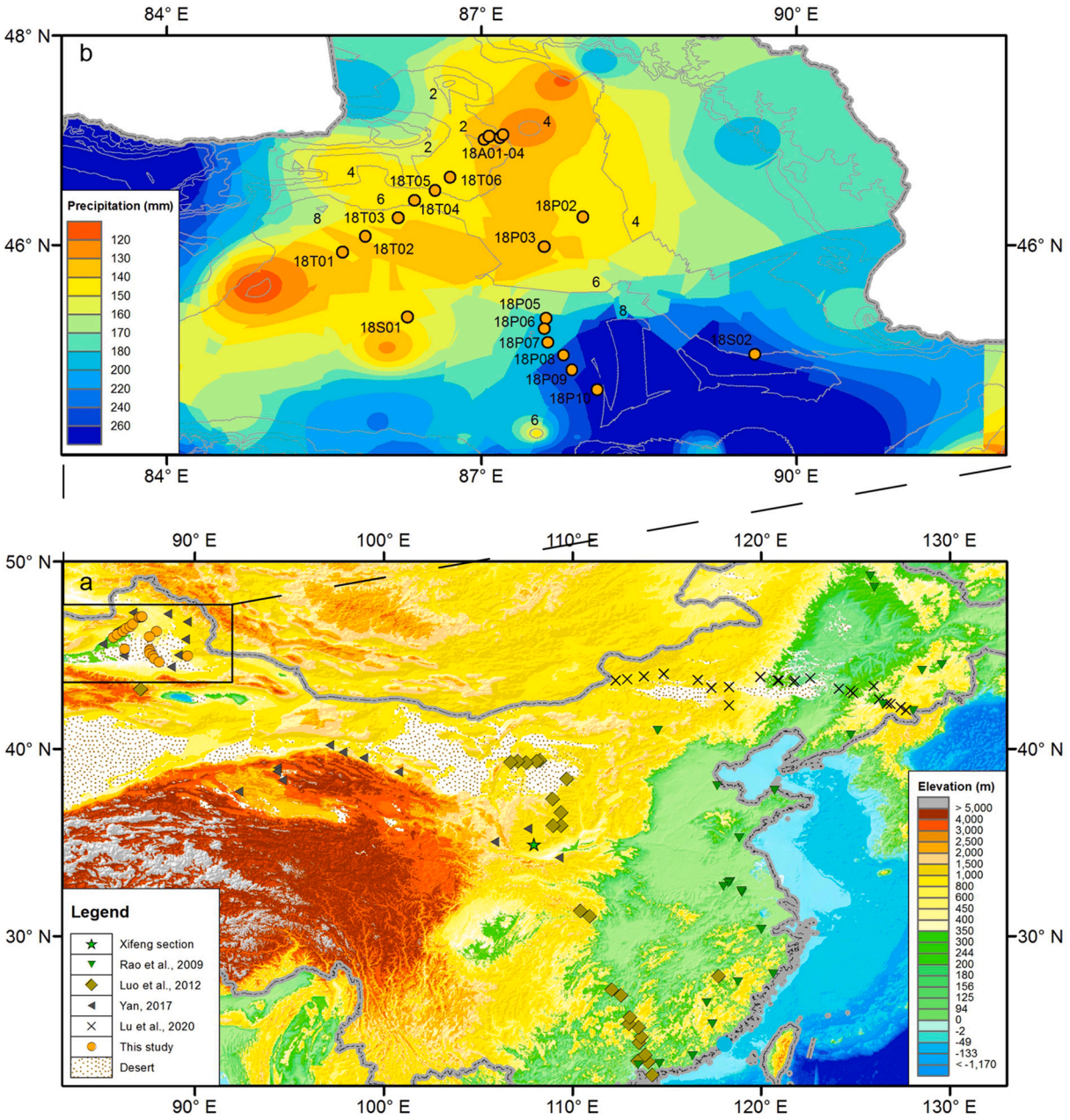

Fig. 1. a) Sampling site distribution of this study and studies for comparison. Distribution of deserts comes from the 1:200,000 desert distribution dataset provided by Environmental and Ecological Science Data Center for West China, National Natural Science Foundation of China (http://westdc.westgis.ac.cn). b) Sampling sites of this study with climatological MAT and MAP contours of the Junggar Basin. Gray lines and accompanying numbers signify the MAT contours with values in degree Celsius, while the MAP contours are the colour-filled ones corresponding to the legend.

oven-dried and re-ground, and then 30–100 mg of subsamples were wrapped in tin boats and analyzed using an Elementar vario PYRO cube elemental analyzer (Xie et al., 2016).

Depending on TOC, ca. 10–20 g of ground soil was extracted in dichloromethane (DCM): methanol (MeOH) (9:1 *v*/v) using a Microwave Accelerated Reaction System (CEM Corporation) with a ramp to 100 °C over 20 min and held for 20 min. The extracts were transferred into 60 ml vials, and combined with repeated solvent washes (repeated at least three times till clear), which were then reduced on a heating plate (38 °C) under nitrogen flow. Non-polar fractions bearing *n*-alkanes were separated by passing the concentrated extracts through LC-NH_2 silica gel columns (1 g, SUPELCO) with 4 ml *n*-hexane, reduced, and transferred into 2 ml vials. For leaves, 0.6–0.8 g of minced samples, spiked with 100 μl androstane (0.088 μg/μl), were extracted ultrasonically for 15 min for at least three times (repeated till the extracts were visually clear). The extracts were similarly combined and reduced. Their non-polar fractions were separated by passing the concentrated extracts through activated silica gel columns (100 mesh, 20*1 cm) with *n*-hexane, and again reduced and transferred into 2 ml vials.

Measurements of soil *n*-alkanes were done using a Hewlett-Packard 9860 gas chromatograph, while those of leaf samples were done using an Agilent 6890 N gas chromatograph. Both extraction and measurement of soil samples were conducted in the Fye Lab, Marine Chemistry & Geochemistry Department, Woods Hole Oceanographic Institution, while those of leaf samples were conducted in the State Key Laboratory of Organic Geochemistry, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences. Quantifications of soil *n*-alkanes were done by comparing with a lab standard containing standard C_{17}, C_{25}, C_{29} and C_{33}

homologues at different concentrations (uncertainty estimated by repeated measurements is less than 5%), while those of leaf *n*-alkanes were done by comparing with the internal androstane standard. Range of identified *n*-alkanes is generally from C_{21} to C_{33} in leaf samples and from C_{16} to C_{35} in soil samples (Table S2, S3). To facilitate comparison, *n*-alkanes of C_{21} to C_{33} were mostly used in the proxy calculations and discussion. ACL and CPI were calculated differently in various literatures (e.g. Luo et al., 2012; Bush and McInerney, 2013; Carr et al., 2014). Here they are calculated as follows:

$$ACL = (21*C_{21} + 23*C_{23} + \ldots + 33*C_{33})/(C_{21} + C_{23} + \ldots + C_{33})$$

$$CPI = ((C_{21} + C_{23} + \ldots + C_{31}) + (C_{23} + C_{25} + \ldots + C_{33}))/2*(C_{22} + C_{24} + \ldots + C_{32})$$

where C_x is the abundance of the *n*-alkane homologue with x carbons. In comparison with results from other studies, different calculation formulas are used and will be noted in the following sections when appropriate.

2.3. Meteorological data and correlation analysis

All sampling sites are in remote desert and Gobi Desert regions, where meteorological data are scarce. We interpolated meteorological data at the sampling sites from two data sources, including the climatological MAT, MAP and aridity (note that higher aridity means drier, which is in contrast to the often used aridity index, e.g. Vogts et al., 2012; Hoffmann et al., 2013) from the aforementioned Chinese background climatology datasets (Xu and Zhang, 2017), and the 2018 summer (June, July and August) averages of ground temperature, air temperature, evaporation, relative humidity and precipitation from the daily surface weather profile dataset provided by the China Meteorological Data Service Center (http://data.cma.cn/).

We systematically explored the correlations between individual abiotic factors, including altitude, SWC, salinity, pH, and the above-mentioned meteorological factors, and *n*-alkanes proxies, including abundances of *n*-alkanes, ACL, CPI, and a series of ratios between important *n*-alkane homologues (Table S4, S5). Correlation of *n*-alkane proxies with abiotic factors in a random group of samples could be suggestive of leading constraints of certain factors, while it could also be the result of compounding influence of multiple factors. Correlation of a proxy with an abiotic factor along its single-factor gradient as we intended should therefore provide a verification or confirmation of the correlation.

Our initial sampling gradients serve as a preliminary grouping of the samples. To further explore sample-abiotic factor relationship, we performed a principal component analysis of abiotic factors, which shows that five high salinity sites contribute the most to the total variance and the first component (explained 51% of the variance), while other sites generally contribute more to the second component (explained 23% of the variance) (Fig. 2). It is also evident that temperature-related factors, including MAT, summer air and ground temperature, and precipitation-related factors, including MAP, aridity, summer precipitation, evaporation and relative humidity, have generally good collinearity, while SWC and pH contributed little to the total variance and have low correlations with other factors. Therefore, in addition to the intended MAT (gradient T1, Gr_{T1}, $n = 6$), MAP (Gr_{P1}, $n = 8$), and altitude gradients (Gr_A, $n = 4$), we grouped the five high salinity samples into a salinity gradient (Table S1). We also reassigned our samples to alternative temperature (Gr_{T2}) and precipitation (Gr_{P2}) gradients by maximizing the ranges of the intended factors while minimizing those of other factors (Table S1). Gr_{T2} includes more samples ($n = 9$) along a 2–7 °C MAT gradient with a much narrower salinity range (0.52–1.73 PSU versus 0.53–5.54 PSU along Gr_{T1}) and a similar MAP range (130.0–148.1 mm). Gr_{P2} ($n = 7$) has a slightly narrower MAP range (142.2–270.5 mm), and far narrower MAT (6.9–7.9 °C), salinity (0.37–0.71 PSU), and even altitude (521–542 m) ranges.

Although single-factor gradients are still not perfectly achieved, given the limitations of available variances of the abiotic factors, we believe that our nested correlation analyses should be able to provide a clear view of the proxy-factor correlations. Generally, only significant correlations ($r^2 > 0.3$, $p < 0.05$) are discussed. For some groups having only 4–5 data points, we only note a significant correlation, if the data also form a monotonic sequence. Because of the high collinearity among abiotic factors, when an *n*-alkane proxy was found to correlate with multiple abiotic factors in a sample subset, a forward-selection multivariate linear regression was applied on standardized *n*-alkane proxy and all abiotic factors to elucidate the dominance of certain factors over the proxy variations, indicated by the highest standard regression coefficients (Legendre and Legendre, 2012).

3. Results

3.1. n-alkane abundances and distribution patterns

Soil *n*-alkanes (C_{16} - C_{35}) have a total abundance ranging from 0.29–1.38 μg/g dwt (dry weight) with an average and standard deviation of 0.80 ± 0.32 ($n = 20$) μg/g dwt, among which C_{16}-C_{20} *n*-alkanes account for 9–36%. The long-chain *n*-alkanes are mostly dominated by C_{29} or C_{31}, with only one exception by C_{27} (Table S2). Abundances and distribution patterns of leaf *n*-alkanes (C_{21} - C_{33}) are obviously dissimilar among different plant species and from soil (Fig. A2). *Artemisia* has the largest *n*-alkane total abundance, ranging from 505 to 2671 μg/g dwt with an average and standard deviation of 1434 ± 633 ($n = 8$) μg/g dwt, while those of *Ephedra* are 56–111 and 80 ± 21 ($n = 5$) μg/g dwt, and those of *Haloxylon* are 42–521 and 150 ± 139 ($n = 14$) μg/g dwt, respectively. Dominant *n*-alkane homologue is exclusively C_{29} for *Artemisia*, and predominantly C_{27} for *Ephedra* and *Haloxylon* (Table S3). Alternatively, there are cases that both *Haloxylon* and *Ephedra* have C_{29} as the dominant *n*-alkane (Table S3). More surprisingly, *Ephedra* also has

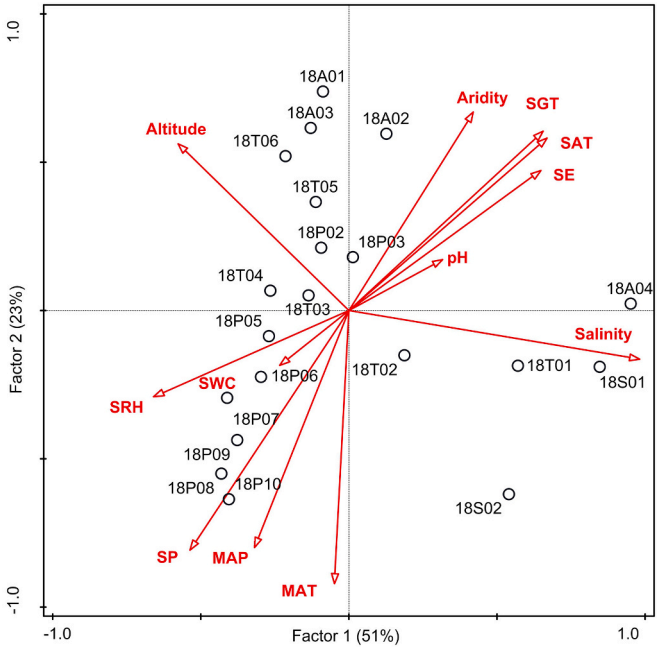

Fig. 2. Abiotic-factor loadings (red arrows) and sample scores (black circles) of the first and second principal components from PCA of abiotic factors at all sampling sites. Additional abbreviations are: SGT - Summer Ground Temperature; SAT - Summer Air Temperature; SE - Summer Evaporation; SP - Summer Precipitation; SRH - Summer Relative Humidity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

C_{21} as the dominant *n*-alkane at one site, and has a generally high abundance of C_{21} throughout (Fig. A2c, Table S3). The range, average and standard deviation of ACL and CPI of soil samples are 28.2–29.8, 29.1 ± 0.4 and 3.0–7.5, 4.6 ± 1.2, respectively, while those are 29.0–29.8, 29.5 ± 0.3 and 9.3–16.3, 13.7 ± 2.1 for *Artemisia*, 25.1–26.8, 26.3 ± 0.6 and 3.3–5.1, 4.3 ± 0.6 for *Ephedra*, and 26.6–28.2, 27.3 ± 0.5 and 3.2–13.9, 9.6 ± 3.3 for *Haloxylon*, respectively (Fig. 3, Table S2, S3).

3.2. *n*-alkane proxies-abiotic factor correlations

Putting data of all soil samples together, we observed no significant correlation between any abiotic factor and *n*-alkane proxy, except for between ACL and salinity (negative correlation, $r^2 = 0.31$; Fig. 4a, Table S4). A close examination found that this correlation is heavily influenced by the monotonic linear trend of the salinity gradient samples ($r^2 = 0.98$; Fig. 4a). After excluding the five high salinity samples, we observed significant strong correlations between CPI and all precipitation-related factors ($r^2 = 0.36$–0.59, positive correlations except for with aridity and evaporation; Table S4). Significant negative correlation between CPI and temperature-related factors, including summer air and ground temperature (both with $r^2 = 0.51$) also exist, while no correlation was found between *n*-alkane proxies and altitude, SWC, salinity or pH. Multivariate linear regression of CPI on all abiotic factors removed all factors but the summer evaporation, suggesting its dominance over the CPI variations (Table S6).

Along Gr_{T1}, Gr_{T2} and Gr_A, we found no correlation between temperature or altitude and the soil *n*-alkane proxies, except for between C_{31}/C_{29} and summer air and ground temperature along Gr_{T1} and Gr_{T2} (both with $r^2 = 0.58$, 0.46, respectively; Table S4). Besides the constraint of salinity on ACL mentioned above, we found no other *n*-alkane proxies having monotonic trends along Gr_S (Table S4). Along Gr_{P1} and Gr_{P2}, correlations between CPI and temperature- and precipitation-related factors persist with higher r^2 values than those of the low-salinity samples, which is again CPI showing generally higher correlation with precipitation related indices ($r^2 = 0.54$–0.64 and 0.67–0.71 for Gr_{P1} and Gr_{P2}, respectively) than with temperature related ones ($r^2 = 0.51$–0.60 and 0.62–0.69, respectively) (Table S4). Multivariate regression of CPI on all factors suggests that summer relative humidity is the sole factor chosen to predict CPI variations along Gr_{P1}, while summer evaporation and salinity both have important contributions to CPI variations along Gr_{P2}, with summer evaporation as the leading constraint (Table S6, Fig. A3a,b). Additional correlation can be found between *n*-alkane proxies and abiotic factors other than the intended factors along their respective gradients, such as between ACL and altitude along the precipitation gradient Gr_{P2} (Table S4). They are considered as by-products of the close coupling of abiotic factors in this region, and will not be discussed further.

Correlations of plant samples are diverse for different species, and quite different from those of soil. Only three plant samples (*Haloxylon* samples) are available along Gr_S (Table S1), from which no reliable observations of *n*-alkane proxy-abiotic factor correlations along the salinity gradient can be made. All their *n*-alkane proxies do not appear as outliers on box-and-whisker plots (not shown). For all samples of single-species groups, no meaningful correlations between *n*-alkane proxies and salinity, pH, or SWC were found, except between pH and C_{27}/C_{21} in the *Ephedra* data (Table S5). For the *Ephedra* dataset, no other meaningful correlations were found.

For *Artemisia*, ACL, C_{31}/C_{29} and C_{31}/C_{27} strongly correlate with several precipitation- (r^2 = 0.59–0.69, 0.66–0.78 and 0.68–0.79, respectively) and temperature-related factors ($r^2 = 0.65$, 0.69–0.80, and 0.62–0.75, respectively; Table S5). A strong negative correlation also exists between altitude and C_{31}/C_{29} ($r^2 = 0.80$). Multivariate regression identified summer evaporation as the leading constraint on ACL and C_{31}/C_{27} changes (exemplified by the ACL-summer evaporation correlation in Fig. 4c), while both MAT and altitude are important constraints of C_{31}/C_{29} variations, with MAT being the leading one (Table S6).

Based on *Haloxylon* data, CPI and C_{27}/C_{21} have weaker correlations with MAT (r^2 = 0.47 and 0.34, respectively) than with precipitation-related factors ($r^2 = 0.54$–0.64 for CPI with MAP and summer precipitation, and 0.32–0.42 for C_{27}/C_{21} with MAP, summer precipitation and evaporation, respectively; Table S5). MAP and ACL, and altitude and C_{31}/C_{29} are also weakly correlated (r^2 = 0.37 and 0.47, respectively; Table S5). Multivariate regression suggests that MAP and summer relative humidity are the dominant constraints of *Haloxylon* CPI and C_{27}/C_{21}, respectively (Table S6). The correlation between CPI and MAP is, however, in strikingly contrary signs to that of soil (Fig. 4d). Note that due to sampling deficiency of *Haloxylon* at sites of higher MAP (Table S1), the CPI-MAP scatter plot shows polarized clustering at the high and low MAP ends. We think the highest r^2 between the two might be an artifact (Table S5). However, the consistent negative correlation between CPI and summer precipitation (the leading constraint in a multivariate regression of CPI on all factors without MAP; Fig. A3c) with no obvious clustering bolstered our confidence in the reliability of the CPI-MAP correlation.

Temperature and precipitation gradients are only possible for the *Haloxylon* dataset. Along the temperature gradient (Gr_{T2}, $n = 8$), CPI shows a significant correlation with MAT ($r^2 = 0.63$) (Table S5). Along the precipitation gradient (all available *Haloxylon* data from Gr_{P1} and Gr_{P2}, $n = 6$), strong correlations are found between precipitation-related factors and all *n*-alkane proxies, except for C_{31}/C_{27} (r^2 = 0.67–0.96), although most of the high r^2 values are due to data clustering (Table S5). The correlation between CPI and summer precipitation becomes insignificant ($p = 0.09$), while that between CPI and MAP becomes slightly stronger ($r^2 = 0.67$) with less clustering (Fig. 4d, A3d). Because of the low sample number and data clustering effect, we did not apply multivariate regression on this sample subset.

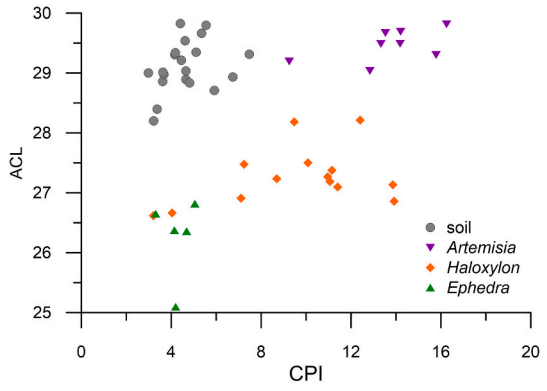

Fig. 3. Scatter plot of ACL versus CPI of soil and leaf samples.

4. Discussion

4.1. Dissimilarities between soil and plant n-alkane patterns

The distribution patterns of plant and soil *n*-alkanes show prominent dissimilarities. The dominant *n*-alkane homologue is mostly C_{29} for the three plant species, while it is C_{31} for almost half of soil samples. ACL of soil *n*-alkanes is generally equivalent to that of *Artemisia*, but systematically higher than that of *Ephedra* and *Haloxylon* (Fig. 3). CPI of soil *n*-alkanes is in contrast systematically lower than that of *Artemisia* and *Haloxylon*, but equivalent to that of *Ephedra* (Fig. 3). Soil *n*-alkanes should represent an average *n*-alkane pattern of the regional biome, but soil samples cluster aside from the linear combination of the three types of plant samples in the ACL-CPI scatter plot (Fig. 3), suggesting soil *n*-alkane patterns cannot be obtained through mixing of those of the three plant species. Under a general assumption that long-chain *n*-alkanes come primarily from higher plants, this suggests that the three "dominant" species do not dominate soil *n*-alkane patterns, and there are more

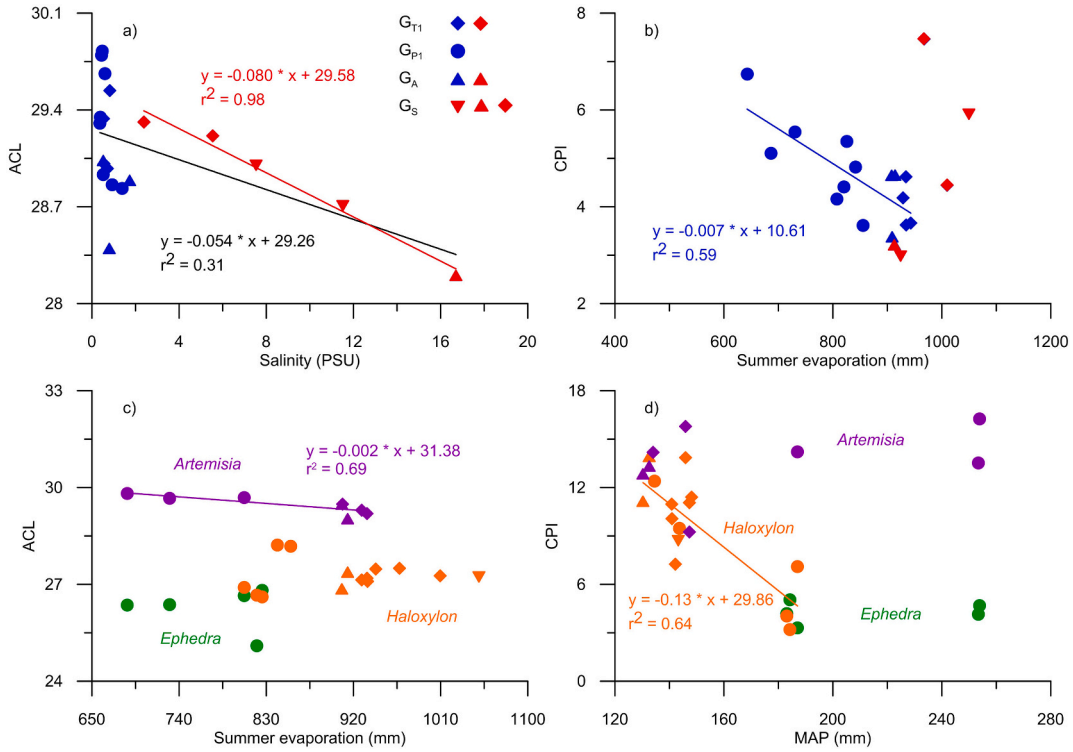

Fig. 4. Representative *n*-alkane proxy - abiotic factor correlations ($p < 0.05$). a) and b) present soil data, while c) and d) present leaf data. Different symbols signify samples from different gradients. The salinity gradient has samples from different gradients, and is thus further differentiated by the red colour with soil data. The same symbols are adopted with additional colour schemes to differentiate data from different plant species as labelled. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

local plant species contributing *n*-alkanes with higher mass production, a longer-chain *n*-alkane dominance, and weaker odd-over-even predominance. If this assumption does not hold, there might be notable long-chain *n*-alkane input from non-photosynthetic bacteria in soil, which has no obvious odd-over-even preference (Han and Calvin, 1969; Albro, 1976; Ladygina et al., 2006). Microbial differential degradation of *n*-alkanes might have played an important role in altering soil *n*-alkane patterns as well (Johnson and Calder, 1973; Walker and Cooney, 1973; Grimalt et al., 1988; Marseille et al., 1990; Buggle et al., 2010). It is also very likely that some or all of these possible scenarios work in consortium. At present, we cannot make further discrimination on these interesting yet unclear possible scenarios. Note that the dissimilarities imply that *n*-alkane patterns are probably constrained through different processes at levels of plant species and averaged biome, as evidenced by the following discussion.

4.2. Constraints of soil and plant n-alkane patterns in cold arid climate

The total abundances of soil and leaf *n*-alkanes (0.29–1.38 and 42–2671 μg/g dwt, respectively) in cold arid climate generally fall in the ranges of those reported in humid to arid regions (e.g. Boom et al., 2014; Bush and McInerney, 2015). The leaf *n*-alkane abundance ranges of all three plant species are within that of leaf samples along the aforementioned MAT and MAP transects in America (61–7115 μg/g dwt, Bush and McInerney, 2015), South Africa (20–22,000 μg/g dwt, Boom et al., 2014), and Australia (10–2818 μg/g dwt, Hoffmann et al., 2013). Similarly, the soil *n*-alkane abundance range falls in that of our previous study in the same region (0.13–4.49 μg/g dwt, Yan, 2017), and also has only two samples with lower abundance than that of the MAT transect in the semihumid America (0.40–27.50 μg/g dwt, Bush and McInerney, 2015). This suggests that in spite of obvious interspecies contrasts, *n*-alkane abundances cannot be used to differentiate plant species or biomes adapted to the cold arid climate and more humid or hot arid climate. It is possible that the majority of plant species in certain climate have systematically high or low *n*-alkane abundances (Carr et al., 2014). Complete survey of leaf *n*-alkane patterns of all plant species at sites in different climate zones could confirm or refute this, which is, however, nearly impossible to implement.

Correlations between *n*-alkane proxies and abiotic factors in the cold arid climate, on the other hand, show prominent contrast with the reported ones in other climate regions, and therefore suggest different leading constraints of abiotic factors on *n*-alkanes in different climatic and/or environmental settings. Although small in sample number, the linear monotonic trend of soil ACL along the salinity gradient strongly supports a salinity control over soil *n*-alkane patterns in high salinity settings (Fig. 4a). Salinity's influence on leaf *n*-alkane patterns is unclear, due to the scarcity of plant samples along the salinity gradient. This negative control of salinity on soil ACL is by the authors' knowledge not reported. It could possibly be affected by three processes. First, degradation of *n*-alkanes by halophilic and halotolerant microorganisms have been well reported by studies of petroleum contamination remediation, with multiple strains identified, isolated and cultured for more in-depth researches (Fathepure, 2014 and the references therein). Especially, differential degradation of *n*-alkanes of various carbon lengths by differing degraders, and the salinity dependence of the degraders were clearly demonstrated in controlled lab experiments (e.g. Ward and Brock, 1978; Al-Mueini et al., 2007; Dastgheib et al., 2011; Al-Mailem et al., 2013). Second, microbes are known to produce a spectrum of *n*-alkane homologues typically with no obvious odd-over-even carbon number predominance (e.g. Blumer et al., 1971; Kolattukudy, 1976; Grimalt et al., 1988; Volkman et al., 1998; Ladygina et al., 2006). Third, vegetation changes might also change the distribution pattern of plant *n*-alkane input to soil. Contribution change of each of the three processes could significantly change distribution patterns of soil *n*-alkane abundances, resulting in ACL and CPI variability. Future studies are needed to further constrain the underlying mechanisms of salinity's influence on

leaf and soil *n*-alkane patterns in these settings.

Consistent and strong correlation between precipitation-related factors and soil CPI for samples in low salinity settings and along different precipitation gradients reveals a strong constraint of water stress on soil *n*-alkanes without the influence of salinity. Yan (2017) noted a positive correlation between soil CPI and relative humidity (RH) in the same region on a smaller gradient (RH = 41–47%, $r^2 = 0.74$, $p < 0.01$), but no significant correlation between CPI and other precipitation-related factors were found, possibly due to interfering factors. CPI of surface sediments of 21 lakes on the Tibetan Plateau also shows a significant positive correlation ($r^2 = 0.37$, $p < 0.01$) with MAP (40–470 mm), although it was not noted in the original work (Tian et al., 2017). In comparison, our results confirmed the validity of the positive correlation between CPI and MAP or RH (on a larger gradient, RH = 44–56%) in the cold arid region, and revealed the pervasive correlations between CPI and precipitation-related factors. One thing worth noting is that the dominant constraints of CPI variations in different sample groups are often the summer evaporation or relative humidity (Section 3.2, Table S4, S5, S6). Considering also the lack of correlation between SWC and other precipitation-related factors (Fig. 2), we argue that this emphasizes the sensitivity to available water and complexity of water-availability for plants in arid regions, where precipitation is readily drained and relocated within hours to 1–2 days instead of being more retained by local soil.

The positive soil CPI-MAP correlation in the cold arid region is in striking contrast to the negative correlation between the two along three humid-semiarid/arid transects in China (Rao et al., 2009; Luo et al., 2012; Lu et al., 2020). If we take the maximum MAP (310 mm) of our sampling sites in the cold arid region as the division between arid and semiarid climates, only 5 samples in Luo et al. (2012) and 9 in Lu et al. (2020) can be classified as arid-region samples, which does not support a negative correlation (Fig. A4, Table S7). The conventional MAP threshold of 200 mm to divide the two climates means even fewer samples of arid regions in each study, which does not lend support to a negative correlation as well (Fig. A4). The contrasting CPI-MAP correlations are thus attributed to differing climatic and environmental settings. While Rao et al. (2009) did not explore the CPI-MAP correlation, both Luo et al. (2012) and Lu et al. (2020) attributed the negative correlation between the two to intensified microbial degradation and/or microbial contribution of long-chain *n*-alkanes in wetter climate, which evidently does not apply here. Soil CPI generally varies within the range of plant CPI suggesting a possible direct inheritance of the correlation from plants (Fig. 3). But, the constraint of precipitation on plant *n*-alkane patterns is species-dependent. The only CPI-MAP correlation was found to be negative with *Haloxylon* (Table S5, S6), which rules out the possibility of a dominant plant governing soil *n*-alkane patterns. We argue that the change of local biome, including plant and microbe communities that are very sensitive to precipitation changes in arid climate, might be the cause for the positive correlation. Varying plant community could contribute a differing mixture of *n*-alkanes from a changing configuration of plant species and their percentages in the local community, and plant physiology itself. Varying microbe community can modulate the intensity of microbial degradation and/or microbial contribution of long-chain *n*-alkanes. It would be intriguing to unveil the different physiological mechanisms governing these processes, which will be a future topic.

Based on our dataset, correlations of *n*-alkane proxies with temperature-related factors are generally absent or weaker than those with precipitation-related ones. Especially, the correlation often reported between ACL and temperature (e.g. Sachse et al., 2006; Tipple and Pagani, 2013; Bush and McInerney, 2015) is absent in soil as well as plant samples. This pattern rejects temperature being a primary constraint on soil and plant *n*-alkane patterns in this climate setting. Because our samples only spanned a very limited MAT gradient (2–8 °C), it is also likely that the influence of temperature might work on a different scale. Especially, the correlations between temperature-related factors or altitude and plant *n*-alkane proxies, such as CPI, C_{31}/C_{29}, and C_{27}/C_{21}, suggest that temperature may still have influence on plant *n*-alkane patterns with special emphasis on certain homologues (Table S5). This, however, needs confirmation from future studies.

Therefore, we think our dataset revealed the sensitivity of *n*-alkane patterns of soil and plants to various abiotic factors in different climate and environment settings, and further validated previous inferences that abiotic factors constrain *n*-alkane patterns of soil and plants independently (Rao et al., 2009; Luo et al., 2012).

4.3. Soil CPI trend reversal along the precipitation gradient

The contrasting CPI-MAP correlations in the cold arid climate and other climate regions led us to wonder if precipitation acts as a first-order constraint of soil *n*-alkane CPI in different climate regions, although with different underlying mechanisms at work. We assess this by compiling available soil *n*-alkane data across China to observe CPI changes along the precipitation gradient. In particular, the compiled data include those from three transect studies in east China covering humid to arid climates (Rao et al., 2009; Luo et al., 2012; Lu et al., 2020), and two in the cold arid region in northwest China (Yan, 2017 and this study; please refer to Fig. 1a for the sample distribution). Note specifically that we did not exclude any samples from these datasets, which means that all interfering abiotic factors are at work on the ranges of soil CPI, including salinity, temperature and vegetation. Corresponding MAP, MAT and aridity data at all the sampling sites were obtained from the same datasets used in this study. Because CPI was calculated differently in these studies, sometimes without providing original *n*-alkane abundance data, we applied the same calculation formulas for $CPI_{25–33}$ as in Luo et al. (2012) to all datasets when possible (Table S7). We had applied different formulas to existing soil *n*-alkane data, which resulted in very strong positive correlation (r^2 is generally higher than 0.9) between different CPIs (not shown). CPIs calculated from different formulas can thus be used interchangeably in respect of their potential climatic or environmental implications.

Plotting $CPI_{25–33}$ against MAP, MAT and aridity, we observed no meaningful distribution pattern for the $CPI_{25–33}$-MAT and $CPI_{25–33}$-aridity pairs. An interesting trend reversal of $CPI_{25–33}$, however, clearly showed up along the MAP gradient, where $CPI_{25–33}$ increases with incrementing MAP from arid to semiarid (without discrimination between semiarid and semihumid climates, semiarid is used

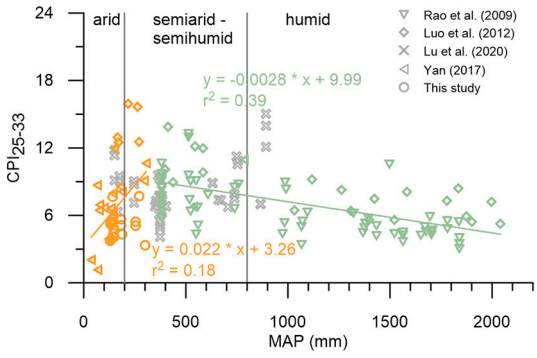

Fig. 5. Soil CPI trend reversal along the precipitation gradient. Arid, semiarid-semihumid and humid climates along the gradient are labelled according to conventional precipitation divisions as a reference, i.e. < 200, 200–800 and > 800 mm, respectively. To include all data collected from the cold arid region, a MAP of 310 mm is adopted to divide the compiled data into the arid (orange symbols) and humid (green symbols) groups. $CPI_{25–33}$ is calculated following Luo et al. (2012) as $((C_{25} + C_{27} + \ldots + C_{33})/(C_{24} + C_{26} + \ldots + C_{32}) + (C_{25} + C_{27} + \ldots + C_{33})/(C_{26} + C_{28} + \ldots + C_{34}))/2$. Data from Lu et al. (2020) is $CPI_{23–33}$, calculated as $((C_{23} + C_{25} + \ldots + C_{31}) + (C_{25} + C_{27} + \ldots + C_{33}))/2^*(C_{24} + C_{26} + \ldots + C_{32})$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

interchangeably with semiarid-semihumid for brevity) climate and then decreases progressively toward humid climate (Fig. 5). We divided the compiled data into arid (arid-semiarid) and humid (semiarid-humid) data groups by choosing MAP of 310 mm as the cut-off to include all soil data from the cold arid region (Table S7). Weak but significant correlation can be observed for both groups (Fig. 5), with correlation coefficient being 0.42 ($p < 0.01$, $n = 50$) for the arid group and − 0.62 ($p < 0.01$, $n = 117$) for the humid group. A MAP of 200 mm as the threshold lowers the correlation strength of both groups without changing the CPI trends (Fig. A5). CPI from Lu et al. (2020) is $CPI_{23–33}$, which is usually within 1 unit away from $CPI_{25–33}$, but can also deviate from it irregularly. We plot them in the background without involving them in correlation analysis, which generally fall within reasonable ranges of $CPI_{25–33}$ in respective climate settings along the MAP gradient, and do not change the CPI trends.

On one hand, the weak but significant correlations imply that precipitation or water stress should be the first-order constraint of soil *n*-alkane distribution patterns along the hot-humid to cold-arid climate gradient in China. The weakness of the correlation is attributable to the sensitivity of leaf wax and soil *n*-alkane distribution patterns to different local factors revealed in Section 4.2. Whether this pattern might hold true along other climate gradients, or in areas with totally different biomes, still needs future studies to elucidate. On the other hand, even though large scatter of CPI of both leaf and soil samples was observed almost everywhere along the precipitation gradient, differing CPI-abiotic factor correlations (with temperature, salinity, relative humidity, etc.) were still confirmed by studies along various abiotic factor gradients (e.g. Bush and McInerney, 2013, 2015; Luo et al., 2012; Yan, 2017; Lu et al., 2020; this study). This suggests that the sensitivity of *n*-alkane patterns to local abiotic factors can often lead to the influence of local factors overriding that of regional factors. We look forward to studies on plant physiology, and microbial *n*-alkane production and degradation in future to reveal the underlying mechanisms.

4.4. Potential of paleoclimate reconstruction using CPI

The significant CPI-MAP correlation provides a possibility to interpret CPI records in terms of paleoclimate reconstruction. If local climate went from hot-humid to cold-arid or vice versa, we should be able to see corresponding CPI changes, possibly modulated by variations of microbial degradation intensity when climate was humid and changes of local biome when climate was arid in response to precipitation changes. To empirically assess this hypothesis, we employed a CPI record of the Xifeng loess-paleosol section on the Chinese Loess Plateau (Fig. 1a, Liu and Huang, 2005). The semiarid-semihumid Chinese Loess Plateau is very sensitive to climate changes in the geological past; therefore CPI trend reversal in local records should be prominent and frequent. In their study, both $CPI_{23–31}$ and δD of leaf wax *n*-alkanes were reported, while only the paleoclimate information encrypted by *n*-alkane δD was presented, and $CPI_{23–31}$ was provided just to verify that the long-chain *n*-alkanes were from higher plants. Representative of all δD of long-chain *n*-alkanes, δD_{C31} is generally higher in loess layers (L_1LL_1, L_1LL_2) corresponding to relatively colder global climate in MIS (Marine Isotope Stage) 2 and 4 that are evidenced by benthic $\delta^{18}O$ (Lisiecki and Raymo, 2005), than in paleosol layers (S_0, L_1LS_1, S_1) when global climate was warmer in MIS 1, 3, and 5 (Fig. 6). Liu and Huang (2005) suggested that this pattern does not support temperature-effect dominance over precipitation δD, i.e. higher δD during warmer periods and vice versa. They then proposed the high δD of *n*-alkanes resulted from evaporation of soil water and evapotranspiration of leaf water during dry periods. This proposition is strongly supported by a systematic study of δD of soil water, soil *n*-alkanes, and leaf *n*-alkanes along a precipitation gradient on the Chinese Loess Plateau (Liu et al., 2019). This evaporation or aridity effect alone, however, cannot explain the whole record (Liu and Huang, 2005), e.g. δD_{C31} was higher in the beginning of S_1 or MIS 5 than L_1LL_2 or MIS 4, when climate is thought to be the warmest or in peak

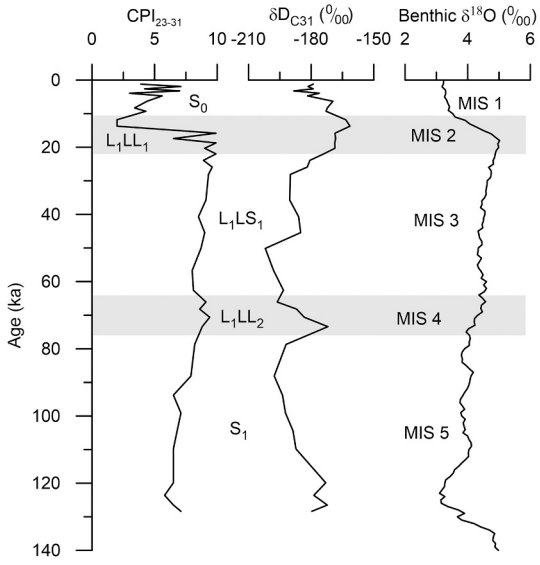

Fig. 6. Time series of $CPI_{23–31}$ and δD_{C31} from the Xifeng section on the Chinese Loess Plateau (Liu and Huang, 2005), and the benthic $\delta^{18}O$ stack (Lisiecki and Raymo, 2005). Shaded layers are L_1LL_1 and L_1LL_2 corresponding to relatively colder global climate in MIS 2 and 4, respectively; unshaded layers are S_0, L_1LS_1 and S_1 corresponding to warmer global climate in MIS 1, 3, and 5, respectively. $CPI_{23–31}$ is calculated as $((C_{23} + C_{25} + \ldots + C_{31})/(C_{22} + C_{24} + \ldots + C_{30}) + (C_{23} + C_{25} + \ldots + C_{31})/(C_{24} + C_{26} + \ldots + C_{32}))/2$.

interglacial conditions. δD_{C31} may thus be interpreted as being affected by multiple factors at different times.

In this context, the $CPI_{23–31}$ record shows a somewhat different trend (Fig. 6), which offers some reasonable insights complementing the interpretation based on δD_{C31} at a few important intervals. In comparison with benthic $\delta^{18}O$, adopting the above interpretation strategy, δD_{C31} variations should have been dominated by temperature-effect prior to MIS 4, when higher temperature (lower $\delta^{18}O$) corresponds to higher δD_{C31} and vice versa, and aridity-effect afterwards, when lower temperature (higher $\delta^{18}O$) corresponds to stronger evaporation and evapotranspiration (higher δD_{C31}) and vice versa. In contrast, $CPI_{23–31}$ shows generally consistent in-phase fluctuations with benthic $\delta^{18}O$ from the beginning of MIS 5 to the Last Glacial Maximum (LGM, 19–21 ka) in MIS 2, and antiphase variations afterwards. Based on our understanding of the CPI trend reversal along the precipitation gradient, this variation pattern suggests local climate went from warmer and more humid to semiarid from MIS 5 to LGM, turned to arid climate at the end of MIS 2, and then became the semiarid-semihumid climate that the Chinese Loess Plateau has now.

Note that systematically higher $CPI_{23–31}$ from MIS 4 to LGM indicates that local climate was essentially semiarid-semihumid, which corresponds to the aridity-effect dominance of δD_{C31}, and thus supports that local climate needs to be sufficiently arid for aridity-effect of δD_{C31} to take over. Note also that $CPI_{23–31}$ peaked at the LGM and jumped to the lowest at the end of MIS 2 corresponding to the δD_{C31} peak. δD_{C31} peak lagging to the LGM was originally interpreted as resulting from combined influence from both temperature and aridity effects (Liu and Huang, 2005). The tight coupling of $CPI_{23–31}$ and δD_{C31}, however, suggests local climate might be the most arid after the LGM. If dating uncertainty of the section is out of question, the lagging between the local climate and global climate extremes might be attributed to local or regional influences. From MIS 2 to MIS 1, local climate turned from arid to semiarid-semihumid. δD_{C31} dropped to its level prior to MIS 2, while $CPI_{23–31}$ increased accordingly without a sudden jump to its peak value representative of a semiarid-semihumid climate. This is a bit baffling. It might be because S_0 was developed under a much warmer and more humid climate than in the LGM, evidenced by the steep decrease of

δD_{C31} and $\delta^{18}O$ from the LGM to MIS 1. Note the similar $CPI_{23–31}$ and $\delta^{18}O$ ranges at the end of MIS 1 and the beginning of MIS 5. The quick transition into this climate in MIS 1 and/or the coarse resolution of the records might have made this slow $CPI_{23–31}$ increase possible. More studies need to be done to corroborate this conjecture.

It is also possible that local climate at certain sites experienced limited fluctuations at the arid or humid climate ends, whereby no trend reversal of CPI would be observed. From the lacustrine deposits of the Qarhan Salt Lake in the Qaidam Basin, Pu et al. (2010) observed higher percentage of woody plant pollens, fewer grass pollens and high CPI during MIS 3. High CPI was attributed to vegetation expansion leading to plant *n*-alkanes outweighing algal and bacterial *n*-alkanes, corroborating a warm-humid climate indicated by the pollen proxies (Pu et al., 2010). In contrast, Cao et al. (2020) observed low CPI corresponding to higher percentages of *Keteleeriaepollenites* and *Podocarpidites* (warm-humid species) pollens in 23.2–21.7 Ma from the fluvial-lacustrine deposits in the Lunpola Basin of Central Tibet. The low CPI was interpreted as resulting from intensified microbial degradation in a more warm-humid climate, again corroborating the pollen records (Cao et al., 2020). Both the high and low CPI being interpreted in favour of more humid climate can be confusing. The CPI trend reversal revealed in this study can help put them into context. Climate at the study site of the Qarhan Salt Lake is in the cold arid region with a MAP of 40 mm now. Although that of the Lunpola Basin is semiarid (MAP = 350 mm) at present, it was well within the humid climate belt in early Miocene (Guo et al., 2008). It is reasonable that the two sites experienced limited climate fluctuations under essentially arid-semiarid and humid-semiarid climates, respectively, which manifested as increasing CPI in the arid region and decreasing CPI in the humid region with elevated precipitation. These inferences again will need future evidence to corroborate.

Our interpretation of the CPI records suggests that CPI might be a valuable paleoclimate proxy. The CPI trend reversal can potentially help differentiate three climate states qualitatively, i.e. arid, semiarid-semihumid and humid climates, although the implicit multiplicity at the low CPI ends means corroborating proxies are always needed for accurate interpretation. As one of the most easily available biomarker proxies, CPI was often reported and interpreted as a biodegradation proxy instead of a paleoclimatic one. We therefore call on scientists in this field to check their old records for possible corroboration or rebuttal of this pattern. Further tests in section records and more regional calibration of this proxy should be warranted.

5. Conclusion

This study surveyed *n*-alkane patterns of soil and leaf wax of three plant species, including *Ephedra distachya*, *Haloxylon ammodendron*, and *Artemisia songarica*, along confined abiotic-factor gradients in the Junggar Basin, a typical cold arid region in northwest China. Our results show that salinity has strong constraint on the ACL of soil *n*-alkanes (negative correlation) in high salinity settings, while precipitation (representative of water availability) controls the CPI of soil *n*-alkanes in low salinity settings (positive correlation). Precipitation also shows strong control over ACL of *Artemisia* (positive correlation), and CPI of *Haloxylon* (negative correlation). These results verified our hypothesis that abiotic factors control *n*-alkane patterns on levels of plant species and averaged biomes, and the *n*-alkane proxy-abiotic factor correlations in cold arid regions are in clear contrast with those reported in other climate regions.

The inconsistency of the *n*-alkane proxy-abiotic factor correlations between soil and plants and among different plant species further revealed that *n*-alkane patterns are very sensitive to, and constrained by abiotic factors independently at levels of individual plant species and averaged biome. Different abiotic factors could thus become leading constraints in various settings, which causes *n*-alkane patterns to respond differently at local to regional scales. The climatic and environmental implications of the well-known *n*-alkane proxies, such as ACL and CPI, might thus vary depending on the discussion context. Especially, the positive soil CPI-MAP correlation in the cold arid region cannot be attributed to microbial degradation that was employed to explain the negative correlation between the two in other climate regions. We argue that the change of local biome, including plant and microbe communities that are very sensitive to precipitation changes in arid climate, might be the cause for the positive correlation.

We further assessed the soil CPI-MAP correlation at regional scales by compiling datasets from previous studies together with our own across the hot-humid to cold-arid regions of China. Despite the sensitivity of *n*-alkane patterns to local factors, our synthesis revealed a striking CPI trend reversal along the precipitation gradient, i.e. CPI increases from arid to semiarid climate and then decreases toward humid climate. Weak but significant CPI-MAP correlations at the arid and humid climate ends suggest that precipitation is the first-order constraint of soil CPI. Further analysis of three existing records showed that the CPI trend reversal pattern can potentially help differentiate three climate states qualitatively, i.e. arid, semiarid-semihumid and humid climates, which could be a valuable complement to existing proxies for paleoclimate reconstructions.

Our study cannot reveal the underlying mechanisms of the *n*-alkane-abiotic factor correlations in soil and plants. Further studies on plant physiology and microbial *n*-alkane production and degradation are warranted. Also, more regional calibration studies of soil and plant *n*-alkane proxies, and section studies of CPI variations in the geological past should be explored to provide corroboration or rebuttal of our findings.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Y. Yan thanks V. Galy for his guidance on lab skills, and for making his lab available for sample processing and analyses during a stay at WHOI, and C. Johnson for his numerous patiently elaborated answers and help during the process. The authors are thankful to Z. Rao for sharing *n*-alkane data from his work. We are also very thankful for the constructive comments from P. Meyers and another anonymous reviewer that have greatly improved this manuscript, and the editorial help from K. Johannesson and D. Kumar. This work was supported by grants from the Natural Science Foundation of Guangdong Province (No. 2018A0303130114), the China Scholarship Council (No. 201804910152), the State Key Laboratory of Loess and Quaternary Geology (No. SKLLQG1839), and the National Natural Science Foundation of China (No. 41673009 and 41977379).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.chemgeo.2021.120402.

References

Albro, P.W., 1976. Bacterial waxes. In: Kolattukudy, P.E. (Ed.), Chemistry and Biochemistry of Natural Waxes. Elsevier, Amsterdam, pp. 419–445.
Al-Mailem, D.M., Eliyas, M., Radwan, S.S., 2013. Oil-bioremediation potential of two hydrocarbon oclastic, diazotrophic Marinobacter strains from hypersaline areas along the Arabian Gulf coasts. Extremophiles 17, 463–470.
Al-Mueini, R., Al-Dalali, M., Al-Amri, I.S., Patzelt, H., 2007. Hydrocarbon degradation at high salinity by a novel extremely halophilic actinomycete. Environ. Chem. 4, 5–7.
Blumer, M., Guillard, R.R.L., Chase, T., 1971. Hydrocarbons of marine phytoplankton. Mar. Biol. 8, 183–189.

Boom, A., Carr, A.S., Chase, B.M., Grimes, H.L., Meadows, M.E., 2014. Leaf wax *n*-alkanes and $\delta^{13}C$ values of cam plants from arid Southwest Africa. Org. Geochem. 67, 99–102.
Bray, E.E., Evans, E.D., 1961. Distribution of n-paraffins as a clue to recognition of source beds. Geochim. Cosmochim. Acta 22 (1), 2–15.
Buggle, B., Wiesenberg, G.L.B., Glaser, B., 2010. Is there a possibility to correct fossil *n*-alkane data for postsedimentary alteration effects? Appl. Geochem. 25, 947–957.
Bush, R.T., McInerney, F.A., 2013. Leaf wax *n*-alkane distributions in and across modern plants: implications for paleoecology and chemotaxonomy. Geochim. Cosmochim. Acta 117, 161–179.
Bush, R.T., McInerney, F.A., 2015. Influence of temperature and C4 abundance on *n*-alkane chain length distributions across the Central USA. Org. Geochem. 79, 65–73.
Cao, M., Sun, J., Liu, W., Hou, J., Tian, Q., Sun, Z., 2020. Paleoclimatic fluctuations inferred from leaf wax *n*-alkane records in Central Tibet in the late Oligocene to early Miocene. Palaeogeogr. Palaeoclimatol. Palaeoecol. 539, 109504.
Carr, A.S., Boom, A., Grimes, H.L., Chase, B.M., Meadows, M.E., Harris, A., 2014. Leaf wax *n*-alkane distributions in arid zone South African flora: environmental controls, chemotaxonomy and palaeoecological implications. Org. Geochem. 67, 72–84.
Castañeda, I.S., Schouten, S., 2011. A review of molecular organic proxies for examining modern and ancient lacustrine environments. Quatern. Sci. Rev. 30, 2851–2891.
Corrigan, D., Kloos, C., O'connor, C.S., Timoney, R.F., 1973. Alkanes from four species of Sphagnum moss. Phytochemistry 12, 213–214.
Dastgheib, S.M.M., Amoozegar, M.A., Khajeh, K., Ventosa, A., 2011. A halotolerant Alcanivorax sp. Strain with potential application in saline soil remediation. Appl. Microbiol. Biotechnol. 90, 305–312.
Diefendorf, A.F., Freimuth, E.J., 2017. Extracting the most from terrestrial plant-derived *n*-alkyl lipids and their carbon isotopes from the sedimentary record: a review. Org. Geochem. 103, 1–21.
Edwards, D., Abbott, G.D., Raven, J.A., 1996. Cuticles of early land plants: A palaeoecophysiological evaluation. In: Kerstiens, G. (Ed.), Plant Cuticles an Integrated Functional Approach. BIOS Scientific Publishers Ltd., Oxford, UK, pp. 1–31.
Eglinton, G., Hamilton, R.J., 1967. Leaf epicuticular waxes. Science 156, 1322–1335.
Eglinton, G., Logan, G.A., 1991. Molecular preservation. Phil. Trans. R. Soc. A 333, 315–328.
Fathepure, B.Z., 2014. Recent studies in microbial degradation of petroleum hydrocarbons in hypersaline environments. Front. Microbiol. 5, 173.
Ficken, K.J., Barber, K.E., Eglinton, G., 1998. Lipid biomarker, $\delta^{13}C$ and plant macrofossil stratigraphy of a Scottish montane peat bog over the last two millennia. Org. Geochem. 28 (3–4), 217–237.
Grimalt, J.O., Torras, E., Albaigés, J., 1988. Bacterial reworking of sedimentary lipids during sample storage. Org. Geochem. 13, 741–746.
Guo, Z., Sun, B., Zhang, Z., Peng, S., Xiao, G., Ge, J., Hao, Q., Qiao, Y., Liang, M., Liu, J., Yin, Q., Wei, J., 2008. A major reorganization of Asian climate by the early Miocene. Clim. Past 4 (3), 153–174.
Han, J., Calvin, M., 1969. Hydrocarbon distribution of algae and bacteria, and microbiological activity in sediments. Proc. Natl. Acad. Sci. U. S. A. 64, 436–443.
Hoffmann, B., Kahmen, A., Cernusak, L.A., Arndt, S.K., Sachse, D., 2013. Abundance and distribution of leaf wax *n*-alkanes in leaves of Acacia and Eucalyptus trees along a strong humidity gradient in northern Australia. Org. Geochem. 62, 62–67.
Jetter, R., Kunst, L., Samuels, A.L., 2006. Composition of plant cuticular waxes. In: Riederer, M., Müller, C. (Eds.), Biology of the Plant Cuticle. Blackwell Publishing, Oxford.
Johnson, R.W., Calder, J.A., 1973. Early diagenesis of fatty acids and hydrocarbons in a salt marsh environment. Geochim. Cosmochim. Acta 37, 1943–1955.
Kerstiens, G., 1996. Plant Cuticles: An Integrated Functional Approach. BIOS Scientific Publishers, Oxford.
Kolattukudy, P.E., 1976. Chemistry and Biochemistry of Natural Waxes. Elsevier Scientific Pub. Co.
Ladygina, N., Dedyukhina, E.G., Vainshtein, M.B., 2006. A review on microbial synthesis of hydrocarbons. Process Biochem. 41, 1001–1014.
Legendre, P., Legendre, L., 2012. Numerical Ecology. Elsevier.
Lisiecki, L., Raymo, M., 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. Paleoceanography 20, PA1003.
Liu, W., Huang, Y., 2005. Compound specific D/H ratios and molecular distributions of higher plant leaf waxes as novel paleoenvironmental indicators in the Chinese Loess Plateau. Org. Geochem. 36 (6), 851–860.
Liu, W., Wang, H., Leng, Q., Liu, H., Yang, H., 2019. Hydrogen isotopic compositions along a precipitation gradient of chinese loess plateau: critical roles of precipitation/evaporation and vegetation change as controls for leaf wax δD. Chem. Geol. 528, 119278.
Lu, J., Zang, J., Meyers, P., Huang, X., Qiu, P., Yu, X., Yang, H., Xie, S., 2020. Surface soil *n*-alkane molecular and δD distributions along a precipitation transect in northeastern China. Org. Geochem. 144, 104015.
Luo, P., Peng, P., Lu, H., Zheng, Z., Wang, X., 2012. Latitudinal variations of CPI values of long-chain *n*-alkanes in surface soils: evidence for CPI as a proxy of aridity. Sci. China Earth Sci. 55 (7), 1134–1146.
Marseille, F., Disnar, J.R., Guillet, B., Noack, Y., 1990. *n*-Alkanes and free fatty acids in humus and A1 horizons of soils under beech, spruce and grass in the Massif-Central (Mont-Lozère), France. Eur. J. Soil Sci. 50, 433–441.
Meyers, P.A., 2003. Applications of organic geochemistry to paleolimnological reconstructions: a summary of examples from the Laurentian Great Lakes. Org. Geochem. 34 (2), 261–289.
Meyers, P.A., Ishiwatari, R., 1993. Lacustrine organic geochemistry—an overview of indicators of organic matter sources and diagenesis in lake sediments. Org. Geochem. 20 (7), 867–900.
Nott, C.J., Xie, S., Avsejs, L.A., Maddy, D., Chambers, F.M., Evershed, R.P., 2000. *n*-Alkane distributions in ombrotrophic mires as indicators of vegetation change related to climatic variation. Org. Geochem. 31 (2), 231–235.
Peters, K.E., Walters, C.C., Moldowan, J.M., 2005. The Biomarker Guide. Cambridge University Press, Cambridge, UK.
Post-Beittenmiller, D., 1996. Biochemistry and molecular biology of wax production in plants. Ann. Rev. Plant Physiol. Plant Mol. Biol. 47, 405–430.
Poynter, J.G., Farrimond, P., Robinson, N., Eglinton, G., 1989. Aeolian-derived higher plant lipids in the marine sedimentary record: Links with palaeoclimate. Paleoclimatology and paleometeorology: modern and past patterns of global atmospheric transport. Springer, Dordrecht, pp. 435–462.
Pu, Y., Zhang, H., Lei, G., Chang, F., Yang, M., Zhang, W., Lei, Y., Yang, L., Pang, Y., 2010. Climate variability recorded by *n*-alkanes of paleolake sediment in Qaidam Basin on the northeast Tibetan Plateau in late MIS3. Sci. China Earth Sci. 53 (6), 863–870.
Rao, Z., Zhu, Z., Wang, S., Jia, G., Qiang, M., Wu, Y., 2009. CPI values of terrestrial higher plant-derived long-chain *n*-alkanes: a potential paleoclimatic proxy. Front Earth Sci. China 3, 266–272.
Riederer, M., Schreiber, L., 2001. Protecting against water loss: analysis of the barrier properties of plant cuticles. J. Exp. Bot. 52, 2023–2032.
Rommerskirchen, F., Plader, A., Eglinton, G., Chikaraishi, Y., Rullkötter, J., 2006. Chemotaxonomic significance of distribution and stable carbon isotopic composition of long-chain alkanes and alkan-1-ols in C_4 grass waxes. Org. Geochem. 37, 1303–1332.
Sachse, D., Radke, J., Gleixner, G., 2006. δD values of individual *n*-alkanes from terrestrial plants along a climatic gradient - implications for the sedimentary biomarker record. Org. Geochem. 37, 469–483.
Sachse, D., Billault, I., Bowen, G.J., Chikaraishi, Y., Dawson, T.E., Feakins, S.J., Freeman, K.H., Magill, C.R., McInerney, F.A., van der Meer, M.T.J., Polissar, P., Robins, R.J., Sachs, J.P., Schmidt, H.L., Sessions, A.L., White, J.W.C., West, J.B., Kahmen, A., 2012. Molecular paleohydrology: interpreting the hydrogen-isotopic composition of lipid biomarkers from photosynthesizing organisms. Annu. Rev. Earth Planet. Sci. 40, 221–249.
Shepherd, T., Wynne Griffiths, D., 2006. The effects of stress on plant cuticular waxes. New Phytol. 171, 469–499.
Thomas, D.S.G., 2011. Arid environments: Their nature and extent. In: Thomas, D.S.G. (Ed.), Arid Zone Geomorphology: Process, Form and Change in Drylands. John Wiley & Sons, Ltd., London, pp. 1–16.
Tian, Q., Fang, X., Wang, M., 2017. Sedimentary *n*-alkanes record of precipitation D/H ratios in arid regions of the Tibetan Plateau (in Chinese). Chin. Sci. Bull. 62, 700–710.
Tipple, B.J., Pagani, M., 2013. Environmental control on eastern broadleaf forest species' leaf wax distributions and D/H ratios. Geochim. Cosmochim. Acta 111, 64–77.
Vogts, A., Moossen, H., Rommerskirchen, F., Rullkotter, J., 2009. Distribution patterns and stable carbon isotopic composition of alkanes and alkan-1-ols from plant waxes of African rain forest and savanna C_3 species. Org. Geochem. 40, 1037–1054.
Vogts, A., Schefuß, E., Badewien, T., Rullkötter, J., 2012. *n*-Alkane parameters from a deep-sea sediment transect off Southwest Africa reflect continental vegetation and climate conditions. Org. Geochem. 47, 109–119.
Volkman, J.K., Barrett, S.M., Blackburn, S.I., Mansour, M.P., Sikes, E.L., Gelin, F., 1998. Microalgal biomarkers: a review of recent research developments. Org. Geochem. 29, 1163–1179.
Walker, J.D., Cooney, J.J., 1973. Aliphatic hydrocarbons of Cladosporium resinae cultured on glucose, glutamic acid, and hydrocarbons. Appl. Environ. Microbiol. 26, 705–708.
Ward, D.M., Brock, T.D., 1978. Hydrocarbon biodegradation in hypersaline environments. Appl. Environ. Microbiol. 35, 353–359.
Xie, L., Spiro, B., Wei, G., 2016. Purification of $BaSO_4$ precipitate contaminated with organic matter for oxygen isotope measurements ($\delta^{18}O$ and $\Delta^{17}O$). Rapid Commun. Mass Spectrom. 30, 1727–1733.
Xu, X.L., Zhang, Y.Q., 2017. Chinese Climatogolical Background Datasets. Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences. http://www.resdc.cn/ (in Chinese).
Yan, Y., 2017. Research Progresses on Organic Provenance Proxies and Deposition Mechanism of East Asian Aeolian Dust (in Chinese). Postdoctoral Report. China University of Geosciences, Wuhan, pp. 43–55.