Palaeogeography, Palaeoclimatology, Palaeoecology 562 (2021) 110138

Contents lists available at ScienceDirect

Palaeogeography, Palaeoclimatology, Palaeoecology

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

Woody dicot leaf traits as a palaeoclimate proxy: 100 years of development and application

Robert A. Spicer[a,b,c,*], Jian Yang[d], Teresa E.V. Spicer[a], Alexander Farnsworth[e]

[a] CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China
[b] Center of Plant Ecology, Core Botanical Gardens, Chinese Academy of Sciences, Mengla 666303, China
[c] School of Environment, Earth and Ecosystem Sciences, The Open University, Milton Keynes MK7 6AA, UK
[d] State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[e] School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, UK

ARTICLE INFO

Keywords:
Foliar physiognomy
Fossil leaves
Climate modelling
CLAMP
Climate change

ABSTRACT

The relationship between plant form and climate has been recognised for more than two millennia, and the idea that fossils can indicate climate change was first recorded almost a thousand years ago. Here we review ways of using plant form to reconstruct, quantitatively, past climates, focussing on techniques that have been developed over the past century. We begin with the relationship between woody dicot leaf margins and temperature, and trace the development of the approach through to the modern multivariate tool known as Climate-Leaf Analysis Multivariate Program (CLAMP), which can derive 24 different climate parameters from 31 macroscopic leaf traits commonly preserved in leaf fossils. We review the complex inter-dependant correlations between leaf traits and climate, recognising that leaves develop and function as integrated systems, and that there is a selective premium on them being well adapted to their immediate climatic environment. We discuss different ways of understanding and decoding climate from leaf form in multidimensional trait space, and provide an over-view of CLAMP applications for tracking climate change, monsoon evolution, and in palaeoaltimetry, from the middle Cretaceous to the Pleistocene.

1. Introduction

The recognition that plant form and climate share an intimate relationship was first recorded around 300 BCE, more than two thousand years ago, by the Greek writer Theophrastus (Hort, 1948). Since then that relationship has been commented upon numerous times (e.g. Seward, 1892; von Humboldt, 1850; Witham, 1833 and references therein). More than 160 years ago there were even attempts to describe the link quantitatively (Heer, 1855, 1856, 1859). The relationship between plant form (physiognomy) and climate is expressed at a range of scales from the level of aggregations of whole plants (Holdridge, 1947; Raunkiaer, 1934), through to individual plant parts, such as wood (e.g. tree ring and histological analysis as in, for example, Baas and Wheeler, 2011; Wheeler et al., 2007) or, as we will focus on here, leaves.

Determining past climate using plant physiognomy is quite different from other plant-based palaeoclimate proxies that rely on the perceived climatic tolerances of specific taxa. The physiognomic approach does not require species to be identified and named, while identification to as close to species level as possible is fundamental to the alternative ’Nearest Living Relative’ (NLR) palaeoclimate proxies. The earliest recorded application of the NLR approach is that of the Chinese natural philosopher Shen Kuo who, in his ’Dream Pool Essays’ written in 1086 CE (Needham, 1986), reasoned that when plant fossils are found in an area where the nearest living relative of that fossil no longer exists, it means that the climate in that area must have changed from when the ancient plant was alive. This reasoning is remarkable, not only for its use of uniformitarian thinking some seven hundred years before this concept was adopted in the west by Hutton (1788) and Lyell (1830-33), but also because Shen Kuo recognised almost 1000 years ago that fossils represent previously living organisms, and that climate is a key determinant as to where plant species grow.

Since then NLR palaeoclimate proxies have been refined considerably, but they still rely on the same basic principles. Firstly, the fossil has to be correctly identified, and it has to have a nearest living relative. For

---

* Corresponding author at: CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China.

E-mail address: r.a.spicer@open.ac.uk (R.A. Spicer).

https://doi.org/10.1016/j.palaeo.2020.110138
Received 25 September 2020; Received in revised form 11 November 2020; Accepted 11 November 2020

long-extinct groups of plants this represents a problem that limits the use of NLR proxies in the deep past. Ideally the NLR should be at the species level, but because species rarely survive unchanged for more than a million years or so, often we have to make do only with identification to the nearest living genus or, in extreme cases, family. Of course, the more distant the relationship the more likely the fossil and living plant will have different climatic tolerance envelopes, and the broader the plant group the larger that envelope becomes.

This brings us to evolution. Plants have to be well-adapted to their local prevailing climate, including its extremes, or they will die. Evolution involves adaptation brought about by random genetic change filtered by natural selection, so we would expect plants to change their climate tolerances over time. Importantly, it is an advantage if their genetics favour such adaptability, and such tolerance changes may not be apparent in the fossil record, particularly when overall plant form is rarely preserved intact. Plants are usually found as isolated detached organs (leaves, pollen, wood etc.), not as complete entities. Potentially, evolutionary change undermines the NLR approach, but the effect can be minimised if we use several NLRs simultaneously. The more taxa are used in the analysis, the more likely it becomes that anomalies due to evolutionary change in one or two taxa will be apparent. They can then be excluded and prevented from distorting the climate estimate. The Coexistence Approach (Mosbrugger and Utescher, 1997) and Bioclimatic Analysis (e.g. Fletcher et al., 2014; Greenwood et al., 2003, 2005; Kershaw and Nix, 1988; Thompson et al., 2012) are examples of such a multi-taxon approach. The main advantage of NLR proxies over other ways of determining past climate from plants is that they can be applied to any identifiable plant part such as pollen, wood, leaves, seeds and fruits. For two contrasting views of the issues associated with such approaches see Grimm and Denk (2012) and Utescher et al. (2014). The focus of this review is not on NLR proxies, however, but the development of our understanding of how leaf physiognomy in woody dicot angiosperms relates to climate.

Because they determine plant success, plant functional traits are the basis for research spanning evolution, conservation, functional ecology, and ecosystem management. Consequently, large, but still incomplete, databases now exist summarising such traits across both herbs and woody plants worldwide (see Kattge et al., 2020, for a recent overview). These functional traits encompass phenology, biochemistry, physiology, anatomy and morphology. Morphology here includes not just some leaf traits such as size, but also whole plant form. Of all these trait groups only morphology, and specifically leaf morphology, is preserved often enough in the fossil record to be useful as a quantitative palaeoclimate proxy. For this reason, here we focus only on the morphological traits of leaves, which in woody dicots are shed in large numbers and encode adaptations to the prevailing atmospheric environment.

2. The leaf/climate relationship

Leaves, particularly those of woody dicotyledonous angiosperms, are highly plastic. That is to say that even on a single plant leaves can display a wide array of macro-and micro-morphologies, textures, and colours. This display of architectural variety extends beyond the individual to across clades and, unsurprisingly, there has been much speculation as to the evolutionary drivers of this plasticity. Some have argued that leaf form is in part the product of plant-animal interactions, (e.g. Brown and Lawton, 1991), and no doubt this may be true for some species, particularly those developing camouflage mottling (Givnish, 1990), feigning disease, or mimicking other, more toxic, leaves (Brown and Lawton, 1991). However, most research into trait plasticity has been in relation to the physical environment, particularly climate.

Leaves have an intimate functional relationship with the atmosphere (e.g. Givnish, 1987). After all, their primary purpose, photosynthesis, involves constantly processing the atmosphere by exchanging carbon dioxide, oxygen, and water vapour between the plant and the surrounding air. To conduct photosynthesis efficiently leaves have to be well adapted to their immediate environment so that they can absorb as much light as possible, while minimising resource investment, preventing overheating, and losing too much water. They do, however, have to keep some water flowing through the plant from the roots to the leaves, and then to the atmosphere, in order to keep nutrients supplied throughout the plant body. When the air becomes saturated with water, losing it from the leaf becomes difficult and many plants resort to actively expelling water from the leaves, usually through glands on the leaf margins situated at the tips of ’teeth’; a process known as ’guttation’ (Fig. 1a). However, leaves that rarely encounter moisture-saturated air for long periods do not possess teeth and their margins are smooth or ’entire’, but may still exhibit other water-related traits such as an attenuate apex, or drip tip, that aids in shedding rainwater from the leaf surface (Fig. 1b).

Because of this complex web of trade-offs and environmental ’tuning’ it is no surprise that leaf form reflects the predominant climate immediately surrounding the parent plant. Leaf form tends to be optimised for efficient functioning under local conditions, and one reason why flowering plants (angiosperms) are so successful is that they have the ability to produce many different leaf forms on an individual plant (i.e. from the same genome) specific to the individual microclimates to which each leaf is exposed (e.g. Ashton et al., 1998; Daly and Gastaldo, 2010; Kouwenberg and Kürschner, 2007; Kürschner, 1997; Maslova et al., 2008a, 2008b, 2018; Tsukaya, 2005; Zalensky, 1904). Palaeoclimate proxies based on leaf physiognomy exploit this adaptability and aim to translate leaf form preserved in fossils into accurate and precise reconstructions of past climate.

While most research has been conducted on leaves of woody flowering plants, the principles just outlined should apply to other major plant groups. Although not all leaf traits have been investigated, it is clear that leaf hydraulics underlies photosynthetic capacity and that across lycopods, ferns, conifers and angiosperms photosynthetic capacity is correlated to the proximity of veins to leaf evaporative surfaces (Brodribb et al., 2007). Minimising the hydraulic pathway within the mesophyll constrains leaf shape and size across clades, with gymnosperms tending to adopt highly dissected leaves (e.g. ferns and cycads), or produce needle-like leaves (e.g. conifers), while angiosperms, following a different route, have developed highly reticulate venation systems within thin, broad leaves (Brodribb et al., 2007). This restricted trait diversity in gymnosperms, but opened up greater potential trait space in angiosperms. Of course, some gymnosperms also have reticulate venation and an obvious example here are the leaves of the predominantly Permo-Triassic Gondwanan glossopterids. These could be investigated as potential palaeoclimate indicators, but would lack comprehensive and reliable climate calibration data.

2.1. Early work

2.1.1. Leaf margins

In the early 20th Century, the relationship between the proportion of woody dicot species with marginal teeth at any given location and prevailing thermal conditions was noted by Bailey and Sinnott (1915, 1916), but it was not until more than 50 years later that it was reexamined and developed further (Wolfe, 1971, 1979). Bailey and Sinnott’s, 1915 article was entitled *’A Botanical index of Cretaceous and Tertiary Climates’*, and they observed that for woody dicotyledonous plants the proportion of toothed leaves versus those without teeth (i.e. with entire margins) varied with the average ambient temperature. From a modern perspective the Bailey and Sinnott (1915) paper was seriously flawed in that nowhere were temperature measurements given. Instead there were groups of locations listed under the headings ’Frigid’, ’Cold Temperate’, ’Warm Temperate’ and ‘Sub-Tropical and Tropical’, and each location was accompanied by the percentage of entire-margined leaves found there. Apart from lacking quantitative measurements of temperature, the geographical regions that formed the sample sites varied enormously in area. So, we find that some leaf

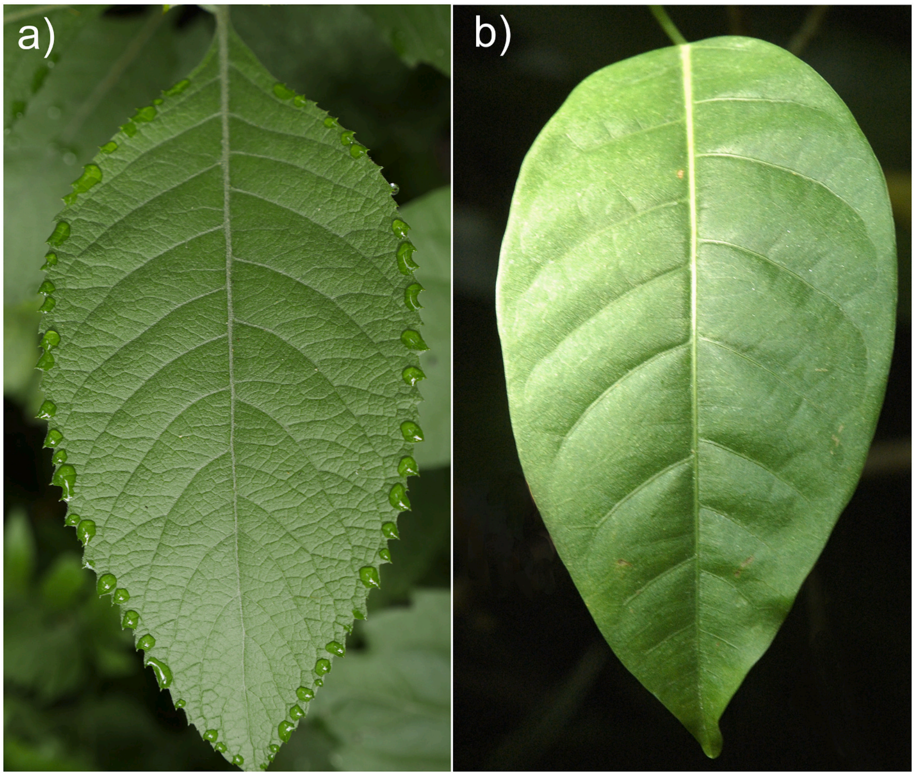


Fig. 1. a) toothed leaf exhibiting guttation; b) entire margined leaf with an apical 'drip tip'.

margin ratios are based on plants from quite small regions such as 'The Los Angeles Area' and 'Hong Kong', while others are extremely large such as 'Brazil' and 'West Central Africa'. Furthermore, no attention was paid to the variations in altitude that exist within those areas.

Another shortcoming was that the percentage values were obtained not from direct field observations, but from floral lists or manuals. Bailey and Sinnott were well aware that this posed a major problem and in 1916 they published a follow-up paper in which they attempted to explore the possible role of phylogeny in determining the form of the leaf margin. They commented "It is to be regretted ...... that so many taxonomists have made their floras and collections representative of political rather than of phytogeographical areas, and their ecological notes usually are meagre or entirely absent" (Bailey and Sinnott, 1916, p. 24). After concluding that the role of phylogeny (or in their terms "factors due to heredity") in determining the leaf form/climate relationship was "highly improbable" compared to that of environment, they went on to say "The correlations between leaf-margin and environment would undoubtedly be even more striking, if it were possible to study the "vegetation" of the temperate and tropical zones rather than their "flora" (Bailey and Sinnott, 1916, p. 29).

Bailey and Sinnott restricted their analysis to the leaves of woody dicotyledonous flowering plants because, as they pointed out, "herbaceous forms are of very infrequent occurrence as fossils below the upper Tertiary." (Bailey and Sinnott, 1916, p. 37). This philosophy of developing a proxy that is suited to the kind of information preserved in the fossil record has underpinned much (but not all) subsequent foliar physiognomic analysis, with the added rationale that plants with above-ground architecture that live for many years (i.e. woody plants) are more strongly influenced by climate than annuals, or those that die back each autumn to underground perennating organs (perennial herbaceous taxa).

The shortcomings of Bailey and Sinnott's observations limited their application until Wolfe (1979) revisited the issue by studying vegetation literature and climate data for eastern Asia. Wolfe's region of interest included parts of northern India, Nepal, Bhutan, Myanmar, northern Thailand, China, Japan, northern Philippines, North and South Korea and eastern Russia. As well as defining the thermal envelopes of different forest types, not unlike the approach of Holdridge (1947) and, to some extent, Axelrod and Bailey (1976), Wolfe also examined the relationship of leaf margin types to both mean annual temperature (Fig. 2) and mean annual range of temperature (Fig. 3). Wolfe restricted these relationships to humid and mesic forests, that is to say those not exposed to extreme drought, and as such he avoided some of the complications arising from leaf form adaptations specific to desert and monsoonal climates.

This technique of leaf margin analysis (LMA), as it is known, has since been revisited by other authors (e.g. Greenwood et al., 2003, 2004; Hinojosa and Villagrán, 2005; Kowalski and Dilcher, 2003; Miller et al., 2006; Wilf, 1997), with some using floral lists instead of direct field observations (e.g. Greenwood et al., 2004; Wilf, 1997). A few authors have sought to offer functional explanations for the relationship between leaf teeth and climate (e.g. Baker-Brosh and Peet, 1997; Feild et al., 2005; Royer and Wilf, 2006), but discreet mechanistic links

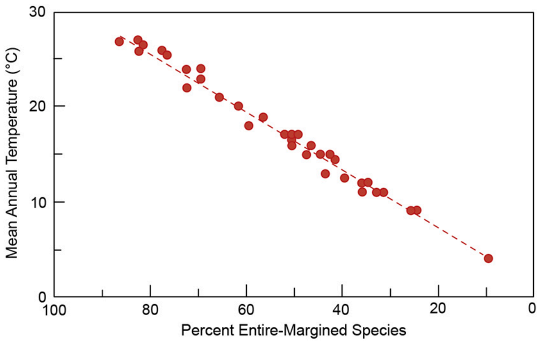


Fig. 2. The relationship between woody dicot leaf margin type and mean annual temperature for humid and mesic East Asian forests modified from Wolfe (1979).

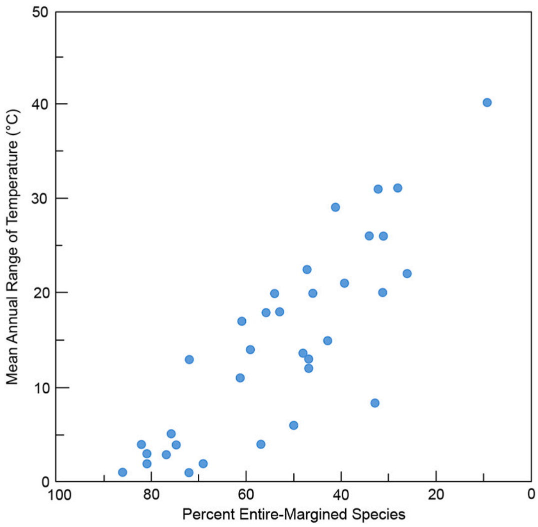


Fig. 3. The relationship between woody dicot leaf margin type and mean annual range of temperatures for humid and mesic East Asian forests modified from Wolfe (1979).

remain elusive (Jordan, 2011). We argue here that this is because teeth have multiple functions and, as we shall see, the whole leaf operates as an integrated system in which individual leaf traits are highly correlated with one another.

Because of the complex interplay between leaf traits and climate, LMA has now been largely superseded by multivariate methods. Moreover, biogeographic history has a profound effect on single trait/climate relationships. This is perhaps best illustrated by Kennedy et al. (2014) who investigated LMA using global data and it became clear that the regressions differ markedly from region to region (Fig. 4.)), and so are likely also to vary through time. This makes LMA highly unreliable as a climate proxy.

2.1.2. Leaf area

Several studies have noted that leaf size, or more strictly area, tends to be correlated with the amount of moisture available to a plant, and traditionally this has been measured in terms of rainfall, either during the growing season or averaged over the year (e.g. Cowling and Campbell, 1980; Dilcher, 1973; Givnish, 1984; Hall and Swaine, 1981; McDonald et al., 2003; Uhl and Walther, 2003; Webb, 1968; Wilf et al., 1998). Unfortunately, rainfall is not the only source of water for a plant. Those plants growing close to a water body, and so most likely to be fossilised, have roots that easily penetrate a year-round groundwater supply; water that may have been sourced many kilometres away in the headwaters of the catchment. Unsurprisingly, leaf size varies with this soil moisture (Givnish, 1979) and not rainfall directly. In addition to responding to water in the soil system, leaf size also varies with soil nutrient status (e.g. Beadle, 1954; Chapin, 1980; Ashton and Hall, 1992; McDonald et al., 2003), elevation (e.g. Halloy and Mark, 1996; Cordell et al., 1998) and temperature (Dolph, 1977; Dolph and Dilcher, 1980; Givnish, 1984), all variables that are to some extent correlated. When growing close to water, the hydrological stress a plant experiences is not due to a lack of rainfall but a dry atmosphere and, as we shall see, a measure of humidity related to leaf water loss, such as vapour pressure deficit (VPD), is more strongly correlated with leaf form than precipitation. Finally, leaf size also varies with light availability so that shade leaves tend to be larger than sun leaves (e.g. Maslova et al., 2018; Rozendaal et al., 2006). In view of the fact that it is affected by so many environmental variables, leaf size by itself such as in leaf area analysis, or LAA, it is a poor and, at best, ambiguous climate proxy.

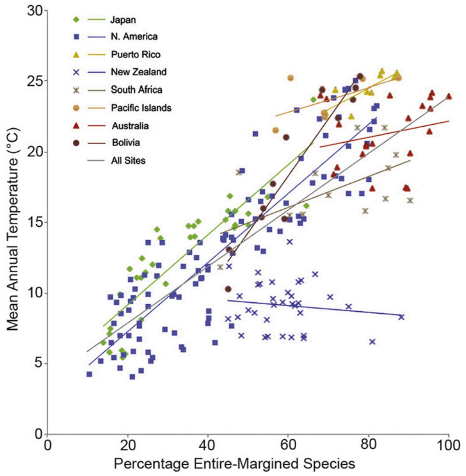


Fig. 4. The relationship between leaf margin type and mean annual temperature varies region by region. Regression lines have markedly different slopes depending on location, and thus biogeography. Modified from Kennedy et al. (2014).

3. CLAMP (Climate-Leaf Analysis Multivariate Program) Development

Climate-Leaf Analysis Multivariate Program (CLAMP) is the most widely applied leaf physiognomy palaeoclimate proxy so far developed. The origins of CLAMP go back to the mid 1970s when the American palaeobotanist Jack Wolfe realised that by using multivariate statistical analysis it might be possible to explore the relationships between multiple leaf traits and several climate variables simultaneously. He set about building a suite of leaf trait observations directly from modern vegetation, and compared those with climate data from nearby meteorological recording stations. He reasoned that to be useful in palaeoclimatology leaf traits had to be those that would be obvious in even poorly preserved impression fossils otherwise insufficient fossil data would be available to explore past climates in any spatial or temporal detail. For this reason, CLAMP traits exclude those found in preserved cuticle, such as stomatal features, surface papillae and trichomes and venation details (Wolfe, 1993), although where they are present such micromorphological traits should be used to help categorise specimens into species-equivalent groups known as ’morphotypes’.

The capacity to analyse numerically multiple traits simultaneously was not available to Bailey and Sinnott when they began their work, but with the advent of computers multivariate statistical analysis of complex data became commonplace and many techniques were developed, each with different strengths and weaknesses. Initially these were most often applied in the social sciences, but later they were used to explore relationships between plant taxa and their environment, i.e. phytosociology.

Two classes of technique have proved popular in phytosociology: classification and ordination (Kent and Coker, 1992). Classification groups samples based on their attribute similarities, or dissimilarities. The underlying assumption is that groups exist and it is just a matter of identifying them. Ordination, however, arranges samples in order based on their attributes, and if groups are present they will automatically ’fall out’ in the analysis. The ordering is done in as many spatial dimensions as there are attributes. Many statistical techniques assume the data are normally distributed (they follow a Gaussian curve and are not skewed) and the attributes are independent of one another. Neither of these conditions exist with leaf traits (e.g. Teodoridis et al., 2011), and because fossils are invariably incomplete, any statistical technique used to analyse them must also be able to accommodate missing data.

One such technique favoured by plant ecologists because it is robust to non-ideal data is Correspondence Analysis (CA) (Benzecri, 1973; Hill, 1973, 1979), which orders vegetation samples based on their attribute data, and simultaneously orders the attribute data based on their occurrences in the vegetation samples. It is possible, therefore, to examine the relationships between traits (attributes) and the locations or sites (samples) within which they occur. In CA climate relationships have to be explored separately and empirically within the physiognomic space defined by the site/trait relationships.

The first use of CLAMP was to investigate thermal changes across the Cretaceous-Paleogene transition in western North America (Wolfe, 1990), where calibration depended upon a mix of direct field sampling and physiognomic data obtained from floral manuals, and a different mix of leaf traits and climate data to that which later became standard for CLAMP (Table 1). These early traits are evident in Fig. 5. Using CA Wolfe ordered the modern vegetation sites using leaf trait data and was able to overlay the trait trends on the same plot.

When the CLAMP technique was published in detail (Wolfe, 1993) the leaf trait data were scored differently and included calibration sites across North America, the Caribbean and Japan, as well as on western Pacific islands. Floral manual data were excluded. Calibration sites were selected for 1) having mature vegetation minimally disturbed by human activity and 2) being close to meteorological stations. Condition 2 could not always be met as we shall see later.

Wolfe (1993) sampled modern vegetation in a methodical way, overcoming deficiencies in previous work. All modern training (calibration) data are now obtained by direct field sampling from relatively small areas, usually no more than a hectare, and within a 0.5 km radius of a central location (for details of the sampling protocols see the CLAMP website:http://clamp.ibcas.ac.cn). At each location all sampling is at approximately the same elevation.

CLAMP sampling protocols also consider the natural sampling that takes place during the formation of a fossil assemblage. Where possible vegetation sites are chosen that are close to a stream or lake, but where this cannot be achieved sampling is permitted along forest tracks that simulate canopy light breaks, such as might occur along a stream. This means that all visible leaf morphological variants within a species are sampled, including sun and shade leaves, those on both vegetative and fertile shoots, and all growth habits of woody taxa are included (vines, shrubs, saplings and mature trees). In the fossil record, where extinct taxa are present, it is rarely possible to infer the habit of the plant from which a leaf originated, or from where on the plant it was derived, so modern vegetation used for calibration has to embrace all these possible sources (Wolfe, 1993), even though this leads to greater calibration 'noise' and reduced predictive precision.

The CLAMP trait categorisations (Fig. 6) were arrived at empirically, and through numerous exploratory analyses Wolfe refined the range of traits and their scoring schemes (Table 1). Unlike in LMA, Wolfe recognised several tooth types based on their geometry and spacing so, as well as the presence or absence of teeth (teeth or no teeth), CLAMP subdivides the tooth trait group into several trait states (Fig. 6). Tooth size is to a large extent linked to leaf size so was not scored separately, but spacing and compound teeth are and these incorporate tooth size to some extent. In CLAMP, spines and teeth are treated separately and because the primary function of spines is to deter predation spines are ignored: a spiny leaf is regarded as entire. Extended (attenuate) leaf apices have long been recognised as indicators of high rainfall and are termed 'drip tips' (Fig. 1 b). Similarly, other trait groups have multiple states (Fig. 6).

Wolfe realised that leaf size measurement had to be categorical, i.e. not measured to the nearest mm^2, but grouped into several size categories. This minimises redundant sampling 'noise' that would otherwise degrade analytical precision, much as log transforms do when measuring abundance (e.g. Spicer and Hill, 1979), and size can get distorted during the fossilisation process. Categorical size scoring is also a lot faster than taking detailed measurements, particularly when preservation is poor, and requires just a simple transparent overlay guide (Fig. 7) that can be used on both living and fossil leaves.

Initially there were only 7 size categories, but then two more were added; Nanophyll at the small end of the size range, and Mesophyll III for the largest leaves (Forest et al., 1999). The Mesophyll III category is open ended in that it incorporates any sizes larger than the minimum for that size category. Again, this is a scoring strategy geared to the fossil record in that leaves larger than Mesophyll III are almost never preserved as fossils because they usually fragment during pre-depositional transport (Spicer, 1981). Their rarity means that when they are encountered such large leaves are likely to generate analytical anomalies with little environmental meaning. Note that the areas that the categories represent do not increase uniformly through the size spectrum, and that size categories differ slightly in terms of area depending on leaf shape, as is evident from the sizing template (Fig. 7). This is because narrow leaves would have to increase greatly in length to accommodate the same area increase as a round leaf when moving from one category to another, and so all narrow leaves would inevitably always be confined to the smaller size categories if no adjustment for shape was incorporated in the size scoring.

3.1. Leaves as 'integrated systems'

Through highly flexible yet integrated developmental pathways and pleiotropy (the production of multiple features by a single gene) (Falconer and Mackay, 1996; Juenger et al., 2005; Pigliucci, 2003; Rodriguez et al., 2014; Schlichting, 1989) leaves operate as complex systems, where all traits are linked to one another and no single trait has a simple relationship with a single climate variable. This means that a change in one feature necessarily means a change in others if the leaf is to remain functionally efficient. At the macro scale this complex web of inter-relationships is illustrated in Fig. 8. Similarly, the number and type of stomata are linked to vein architecture (Brodribb and Jordan, 2012), which must in turn be linked to overall leaf form if the functional integrity of the leaf is to remain intact. This complex interaction of leaf characters and climate means that single character to single climate variable links must be spurious, and so unreliable for use as a climate predictor in 'deep time' (Kennedy et al., 2014; Yang et al., 2015).

Similarly, all climate variables are interlinked through the laws of thermodynamics. For example, temperature and humidity are linked through phase transitions and latent heat. This means that any analytical technique used to decode climate signals from leaf physiognomy must be insensitive to correlations between both attributes and environmental variables. The requirement for tolerance of correlated attributes and variables, non-normality and missing data eliminates many multivariate statistical techniques such as Principal Components Analysis and Multiple Regression (Wolfe, 1995), but one technique that does possess these properties, and has found widespread application in ecology precisely because of this resilience to correlated, incomplete and non-normal data, is Canonical Correspondence Analysis (CCA).

3.2. Canonical correspondence analysis - the statistical engine within CLAMP

Developed by ter Braak (1986, 1987), CCA is the statistical engine in CLAMP (Kovach and Spicer, 1996; Wolfe, 1995) and a derivative of CA. Like CA, CCA in CLAMP orders the modern vegetation sites in multidimensional space based on their aggregate trait spectrum score. The important innovation in CCA is that it combines the trait data with a separate climate data array to position climate vectors in physiognomic space, and not, as in Wolfe (1993), by treating climate data as if it were a vegetation sample.

A standard CCA CLAMP analysis combines three data arrays (Fig. 9). The first consists of the physiognomic scores that describe leaf form in stands of modern natural or naturalised vegetation (Fig 9a). The scores comprise a string of 31 numbers for each site, each number being the percentage aggregate score for each trait across a minimum of 20 species. This minimum was first arrived at empirically (Povey et al., 1994) (Fig. 10) and subsequently confirmed mathematically (Forest et al.,

Table 1
A timeline of CLAMP development.

Innovation	Description	Date/Source
Introduction of CLAMP and first application to temperature changes across the K/Pg interval.	An early form of CLAMP using data from 18 field collections plus floral manual data for Nigeria and Borneo. Correspondence Analysis (CA) used to position modern vegetation based on 18 leaf traits. Climate data was added to the analysis one variable at a time and treated as a leaf trait in order to identify climate trends (Fig. 5).	Wolfe, 1990
CLAMP methodology published	CA used to create physiognomic space, from 29 traits exhibited in 106 modern vegetation sites scored from N. America, Japan, western Pacific and Caribbean spanning 9°N to 62°N and altitudes up to 2700 m. Field sampling protocols established. Climate data used 7 parameters of variable quality over different 30-year intervals as measured at recording stations at different distances from the vegetation sites. Elevation corrections done using free air mean lapse rates. Climate vectors positioned as in Wolfe (1990) and 2nd order polynomial predictive regressions in 2D (axes 1-2 space). Adding Climate variables or fossil sites distorts physiognomic space.	Wolfe, 1993
Minimum morphotype (species) size of 20	Using variously sized random subsets of 143 fossil taxa from the Republic Flora a minimum of 20 taxa were shown to give the greatest precision, while being realistic about fossil assemblage diversity. By setting a minimum number of 20 taxa for both modern and fossil physiognomic data uncertainties are minimised. Two calibration data sets of 80 and 106 sites used.	Povey et al., 1994
Canonical Correspondence Analysis used as the statistical engine for CLAMP	Initial exploration of Canonical Correspondence Analysis (CCA) instead of CA as the CLAMP statistical engine. This allows climate vectors to be assigned explicitly using a separate climate data array (not as surrogate vegetation plots as in CA) (Fig. 9). This minimises (but does not eliminate entirely) distortions of physiognomic space by the addition of climate data. Multiple fossil samples can now be added simultaneously as 'passive' samples without physiognomic space distortion. Calibration similar to that in Wolfe (1993), but mostly using 82 modern vegetation samples.	Wolfe, 1995
Moist enthalpy added to the CLAMP climate variables	Moist enthalpy explored as a palaeoaltimetry proxy using CLAMP initially calibrated with 29 leaf traits from 92 North American locations	Forest et al., 1995
Canonical Correspondence Analysis further explored as the basis for CLAMP	Canonical Correspondence Analysis (CCA) in CLAMP is explored further.	Kovach and Spicer, 1996
Calibration extended into Southern Hemisphere using samples from New Zealand.	21 Modern New Zealand vegetation samples added. Leaf traits remain at 29.	Stranks, 1996
Local regressions suggested instead of CCA vectors	Recognising the complex structure of physiognomic space and the limitations of simple vectors as predictors the use of a resemblance function is applied to derive more local climate predictions (Fig. 13).	Stranks and England, 1997
Calibration extended to 31 traits.	Nanophyll and Mesophyll 3 size categories introduced, 103 vegetation sites used with 8 climate variables.	Herman and Spicer, 1996
Completeness statistic introduced	The completeness statistic (C) is a measure of how complete fossil trait data are, and is calculated as: $C = F\text{-}M/P$ where F is the actual number of data matrix cells filled when scoring the fossil assemblage, M is the minimum number of data matrix cells unfilled and P is the minimum number of data matrix cells that would be filled if all character states were to be scored for all taxa.	Herman and Spicer, 1997
Further moist enthalpy explorations, leaf traits expanded to 31	Nanophyll and Mesophyll 3 added to the standard trait spectrum. Twenty taxa confirmed as the minimum taxon count using 123 vegetation sites. Moist enthalpy interpolated to sampling sites.	Forest et al., 1999
First CLAMP website goes live	CLAMP collecting and scoring protocols formalised. CLAMP calculations required third party software to conduct the CCA analysis, but spreadsheets used to calculate fossil palaeoclimates. Physg3arAZ and Phys3brAZ with 11 climate variables based on recording station data (MET files) become the standard calibrations.	2004
Regressions move from 2D to 4D space	Regressions formed 'by at least 3 axes' is suggested in Wolfe and Spicer (1999, p.238) and regressions based on vectors positioned in axes1-4 space become standard on the CLAMP website.	2004
Modern vegetation calibration data extended across Russia	20 modern Russian vegetation sites are added including some with cold month mean temperatures of -41.5°C. The effect of seasonal dormancy on CLAMP is explored.	Spicer et al., 2004
The effect of taphonomic trait loss on CLAMP explored	Physg3arAZ (173 sites) and Physg3brAZ (144 sites) calibrations used to test the effect of both taxon loss and character loss from taphonomic processes.	Spicer et al., 2005
Gridded Climate data (GRIDMET file) introduced alongside a standardised Physg3brAZ and Physg3arAZ physiognomic files	GRIDMET files with 0.5° x 0.5° (latitude/longitude) gridded climate data introduced based on New et al. (1999) global land observations between 1961 and 1990 modified to accommodate elevation differences by using the ERA40 re-analysis data for upper level atmospheric temperatures. Data were then corrected to the actual sample location by interpolation. This approach was later extended to add the higher resolution (0.16° x 0.16°) data of New et al. (2002).	Spicer et al., 2009
CLAMP data expanded to South Africa	23 sites added including those of the highly endemic Fynbos flora.	Steart et al., 2010
Automated CLAMP scoresheet introduced	Scoresheet_2010 was introduced by Yang Jian on the CLAMP website. This removed the need to calculate trait scores by hand when filling out the scoresheet, and this reduced scorer error.	2010
Modern vegetation calibration data extended across China	45 modern Chinese vegetation sites are added to Physg2brc to form the PhysgAsia1 calibration. This included extremely cold sites from northern China as well as those from tropical conditions in the south and comprised sites heavily impacted by the East Asian Monsoon. The introduction of monsoon-affected sites is explored.	Jacques et al., 2011
CLAMP calculations done online via the CLAMP website.	An automated CLAMP analysis (CLAMP Online) is offered on the CLAMP website using Physg3ar, Physg3br, with gridded (GRIDMET files) and ungridded (MET files) calibration options for 11 climate variables. The system is tested using an analysis of 82 fossil sites simultaneously.	Yang et al., 2011

(continued on next page)

Table 1 (*continued*)

Innovation	Description	Date/Source
'Correction coefficients' introduced	These coefficients offer locally specific 'corrections' to try to overcome the complex nature of physiognomic space, similar to, but distinct from, the local regressions of Stranks and England (1997).	Teodoridis et al., 2011
India added to the calibration data	Physg3brc with 18 sites added from across India to form PhysgIndia1.	Srivastava et al., 2012
Modern South America (Bolivia) sites added	12 modern vegetation sites from Bolivia added to the CLAMP calibration.	Gregory-Wodzicki, 2000
Southern Hemisphere calibration explored	Additional samples from New Zealand, Australia and Argentina added to the CLAMP to form a 90 site SH calibration. The use of different climate datasets explored.	Kennedy et al., 2014
PhysgAsia2 dataset introduced	PhysAsia2 comprises 177 sites obtained by adding Chinese, Indian and Thai sites to those of Physg3brc. The dataset was validated against previous results for the mid Miocene Namling (Tibet) site (Spicer et al., 2003) and stable isotope analyses (Currie et al., 2005; Polissar et al., 2009). It appears robust for Asian sites subject to monsoon climates.	Khan et al., 2014
CLAMP data tested using a Neural Network Approach	General Regression Neural Networks (GRNN) applied to CLAMP calibration data to try and overcome the limitations of the CCA vector approach as calibrations include more climate space.	Li et al., 2016
A global calibration of 378 sites assembled and explored	The global calibration is extended to all continents except Antarctica across on altitudinal range of 0.5 m to >3000m. This allows an exploration of the roles of biogeographic history, and the complex structure of physiognomic space using General Additive Models (Figs. 14 & 15).	Yang et al., 2015
WorldClim2 Climate data added for Physg3brc	WorldClim2 data (Fick and Hijmans, 2017; www.worldclim.org) gridded climate data at ~1 km^2 spatial resolution and with an underlying DEM used to investigate Late Cretaceous Arctic floras using the Physg3brc calibration. Climate variables extended to 26.	Spicer et al., 2019; Zolina et al., 2020.
WorldClim2 Climate data added for PhysgAsia2	WorldClim2 data expanded to PhysgAsia2 to explore humidity (VPD and PET) in monsoon environments.	Hazra et al., 2020; Song et al., 2020

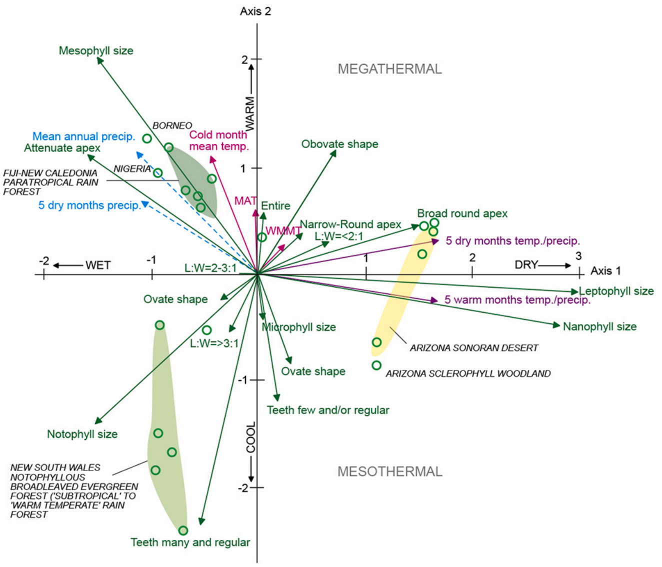


Fig. 5. An early version of CLAMP in which Correspondence Analysis (CA) was used to arrange 20 modern vegetation sites (green open circles) based on a range of leaf traits (green text). In CA the attributes (leaf traits) and objects (vegetation sites) are similarly scaled so, for example, this plot tells us that leaves from Borneo and Nigeria are likely to be large (Mesophylls) with attenuate apices, while cool temperate forests in New South Wales are likely to have toothed leaves. Wolfe did not have a direct way of including climate data using CA so to find climate trends in the data he added climate variables, one at a time, to the physiognomic data as it they were leaves. Modified from Wolfe (1990).

1999). The name of this array begins with 'Physg' followed by a string of letters and numbers that identify the composition of the array. Currently, the number of modern vegetation sites in published arrays varies from 144 in Physg3brc to 378 in PhysgGlobal378. This is matched by a climate data array (Fig. 9b) consisting of 30-year average observations for up to 23 different climate variables for each of the vegetation sites in the Physg array. This climate data array is identified by 'Met-' or 'GridMet-' followed by the same identifiers as for the Physg files. Lastly there is an array containing the trait scores for any fossil assemblages to be analysed (Fig. 9c).

The modern vegetation sites are arranged in 31-dimensional space based only on their trait scores; samples with similar trait scores plot close together, those with dissimilar scores plot apart (Fig. 9d). The climate array provides data to locate climate trends (the vectors) through the cloud of vegetation sites that make up physiognomic space. The fossil array provides the trait scores that allow the fossil sites to be positioned within the physiognomic space as defined by the Physg trait array, and the position of the fossil assemblage when projected normally onto the climate vectors gives the predicted palaeoclimate. In a standard CLAMP analysis this projection is done in 4-dimensional (axes 1-4) space.

This whole process is conducted automatically in the online CLAMP analysis tool (http://clamp.ibcas.ac.cn) (Yang et al., 2011) using underlying code written in R (R Core Team, 2012). However, spreadsheets that show how the calculations are performed can also be accessed via the CLAMP website.

Axes 1-3 plots for the PhysgAsia2 calibration (capturing ~70% of the total variation) are shown in Figs 11 a-c. and the regressions for MAT, mean annual moist enthalpy and mean annual VPD are shown in Figs 12 a-c.

3.3. The structure of CLAMP physiognomic space

The structure of global physiognomic space is complex, and this complexity contributes to increasing the statistical uncertainties simply because the leaf trait/climate relationship cannot be represented adequately by a straight-line vector. As an example, Fig. 13a shows an Axis 1-2 CCA plot of modern global (378) vegetation calibration sites, coded for their geographic origins, overlain by the warm month mean temperature (WMMT) isotherms as determined by observations at those locations. The WMMT vector (blue line) clearly reflects the general trend, but does not capture the different shape of the isotherms that exist between northern America and the stronger monsoon climates predominating over Asia/Australia. Consequently, any sites at points X (on the 24 °C contour) and Y (on the 20 °C contour) have the same WMMT vector score, and any fossil site projecting to the vector at this point could have a predicted palaeo-WMMT of either 20 or 24 °C.

To try to overcome this problem we could use vectors that reflect local trends. In Fig. 13b the local vector at *y* is shown within the pink shaded circle, and clearly is oriented in a different direction from that for

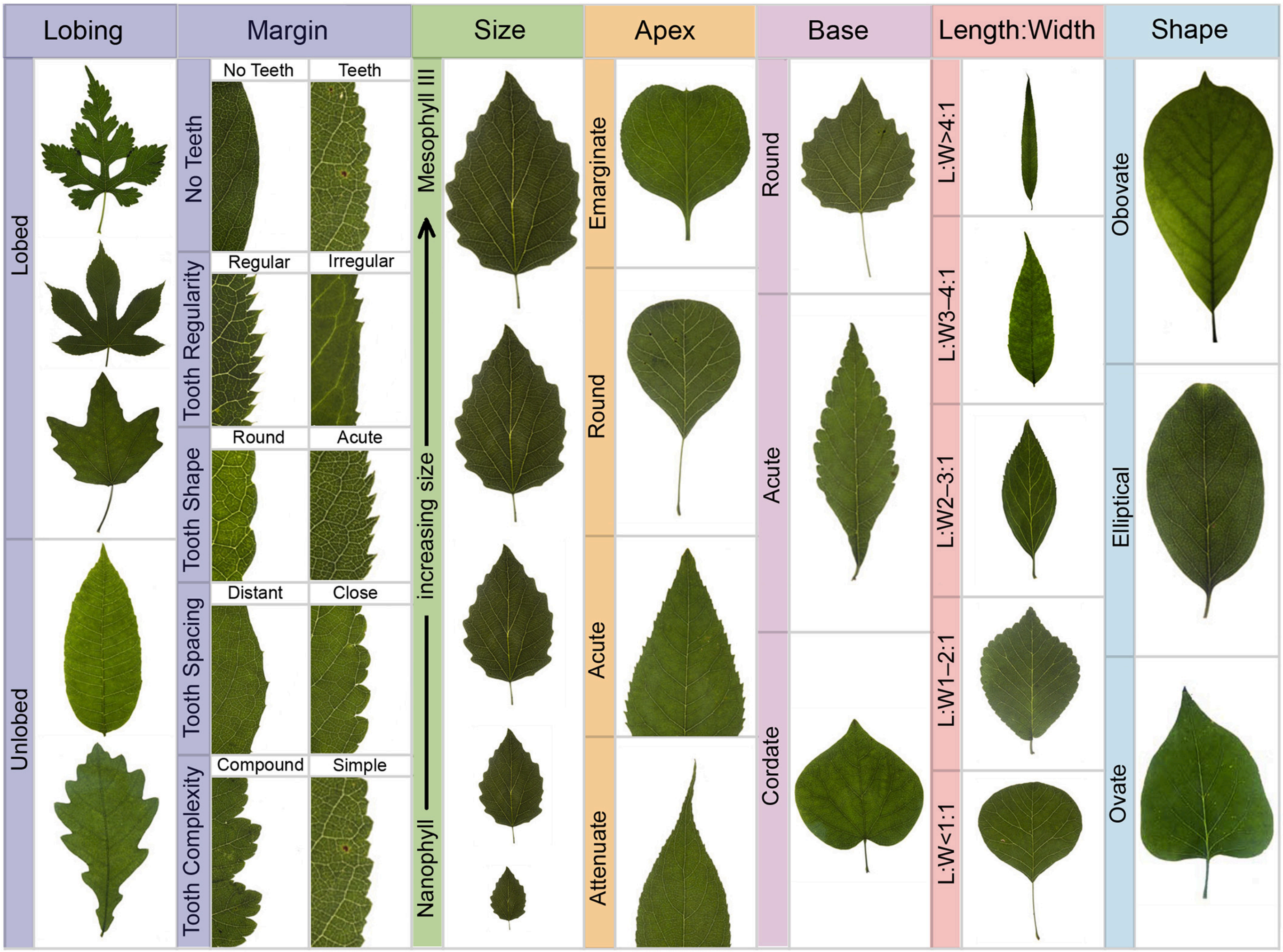


Fig. 6. Overview of CLAMP leaf traits groups and trait states. Colours denote trait groups within which different trait states are recognised. The colour coding matches that of the CLAMP Scoresheet. For numerical scoring protocols see the CLAMP website and scoresheet. Modified from Spicer (2018).

the entire dataset. This approach is at the heart of a proposal made by Stranks and England (1997) in which local regressions would be used, and to local 'correction coefficients" as explored by Teodorides et al. (2011). These approaches are not without their own problems, however, particularly in determining how large the local volume should be in multidimensional space. This volume will change with calibration site density, and is difficult to define at the margins of physiognomic space.

A derivative of this approach is to use an algorithm to fit isopleth surfaces in multidimensional space, and use the position of a fossil site on any given surface to derive a palaeoclimate estimate. Mathematically a surface can exist in any number of dimensions but visualising it becomes difficult beyond three, so Figs. 13 and 14 just illustrate such surfaces in CCA axes 1-3 space. The surfaces are fitted using a General Additive Model (GAM) (Wood, 2003) using R code (Yang et al., 2015). What they show is that with global CLAMP data MAT (Fig. 14a) and moist enthalpy surfaces (Fig. 14b) appear as near parallel sheets normal to the MAT and moist enthalpy vectors respectively, so in these cases the vectors summarise well the climate trends across physiognomic space.

However, if we look at the WMMT surfaces (Fig. 15a) we see they are more complex, just as in Fig. 13. For precipitation in the 3 driest months (Fig. 15b) there is even a 30 cm isopleth 'egg' located within a 20 cm isopleth 'tunnel'. This means that the vectors cannot adequately summarise the principal climate trends. The shape of these surfaces depends in part on the number of data points within a specified volume used to position each node of the surface, but for the default options we see that using such surfaces (Fig. 16a) instead of vectors (Fig. 16b) improves the predictive capability of CLAMP with regard to WMMT. Improvements are also possible with other climate variables such as moist enthalpy (Figs. 17a and b).

3.4. *The issue of 'extinct climates'*

It is conceivable that during Earth's complex climate evolution particular combinations of environmental variables might have arisen that are not represented in today's world. The physics of energy conservation and motion will constrain the combinations possible, but unique leaf adaptations might arise for climate regimes not present in today's world. For example, in the Cretaceous and Paleogene high latitude winter warmth was coupled with extended periods of winter darkness and not the frigid conditions seen today. If extinct climates did exist in the past then it would have been likely that through natural selection plants would have acquired trait adaptations to occupy that climate space, but this seems not to have happened. This is best illustrated by plotting all known fossil assemblages within the CLAMP global physiognomic space (Fig. 18), which shows 170 fossil assemblages ranging in age from mid Cretaceous to Pleistocene plotted in CCA axis 1 versus axis 2 Global378 physiognomic space using WorldClim2 climate data. No fossil assemblages plot outside this space, suggesting a lack of novel trait combinations within the fossil leaf assemblages. From this we infer from the mid Cretaceous onwards there have been no climate

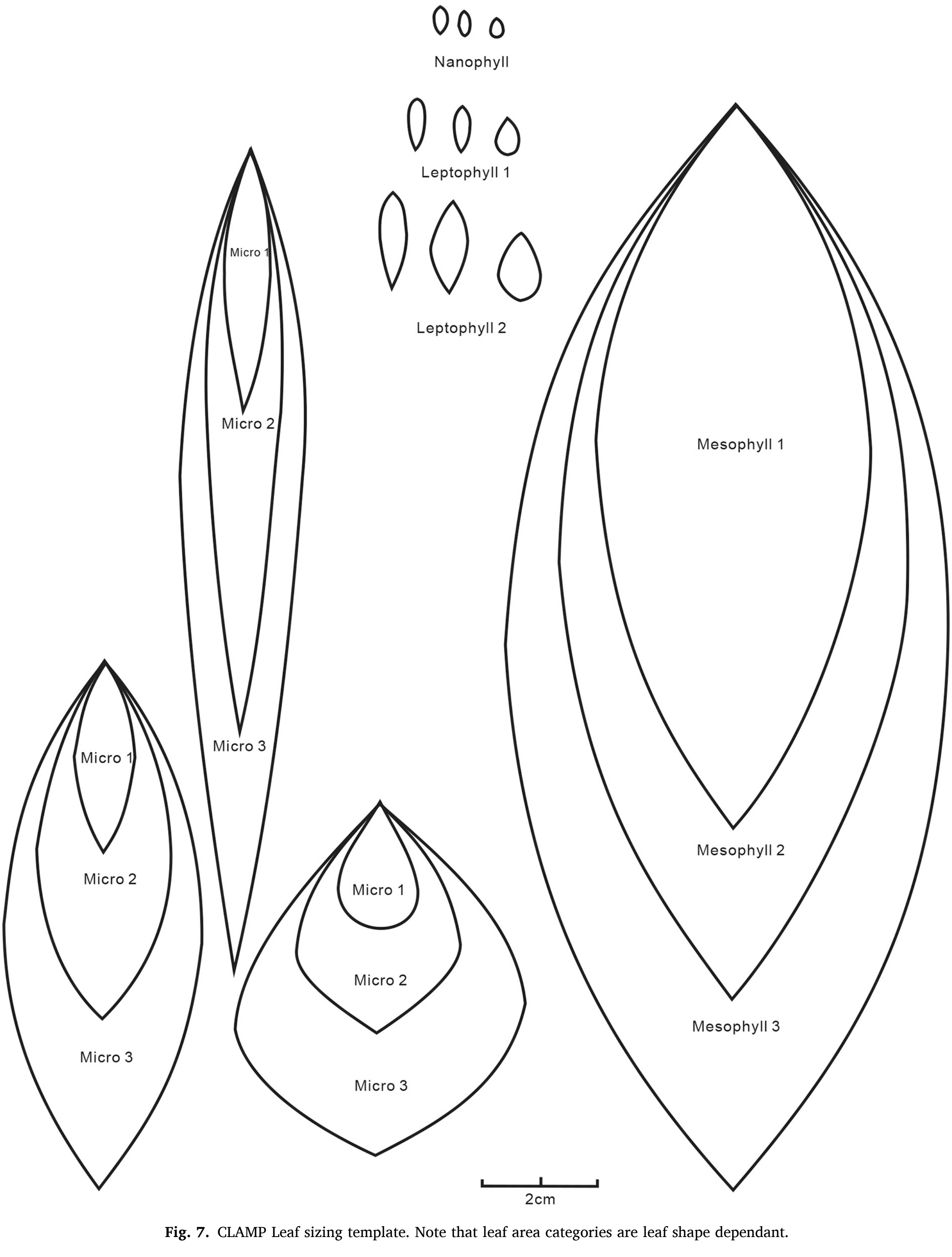


Fig. 7. CLAMP Leaf sizing template. Note that leaf area categories are leaf shape dependant.

adaptations that are not represented in the modern world, and so no 'extinct climates'. Fortunately, the modern world is characterised by a steeper equator-to-pole temperature gradient than existed in much of the geological past, so a much greater range of climate types, both latitudinally and altitudinally, exists today to which plants have adapted. It seems that modern woody dicot leaf/climate relationships encompass most, if not all, such relationships that have existed in the last 100 million years.

3.5. *Calibration data sets*

These explorations of the structure of physiognomic space are based on global calibration data, but such large datasets clearly cannot offer the most precise palaeoclimate predictions because, as we have seen, the

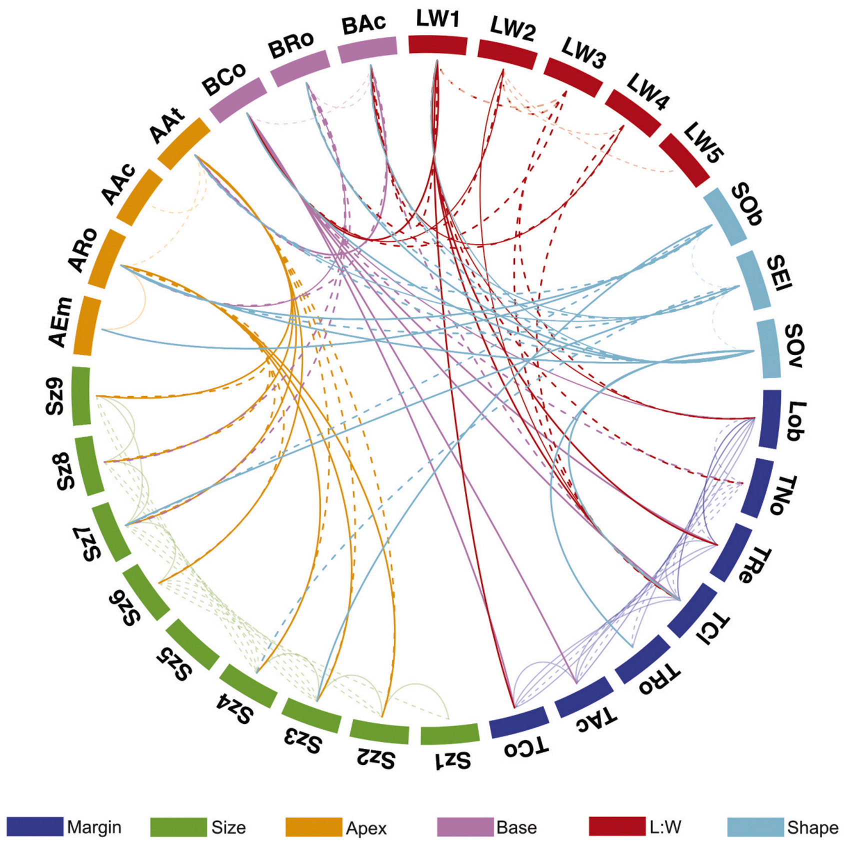


Fig. 8. Correlations between leaf traits used in CLAMP. Margin traits are colour-coded in dark blue, Size in green, Apex orange, Base purple, Length:Width red, and Shape light blue. Abbreviations are as follows: Lob - lobing, TNo - no teeth, TRe - teeth regular, TCl - teeth close, TRo - teeth round, TAc - teeth acute, TCo - teeth compound, Sz1 - nanophyll, Sz2, leptophyll 1, Sz3 - leptophyll 2, Sz4 - microphyll 1, Sz5 - microphyll 2, Sz6 - microphyll 3, Sz7 - mesophyll 1, Sz8 - mesophyll 2, Sz9 - mesophyll 3., AEm - apex emarginate, AAc - apex acute, AAt - apex attenuate, BCo - base cordate, BRo - base round, BAc, base acute, LW1 - length to width ratio <1:1, LW2 - length to width ratio 1-2:1, LW3 - length to width ratio 2-3:1, LW4 - length to width ratio 3-4:1, LW5 - length to width ratio >4:1, SOb - shape obovate, SEl - shape elliptic, SOv - shape ovate. Solid curved lines represent positive correlations, dotted lines negative correlations. Note that within-trait category correlations are spurious artefacts of the scoring scheme (e.g. traits are mutually exclusive or co-dependant) and such correlations are indicated by faded links. Modified from Yang et al. (2015).

structure of physiognomic space within them is complex. The approach adopted in current versions of CLAMP is to use subsets of the global data that have simpler structures. The disadvantage of this approach is that it reduces the number of sites available to define the surfaces, and so risks producing spurious outcomes. This means that while the GAM approach improves the predictive capabilities of the global calibration, the reduced number of samples in the calibration subsets means that the GAM surfaces are not so well fitted, which decreases precision. Table 1 summarises the evolution of CLAMP and the different calibrations developed during CLAMP evolution. The calibration data currently offered in a CLAMP online analysis are: 1) Physg3arcAZ with Met3arcAZ or GRIDMet3arAZ, Physg3brcAZ with Met3brcAZ or GRIDMet3brAZ or, 2) PhysgAsia1 with HiResGridMetAsia1, 3) PhysgAsia2 with HiResGridMetAsia2, 4) PhysgGlobal378 with HiResGRIDMet378, (Table 2). All these data sets are available on the CLAMP website and in supplementary files associated with this paper. Future offerings will include the Southern Hemisphere calibration (Kennedy et al., 2014) (Table 2), but ideally this needs to be expanded by including more samples from South America and Western Australia.

The climate variables returned using CLAMP Online are currently restricted to 11: Mean Annual Temperature (MAT), Warm Month Mean Temperature (MAT), Cold Month Mean Temperature (CMMT), Length of the growing Season (LGS), mean Growing Season Precipitation (GSP), Mean Monthly Growing Season Precipitation (MMGSP), precipitation during the 3 consecutive wettest months (3-WET), precipitation during the 3 consecutive driest months (3-DRY), mean annual Relative Humidity (RH), mean annual Specific Humidity (SH) and mean annual moist Enthalpy (ENTHAL).

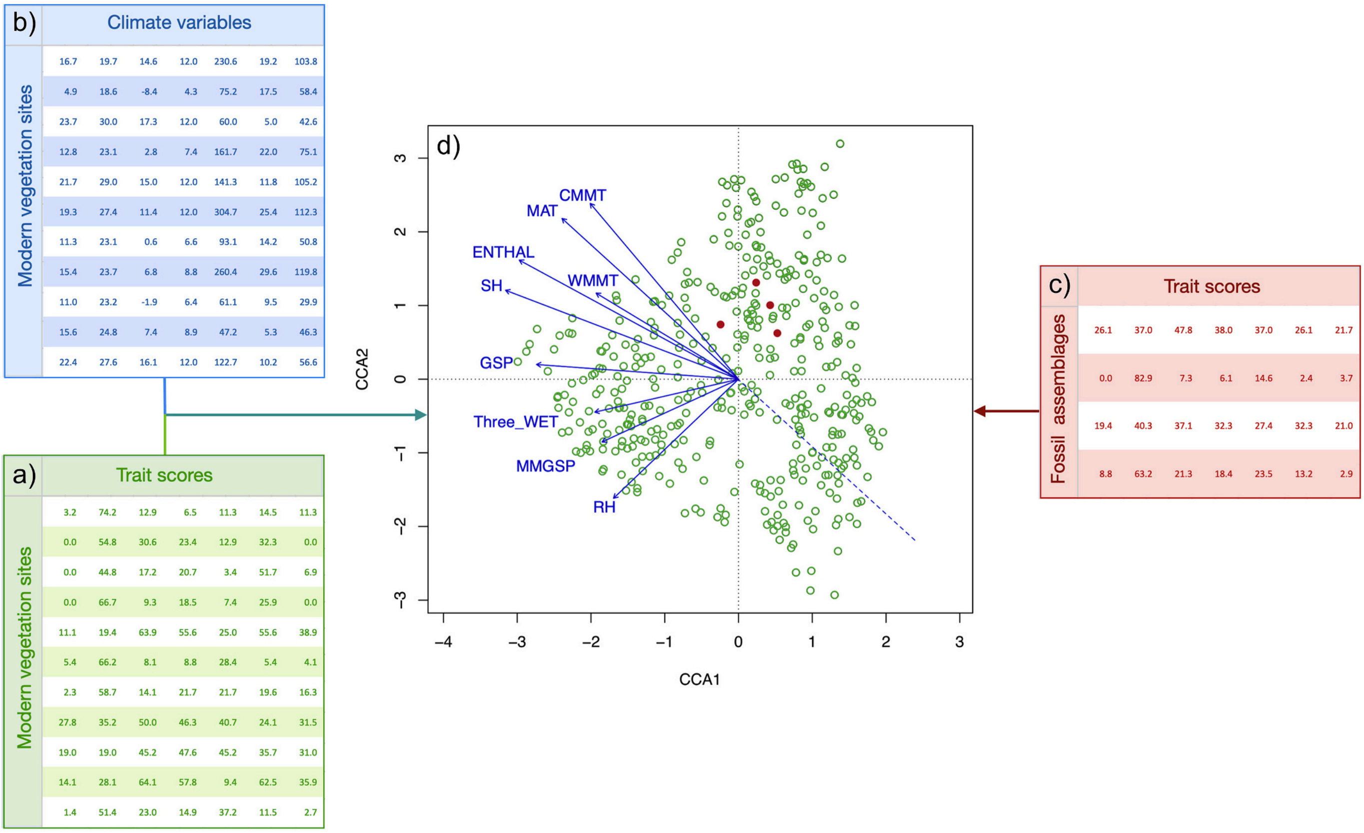
Using WorldClim2 gridded climate (Fick and Hijmans, 2017) also allows us to increase the number of variables to 24 by adding mean annual Vapour Pressure Deficit (VPD_ANN), mean summer Vapour Pressure Deficit (VPD_SUM), mean winter Vapour Pressure Deficit (VPD_WIN), mean spring Vapour Pressure Deficit (VPD_SPR) and mean autumn Vapour Pressure Deficit (VPD_AUT), mean annual Potential EvapoTranspiration (PET_ANN), mean Potential EvapoTranspiration of the warmest quarter (PET_WARM), mean Potential EvapoTranspiration of the coldest quarter PET_COLD), mean minimum temperature of the warmest month (MIN_T_W), mean maximum temperature of the coldest month (MAX_T_C), the mean compensated Thermicity index (THERM), mean Growing Degree Days above 0 °C (GDD_0) and mean Growing Degree Days above 5 °C (GDD_5) (Hazra et al., 2020; Song et al., 2020; Spicer et al., 2019). Although these datasets are not currently available for CLAMP Online, they will soon be added. They are described in Table 2 and are given as supplementary files.

3.6. Uncertainties in CLAMP

CLAMP uncertainties are explored on the CLAMP website so will only be reviewed briefly here. Clamp uncertainties arise from four sources:

1) Taphonomic (how well the traits in the fossil assemblage reflect those in the original living vegetation). Taphonomic processes inevitably degrade leaf trait information and because information is lost it is difficult to quantify. Nevertheless, because CLAMP incorporates numerous traits across numerous taxa it is statistically robust and empirical tests so far conducted confirm this (Spicer et al., 2005, 2011). As a guide to when too much information has been lost, the ’completeness statistic’ was introduced (Herman and Spicer, 1997) (Table 1) and is automatically computed as data are entered into the CLAMP scoresheet. A completeness statistic of below 0.6 means that the analysis is likely to be unreliable.

2) Climate measurement. Climate, being a description of (30 year) average weather, is based on measurements made across thousands of stations worldwide, often hourly, and inevitably records depend on instrument calibration, and proper continuous functioning. Precision is operator dependant and automated recording is different from manual measurements. Importantly, recording stations are located in open



b)	Climate variables						
Modern vegetation sites	16.7	19.7	14.6	12.0	230.6	19.2	103.8
	4.9	18.6	-8.4	4.3	75.2	17.5	58.4
	23.7	30.0	17.3	12.0	60.0	5.0	42.6
	12.8	23.1	2.8	7.4	161.7	22.0	75.1
	21.7	29.0	15.0	12.0	141.3	11.8	105.2
	19.3	27.4	11.4	12.0	304.7	25.4	112.3
	11.3	23.1	0.6	6.6	93.1	14.2	50.8
	15.4	23.7	6.8	8.8	260.4	29.6	119.8
	11.0	23.2	-1.9	6.4	61.1	9.5	29.9
	15.6	24.8	7.4	8.9	47.2	5.3	46.3
	22.4	27.6	16.1	12.0	122.7	10.2	56.6

c)	Trait scores						
Fossil assemblages	26.1	37.0	47.8	38.0	37.0	26.1	21.7
	0.0	82.9	7.3	6.1	14.6	2.4	3.7
	19.4	40.3	37.1	32.3	27.4	32.3	21.0
	8.8	63.2	21.3	18.4	23.5	13.2	2.9

a)	Trait scores						
Modern vegetation sites	3.2	74.2	12.9	6.5	11.3	14.5	11.3
	0.0	54.8	30.6	23.4	12.9	32.3	0.0
	0.0	44.8	17.2	20.7	3.4	51.7	6.9
	0.0	66.7	9.3	18.5	7.4	25.9	0.0
	11.1	19.4	63.9	55.6	25.0	55.6	38.9
	5.4	66.2	8.1	8.8	28.4	5.4	4.1
	2.3	58.7	14.1	21.7	21.7	19.6	16.3
	27.8	35.2	50.0	46.3	40.7	24.1	31.5
	19.0	19.0	45.2	47.6	45.2	35.7	31.0
	14.1	28.1	64.1	57.8	9.4	62.5	35.9
	1.4	51.4	23.0	14.9	37.2	11.5	2.7

Fig. 9. In a CLAMP analysis three data arrays: a) a trait score for each modern calibration vegetation site array, b) a climate variable record for each modern vegetation site array, and c) a trait score for each fossil assemblage array. d) These are combined in a Canonical Correspondence Analysis (CCA) to give a plot of all the modern calibration vegetation sites (green open circles) positioned relative to one another based on their trait data (array a). These sites create a cloud of sites known as physiognomic space through which climate vectors are positioned (blue arrows) using the climate data (b). In this example the Mean Annual Temperature (MAT) vector is projected past the origin to show negative vector scores. The fossil traits (c) define the position of the fossil assemblages (red closed circles) within physiognomic space. The position of the fossils along each climate vector is known as their vector score and this is used to estimate the ancient climate for that fossil assemblage. For simplicity the full data arrays are not shown and only axes 1-2 space are represented in the plot (d).

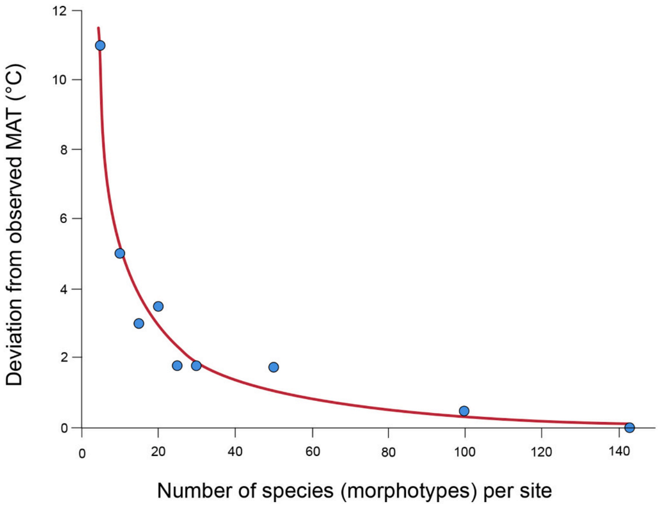


Fig. 10. Showing the relationship between the number of species (morphotypes) per sample and deviation from the observed mean annual temperature as predicted by CLAMP. Below 20 species the predictive uncertainty in CLAMP rises sharply. This is why 20 species is the minimum used in each calibration site.

standardised settings and so not typical of the conditions in the within- and sub-canopy space of forests. Moreover, recording stations are not uniformly distributed across the globe and tend to be located away from natural vegetation in towns, agricultural landscapes and airports, where the climate can be quite different from that in a naturally vegetated landscape (Spicer et al., 2011). To obtain climate data for modern vegetation calibration sites requires some form of interpolation. The degree to which these interpolations lead to data that match reality depends on recording station position and density, the interpolation scheme, and the spatial resolution of the interpolated grid. Further interpolation is required to translate the gridded data to actual sampling site location data. Unsurprisingly different climate data sets give rise to different CLAMP calibrations; a phenomenon that applies to all palaeoclimate proxies, but is seldom fully appreciated when compiling and comparing reconstructions of past climate.

3) Spatial and functional heterogeneity are inherent characteristics of all ecosystems, which often manifests as ’noise’ (Legendre, 1993). CLAMP attempts to accommodate variation within leaf form by including all leaf morphological variants, but it also tries to minimise ’noise’ by categorical scoring (for example for leaf size). This approach captures significant trait differences while downplaying small variations in measurements. However, noise due to factors other than those of interest (e.g. climate variables) cannot be eliminated altogether and this contributes to predictive uncertainty.

4) Calibration uncertainties can arise from errors in sampling vegetation and/or trait scoring. CLAMP sampling protocols require inclusion of all woody growth forms (shrubs, trees, vines etc.) and exclusion of any

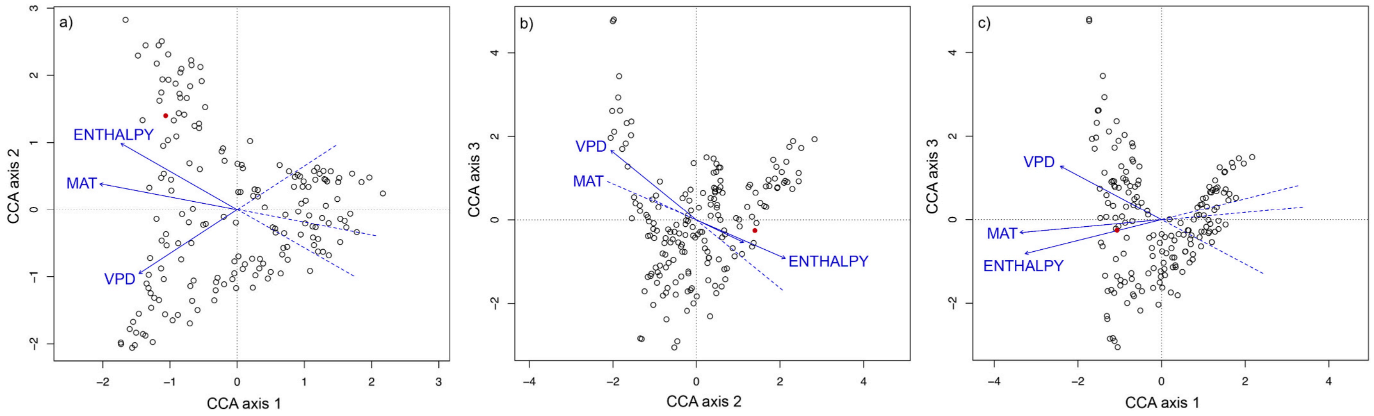


Fig. 11. CCA biplots for the PhysgAsia2 modern vegetation CLAMP calibration, with the physiognomic space occupied by modern vegetation sites shown by open circles and mean annual temperature (MAT), annual mean moist enthalpy (ENTHAL) and mean annual vapour pressure deficit (VPD) vectors shown as blue arrows. A fossil site is shown as the red filled circle. a) axes 1-2 space, b) axes 2-3 space and c) axes1-3 space.

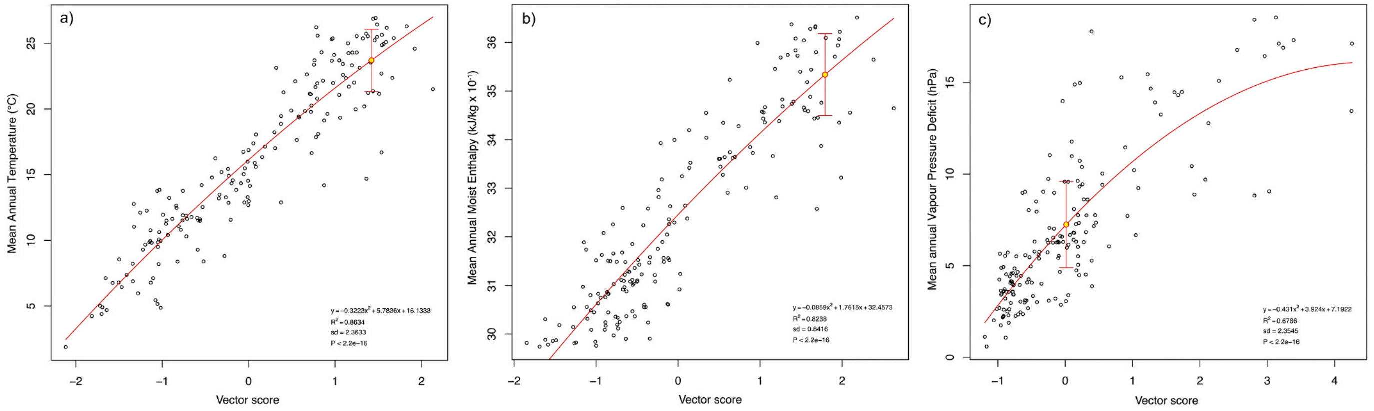


Fig. 12. Regression plots (2nd order polynomial) for a) mean annual temperature, b) annual mean moist enthalpy, c) mean annual vapour pressure deficit relating the vector score (relative position along the climate vector) to the observed value for the climate variable. The fossil site (the same as in Fig. 11) has a vector score, but no observed climate value so its position on the regression curve, shown by a yellow circle with 1 sd uncertainty bars, predicts the palaeoclimate on the Y axis. Regression plots for all CLAMP climate variables are given as Supplementary data.

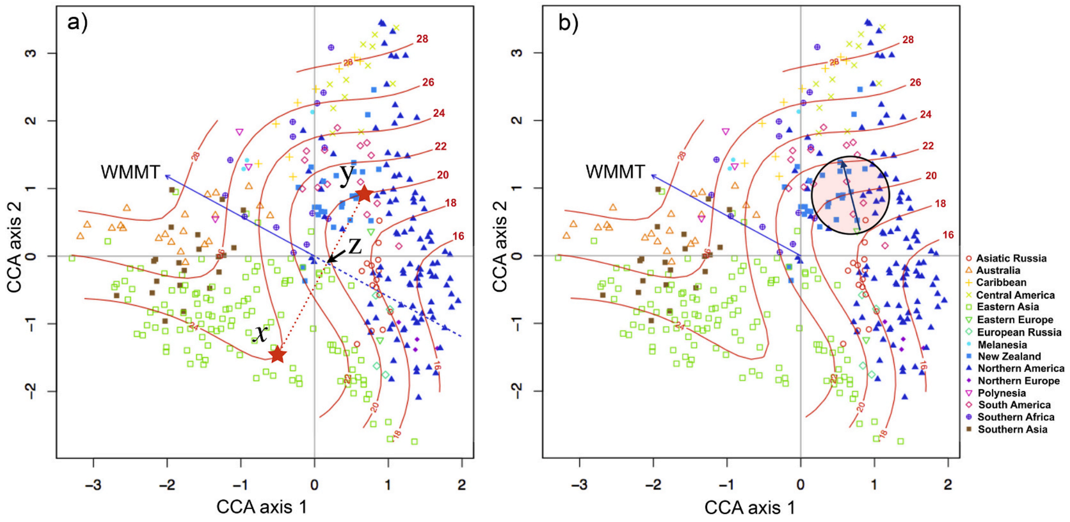


Fig. 13. CLAMP CCA axes 1-2 space showing the positions of 378 modern global vegetation sites coded for their geographic origins overlain by isotherms (°C) for the observed warm month mean temperatures (WMMT) at these sites and the direction of the CCA assigned WMMT vector. a) any fossil sites at either x or y locations will have the same vector score (z) but because the isotherms are curved the palaeo-WMMT predictions are ambiguous: they could be either 20 or 24 °C. b) a more accurate prediction could be obtained for y (as an example) if only the modern observations local to y in physiognomic space (i.e. within the circle) were considered. Within that circle the local WMMT has a different direction from that of the global set. Modified from Yang et al. (2015).

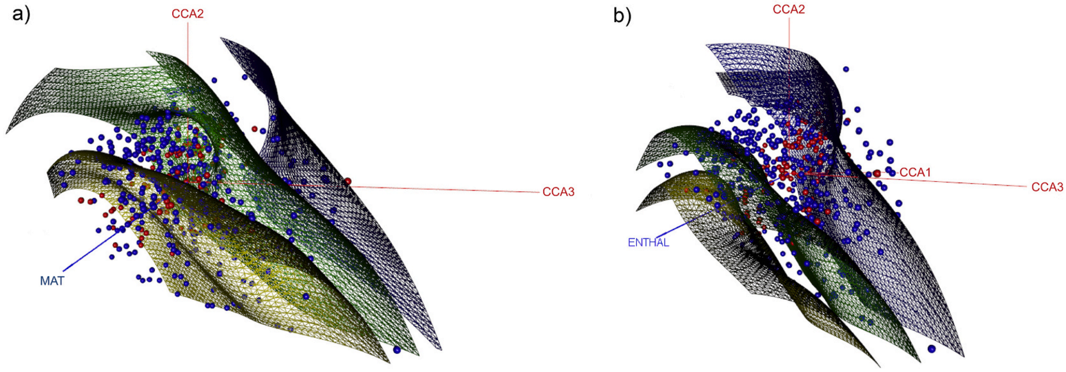


Fig. 14. a) Isopleth surfaces for CLAMP global 378 modern vegetation sites (blue spheres) in CCA axes 1-3 space using a General Additive Model (Wood, 2003). Red spheres represent a collection of fossil sites that do not contribute to defining the surfaces. a) Surfaces showing the 0 °C mean annual temperature (MAT) isotherm surface (blue), 10 °C MAT isotherm surface (green) and 20 °C isotherm surface (yellow). In practice there are numerous such surfaces each representing different temperatures between those shown. The CCA assigned MAT vector is shown as a blue line. b) isopleth surfaces for mean annual moist enthalpy showing the 300 kJ/kg surface (blue), 330 kJ/kg surface (green) and the 350 kJ/kg surface (yellow). In both cases the surfaces approximate to flat parallel plates (except near the margins of physiognomic space where they are less constrained) with the vectors normal to the plates. This indicates that the vectors are summarising the major trends reliably. Modified from Yang et al. (2015).

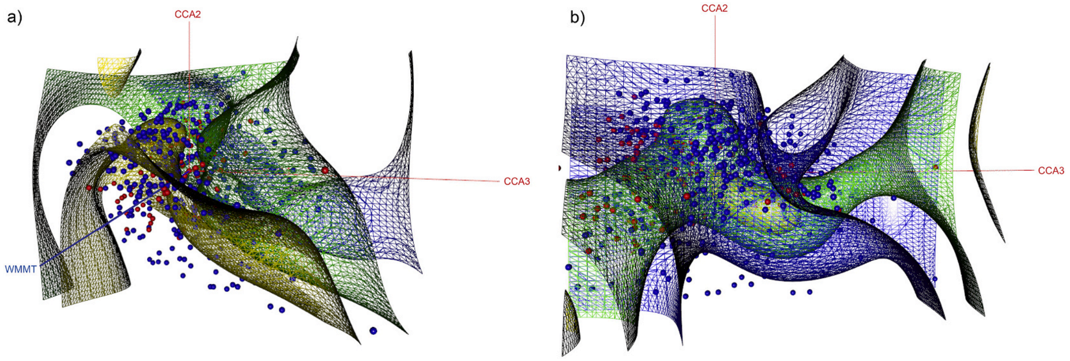


Fig. 15. General Additive Model isopleth surfaces for a) warm month mean temperature (WMMT - surfaces: blue 15 °C, green 20 °C and yellow 25 °C) and b) precipitation during the three consecutive driest months (surfaces: blue 10 cm, green 20 cm, yellow 30 cm). In these cases, vectors cannot adequately represent the complex primary trends in physiognomic space, so their predictive capacity is limited. Modified from Yang et al. (2015).

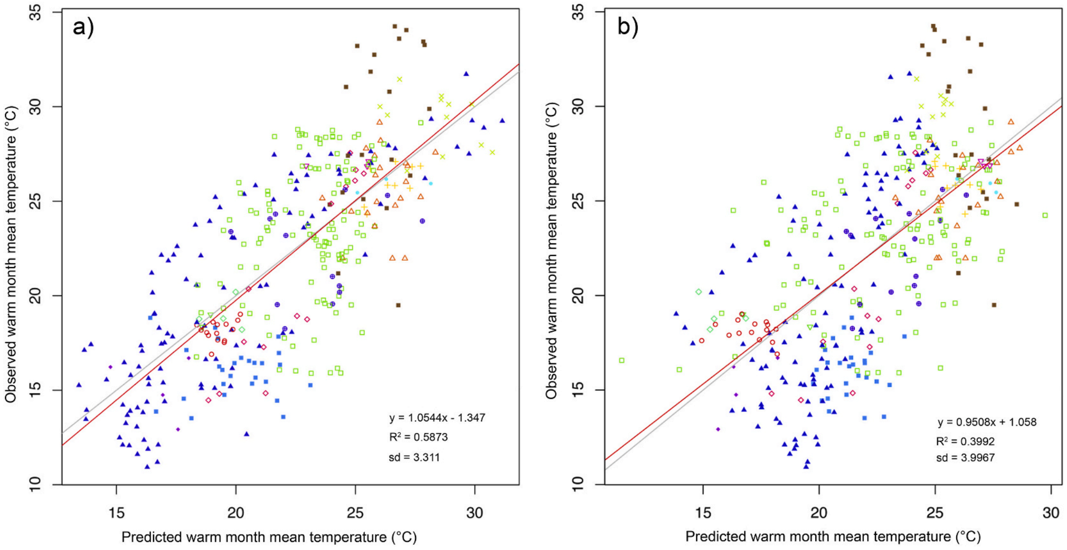


Fig. 16. Comparison of General Additive Model (GAM) versus CLAMP vector-based WMMT predictions. a) GAM predictions, b) conventional CLAMP vector-based predictions. The modern vegetation global calibration sites are coded as in Fig. 13. The grey line represents 1:1 agreement between the predicted and observed, the red line is the relationship yielded by the analysis. Note the slight improvement in R^2 and the reduction in scatter using GAM. Modified from Yang et al. (2015).

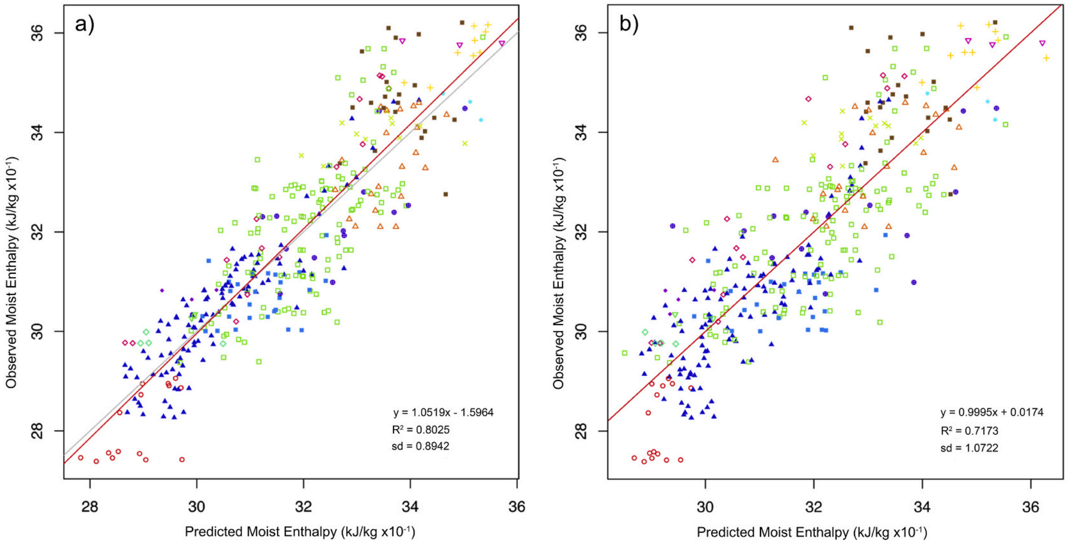


Fig. 17. Comparison of General Additive Model (GAM) versus CLAMP vector-based mean annual moist enthalpy predictions. a) GAM predictions, b) conventional CLAMP vector-based predictions. The modern vegetation global calibration sites are coded as in Fig. 13. The grey line represents 1:1 agreement between the predicted and observed, the red line is the relationship yielded by the analysis. Note marked improvement in R^2 and the reduction in scatter using GAM, but the introduction of a small slope divergence between the grey 1:1 line and the red line, which shows the relationship between observed and predicted. Modified from Yang et al. (2015).

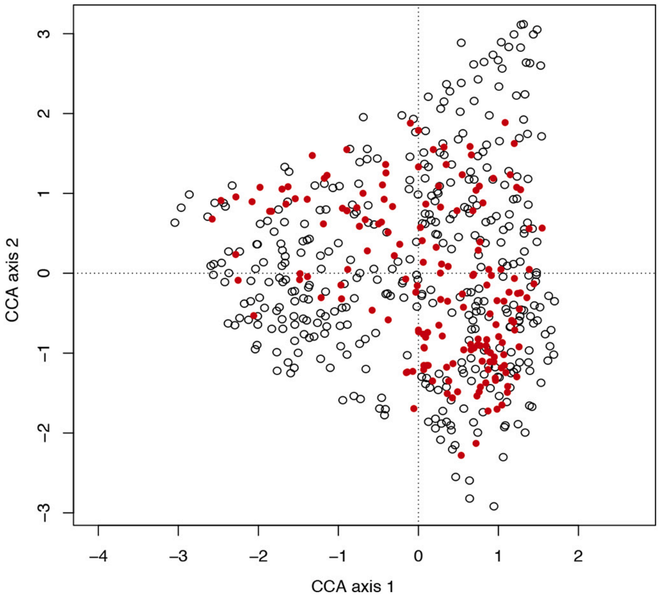


Fig. 18. CCA axes 1-2 plot of modern physiognomic space defined by the PhysgGlobal378 vegetation samples (black open circles) and WorldClim2-GRIDMET_Global378AZ_24var climate data, showing 170 fossils assemblages (closed red circles,) ranging in age from the mid Cretaceous to Pleistocene, all plotting within that space. The same is true for axis 3 (not shown), which displays no anomalous fossil outliers. This demonstrates no unusual trait combinations occurred in the past, suggesting no 'extinct climates' existed for which novel trait combinations were required.

of these growth forms contributes to calibration error (Burnham et al., 2001). This demands direct field sampling from limited areas following strict protocols. Gregory-Wodzicki (2000) pointed out the wisdom of adopting this strategy in contrast to less constrained data derived from floral manuals that include taxa distributed over large, often undefined regions (e.g. Wilf et al., 1998, 1999). If the modern vegetation sites used to calibrate CLAMP are not appropriately sampled, then this will lead to increased uncertainty or bias. Similarly, errors in trait scoring will degrade predictive ability. Ideally all scoring should be done as a collective exercise involving several scorers. Scorer identities should be entered on the score sheets.

While evergreen taxa have leaf cohorts that are exposed to climate year-round, and so can be expected to yield reliable annual averages for climate parameters, deciduous taxa typically have a dormant period when the plant may not code for local climate. However, during their expansion phase leaves do have to be able to cope with rapidly warming spring conditions and so tend to encapsulate, to some extent, winter cold (Spicer et al., 2004). Similarly, leaves adapted to spring soil hydrological conditions have to be able to survive changed circumstances (droughts or excess water) that may occur later in the year. This means that over time selection gives rise to leaf adaptations capable of dealing with typical conditions normally experienced at different times of the year.

Naturally, as with NLR proxies, some taxa may experience mutations, which change their trait integrations over time and thus, potentially, their environmental optimisation. If the leaf cohort ceases to perform optimally it may ultimately lead to extinction of that genome from that location, but in less severe cases it will just result in a subtle shift in that taxon's contribution to the overall trait spectrum at that location. By incorporating a minimum of 20 morphotypes (= species) for any given vegetation sample (living or fossil) these effects are minimised and fall within the normal statistical uncertainties considered above. This means they have negligible impact on CLAMP's predictive outcome.

3.7. Limitations of CLAMP

Restricting CLAMP sampling of modern vegetation to small areas at the same elevation improves predictive precision (Wolfe and Uemura, 1999), but this, coupled with the observation that leaf assemblages also tend to reflect plants growing close to the site of deposition, means that CLAMP tends to return local or even microclimate conditions. This is in contrast to NLR-derived palaeoclimates when they are based on palynomorphs, which are traditionally thought of as representing regional climate. Moreover, CLAMP can only be used on fossil leaf assemblages, and then only when the assemblages comprises 20 morphotypes or more. This limits its application compared to the NLR approach, which can be applied to any identifiable plant parts. The exclusion of

Table 2
A summary of CLAMP calibration data sets, their Supplementary file numbers, and their properties.

Calibration Dataset Name	Description
Physg3arAZ.csv (File_S1)	Alphabetically ordered, comma delimited, physiognomic file of 173 modern vegetation sites mostly from North America, Japan, Caribbean, western Pacifica.
MET3arAZ_11var.csv (File_S2)	Meteorological point data file for 11 climate variables derived from weather recording stations proximal to the vegetation sites. No corrections for missing data or equipment failure. Variable measurement periods. Useable with Physg3arAZ.csv.
GRIDMet_3arAZ_11var.csv (File_S3)	Gridded meteorological data based on New et al. (1999) for 11 variables measured between 1961 and 1990. The spatial resolution of the grid is 0.5° with interpolation to the exact vegetation site co-ordinates (Spicer et al., 2009). Useable with Physg3arAZ.csv.
CRU2160_GRIDMet_3arAZ_11var.csv (File_S4)	High spatial resolution (10’) gridded meteorological data based on New et al. (2002) for 11 variables measured between 1961 and 1990 with interpolation to the exact vegetation site co-ordinates. Useable with Physg3arAZ.csv.
WorldClim2_GRIDMet_3arAZ_11var.csv (File_S5)	Worldclim2 meteorological data for 11 climate variables measured between 1970 and 2000 at a spatial resolution of 30" (~ 1km$^{2)}$ (Fick and Hijmans, 2017) with interpolation to the exact vegetation site co-ordinates. Useable with Physg3arAZ.csv.
WorldClim2_GRIDMet_3arAZ_24var.csv (File_S6)	Worldclim2 meteorological data for 24 climate variables, otherwise as above.
Physg3brAZ.csv (File_S7)	Alphabetically ordered, comma delimited, physiognomic file of 144 modern vegetation sites mostly from North America, Japan, Caribbean, western Pacifica. This excludes the so called ’alpine nest’ vegetation adapted to cold temperatures.
MET3brAZ_11var.csv (File_S8)	Meteorological point data file for 11 climate variables derived from weather recording stations proximal to the vegetation sites. No corrections for missing data or equipment failure. Variable measurement periods. Useable with Physg3brAZ.csv.
GRIDMet_3brAZ_11var.csv (File_S9)	Gridded meteorological data based on New et al. (1999) for 11 variables measured between 1961 and 1990. The spatial resolution of the grid is 0.5° with interpolation to the exact vegetation site co-ordinates (Spicer et al., 2009). Useable with Physg3brAZ.csv.
CRU2160_GRIDMet_3brAZ_11var.csv (File_S10)	High spatial resolution (10’) gridded meteorological data based on New et al. (2002) for 11 variables measured between 1961 and 1990 with interpolation to the exact vegetation site co-ordinates. Useable with Physg3brAZ.
WorldClim2_GRIDMet_3brAZ_11var.csv (File_S11)	Worldclim2 meteorological data for 11 climate variables measured between 1970 and 2000 at a spatial resolution of 30" (~ 1km^2) (Fick and Hijmans, 2017) with interpolation to the exact vegetation site co-ordinates. Useable with Physg3brAZ.csv.
WorldClim2_GRIDMet_3brAZ_24var.csv (File_S12)	Worldclim2 meteorological data for 24 climate variables, otherwise as above.
PhysgAsia1AZ.csv (File_S13)	Alphabetically ordered, comma delimited, physiognomic file of 189 modern vegetation sites mostly from North America, Japan, Caribbean, western Pacifica, China, India, Thailand. It includes monsoon adapted vegetation.
CRU2160_GRIDMet_Asia1AZ_11var.csv (File_S14)	High spatial resolution (10’) gridded meteorological data based on New et al. (2002) for 11 variables measured between 1961 and 1990 with interpolation to the exact vegetation site co-ordinates. Useable with PhysgAsia1AZ.csv.
PhysgAsia2AZ.csv (File_S15)	Alphabetically ordered, comma delimited, physiognomic file of 177 modern vegetation sites mostly from North America, Japan, Caribbean, western Pacifica, China, India, Thailand. It includes monsoon adapted vegetation, but excludes cold-adapted vegetation.
CRU2160_GRIDMet_Asia2AZ_11var.csv (File_S16)	High spatial resolution (10’) gridded meteorological data based on New et al. (2002) for 11 variables measured between 1961 and 1990 with interpolation to the exact vegetation site co-ordinates. Useable with PhysgAsia2AZ.csv.
WorldClim2_GRIDMet_Asia2AZ_11var.csv (File_S17)	Worldclim2 meteorological data for 11 climate variables measured between 1970 and 2000 at a spatial resolution of 30" (~ 1km^2) (Fick and Hijmans, 2017) with interpolation to the exact vegetation site co-ordinates. Useable with PhysgAsia2AZ.csv.
WorldClim2_GRIDMet_Asia2AZ_24var.csv (File_S18)	Worldclim2 meteorological data for 24 climate variables, otherwise as above.
PhysgSHAZ.csv (File_S19)	Alphabetically ordered, comma delimited, physiognomic file of 90 modern vegetation sites mostly from Australia, New Zealand Bolivia and Argentina (Kennedy et al., 2014).
CRU2160_GRIDMet_SH_11var.csv (File_S20)	High spatial resolution (10’) gridded meteorological data based on New et al. (2002) for 11 variables measured between 1961 and 1990 with interpolation to the exact vegetation site co-ordinates. Useable with PhysgAsia2AZ.csv.
WorldClim2_GRIDMet_SH_11var.csv (File_S21)	Worldclim2 meteorological data for 11 climate variables measured between 1970 and 2000 at a spatial resolution of 30" (~ 1km^2) (Fick and Hijmans, 2017) with interpolation to the exact vegetation site co-ordinates. Useable with PhysgSH.csv.
WorldClim2_GRIDMet_SH_24var.csv (File_S22)	Worldclim2 meteorological data for 24 climate variables, otherwise as above.
PhysgGlobal378AZ.csv (File_S23)	Alphabetically ordered, comma delimited, physiognomic file of 378 modern vegetation sites mostly from all continents except Antarctica and excluding equatorial rain forests.
CRU2160_GRIDMet_Global 378AZ_11var.csv (File_S24)	High spatial resolution (10’) gridded meteorological data based on New et al. (2002) for 11 variables measured between 1961 and 1990 with interpolation to the exact vegetation site co-ordinates. Useable with PhysgGlobal378AZ.csv.
WorldClim2_GRIDMet_Global378AZ_11var.csv (File_S25)	Worldclim2 meteorological data for 11 climate variables measured between 1970 and 2000 at a spatial resolution of 30" (~ 1km^2) (Fick and Hijmans, 2017) with interpolation to the exact vegetation site co-ordinates. Useable with Physg Global378AZ.csv.
WorldClim2_GRIDMet_ Global378AZ _24var.csv (File_S26)	Worldclim2 meteorological data for 24 climate variables, otherwise as above.

micromorphological characters from CLAMP means, however, that it can be applied to even imperfectly preserved impression fossils. Other limitations are those that arise because there are still gaps in the CLAMP calibration space. At present there are too few calibration samples from South America and equatorial forests worldwide.

European vegetation is also generally under-represented. Most of the natural lowland vegetation across Europe no longer exists or is highly disturbed by management practices dating back centuries, such as coppicing, and in mountain regions where appropriate vegetation does survive the taxon count is often less than the 20 required. To try and develop a CLAMP-like physiognomic tool for Europe Traiser (2004); Traiser et al. (2004), Traiser and Mosbrugger, 2004, and Traiser et al. (2007) assembled a ’synthetic’ data set of hardwood leaf trait data from European floral records (i.e. not directly field-collected following CLAMP protocols), and used a scoring scheme similar, but not identical, to that of CLAMP. They paired this physiognomic array with gridded climate data at a 0.5° x 0.5° spatial resolution. Using simple and multiple linear regressions, they found strong correlations between leaf traits and the thermal regime, but no correlation with precipitation variables such as mean annual precipitation. This broadly matches CLAMP data where the ability to estimate precipitation carries large uncertainties. Innovatively, they then used cluster analysis to identify ’leaf physiognomic communities’ or ’classes’, and showed an increase in trait diversity in more variable climate regimes. For example, Mediterranean vegetation showed higher trait diversity than vegetation in more northerly settings where the hydrological and thermal regime is more uniform.

3.8. *Criticisms of CLAMP*

The development of CLAMP has been a collective effort involving researchers worldwide. Early CLAMP calibration data involved colleagues in Japan and Puerto Rico (Wolfe, 1993), and since then the team creating CLAMP has expanded considerably, as is acknowledged on the CLAMP website. As the number of researchers contributing to the technique has grown so too has the intellectual capital invested in the CLAMP, and scrutiny by those contributing to its development. Nevertheless, in science it is essential that ideas are tested and understandably CLAMP has been challenged from time to time. Typically, these challenges are in the form of claims of some form of bias in CLAMP's predictive capabilities (e.g. Peppe et al., 2010), and while there are acknowledged improvements to be made in terms of precision (likely achieved through analytical advances), and the calibration data still have significant gaps (e.g. the lack of equatorial rainforest samples), the testing of CLAMP has to be compliant with the methodology. To-date CLAMP challenges have not gained traction because of flawed implementation or violation of the CLAMP protocols.

An early claim was that many of the CLAMP traits should be eliminated because they exhibited statistically insignificant relationships with climate (e.g. Gregory, 1994; Gregory and Chase, 1992; Wing and Greenwood, 1993), but this conclusion was reached using multiple regression, which assumes trait and climate variables are both independent and normally distributed. As we have shown, this is not the case. Traits have complex and inter-dependant relationships with climate variables, which themselves are co-dependant through the laws of thermodynamics and conservation of energy. Moreover, these individual trait/climate relationships likely change with time and place, but the aggregate trait spectrum/climate relationship must remain optimised or evolutionary fitness will be compromised, and less well adapted (poorly optimised) leaf forms will be eliminated by natural selection. CLAMP examines aggregate trait spectra, not individual traits, and to test CLAMP by ignoring this is to misunderstand CLAMP's underlying premise that leaves are integrated systems. It is also worth remembering that some apparent redundancy is beneficial because it confers resilience to taphonomic data loss (Spicer et al., 2005).

Other challenges have been flawed because they violated the CLAMP protocols, usually by excluding important woody dicot growth forms (e.g. vines and/or shrubs) from test data (Wilf, 1997), not collecting test data directly from the field but using a mix of field sampling and floral lists/manuals (Traiser et al., 2004, 2007; Wilf, 1997; Wilf et al., 1998 - challenged by Wolfe and Uemura, 1999), or deliberately not following the CLAMP scoring scheme for test data, while at the same time analysing such data using the standard CLAMP calibration data sets (Peppe et al., 2010). In this instance Peppe et al. (2010) used continuous (mm^2) measurement of leaf size in test data, but did not recalibrate CLAMP by rescoring the modern vegetation training data the same way. Inevitably CLAMP analysis calibrated using categorical size data gave erroneous results for the differently scored test data. A more detailed critique of this, and other shortcomings in the work, was given in Spicer and Yang (2010).

A more understandable form of scoring error is, however, of genuine concern. This is error that arises when novice scorers misinterpret the scoring scheme. A few authors have expressed the opinion that in some instances the character definitions, as originally expressed, were ambiguous (Green, 2006; Wiemann et al., 1998; Wilf, 1997). As a result of these concerns the character definitions were rewritten and the scoring procedure made more explicit on the CLAMP website. Subsequent tests using multinational groups of novice scorers (Spicer and Yang, 2010), and the application by authors independent of the CLAMP development group (e.g. Clifton, 2012; Demers-Potvin and Larsson, 2019; Uhl et al., 2007), suggest this now is less of an issue, and the introduction of the automated scoresheet has further lessened the opportunity for scoring error.

3.9. CLAMP applications

Early applications of CLAMP were mostly confined to temperature and precipitation determinations for fossil assemblages in western North America. As we have seen the first application was to study climate change across the Cretaceous-Paleogene interval (Wolfe, 1990), followed by Paleogene and Neogene assemblages from the western USA including Alaska (Wolfe, 1994, 1994b, 1994c; Wolfe, 1995). These assemblages were subsequently revisited when CLAMP Online was launched (Table 1) and the relevant calibrations had stabilised (Yang et al., 2011). Subsequently CLAMP has been applied widely to determine the thermal and hydrological regimes of fossil assemblages ranging from the Cenomanian to Pleistocene (Demers-Potvin and Larsson, 2019; Fletcher et al., 2014; Godefroit et al., 2009; Golovneva, 2000; Hazra et al., 2020; Herman et al., 2016; Herman and Spicer, 1996, 1997, 2010; Kennedy et al., 2002; Kvaček et al., 2019; Moiseeva et al., 2009; Parrish et al., 1998; Spicer and Herman, 2010; Spicer et al., 2008; Spicer et al., 2002; Spicer et al., 2014; Spicer et al., 2019; Tomsich et al., 2010; Uhl et al., 2007; Velasco-de Leon et al., 2010; Yang et al., 2007; Zolina et al., 2020).

3.9.1. Palaeoaltimetry

As well as reconstructing past temperature and hydrological regimes, CLAMP has found significant application in palaeoaltimetry (e.g. Ding et al., 2017; Jacques et al., 2014; Khan et al., 2014; Song et al., 2020; Spicer et al., 2003; Wolfe et al., 1998). The history of using plants to determine the heights of ancient land surfaces is reviewed in Spicer (2018) and can be divided into two approaches: 1) using thermal lapse rates and, 2) applying conservation of energy principles.

The thermal lapse rate approach exploits the observation that, in general, the average temperature of the atmosphere declines with increasing altitude. For geological proxies it is the altitudinal thermal gradient near the ground, not in the free atmosphere, that is preserved, but this 'terrestrial lapse rate' is very much location dependant. Unless modelling using different palaeogeographies and topographies is used to explore landscape-related lapse rate variation, surface heights derived from thermal lapse rates are likely to be erroneous.

The energy conservation approach was devised by Forest et al. (1995), who then explored its application in CLAMP (Forest, 1996; Forest et al., 1999) using fossil assemblages in western North America. The basis of this approach is that as a parcel of air rises over a mountain front it both cools and its humidity increases, but overall the moist static energy (h) it contains remains the same. Moist static energy is made up of two components: moist enthalpy (H) and potential energy (gZ), where g is the acceleration due to gravity (a constant) and Z is height:

$$h = H + gZ \tag{1}$$

To find the height of an unknown location above sea level this expression (1) can be re-arranged:

$$Z = \left(H_{sea\ level} - H_{high}\right)/g \tag{2}$$

This means that the elevation of the H_{high} location can be determined if we know the moist enthalpy there, and we know the moist enthalpy at sea level nearby. For the caveats for this approach see Spicer (2018). As we have seen (Figs. 12, 14 and 17) CLAMP regressions show that moist enthalpy is well coded for in leaf form and is now often used in palaeoaltimetry.

3.9.2. Monsoons

Monsoon climates present plants with severe challenges; for part of the year there is an over-abundance of water potentially drowning root systems and creating the continual threat of bacterial and fungal entry to the leaf if its surface is not kept dry, while in other parts of the year severe drought and high temperatures require water retention and leaf cooling strategies. For evergreen plants, balancing these opposing demands is particularly acute. Unsurprisingly then, low latitude monsoon-adapted leaves display trait spectra that are characteristic of different monsoon types (Spicer et al., 2016, 2017). Early CLAMP calibrations largely ignored monsoon climates because they were focussed on Paleogene and Neogene palaeoclimate determinations in North America (Wolfe, 1990; Wolfe, 1994, 1994b, 1994c; Wolfe, 1995); and the Arctic (Golovneva, 2000; Herman and Spicer, 1996, 1997; Parrish et al., 1998). However, when CLAMP was used to determine the palaeoelevation of Tibet (Spicer et al., 2003), and lowlands sites more generally across China (e.g. Yang et al., 2007), it became clear that monsoon-adapted vegetation had to be included in the CLAMP calibration system. Largely through the work of Su Tao and Frédéric Jacques, in Kunming at that time, and Yang Jian in Beijing, a collection of over 50 modern vegetation samples across China was added to the Physg3arc and Physg3brc datasets and the first explorations of CLAMP in the context of monsoon climates using a new PhysgAsia1 calibration were undertaken (Jacques et al., 2011). Subsequently sites from India and Thailand were added to form PhysgIndia1 and PhysgAsia2 calibrations and applied to Asian fossil floras (Bhatia et al., 2021; Ding et al., 2017; Hazra et al., 2020; Herman et al., 2017; Jacques et al., 2014; Jin et al., 2017; Khan et al., 2014, 2018, 2019; Shukla et al., 2014; Song et al., 2020; Spicer et al., 2016, 2017; Srivastava et al., 2012; Su et al., 2019; Xu et al., 2017). Because most deep time Asian climates are warmer than those of the present, the PhysgAsia2 calibrations is the most widely applied because this excludes cold sites from northern China, and delivers results consistent with those derived from stable isotopes (Currie et al., 2005, 2016; Pollisar et al., 2009; Xu et al., 2017).

4. Future work

As the number of CLAMP calibration sites increase and represent more diverse climates it becomes increasingly problematic to use vectors to summarise major trends in the leaf form/climate relationship. Ideally CLAMP would consist of a single calibration dataset, but when global calibration data are used with vectors the predictive precision declines markedly compared to more local calibrations. The current number of modern vegetation sites in the total CLAMP database is 424 (not presented here), yet there are still key areas of the world not represented such as the equatorial forests of the Amazon, Africa and Southeast Asia, western Australia, and South America other than a few sites in Argentina and Bolivia. Hopefully modern vegetation from these areas will be sampled in the future before all the primary vegetation is destroyed.

As the CLAMP calibration database increases it is also being applied in different ways. Recently it has been used to investigate spatial and temporal patterns of stomatal conductance and water use efficiency (Murray et al., 2019, 2020; Soh et al., 2019). All CLAMP data that have been checked for accuracy and adherence to sampling protocols are publicly available via the CLAMP website. Future developments of the CLAMP Online site will include the addition of the WorldClim2 data with its expansion of the number of climate variables, and the addition of new calibration data as they become available.

It has taken more than a century to develop protocols for understanding leaf form/climate relationships that allow meaningful comparisons across different vegetation types and climate space. As far as we know CLAMP represents the largest foliar physiognomic dataset collected directly from the field using well-defined sampling protocols and consistent trait scoring, yet more sites need to be added urgently before more climate change and human disturbance compromises our ability to make consistent collections.

Further improvements are likely to come from developing CLAMP in tandem with numerical climate and vegetation modelling. Climate modelling has long been used in conjunction with CLAMP to enabled novel determinations of palaeoelevation (Khan et al., 2014; Song et al., 2020; Spicer et al., 2003), but ideally this could be extended and used to better understand the application of thermal lapse rates, which would then allow combined CLAMP and NLR surface height determinations. It may also be possible to use models to predict foliar physiognomy directly from climate parameters, which would also provide new tools for model validation and CLAMP refinement.

5. Conclusions

Although it has been recognised for over a century that functionally optimised macroscopic leaf form (physiognomy) can be used as a palaeoclimate proxy, the methodology by which this can be achieved has evolved considerably. Early attempts focused on the single trait of the presence or absence of margin teeth as an indicator of the prevailing thermal regime, at fist qualitatively and then quantitatively. Leaf size was also explored as a proxy for water availability. However, such relationships have proved unreliable because of spatial (and presumably temporal) variability tied to biogeographic history.

The recognition that leaves function as integrated systems where all traits are highly correlated demands a more multivariate approach that over 30 years or so has seen the progressive development of the Climate-Leaf Multivariate Program (CLAMP). The introduction of field sampling protocols, honing of trait scoring and improvements in climate data processing, as well as improvements in the underlying statistical methodology have resulted in a palaeoclimate tool that can yield climate and altitudinal data comparable with geochemical proxies, but without problems linked to diagenetic alteration. Future CLAMP development needs to centre on 1) increasing foliar physiognomic data from South America and equatorial forests before they become too degraded, and 2) integrating foliar physiognomy with quantitative environmental (climate) modelling. This will refine the predictive capabilities of both approaches for understanding better the evolution of Earth systems and their different sensitivities to perturbation.

Funding information

This work was funded by the joint UK-China Project administered by the UK Natural Environment Research Council (NE/P013805/1) and the Natural Science Foundation of China Project (No. 41661134049) and an XTBG Visiting Scholarship to RAS.

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgements

We thank all the contributors to collecting the CLAMP calibration data (listed on the CLAMP website), and students who have helped with testing the scoring procedures.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.palaeo.2020.110138.

References

Ashton, P.S., Hall, P., 1992. Comparisons of structure among mixed dipterocarp forests of north-western Borneo. J. Ecol. 80, 459–481.

Ashton, P.M.S., Olander, L.P., Berlyn, G.P., Thadani, R., Cameron, I.R., 1998. Changes in leaf structure in relation to crown position and tree size of *Betula papyrifera* within fire-origin stands of interior cedar-hemlock. Can. J. Bot. 76, 1180–1187.

Axelrod, D.I., Bailey, H.P., 1976. Tertiary vegetation, climate, and altitude of the Rio Grande depression, New Mexico–Colorado. Paleobiology 2, 235–254.

Baas, P., Wheeler, E.A., 2011. Wood anatomy and climate change. In: Hodkinson, T.R., Jones, M.B., Waldren, S., Parnell, J.A.N. (Eds.), Climate Change, Ecology and Systematics. Cambridge University Press. Cambridge, pp. 141–155.

Bailey, I., Sinnott, E., 1915. A botanical index of Cretaceous and Tertiary climates. Science 41, 831–834.

Bailey, I., Sinnott, E., 1916. The climatic distribution of certain types of angiosperm leaves. Am. J.Bot. 3, 24–39.

Baker-Brosh, K.F., Peet, R.K., 1997. The ecological significance of lobed and toothed leaves in temperate forest trees. Ecology 78, 1250–1255.

Beadle, N.C.W., 1954. Soil phosphate and the delimitation of plant communities in eastern Australia. Ecology 35, 370–375.

Benzecri, J.P., 1973. L'analyse des données: L'analyse des correspondences. Dunod, Paris.

Bhatia, H., Srivastava, G., Spicer, R.A., Farnsworth, A., Spicer, T.E.V., Mehrotra, R.C., Paudayal, K., Valdes, P.J., 2021. Leaf physiognomy records the Miocene intensification of the South Asia Monsoon. Glob. Planet. Chang. 196 https://doi.org/10.1016/j.gloplacha.2020.103365, 103365.

Brodribb, T.J., Jordan, G.J., 2012. Water supply and demand remain balanced during leaf acclimation of *Nothofagus cunninghamii* trees. New Phytol. 192, 437–448.

Brodribb, T.J., Feild, T.S., Jordan, G.J., 2007. Leaf maximum photosynthetic rate and venation are linked by hydraulics. Plant Physiol. 144, 1890–1898.

Brown, V.K., Lawton, J.H., 1991. Herbivory and the evolution of leaf size and shape. Philosophical Transactions: Biological Sciences 333, 265–272.

Burnham, R.J., Pitman, N.C.A., Johnson, K.R., Wilf, P., 2001. Habitat-related error in estimating temperatures from leaf margins in a humid tropical forest. Am. J. Bot. 88, 1096–1102.

Chapin, F.S., 1980. The mineral nutrition of wild plants. Annu. Rev. Ecol. Syst. 11, 233–260.

Clifton, A., 2012. The Eocene flora of Svalbard and its climatic significance. PhD thesis University of Leeds 1–401.

Cordell, S., Goldstein, G., Mueller-Dombois, D., Webb, D., Vitousek, P.M., 1998. Physiological and morphological variation in *Metrosideros polymorpha*, a dominant Hawaiian tree species, along an altitudinal gradient: the role of phenotypic plasticity. Oecologia 113, 188–196.

Cowling, R.M., Campbell, B.M., 1980. Convergence in vegetation structure in the Mediterranean communities of California, Chile and South Africa. Vegetatio 43, 191–197.

Currie, B.S., Rowley, D., Tabor, N.J., 2005. Middle Miocene paleoaltimetry of southern Tibet: implications for the role of mantle thickening and delamination in the Himalayan orogen. Geology 33, 181–184.

Currie, B.S., Polissar, P.J., Rowley, D.B., Ingalls, M., Li, S., Olack, G., Freeman, K.H., 2016. Multiproxy palaeoaltimetry of the late Oligocene-Pliocene Oiyug Basin, Southern Tibet. Am. J. Sci. 316, 401–436.

Daly, R.G., Gastaldo, R.A., 2010. The effect of leaf orientation to sunlight on stomatal parameters of *Quercus rubra* around the Belgrade Lakes, central Maine. Palaios 25, 339–346.

Demers-Potvin, A., Larsson, H.C.E., 2019. Palaeoclimatic reconstruction for a Cenomanian-aged angiosperm flora near Schefferville. Labrador. Palaeontology 62, 1027–1048.

Dilcher, D.L., 1973. A paleoclimatic interpretation of the Eocene floras of south-eastern North America. In: Graham, A. (Ed.), Vegetation and vegetational history of northern Latin America. Elsevier, Amsterdam, pp. 39–59.

Ding, L., Spicer, R.A., Yang, J., Xu, Q., Cai, F., Li, S., Lai, Q., Wang, H., et al., 2017. Quantifying the rise of the Himalaya orogen and implications for the South Asian monsoon. Geology 45, 215–218.

Dolph, G.E., 1977. The effect of different calculational techniques on the estimation of leaf area and the construction of leaf size distributions. Bull. Torrey Botanical Club 104, 264–269.

Dolph, G.E., Dilcher, D.L., 1980. Variation in leaf size with respect to climate in Costa Rica. Biotropica 12, 91–99.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to quantitative genetics. Addison Wesley. Longman, Harlow.

Feild, T.S., Sage, T.L., Czerniak, C., Iles, W.J.D., 2005. Hydathodal leaf teeth of *Chloranthus japonicus* (Chloranthaceae) prevent guttation-induced flooding of the mesophyll. Plant Cell Environ. 28, 1179–1190.

Fick, S.E., Hijmans, R.J., 2017. WorldClim2: new 1-km spatial resolution climate surfaces for global land surfaces. Int. J. Climatol. 37, 4302–4315.

Fletcher, T.L., Greenwood, D.R., Moss, P.T., Salisbury, S.W., 2014. Paleoclimate of the Late Cretaceous (Cenomanian-Turonian) portion of the Winton Formation, Central-Western Queensland, Australia: New Observations based on CLAMP and Bioclimatic Analysis. Palaios 29, 121–128. http://www.bioone.org/doi/full/10.2110/palo.2013.080.

Forest, C.E., 1996. Paleoaltimetry incorporating atmospheric physics and botanical estimates of paleoclimate. Meteorology PhD dissertation. Massachusetts Institute of. Technology. 152.

Forest, C.E., Molnar, P., Emanuel, K.A., 1995. Palaeoaltimetry from energy conservation principles. Nature 374, 347–350.

Forest, C.E., Wolfe, J.A., Molnar, P., Emmanuel, K.A., 1999. Paleoaltimetry incorporating atmospheric physics and botanical estimates of paleoclimate. Geol. Soc. Am. Bull. 111, 497–511.

Givnish, T.J., 1979. On the adaptive significance of leaf form. In: Solbrig, O.T., et al. (Eds.), Topics in plant population biology. Columbia University Press, New York, pp. 375–407.

Givnish, T.J., 1984. Leaf and canopy adaptations in tropical forests. In: Medina, E, Mooney, HA, Vazquez-Yanes, C, eds. Physiological Ecology of Plants of the Wet Tropics. The Hague, the Netherlands: Junk, 51–84.

Givnish, T.J., 1987. Comparative studies of leaf form: Assessing the relative roles of selective pressures and phylogenetic constraints. New Phytol. 106, 131–160.

Givnish, T.J., 1990. Leaf mottling: relation to growth form and leaf phenology and possible role of camouflage. Funct. Ecol. 4, 463–474.

Godefroit, P., Golvneva, L.B., Shchepetov, S., Garcia, G., Alekseev, P., 2009. The last polar dinosaurs: high diversity of latest Cretaceous Arctic dinosaurs in Russia. Naturwissenschaften 96, 495–501.

Golovneva, L.B., 2000. The Maastrichtian (Late Cretaceous) climate in the Northern Hemisphere. In: Hart, M.B. (Ed.), Climates: Past and Present, 181, pp. 43–54. Geological Society of London, Special Publications.

Green, W.A., 2006. Loosening the CLAMP: An exploratory graphical approach to the Climate Leaf Analysis Multivariate Program. Palaeontol. Electron. 9, 1–17.

Greenwood, D.R., Moss, P., Rowett, A., Vadala, A.J., Keefe, R., 2003. Plant communities and climate change in southeastern Australia during the Early Paleogene. In: Wing, S., Gingerich, P., Schmitz, B., Thomas, E. (Eds.), Causes and Consequences of Globally Warm Climates in the Early Paleogene, 369, pp. 365–390. Geological Society of America Special Paper, Boulder, Colorado.

Greenwood, D.R., Wilf, P., Wing, S.L., Christophel, D.C., 2004. Paleotemperature estimation using leaf margin analysis: is Australia different? Palaios 19, 129–142.

Greenwood, D.R., Archibald, S.B., Mathewes, R.W., Moss, P.T., 2005. Fossil biotas from the Okanagan Highlands, southern British Columbia and northeastern Washington State: climates and ecosystems across an Eocene landscape. Can. J. Earth Sci. 42 (2), 167–185.

Gregory, K.M., 1994. Paleoclimate and paleoelevation of the 35 Ma Florissant flora, Front Range, Colorado. Paleoclimates 1, 23–57.

Gregory, K.M., Chase, C.G., 1992. Tectonic significance of paleobotanically estimated climate and altitude of the late Eocene erosion surface, Colorado. Geology 20, 581–585.

Gregory-Wodzicki, K.M., 2000. Relationships between leaf morphology and climate, Bolivia: implications for estimating paleoclimate from fossil floras. Paleobiology 26, 668–688.

Grimm, G.W., Denk, T., 2012. Reliability and resolution of the coexistence approach — a revalidation using modern-day data. Rev. Palaeobot. Palynol. 172, 33–47.

Hall, J.B., Swaine, M.D., 1981. Distribution and ecology of vascular plants in a tropical rain forest. Dr. W. Junk, The Hague.

Halloy, S.R.P., Mark, A.F., 1996. Comparative leaf morphology spectra of plant communities in New Zealand, the Andes and the European Alps. J. R. Soc. N. Z. 26, 41–78.

Hazra, T., Spicer, R.A., Hasra, M., Mahato, S., Spicer, T.E.V., Bera, S., Valdes, P.J., Farnsworth, F., Hughes, A.C., Yang, J., Khan, M.A., 2020. Latest Neogene monsoon of the Chotanagpur Plateau, eastern India, as revealed by fossil leaf architectural signatures. Palaeogeogr. Palaeoclimatol. Palaeoecol. 545, 109641 https://doi.org/10.1016/j.palaeo.2020.109641.

Heer, O., 1855. Flora Tertiaria Helvetiae. 1. Verlag Wurster & Co., Winterthur.

Heer, O., 1856. Flora Tertiaria Helvetiae. 2. Verlag Wurster & Co., Winterthur.

Heer, O., 1859. Flora Tertiaria Helvetiae. 3. Verlag Wurster & Co., Winterthur.

Herman, A.B., Spicer, R.A., 1996. Palaeobotanical evidence for a warm Cretaceous Arctic Ocean. Nature 380, 330–333.

Herman, A.B., Spicer, R.A., 1997. New quantitative palaeoclimate data for the Late Cretaceous Arctic: Evidence for a warm polar ocean. Palaeogeogr. Palaeoclimatol. Palaeoecol. 128, 227–251. https://doi.org/10.1016/S0031-0182(96)00080-6.

Herman, A.B., Spicer, R.A., 2010. Mid-Cretaceous floras and climate of the Russian high Arctic (Novosibirsk Islands, Northern Yakutia). Palaeogeogr. Palaeoclimatol. Palaeoecol. 295, 409–422.

Herman, A.B., Spicer, R.A., Spicer, T.E.V., 2016. Environmental constraints on terrestrial vertebrate behaviour and reproduction in the high Arctic of the Late Cretaceous. Palaeogeogr. Palaeoclimatol. Palaeoecol. 441, 317–338.

Herman, A.B., Spicer, R.A., Aleksandrova, G.N., Yang, J., Kodrul, T., Masolva, N.P., Spicer, T.E.V., Chen, G., Jin, J.-J., 2017. Eocene-Oligocene climate and vegetation change in southern China: Evidence from the Maoming Basin. Palaeogeogr. Palaeoclimatol. Palaeoecol. 479, 126–137.

Hill, M.O., 1973. Reciprocal averaging — an eigenvector method of ordination. J. Ecol. 61, 237–249.

Hill, M.O., 1979. Correspondence analysis — a neglected multivariate method. Appl. Stat. 23, 340–354.

Hinojosa, L.F., Villagrán, C., 2005. Did South American mixed paleofloras evolve under thermal equability or in the absence of an effective Andean barrier during the Cenozoic? Palaeogeogr. Palaeoclimatol. Palaeoecol. 217, 1–23.

Holdridge, L.R., 1947. Determination of world plant formations from simple climatic data. Science 105, 367–368.

Hort, A., (Translator), 1948. Enquiry into Plants Vol. 1 by Theophrastus. Harvard University Press, Cambridge, Mass.

Hutton, J., 1788. Theory of the Earth; or an investigation of the laws observable in the composition, dissolution, and restoration of land upon the Globe. Trans. Roy. Soc. Edin. 1 (2), 209–304.

Jacques, F.M.B., Tao, S., Spicer, R.A., Xing, Y., Huang, Y., Wang, W., Zhou, Z., 2011. Leaf physiognomy and climate: are monsoon systems different? Glob. Planet. Chang. 76, 56–62.

Jacques, F.M.B., Su, T., Spicer, R.A., Xing, Y.-W., Huang, Y.-J., Zhou, Z., 2014. Late Miocene southwestern Chinese floristic diversity shaped by the southeastern uplift of the Tibetan Plateau. Palaeogeogr. Palaeoclimatol. Palaeoecol. 411, 208–215.

Jin, J.-H., Herman, A.B., Spicer, R.A., Kodrul, T.M., 2017. Palaeoclimate background of the diverse Eocene floras of South China. Sci. Bull. 62, 1501–1503.

Jordan, G.J., 2011. A critical framework for the assessment of biological palaeoproxies: predicting past climate and levels of atmospheric CO_2 from fossil leaves. New Phytol. 192, 29–44.

Juenger, T., Pérez-Pérez, J.M., Micol, J.L., 2005. Quantitative trait loci mapping of floral and leaf morphology traits in *Arabidopsis thaliana*: evidence for modular genetic architecture. Evol. Dev. 7, 259–271.

Kattge, J., Börnisch, G., Diaz, S., et al., 2020. TRY plant trait database - enhanced coverage and open access. Glob. Chang. Biol. 26, 119–188. https://doi.org/10.1111/gcb.14904.

Kennedy, E.M., Spicer, R.A., Rees, P.M., 2002. Quantitative paleoclimate estimates from Late Cretaceous and Paleocene leaf floras in the northwest of the South Island, New Zealand. Palaeogeogr. Palaeoclimatol. Palaeoecol. 184, 321–345.

Kennedy, E.M., Arens, N.C., Reichgelt, T., Spicer, R.A., Spicer, T.E.V., Stranks, L., Yang, J., 2014. Deriving temperature estimates from Southern Hemisphere leaves. Palaeogeogr. Palaeoclimatol. Palaeoecol. 412, 80–90.

Kent, M., Coker, P., 1992. Vegetation Description and Analysis: A Practical Approach. Belhaven Press, London.

Kershaw, A.P., Nix, H.A., 1988. Quantitative palaeoclimatic estimates from pollen data using bioclimatic profiles of extant taxa. J. Biogeogr. 15, 589–602.

Khan, M.A., Spicer, R.A., Bera, S., Ghosh, R., Yang, J., Spicer, T.E.V., Guo, S.-X., Su, T., Jacques, F.M.B., Grote, P.J., 2014. Miocene to Pleistocene floras and climate of the Eastern Himalayan Siwaliks, and new palaeoelevation estimates for the Namling–Oiyug Basin. Tibet. Global. Planet. Change 113, 1–10.

Khan, M.A., Bera, M., Spicer, R.A., Spicer, T.E.V., Bera, S., 2018. Floral diversity and environment during the middle Siwalik sedimentation (Pliocene) in the Arunachal sub-Himalaya. Palaeobio. Palaeoenv. 99, 401–424. https://doi.org/10.1007/s12549-018-0351-2.

Khan, M.A., Bera, M., Spicer, R.A., Spicer, T.E.V., Bera, S., 2019. Palaeoclimatic estimates for the late Cenozoic (Mio-Pliocene) Siwalik Flora of Bhutan: Evidence for the development of the South Asian Monsoon in the eastern Himalaya. Palaeogeogr. Palaeoclimatol. Palaeoecol. 514, 326–335.

Kouwenberg, L.L., Kürschner, W.M., McElwain, J.C., 2007. Stomatal frequency change over altitudinal gradients: Prospects for paleoaltimetry. Rev. Mineral. Geochem. 66, 215–241.

Kovach, W.L., Spicer, R.A., 1996. Canonical Correspondence Analysis of Leaf Physiognomy: A Contribution to the development of a new palaeoclimatological tool. Palaeoclimates 2, 125–138.

Kowalski, E.A., Dilcher, D.L., 2003. Warmer Paleotemperatures for Terrestrial Ecosystems. P. Natl. Acad. Sci. USA 100, 167–170.

Kürschner, W.M., 1997. The anatomical diversity of recent and fossil leaves of the durmast oak (*Quercus petraea* Lieblein/*Q. pseudocastanea* Goeppert)—implications for their use as biosensors of palaeoatmospheric CO_2 levels. Rev. Palaeobot. Palynol. 96, 1–30.

Kvaček, Z., Teodorides, V., Denk, T., 2019. The Pliocene flora of Frankfurt am Main, Germany: taxonomy, palaeoenvironments and biogeographic affinities. Palaeobio. Palaeoenv. 100, 647–703. https://doi.org/10.1007/s12549-019-00391-6.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74, 1659–1673.

Li, S.–F., Jacques, F.M.B., Spicer, R.A., Su, T., Spicer, T.E.V., Yang, J., Zhou, Z., 2016. Artificial neural networks reveal a high-resolution climatic signal in leaf physiognomy. Palaeogeogr. Palaeoclimatol. Palaeoecol. 442, 1–11.

Lyell, C., 1830. 1833. Principles of geology, being an attempt to explain the former changes of the Earth's surface, by reference to causes now in operation. 1-3. John Murray, London.

Maslova, N.P., Volkova, L.D., Gordenko, N.V., 2008a. Structural epidermal features of leaves of *Platanus acerifolia* Willd. (Platanaceae) and significance of cuticular analysis for the determination of Cretaceous *Platanus*-like leaves. Botanicheskii Zhurnal 93, 982–995 (in Russian).

Maslova, N.P., Volkova, L.D., Gordenko, N.V., 2008b. Morphological variability of leaves of *Platanus acerifolia* Willd. (Platanaceae) and approaches to the determination of dispersed Cretaceous *Platanus*-like leaves. Botanicheskii Zhurnal 93, 825–839 (in Russian).

Maslova, N.P., Karasev, E.V., Kodrul, T.M., Spicer, R.A., Volkova, L.D., Spicer, T.E.V., Jin, J.-H., Liu, X.Y., 2018. Sun and shade leaf variability in *Liquidambar chinensis* and *Liquidambar formosana* Hanse (Altingiaceae): implications for palaeobotany. Bot. J. Linn. Soc. 188, 296–315.

McDonald, P.G., Fonseca, C.R., Overton, J., Mc, C., Westoby, M., 2003. Leaf-size divergence along rainfall and soil-nutrient gradients: is the method of size reduction common among clades? Funct. Ecol. 17, 50–57.

Miller, I.M., Brandon, M.T., Hickey, L.J., 2006. Using Leaf Margin Analysis to estimate the mid-Cretaceous (Albian) paleolatitude of the Baja BC block. Earth Planet. Sci. Lett. 245, 95–114.

Moiseeva, M.G., Herman, A.B., Spicer, R.A., 2009. Late Paleocene Flora of the Northern Alaska Peninsula: The Role of Transberingian Plant Migrations and Climate Change. Palaeontol. J. 43, 1298–1308.

Mosbrugger, V., Utescher, T., 1997. The coexistence approach — a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils. Palaeogeogr. Palaeoclimatol. Palaeoecol. 134, 61–86.

Murray, M., Soh, W.K., Yiotis, C., Batke, S., Parnell, A., Spicer, R., Lawson, T., Caballero, R., Wright, I., Purcell, C., McElwain, J., 2019. Convergence in Maximum Stomatal Conductance of C3 Woody Angiosperms in Natural Ecosystems Across Bioclimatic Zones. Front. Plant Sci. 10, 558 https://doi.org/10.3389/fpls.2019.00558.

Murray, M., Soh, W.-K., Yiotis, C., Spicer, R.A., Lawson, T., McElwain, J.C., 2020. Consistent Relationship between Field-Measured Stomatal Conductance and Theoretical Maximum Stomatal Conductance in C_3 Woody Angiosperms in Four Major Biomes. Int. J. Plant Sci. 181, 142–154. https://doi.org/10.1086/706260.

Needham, J., 1986. Science and Civilization in China: Volume 3, Mathematics and the Sciences of the Heavens and the Earth. Caves Books Ltd., Taipei.

New, M., Hulme, M., Jones, P., 1999. Representing Twentieth-Century Space–Time Climate Variability. Part I: Development of a 1961–90 Mean Monthly Terrestrial Climatology. J. Clim. 12, 829–856.

New, M., Lister, D., Hulme, M., Makin, I., 2002. A high-resolution data set of surface climate over global land areas. Clim. Res. 21, 1–15.

Parrish, J.T., Daniel, I.L., Kennedy, E.M., Spicer, R.A., 1998. Paleoclimatic Significance of Mid-Cretaceous Floras from the Middle Clarence Valley, New Zealand. Palaios 13, 146–156.

Peppe, D.J., Royer, D.L., Wilf, P., Kowalski, E.A., 2010. Quantification of large uncertainties in fossil leaf paleoaltimetry. Tectonics 29. https://doi.org/10.1029/2009TC002549. TC3015.

Pigliucci, M., 2003. Phenotypic integration: studying the ecology and evolution of complex phenotypes. Ecol. Lett. 6, 265–272.

Polissar, P., Freeman, K., Rowley, D., McInerney, F., Currie, B.S., 2009. Paleoaltimetry of the Tibetan Plateau from D/H ratios of lipid biomarkers. Earth Planet. Sci. Lett. 287, 64–76.

Povey, D.A.R., Spicer, R.A., England, P.C., 1994. Palaeobotanical investigation of early Tertiary palaeoelevations in northeastern Nevada: initial results. Rev. Palaeobot. Palynol. 81, 1–10.

R Core Team, 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raunkiaer, C., 1934. The life forms of plants and statistical plant geography. Oxford University Press, Oxford.

Rodriguez, R.E., Debernardi, J.M., Palatnik, J.F., 2014. Morphogenesis of simple leaves: regulation of leaf size and shape. WIREs Dev. Biol. 3, 41–57.

Royer, D., Wilf, P., 2006. Why do toothed leaves correlate with cold climates? Gas exchange at leaf margins provides new insights into a classic paleotemperature proxy. Int. J. Plant Sci. 167, 11–18.

Rozendaal, D.M.A., Hurtado, V.H., Poorter, L., 2006. Plasticity in leaf traits of 38 tropical tree species in response to light; relationships with light demand and adult stature. Funct. Ecol. 20, 207–216.

Schlichting, C.D., 1989. Phenotypic integration and environmental change. BioScience 39, 460–464.

Seward, A.C., 1892. Fossil Plants as Tests of Climate. C.J. Clay and Sons. Cambridge University Press, London.

Shukla, A., Mehrotra, R.C., Spicer, R.A., Spicer, T.E.V., Kumar, M., 2014. Cool equatorial terrestrial temperatures and the South Asian monsoon in the Early Eocene: Evidence from the Gurha Mine, Rajasthan, India. Palaeogeogr. Palaeoclimatol. Palaeoecol. 412, 187–198.

Soh, W.-K., Yiotis, C., Murray, M., Parnell, A., Wright, I.J., Spicer, R.A., Lawson, T., Caballero, R., McElwain, J.C., 2019. Rising CO_2 rise drives divergence in water-use efficiency of evergreen and deciduous plants. In: Science Advances 5: eaax7906.

Song, B., Spicer, R.A., Zhang, K., Ji, J., Farnsworth, A., Hughes, A.C., Yang, Y., Han, F., Xu, Y., Spicer, T.E.V., Shen, T., Lunt, D.J., Shi, G., 2020. Qaidam Basin leaf fossils show northeastern Tibet was high, wet and cool in the early Oligocene. Earth Planet. Sci. Lett. 537, 116175 https://doi.org/10.1016/j.epsl.2020.116175.

Spicer, R.A., 1981. The sorting and deposition of allochthonous plant material in a modern environment at Silwood Lake, Silwood Park, Berkshire, England. U.S. Geological Survey Professional Paper 1143, 77p.

Spicer, R.A., 2018. Phytopaleoaltimetry: Using Plant Fossils to Measure Past Land Surface Elevation. In: Hoorn, C., Perrigo, A., Antonelli, A. (Eds.), Mountains. Climate and Biodiversity, Wiley-Blackwell, pp. 96–109.

Spicer, R.A., Herman, A.B., 2010. The Late Cretaceous Environment of the Arctic: A quantitative reassessment using plant fossils. Palaeogeogr. Palaeoclimatol. Palaeoecol. 295, 423–442.

Spicer, R.A., Hill, C.R., 1979. Principal components and Correspondence Analyses of quantitative data from a Jurassic plant bed. Rev. Palaeobot. Palynol. 28, 273–299.

Spicer, R.A., Yang, J., 2010. Quantification of uncertainties in fossil leaf paleoaltimetry: Does leaf size matter? Tectonics 29, TC6001, doi:10 1029/2010TC002741.

Spicer, R.A., Ahlberg, A., Herman, A.B., Kelley, S.P., Raikevich, M.I., Rees, P.M., 2002. Palaeoenvironment and ecology of the middle Cretaceous Grebenka flora of northeastern Asia. Palaeogeogr. Palaeoclimatol. Palaeoecol. 184, 65–105.

Spicer, R.A., Harris, N.B.W., Widdowson, M., Herman, A.B., Guo, S.-X., Valdes, P.J., Wolfe, J.A., Kelley, S.P., 2003. Constant elevation of southern Tibet over the past 15 million years. Nature 421, 622–624.

Spicer, R.A., Herman, A.B., Kennedy, E.M., 2004. The foliar physiognomic record of climatic conditions during dormancy: CLAMP and the cold month mean temperature. J. Geol. 112, 685–702.

Spicer, R.A., Herman, A.B., Kennedy, E.M., 2005. The sensitivity of CLAMP to taphonomic loss of foliar physiognomic characters. Palaios 20, 429–438.

Spicer, R.A., Ahlberg, A., Herman, A.B., Hofmann, C.-C., Raikevich, M., Valdes, P.J., Markwick, P.J., 2008. The Late Cretaceous continental interior of Siberia: A challenge for climate models. Earth Planet. Sci. Lett. 267, 228–235.

Spicer, R.A., Valdes, P.J., Spicer, T.E.V., Craggs, H.J., Srivastava, G., Mehrotra, R.C., Yang, J., 2009. New development is CLAMP: calibration using global gridded meteorological data. Palaeogeogr. Palaeoclimatol. Palaeoecol. 283, 91–98.

Spicer, R.A., Bera, S., De Bera, S., Spicer, T.E.V., Srivastava, G., Mehrotra, R.C., Mehrotra, N.C., Yang, J., 2011. Why do foliar physiognomic climate estimates sometimes differ from those observed? Insights from taphonomic information loss and a CLAMP case study from the Ganges Delta. Palaeogeogr. Palaeoclimatol. Palaeoecol. 299, 39–48.

Spicer, R.A., Herman, A.B., Liao, W., Spicer, T.E.V., Kodrul, T.M., Yang, J., Jin, J.-J., 2014. Cool tropics in the Middle Eocene: Evidence from the Changchang Flora, Hainan Island. China. Palaeogeography, Palaeoclimatology, Palaeoecology 412, 1–16.

Spicer, R.A., Yang, J., Herman, A.B., Kodrul, T., Maslova, N., Spicer, T.E.V., Aleksandrova, G., Jin, J.-J., 2016. Asian Eocene Monsoons as revealed by leaf architectural signatures. Earth Planet. Sci. Lett. 449, 61–68.

Spicer, R.A., Yang, J., Herman, A.B., Kodrul, T., Aleksandrova, G., Maslova, N., Spicer, T., Ding, L., Xu, Q., Shukla, A., Srivastava, G., Mehrotra, R., Liu, X.-Y., Jin, J.-J., 2017. Paleogene monsoons across India and South China: Drivers of biotic change. Gondwana Res. 49, 350–363.

Spicer, R.A., Valdes, P.J., Hughes, A.C., Yang, J., Spicer, T.E.V., Herman, A.B., Farnsworth, A., 2019. New insights into the thermal regime and hydrodynamics of the early Late Cretaceous Arctic. Geol. Mag. 157, 1729–1746. https://doi.org/10.1017/S0016756819000463.

Srivastava, G., Spicer, R.A., Spicer, T.E.V., Yang, J., Kumar, M., Mehrotra, R.C., Mehrotra, N.C., 2012. Megaflora and palaeoclimate of a Late Oligocene tropical delta, Makum Coal field, Assam: evidence for the early development of the South Asia Monsoon. Palaeogeogr. Palaeoclimatol. Palaeoecol. 343, 130–142.

Steart, D.C., Spicer, R.A., Bamford, M.K., 2010. Is Southern Africa different? An investigation of the relationship between leaf physiognomy and climate in southern African mesic vegetation. Rev. Palaeobot. Palynol. 162, 607–620.

Stranks, L., 1996. Physiognomic and taphonomic studies in New Zealand and Australia: implications for the use of palaeobotany as a tool for palaeoclimate estimation. University of Oxford, Oxford, D. Phil. dissertation, 135 p.

Stranks, L., England, P., 1997. The use of a resemblance function in the measurement of climatic parameters from the physiognomy of woody dicotyledons. Palaeogeogr. Palaeoclimatol. Palaeoecol. 131, 15–28.

Su, T., Spicer, R.A., Li, S.-H., Xu, H., Huang, J., Sherlock, S., Huang, Y.-J., Li, S.-F., Wang, L., Jia, L.-B., Deng, W.-Y.-D., Deng, C.-L., Zhang, S.-T., Valdes, P.J., Zhou, Z.-K., 2019. Uplift, Climate and Biotic Changes at the Eocene-Oligocene Transition in Southeast Tibet. Natl. Sci. Rev. 6 https://doi.org/10.1093/nsr/nwy062, 495–504.

Teodoridis, V., Mazouch, P., Spicer, R.A., Uhl, D., 2011. Refining CLAMP – investigations towards improving the climate leaf analysis multivariate program. Palaeogeogr. Palaeoclimatol. Palaeoecol. 299, 39–48.

ter Braak, C.J.F., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67, 1167–1179.

ter Braak, C.J.F., 1987. The analysis of vegetation-environment relationships by canonical correspondence analysis. Vegetatio 69, 69–77.

Thompson, R.S., Anderson, K.H., Pelltier, R.T., Strickland, L.E., Bartlein, P.J., Shafer, S. L., 2012. Quantitative estimation of climatic parameters from vegetation data in North America by the mutual climatic range technique. Quat. Sci. Rev. 51, 18–39.

Tomsich, C.S., McCarthy, P.J., Fowell, S.J., Sunderlin, D., 2010. Paleofloristic and Paleoenvironmental information from a Late Cretaceous (Maastrichtian) flora of the lower Cantwell Formation near Sable Mountain, Denali National Park, Alaska. Palaeogeogr. Palaeoclimatol. Palaeoecol. 295, 389–408.

Traiser, C., 2004. Blattphysiognomie als Indikator für Umweltparameter: Eine Analyse rezenter und fossiler Floren (Leaf physiognomy as environmental indicator: An analysis of extant and fossil floras). PhD Thesis, Geowissenschaftlichen Fakultät der Eberhard-Karls Universität Tübingen, Germany, 113 pp, urn:nbn:de:bsz:21-opus-14947.

Traiser, C., Klotz, S., Uhl, D., Mosbrugger, V., 2004. Environmental signals from leaves – a physiognomic analysis of European vegetation. New Phytol. 166, 465–484. https://doi.org/10.1111/j.1469-8137.2005.01316.x.

Traiser, C., Mosbrugger, V., 2004. ELPA (European Leaf Physiognomic Approach): Grid data set of the leaf physiognomic composition of the extant hardwood vegetation. PANGAEA. doi:10.1594/PANGAEA.218109.

Traiser, C., Uhl, D., Klotz, S., Mosbrugger, V., 2007. Leaf physiognomy and palaeoenvironmental estimates – an alternative technique based on a European calibration. Acta Palaeobotanica 47, 183–201.

Tsukaya, H., 2005. Leaf shape: genetic controls and environmental factors. Int. J. Dev. Biol. 49, 547–555.

Uhl, D., Walther, H., 2003. Aderungsdichte und Blattgröße bei vier Arten der Gattung *Acer* L. aus SW-Deutschland. Botanische Jahrbücher für Systematik. Pflanzengeschichte und PflanzenGeographyraphie 124, 449–460.

Uhl, D., Traiser, C., Griesser, U., Denk, T., 2007. Fossil leaves as palaeoclimate proxies in the Palaeogene of Spitsbergen (Svalbard). Acta Palaeobotanica 47, 89–107.

Utescher, T., Bruch, A.A., Erdei, B., François, L., Ivanov, D., Jacques, F.M.B., Kern, A.K., Liu, Y-S.(C), Mosbrugger, V., Spicer, R.A., 2014. The Coexistence Approach—Theoretical background and practical considerations of using plant fossils for climate quantification. Palaeogeogr. Palaeoclimatol. Palaeoecol. 410, 58–73.

Velasco-de Leon, M.P., Spicer, R.A., Steart, D.C., 2010. Climatic reconstruction of two Pliocene floras from Mexico. Palaeobio. Palaeoenv. 90, 99–110.

von Humboldt, A., 1850. Aspects of Nature in Different Lands and different Climates; with Scientific Elucidations. In: Lea & Blanchard. Pennsylvania, Philadelphia.

Webb, L.J., 1968. Environmental relationships of the structural types of Australian rain forest vegetation. Ecology 49, 296–311.

Wheeler, E.A., Baas, P., Rodgers, S., 2007. Variations in dicot wood anatomy: a global analysis based on the Inside Wood database. IAWA J. 28, 229–258.

Wiemann, M.C., Manchester, S.R., Dilcher, D.L., Hinojosa, L.F., Wheeler, E.A., 1998. Estimation of temperature and precipitation from morphological characters of dicotyledonous leaves. Am. J. Bot. 85, 1796–1802. https://doi.org/10.2307/2446514.

Wilf, P., 1997. When are leaves good thermometers? A new case for leaf margin analysis. Paleobiology 23, 373–390.

Wilf, P., Wing, S.L., Greenwood, D.R., Greenwood, C.L., 1998. Using fossil leaves as paleoprecipitation indicators: an Eocene example. Geology 26, 203–206.

Wilf, P., Wing, S.L., Greenwood, D.R., Greenwood, C.L., 1999. Using fossil leaves as paleoprecipitation indicators: An Eocene example—Reply. Geology 27, 92.

Wing, S.L., Greenwood, D.R., 1993. Fossils and fossil climate: the case for equable continental interiors in the Eocene. Philos. T. R. Soc. Lond. B 341, 243–252.

Witham, W., 1833. The internal structure of fossil vegetables found in the Carboniferous and Oolitic deposits of Great Britain. Edinburgh.

Wolfe, J.A., 1971. Tertiary climatic fluctuations and methods of analysis of Tertiary floras. Palaeogeogr. Palaeoclimatol. Palaeoecol. 9, 27–57.

Wolfe, J.A., 1979. Temperature parameters of humid to mesic forests of Eastern Asia and relation to forests of other regions in the northern hemisphere and Australasia. U.S. Geol. Surv. Prof. Pap. 1106, 37.

Wolfe, J.A., 1990. Palaeobotanical evidence for a marked temperature increase following the Cretaceous/Tertiary boundary. Nature 343, 153–156.

Wolfe, J.A., 1993. A method of obtaining climatic parameters from leaf assemblages. U.S. Geol. Den. Surv. Bull. 2040, 73.

Wolfe, J.A., 1994. a Tertiary climatic changes at middle latitudes of western North America. Palaeogeogr. Palaeoclimatol. Palaeoecol. 108, 195–205.

Wolfe, J.A., 1994b. Alaskan Palaeogene climates as inferred from the CLAMP database. In: Boulter, M.C., Fisher, H.C. (Eds.), Cenozoic Plants and Climate of the Arctic. Springer-Verlag, Berlin, pp. 223–237.

Wolfe, J.A., 1994c. An analysis of Neogene climates in Beringia. Palaeogeogr. Palaeoclimatol. Palaeoecol. 108, 207–216.

Wolfe, J.A., 1995. Paleoclimate estimates from Tertiary leaf assemblages. Annu. Rev. Earth Pl. Sc. 23, 119–142.

Wolfe, J.A., Spicer, R.A., 1999. Fossil leaf character states: multivariate analysis. In: Jones, T.P., Rowe, N.P. (Eds.), Fossil plants and spores: modern techniques. Geological Society, London, pp. 233–239.

Wolfe, J.A., Uemura, K., 1999. Using fossil leaves as paleoprecipitation indicators: An Eocene example: Comment and Reply. Geology 27, 91–92.

Wolfe, J.A., Forest, C.E., Molnar, P., 1998. Paleobotanical evidence of Eocene and Oligocene paleoaltitudes in midlatitude western North America. Geol. Soc. Am. Bull. 110, 664–678.

Wood, S.N., 2003. Thin plate regression splines. J. R. Stat. Soc. B 65, 95–114.

Xu, Q., Ding, L., Spicer, R.A., Liu, X., Li, S., Wang, H., 2017. Stable isotopes reveal southward growth of the Himalayan-Tibetan Plateau since the Paleocene. Gondwana Res. 54, 50–61. https://doi.org/10.1016/j.gr.2017.10.005.

Yang, J., Wang, Y.F., Spicer, R.A., Mosbrugger, V., Li, C.S., Sun, Q.G., 2007. Climatic Reconstruction of Miocene Shanwang Basin, China, by using Leaf Margin Analysis, CLAMP, Co- existence Approach and Overlapping Distribution Analysis. Am. J. Bot. 94, 599–608.

Yang, J., Spicer, R.A., Spicer, T.E.V., Li, C.-S., 2011. 'CLAMP Online': a new web-based palaeoclimate tool and its application to the terrestrial Paleogene and Neogene of North America. Palaeobio. Palaeoenv. 91, 163–183.

Yang, J., Spicer, R.A., Spicer, T.E.V., Arens, N.C., Jacques, F.M.B., Su, T., Kennedy, E.M., Herman, A.B., Steart, D.C., Srivastava, G., Mehrotra, R.C., Valdes, P.J., Mehrotra, N. C., Zhou, Z.K., Lai, J.S., 2015. Leaf Form-Climate Relationships on the Global Stage: An Ensemble of Characters. Glob. Ecol. Biogeogr. 10, 1113–1125.

Zalensky, V.R., 1904. Materials to numerical anatomy of various leaves in the same plants. Izvestiya Kievskogo Politekhnicheskogo Instituta 4, 1–209 (In Russian).

Zolina, A.A., Golovneva, L.B., Spicer, R.A., 2020. Latest Cretaceous (Maastrichtian) climate of the Koryak Upland of North- East Russia based on a quantitative analysis of a palaeo-polar flora. Palaeogeogr. Palaeoclimatol. Palaeoecol. 560, 109997 https://doi.org/10.1016/j.palaeo.2020.109997.