Palaeogeography, Palaeoclimatology, Palaeoecology 678 (2025) 113236

Contents lists available at ScienceDirect

Palaeogeography, Palaeoclimatology, Palaeoecology

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

Normalization of fossil plant megafossil databases for diversity and palaeobiogeography analyses by filtering taxonomic duplication: Principles, methods, examples, and recommendations

Zhen Xu[a,b,*], Emma Bateson[c], Christopher J. Cleal[d], Reece Hutton[c], Jianxin Yu[b], Shi-Jun Wang[e], Andrew H. Knoll[f], Benjamin J.W. Mills[a], Jason Hilton[c,g]

[a] *School of Earth and Environment, University of Leeds, Leeds LS29JT, United Kingdom*
[b] *State Key Laboratory of Geomicrobiology and Environmental Changes, School of Earth Sciences, China University of Geosciences, Wuhan 430074, PR China*
[c] *School of Geography, Earth and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom*
[d] *School of Earth Sciences, University of Bristol, Bristol BS8 1TQ, United Kingdom*
[e] *State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China*
[f] *Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, United Kingdom*
[g] *Birmingham Institute of Forest Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom*

ARTICLE INFO

Editor: H Falcon-Lang

Keywords:
Diversity
Species richness
Palaeobiogeography
Plant fossil occurrences
Whole-plant concepts
Taxonomic duplication
Nomenclature

ABSTRACT

Fossil plants are key to many palaeobiogeographic and deep time diversity studies, but correctly interpreting them can be fraught with problems due to fragmentation in the fossil record. A typical vascular plant comprises 10–12 separate organs depending on its systematic affinity, but complete individuals are exceptionally rare. Fragmentation can result from multiple processes including ontogeny during the plant's life-cycle, or from post-mortem taphonomic processes in fluvial systems. In traditional approaches where raw data is amassed directly in the field, from existing physical collections or electronic databases, duplication is inevitable in that different organs of the same plant species may be inadvertently counted independently, skewing results. Here we outline normalization methods for filtering the palaeobotanical data to remove taxonomic duplications, with examples provided from different types of preservation. We use two case studies to highlight the impact of normalization by analysing raw (unfiltered) versus normalized (filtered) data. The first case study evaluates plant data from the late Permian and Triassic compression/impression floras of South China, focussing on species richness/diversity assessments through the Permian-Triassic mass extinction and its recovery. In this case study, normalization reduced the number of taxa but revealed more detailed evolutionary patterns including the magnitude of floristic turnover, previously obscured by the fragmental preservation typical of plant fossils and nomenclature. The second case study evaluates Carboniferous to Permian anatomically preserved coal-ball floras from Europe, North America and China, focussing on palaeobiogeography and floral provinciality. Normalization reduced the number of coal-ball assemblages when surveyed at both genus and species level but revealed differences in relationships and floristic endemism. We conclude that normalized results should be considered alongside raw data, as they show important and complementary information which can greatly aid in overall interpretation.

1. Introduction

Finding complete remains of plants in the fossil record is exceptionally rare, the vast majority of plant fossils being fragmented parts of various shapes and sizes (e.g., Knoll, 1986; Bateman, 1991; Spicer, 1991; Taylor et al., 2009; Cleal and Thomas, 2019). A typical vascular plant may be represented as fossils by as many as 10–12 separate organs (e.g., roots, stem, leaves, cones, seeds, pollen) depending on its systematic affinity (Bateman and Hilton, 2009; Taylor et al., 2009; Cleal et al., 2021). This fragmentation can result from multiple processes, including ontogeny and other aspects of the plant's life-cycle (e.g., seed dispersal, leaf abscission), or from post-mortem taphonomic processes, such as breaking and size sorting in fluvial systems during transport prior to the final deposition (Spicer and Greer, 1986; Bateman, 1991; Spicer, 1991;

---

* Corresponding author at: School of Earth and Environment, University of Leeds, Leeds LS29JT, United Kingdom
E-mail address: Z.xu@leeds.ac.uk (Z. Xu).

https://doi.org/10.1016/j.palaeo.2025.113236
Received 7 July 2025; Received in revised form 25 August 2025; Accepted 25 August 2025

Gastaldo, 1992; Thomas and Cleal, 1999; Behrensmeyer et al., 2000; Ferguson, 2005). There is also the added complication that fossils may be found in different modes of preservation that can yield varying levels of taxonomic information, as this can affect the way they are classified and named (e.g., Chaloner, 1986; Bateman and Hilton, 2009; Thomas and Cleal, 2020).

Although the aim of palaeobotany is normally to interpret the parent vegetation, reliably reconstructed whole plants based on the fossil record are rare; there are a few examples (e.g., Rothwell, 1981; Liu et al., 2021; Li et al., 2025) but they are very much the exceptions. Consequently, since the pioneering days of the subject (e.g., Sternberg von Sternberg, 1820; Brongniart, 1822) it has been recognised that the best way to develop an objective palaeobotanical taxonomy is to treat the plant fossils as intrinsic objects. To this end, plant fossils are usually classified and named as what are now called fossil-taxa as defined in the International Code of Botanical Nomenclature (ICN – Turland et al., 2018), with each fossil-taxon representing a particular plant part in a particular mode of preservation (for further discussions see Cleal, 1986; Thomas, 1989; Cleal and Thomas, 2010, 2019, 2021). The aim is that the resulting taxa reflect as far as possible both the systematics of the original plants and the constraints introduced by taphonomy.

Whilst this approach has much merit in terms of objectivity for many aspects of palaeobotany, it has the drawback that it may significantly distort the apparent diversity structure of the original vegetation (Knoll et al., 1979); not only can an individual plant taxon be represented by several fossil-taxa, the number of fossil-taxa per plant taxon varies according to plant group (e.g., Bateman and Hilton, 2009; Cleal et al., 2012). This has become especially important in discussions about the impact and scale of floristic response during the global biotic crises known as mass extinctions (Boulter et al., 1986; Niklas et al., 1986; Niklas, 1988; Niklas and Tiffney, 1994). Mass extinctions were originally identified in the faunal fossil record, mainly from the marine realm (e.g., Newell, 1963; Raup and Sepkoski, 1982; Sepkoski, 1986; Bambach et al., 2004; Bambach, 2006; Marshall, 2023), where an organism is represented by only one taxonomic name at any particular rank. However, when the palaeobotanical evidence for mass extinctions was investigated, rather different diversity patterns were revealed (e.g., McElwain and Punyasena, 2007; Cascales-Miñana and Cleal, 2014). This begged the question: is this due to the use of fossil-taxa in palaeobotany, or were there real differences in the taxonomic diversity dynamics in terrestrial vegetation and marine faunas?

In this paper we develop and implement a method that we term normalization to estimate the true numbers of species in a plant fossil assemblage, given multiple organ taxa that can originate from a single biological source. While several previous studies have employed similar approaches (Cleal and Wang, 2002, Cleal, 2005, Cleal, 2008; Hilton and Cleal, 2007; Uhl and Cleal, 2010; Cleal et al., 2012; Opluštil et al., 2017; Molina-Solís et al., 2023), the broader implications and potential side-effects of this methodology have not yet been systematically explored. To address this gap, we compare raw (unfiltered) and normalized (filtered) plant fossil datasets from two case studies, each designed to answer distinct research questions.

2. Sources of data

This study is based on two datasets that are each first filtered to ensure their taxonomic accuracy before being normalized. The first dataset records adpressions from the Permian – Triassic of the South China Plate representing lowland, coastal vegetation growing on the western edge of the Tethys Ocean (Figs. 1, 2). The data were taken from a study of the effects on vegetation of the Permian – Triassic mass extinction (PTME) (Xu et al., 2022, 2025) and include data from the uppermost Permian Xuanwei and Dalong formations (Yao et al., 1980; Li et al., 1995; Yu, 2008; Yu et al., 2015, 2022; Li et al., 2019; Feng et al., 2020), the Permian – Triassic transitional Kayitou Formation (Yu, 2008; Yu et al., 2010, 2022; Chen et al., 2011; Feng et al., 2020), the upper Lower Triassic Feixianguan (Xue, 2019; Broutin et al., 2020; Xu et al., 2022) and Lingwen formations (Zhou and Li, 1979; Zhang et al., 1992; Xu et al., 2022), and the Middle Triassic Badong Formation (Meng et al., 1995; Ran, 2021; Xu et al., 2022). It also considers background pre-extinction floras from the lower to upper Permian Liangshan, Maokou, and Longtan Formations (Li et al., 1995; Yao et al., 1980), and the radiation stage Upper Triassic data from the Jiuligang (Meng, 1982), Daqiaodi, and Dajing formations (Xu et al., 1979), and the Yangbaichong Formation (Zhou, 1989; Liu et al., 2009).

The second dataset records anatomically preserved plant fossils preserved in Carboniferous – Permian coal-balls from Europe, North America and China, that were intended to throw light on the biome migration that took place at this time between Euramerica and Cathaysia (Hilton and Cleal, 2007) (Fig. 3). The data are updated from the collations in Phillips (1980), Tian et al. (1996), Galtier (1997), and Wang et al. (2009), including the addition of the North American Stuebenville flora (Rothwell, 1988) that is now amongst the most comprehensively known coal-ball floras (Fig. 4).

3. Methods

3.1. Plant macro fossil data normalization

Analysing large taxonomic datasets has become increasingly common to determine diversity patterns in both present and past biotas. Even with modern-day, living biotas, however, it has become evident that simply using raw data for these analyses is fraught with problems (e.g., Cai and Zhu, 2015; Isaac and Pocock, 2015; Ribeiro et al., 2022). In a palaeontological context, simply extracting and analysing raw data such as from the Paleobiology Database (PBDB) can be highly misleading unless the taxonomic lists have been suitably normalized to minimize issues such as nomenclatural synonymies and taxonomic revisions; to be

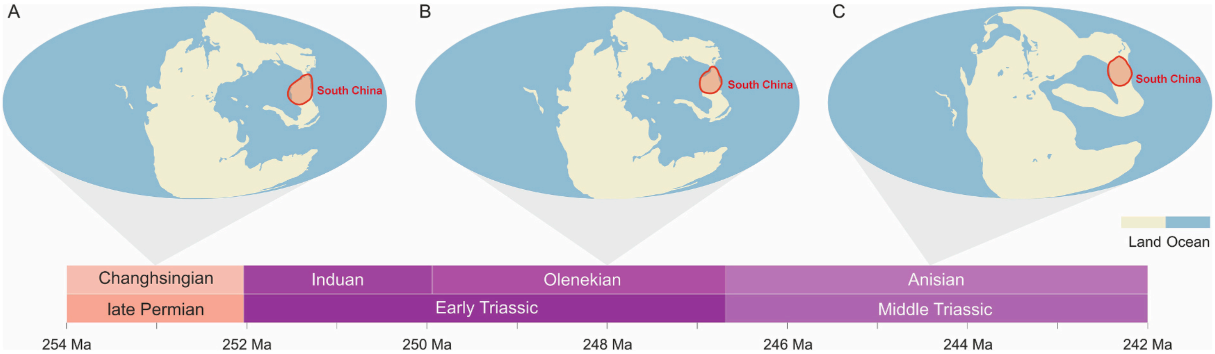


Fig. 1. World palaeogeography from the late Permian to Middle Triassic. emphasising differing continental configurations of the low latitude South China Plate. **(A)** Permian-Triassic boundary transition ~252 Ma. **(B)** Early Triassic ~250 Ma. **(C)** Middle Triassic ~244 Ma.

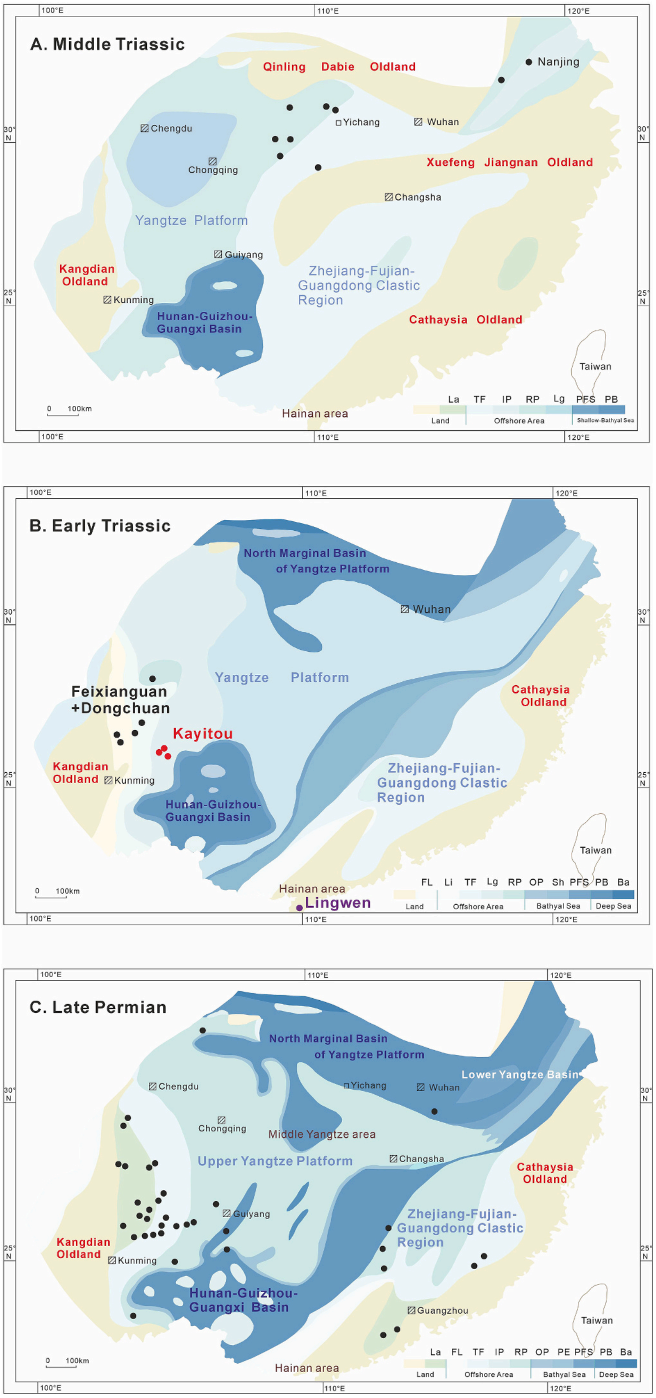


Fig. 2. Palaeogeographic maps of South China from the late Permian **(A)**, Early Triassic **(B)** and Middle Triassic **(C)** showing position of fossil plant localities. Black circles in A show location of the combined Xuanwei and Dalong formation outcrops. Black circles in B show location of the combined Feixianguan and Dongchuan formation sections, red dots show Kayitou Formation sections. Black circles in C show locations of Badong Formation sections, purple circles show position of Lingwen Formation sections. Locality details from Xu et al. (2022). La = Lacustrine; FL = Flood land; TF = Tide flat; IP = Isolate platform; RP = Regional platform; OP = Open platform; PE = Platform edge; PFS = Carbonate platform fore slope; PB = Platform basin; Ba = Bathyal sea; modified from Zheng (2011). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

done robustly, this of course requires at least some understanding of the taxonomy of the group being investigated. In palaeobotany, there is the added complication introduced using fossil-taxa. Since we are usually interested in the diversity of the original vegetation rather than of the fossils themselves, we need some means of translating the list of fossil-taxa into something approaching a list of the original plant taxa.

In an early attempt to deal with the occurrence of multiple organ taxa within fossil floras, Knoll (1986) adopted something of a “minimum species” concept. Within a given higher taxon, a given flora might contain six leaf species, four sporangial species, and three root species. Counted individually, this would suggest as many as thirteen species within the more inclusive clade; however, as different organ taxa may have been produced by the same biological species, no more than six actual species are necessary to account for the observed diversity. A more recent, and arguably better strategy to translate lists of fossil-taxa into something approaching the taxonomic diversity of the parent plants is to select the fossil-taxa of just one plant part per plant group (Cleal et al., 2021). The choice of plant part to use is to an extent subjective, but from experience working with adpression floras representing Carboniferous and Permian swamp vegetation (Cleal, 2005, 2007; Hilton and Cleal, 2007; Opluštil and Cleal, 2007; Cleal et al., 2012; Molina-Solís et al., 2023) it has been found that foliage taxa provide the best indication of plant species diversity; foliage remains represent by far the bulk of the adpression record and they usually preserve more taxonomically useful morphological features compared to say seeds / ovules and stems at the rank of species. The main exception to this is the arborescent lycopsids, for which the stem taxa are normally used, but the main taxonomic characters of these stems at the species rank are the leaf cushions and scars (Thomas, 1970, 1977; Cleal and Thomas, 1994, 2019) and so also in part reflect the foliage of the plants. Cordaitalean foliage was used in many early diversity studies (e.g., Cleal, 2005, 2007), but these leaves are difficult to identify based purely on their morphology (Šimůnek, 2007) and so Molina-Solís et al. (2023) used instead the reproductive structures. For detailed examples and applications involving end Permian to Middle Triassic plant macro fossils, please see Xu et al. (2022) and the normalization information for each taxon provided in the supplementary material 1 of this paper.

Coal-ball floras reveal anatomical rather than gross-morphological data compared to adpressions and so require a different data filtration strategy. As with adpressions, stem taxa were collated for lycopsids and foliage taxa for the ferns. However, with the pteridospermous (seed-fern) seed-plant adpressions, the fossil-taxa are mainly based on pinnule shape and venation, and frond architecture, which are difficult to determine in coal-balls. Consequently, pteridosperm diversity was based on the seed / ovule taxa as these are identified based on internal anatomy which can be accurately determined in peels and thin sections (Bateman et al., in reviewb); these features can rarely be seen in adpressions. Since they are based on different plant parts, it is impossible to integrate the adpression and coal-ball data-sets, but it should nevertheless be possible to compare the two diversity trends revealed internally within each data-set. The normalization information of each taxon is provided in the supplementary material 2.

Paleozoic swamp habitats were characterised by a strong ecological partitioning of the major plant groups (DiMichele et al., 2001, 2005) and so previous studies on this biome have tended to look separately at the diversity patterns of each major plant group (lycopsids, calamitaceans, sphenophyllaceans, marattialean ferns, non-marattialean ferns, lyginopteridaleans, medullosaleans, cordaitaleans); this approach has been followed here. Although the fossil-taxa being compiled as part of the study each represent a part of the parent plant, there was no difficulty in assigning them to these major plant groups (for further discussion see Cleal et al., 2012).

3.2. Data collation

This study is based on presence / absence records of selected fossils-

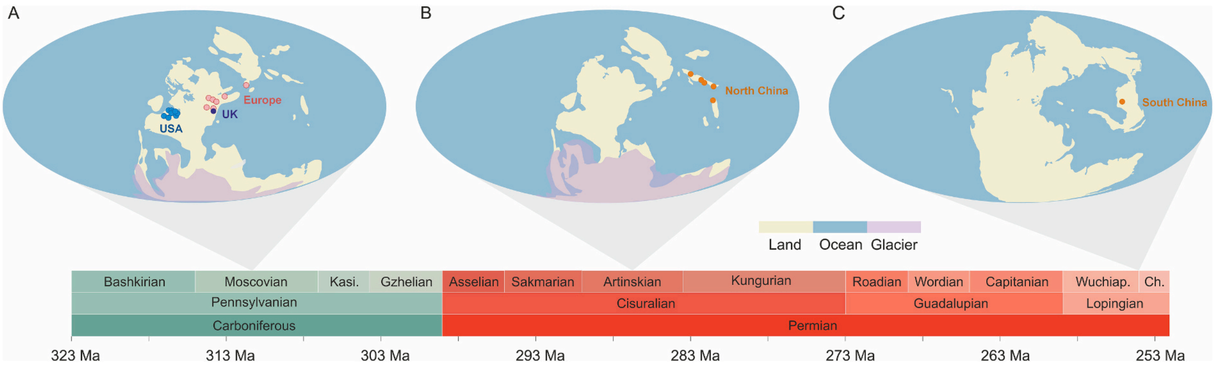


Fig. 3. Representative world palaeogeography maps showing locations of coal ball floras during the Pennsylvanian **(A)**, early Permian **(B)**, and late Permian **(C)**. Abbreviation: Wuchiap. = Wuchiapingian, Ch. = Changhsingian. Adapted from Hilton and Cleal (2007), Wang et al. (2009), and Montañez et al. (2016).

species through relevant stratigraphical succession collated from the references above. Occurrences were subdivided by stages of the current IUGS Global Chronostratigraphical Scale (Cohen et al., 2025). Because of the sampling problems often inherent with such data, the Range-Through Method (sensu Boltovskoy, 1988) was used where a taxon is recorded as being present at a stratigraphical level, even if it has not been found there, if it has been recorded at both higher and lower stratigraphical levels. Although this means that some of the fine details of the diversity dynamics are lost, the resulting smoothed data allow broad, landscape level changes to be readily identified (Cleal et al., 2021; Pardoe et al., 2021). The results were a binary data matrix in a spreadsheet, which could then be subjected to multivariate numerical analysis. The datasets are available in the supplementary data files 1 and 2.

3.3. *Methods of analysis*

After extending the distribution ranges of normalized plant fossils, we assessed the regional taxonomic turnover by identifying the last and first occurrences of taxa within each stage in the South China flora. Here, we counted the number of taxa that disappeared relative to the subsequent stage and those that newly appeared compared to the previous stage, defining these as local extinction and origination events. These reflect the first and last appearances of taxa in South China without reference to their global stratigraphic ranges. Since South China served as the final refugium of the Paleozoic Cathaysian flora and one of the earliest sites of Triassic pioneer lycopod emergence, these local extinction and origination rates may also approximate global-scale patterns (Yu et al., 2022; Xu et al., 2025). Nevertheless, we emphasize that these patterns are based on regional presence–absence data. Extinction and origination rates were calculated by dividing the number of disappearing and appearing taxa by the total number of normalized taxa within each stage. These rates serve as a measure of floral turnover. The difference between the extinction and origination rates yields the net turnover rate, commonly used to assess the magnitude of biotic change. A positive net turnover rate indicates a phase of net loss (interpreted as extinction-dominated), whereas a negative value suggests net recovery or floral diversification.

Distributional patterns within the binary (presence / absence) datasets were also investigated using ordination. Since we are looking at stratigraphic trends, the data has essentially linear structures and so the widely-used principal component analysis (PCA) algorithm was suitable (see discussion in Pardoe et al., 2021). In the PCA analysis, presence / absence data were converted into a binary matrix (1/0) and saved in the supplementary files. These matrices can be read and analysed using the free access software PAST (https://palaeo-electronica.org/2001_1/past/pastprog/index.html), employing the Euclidean distance method. The resulting PCA plots are colour-edited in CorelDRAW to ensure stylistic consistency across figures. Since the Permian–Triassic data from South China integrate evidence from multiple basins, all sections and outcrops within each stage are combined into a single representative flora, named after the corresponding lithological formation. This approach emphasizes stratigraphic trends while minimizing the "noise" introduced by palaeogeographic variation and the "Signor–Lipps" effect (Rees, 2002). In contrast, the coal-ball data were categorised into three distinct provinces (the Bashkirian to Gzhelian of Europe, the Moscovian to Gzhelian of North America, and Gzhelian to Lopingian of China) each of which was analysed separately by PCA at both species and genus rank, with the aim of revealing the phytogeographic evolution of floral the provinces.

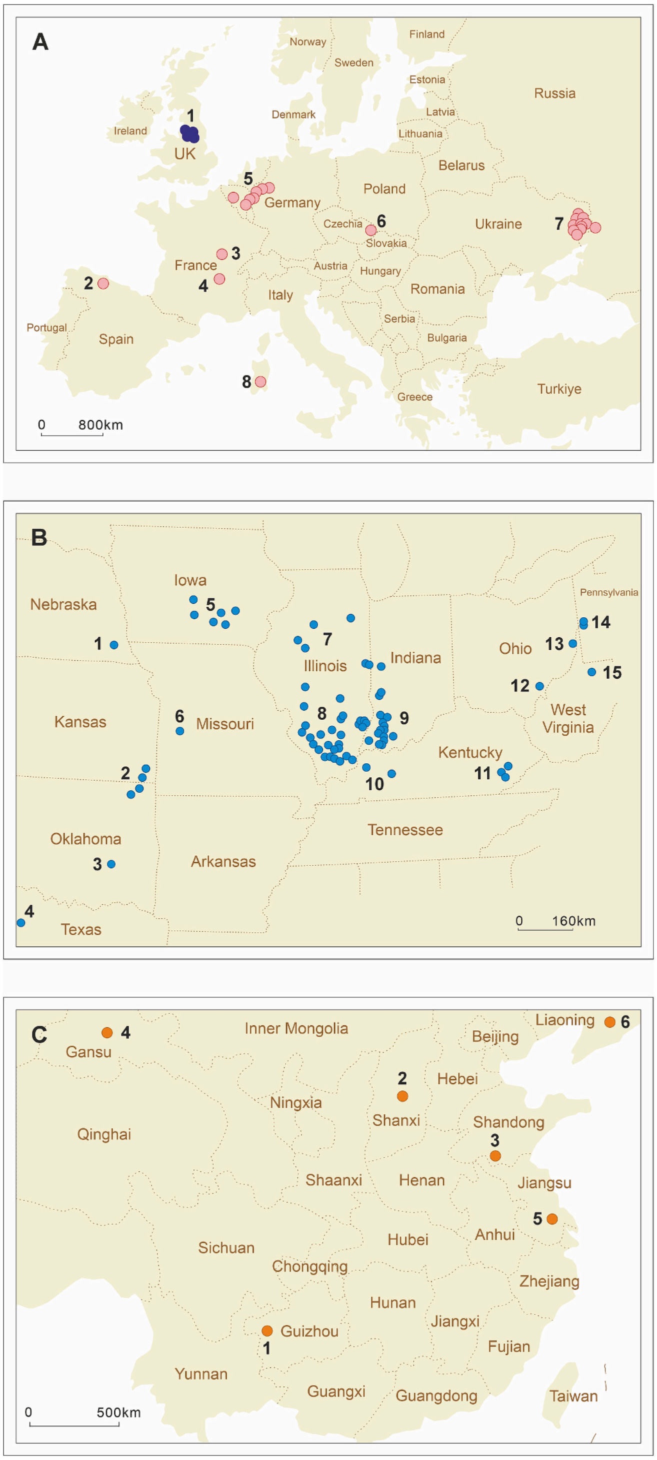
4. Case studies

4.1. *Taxonomic richness and palaeofloristics of Permian-Triassic floras in South China*

The aim here was to estimate the magnitude of vegetation change through the Permian – Triassic Mass Extinction (PTME), as detailed in Xu et al. (2022, 2025). Using the normalized plant macrofossil distributions with extirpated ranges, we calculated the number of extinct and originating taxa, as well as the extinction and origination rates for each stratigraphical interval from extinction to recovery.

Normalization resulted in the removal of up to 33 fossil-species (21.7 %) from the end-Permian Changhsingian Xuanwei and Dalong floras, reducing species richness across the floras (see Fig. 5 for details). Nevertheless, the broad pattern of taxonomic diversity was similar with both raw and normalized data, with a progressive increase through much of the Permian, a reduction in the uppermost Permian (Changhsingian), and an even more dramatic fall after the Permian – Triassic boundary (Fig. 5). Diversities remain relatively low until the Upper Triassic, where they recover in the Carnian / Norian (Fig. 5). However, there are clear differences in the extinction and origination rates, particularly in the Lower Triassic (Induan) where there is a significantly reduced extinction rate seen in the normalized data but not the raw data (Fig. 6). This has potentially important implications for understanding the post-PTME recovery of floral taxonomic diversity.

The two-dimensional PCA of the raw species data was unable to separate the Permian from the Triassic floras (Fig. 7A). Notably, the distance within the Paleozoic floras—such as Xuanwei and Longtan/Maokou—is even larger than the distance between the Paleozoic and Mesozoic floras in the raw species data PCA result, which contradicts the established classification of these floras (Fig. 7A). Although the normalization of the species data can roughly separate the Paleozoic from the Mesozoic floras, the result did not reveal a clear meaningful pattern (Fig. 7B). The PCA of the raw and normalized genus data, in contrast, provides a broadly similar evolutionary picture of the vegetation and its relation to palaeogeography (Fig. 7C, D). In the normalized genus level data result (Fig. 7D) the floras from each stratigraphic stage



(caption on next column)

Fig. 4. Present day maps showing locations of coal-ball floras from the late Pennsylvanian of Europe **(A)**, late Pennsylvanian of North America **(B)**, and Permian of China **(C)**. Localities in (A) are from Phillips (1980), including: 1. First, Union, Upper foot, and Halifax in United Kingdom (UK); 2. Truebano, Spain; 3. Autun, France; 4. Grand'Croix, France; 5. Bouxharmont in Belgium and Finefrau, Aegir, Hauptfloz, Katharina in Netherlands and Germany; 6. Koksfloz in Czech-Republic; 7. Kalmious, K8, Un-designated K5 in the Donets Basin (Ukraine); 8. Sardinia. Italy. Localities in (B) are from Phillips (1980), including: 1. Nebraska coals; 2. Iron-Post coals, Mineral-Flemming coal, and Weir-Pittsburgh coal; 3. Secor coal; 4. Newcast; 5. Urbandale & Shuler (Rock Island), Indianola (Cherokee Group coal), Lovilla (Buffaloville), Oskaloosa (Rock Island & Buffaloville), Ottumwa (Cherokee Group coal), and What Cheer (Cherokee Group coal); 6. Appleton City (above Tebo coal); 7.Herrin coal, Colchester coal, Summum coal; 8. Herrin coal, Danville coal, Springfield coal, Briar Hill coal, Opdyke coal, Calhoun coal, Unnamed coal member of the Shumway Cyclothem, Unamed coal, Friendsvulle coal; 9. Lower Block coal, Buffaloville coal, Murphysboro equivalent, DeKoven coal, Springfield coal, Briar Hill coal, Parker coal; 10. Herrin coal, Baker coal; 11. Shack Branch (Copland coal), Lewis Creek (Copland coal), Bear Branch (Copland coal); 12. Redstone/Pittsburgh coal; 13. Ames/Duquesne coal; 14. Above Middle Kittanning; Morgantown (Pittsburgh coal). Localities in (C) are from Hilton and Cleal (2007) and Wang et al. (2009), including: 1. Wangjiazhai coal, Shuicheng, Guizhou; 2. Xishan coal, Taiyuan, Shanxi; 3. Taozhuang, Zaozhuang, Zouxian, and Xintai of Shandong; 4. Jingyuan, Gansu; 5. Xuzhou, Jiangsu; 6. Chaoyang, Liaoning.

cluster more tightly and are distinctly separated from those of other stages, indicating pronounced floral turnover across time in South China. There is clear trend along Axis 1 from the Permian floras (high scores) to the Triassic floras (low scores, with the transitional Kayitou flora having intermediate scores. The Triassic floras also show a broad trend on PCA Axis 2 corresponding to stratigraphical position, with the Lower and Middle Triassic floras having low scores and the Upper Triassic having high scores.

4.2. Carboniferous-Permian coal-ball palaeobiogeography

This study analysed the data at both species and genus ranks. Normalization retained only 23 % of the original (raw) data for both genus and species level analyses (Fig. 8), rendering 77 % of the raw data as "systematic noise" in each dataset. PCA results of the genus level analysis for the raw data (Fig. 9A) shows considerable overlap between the coal-ball floras from Europe (EU), North America (US) and China (CN), and a minor overlap between Continental EU and the United Kingdom (UK). As all the ordination areas overlap, it suggests a broad coal-ball floral distribution but with individual floras occupying only part of the total range. In this result the coal-ball floras of China represent a distinct subset of the stratigraphically earlier Pennsylvanian floras of Europe and America, and the floras of the UK and US are distinct from each other but each overlap with a different part of the EU ordination area. Although the coal-ball floras of Europe and America have substantial overlap, they also have their own unique ordination areas.

PCA of the species coal-ball data (Fig. 9A, B) are broadly similar at the generic rank but with important differences. The raw species data (Fig. 9A) show slight overlap between the floras of EU and the UK, and extensive overlap between the floras of EU and US, the latter completely enclosing the CN data. The normalized species data (Fig. 9B) show a broadly similar pattern along Axis 1, but along Axis 2 there is an inverted arrangement of the main areas and, more importantly, a greater separation of the CN data, suggesting a greater degree of Cathaysian endemism.

The normalized genus level PCA result (Fig. 9D) shows broadly the same relationships as the raw data (Fig. 9C), but with the Chinese flora having less overlap with the floras from US, and the UK flora having greater overlap with the EU flora. In both the raw and normalized data, the coal-ball floras show a global distribution with each region sharing

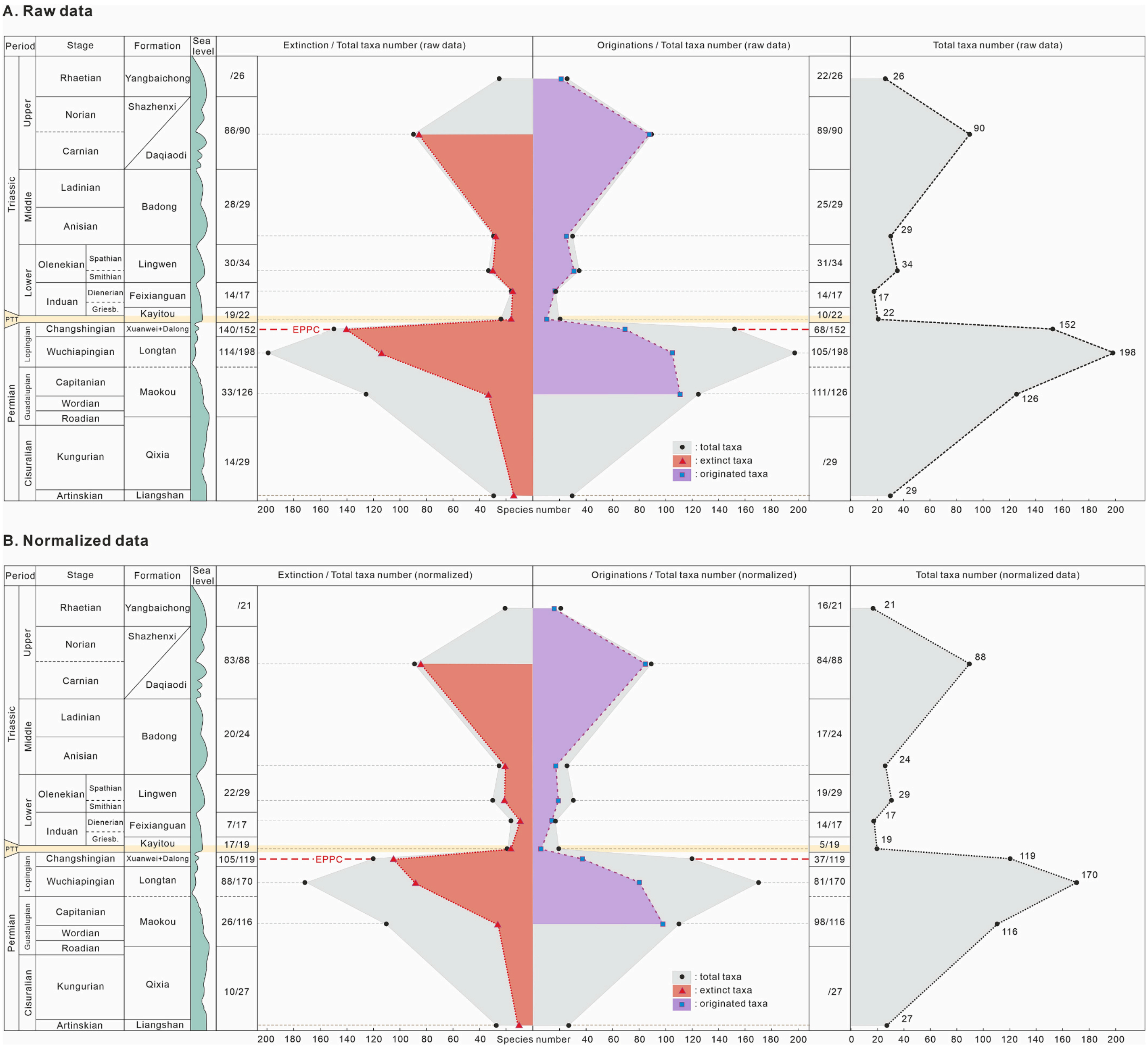


Fig. 5. Total, extinction and origination taxon number from the middle Permian to Late Triassic floras of South China using raw **(A)** and normalized data **(B)**. In each number combination, the number to the left of the slash indicates the extinction or origination taxon number, and to the right of the slash is the total taxa number. See raw and normalized data in the Supplementary Materials.

morphospace with one or more other region and having its own unique morphospace.

5. Discussion

Both the raw and normalized data for the Permian – Triassic plant macrofossils of equatorial South China show broadly similar diversity patterns (Figs. 5, 6). In both datasets there are progressive increases in diversity through most of the Permian, representing the proliferation of the Cathaysian Flora (Paleophytic sensu Cascales-Miñana and Cleal, 2014, Cleal and Cascales-Miñana, 2014, Cleal and Cascales-Miñana, 2021), followed by a major diversity decline through the Permian – Triassic transition. This is initially seen in the Changhsingian Xuanwei and Dalong formations, with the extinction of a ca. 88 % of the lowland taxa (based on normalized data), followed by the extinction of all remaining Cathaysian survivors by the Kayitou flora; the latter becomes dominated (50 %) by pioneer herbaceous lycopods and *Peltaspermum*. Diversity initially remains low but starts to recover higher up in the Lower Triassic, mainly due to a reduction in extinction rates in the Feixianguan Formation, although this is only shown by the normalized data. The succeeding Lingwen and Badong floras are reported to show an increase in vegetation biomass (Xu et al., 2022, 2025), but extinction and origination rates remain approximately constant and balance each other out, resulting in stable diversity figures in both raw and normalized data for the Olenekian and Anisian Stages. There is a major change in the Carnian – Norian Stages, although the absence of intermediate floras in our dataset makes it difficult to meaningfully interpret the details of this floristic change, a diversification in Mesozoic gymnosperm-dominated floras. However, the timing and magnitude of these changes differ between these datasets. Xu et al. (2022) identified the parent plant groups of these spore and pollen data that came from the same strata as the plant macro fossils. The result indicates

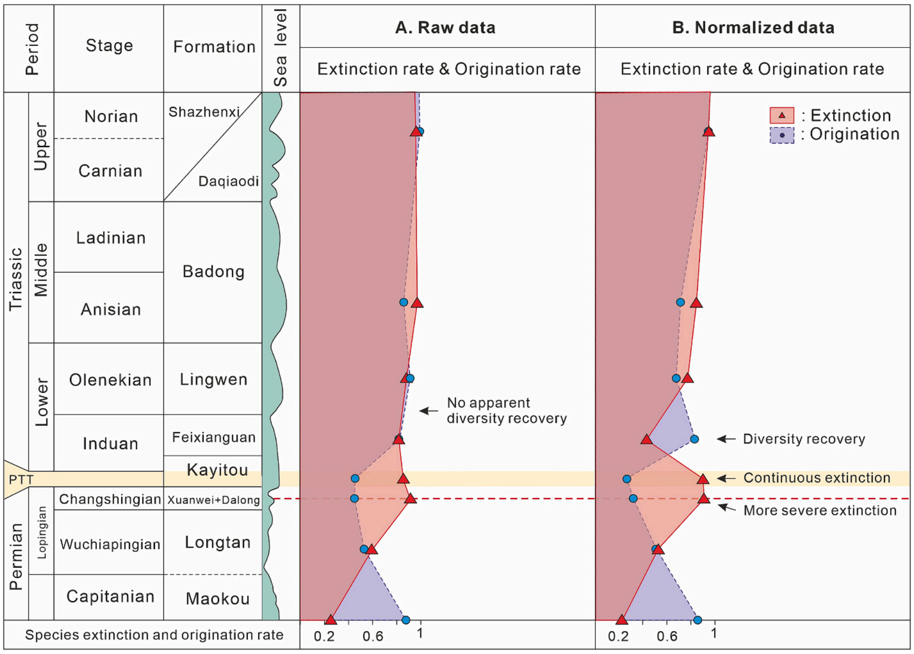


Fig. 6. Extinction and origination rates from the middle Permian to Late Triassic floras in South China calculated using raw **(A)** and normalized data **(B)**. See raw and normalized data in the Supplementary Materials. Arrows point out the main difference between sub-fig. A and B.

macrofossil assemblages predominantly represent lowland vegetation, while the palynology data capture signals from the “hidden” floras outside of the lowland preservation window (DiMichele et al., 2020). Notably, 30–50 % spores and pollen taxa derived from the Early Triassic already existed in end-Permian strata—prior to the Permian–Triassic Mass Extinction (PTME)—supporting the hypothesis that low-latitude and extrabasinal regions functioned as evolutionary cradles and refugia for post-extinction pioneer lineages (Xu et al., 2022). The higher environmental and climatic perturbation stresses in the extrabasinal regions prior to the main extinction event likely pre-adapted these lineages to withstand the subsequent extreme warming episodes (Looy et al., 2014; Blomenkemper et al., 2018). This pattern explains why palynological records typically exhibit lower extinction magnitudes and more immediate recoveries—similar to patterns observed in Devonian vegetation records (Knoll et al., 1984). However, the palynological record should be interpreted with caution as palynomorphs are readily recycled in sedimentary systems over extended time intervals, mixing their provenance and diluting their reliability as direct proxies of the source flora (e.g., Truswell and Drewry, 1984; Traverse, 1990; Batten, 1991).

Nevertheless, the discovery of rare macrofossil remains of the lycopsid *Tomiostrobus* (*Annalepis*) and the seed fern *Germaropteris* (vegetative leaf: *Lepidopteris* and cone: *Peltaspermum*) in the Early Triassic of South China suggests the lowland preservation window was not fully closed after the PTME. Instead, the rarity of macrofossils and the absence of macro remains for many taxa known from palynology likely reflect extremely low biomass in lowland ecosystems—the centre of terrestrial carbon sequestration. This supports the interpretation that the PTME did not necessarily constitute a complete loss of terrestrial plant diversity, but rather a dramatic collapse in vegetation biomass, as mirrored by the plant macro fossils records. Recent Earth system modelling shows that the extinction of land plant productivity during the PTME substantially reduced organic carbon burial linked to photosynthesis and biologically driven weathering processes, thereby prolonging Early Triassic greenhouse conditions for nearly five million years (Xu et al., 2025). In summary, palynological data offer broader insights into regional vegetation composition, while plant macrofossils provide crucial information on ecosystem structure and plant population abundance on land (e.g., Knoll et al., 1984; Xu et al., 2022).

Raw species data (Fig. 7A) provides limited resolution in distinguishing between Paleozoic and Mesozoic floras. In contrast, only the analyses based on normalized species and genus data (Fig. 7B and D) clearly separate these groups. The most distinct and interpretable pattern emerges from the normalized genus data (Fig. 7D), particularly in the placement of the transitional Kayitou flora, which now clearly occupies an intermediate position between the Permian (Cathaysian) and Triassic floras. Additionally, this dataset more precisely differentiates floras at the stage level. Overall, genus-level data appear to better capture evolutionary changes in floral composition, whereas species-level data may be more affected by taxonomic uncertainties and nomenclatural inconsistencies, potentially obscuring true evolutionary patterns.

With the coal-ball data, PCA reveals relatively little differentiation at the genus rank with both raw and normalized data, except to an extent for the stratigraphically older UK data. At the species rank, however, a clearer pattern is shown, with a clear transition from the UK to the EU to the US floras. This is most unequivocally seen in the normalized data, which also shows a clearer positioning of the CN data low on Axis 2. The relative uniformity of all these floras at the genus rank appears to confirm the general similarity of the vegetation of this biome through the Pennsylvanian and Permian, across palaeotropical Euramerica and Cathaysia, suggesting that we are seeing evidence of large-scale biome migration at this time in response to palaeogeographic changes. There are nevertheless major microevolutionary floristic changes at the rank of species that are essential to document for a full understanding of the evolution of this biome, but this can only be done reliably using normalized data.

It is important to recognize that the results presented above for each case study is based on datasets we have established and normalized for quality and consistency. These represent important analyses in their own rights but also serve as base-level data for future studies to expand the coverage to include additional occurrences as they are discovered. We also note the potential in the future to merge adpression data with coal-ball data based on unique assemblages such as the Wuda tuff flora (Wang et al., 2021a) that contain both adpression and anatomically preserved plants. While the taphonomic context of the Wuda tuff clearly differs from that of coal-ball forming environments, taxa recognised in coal-balls are increasingly found in the tuff, including *Psaronius* stems

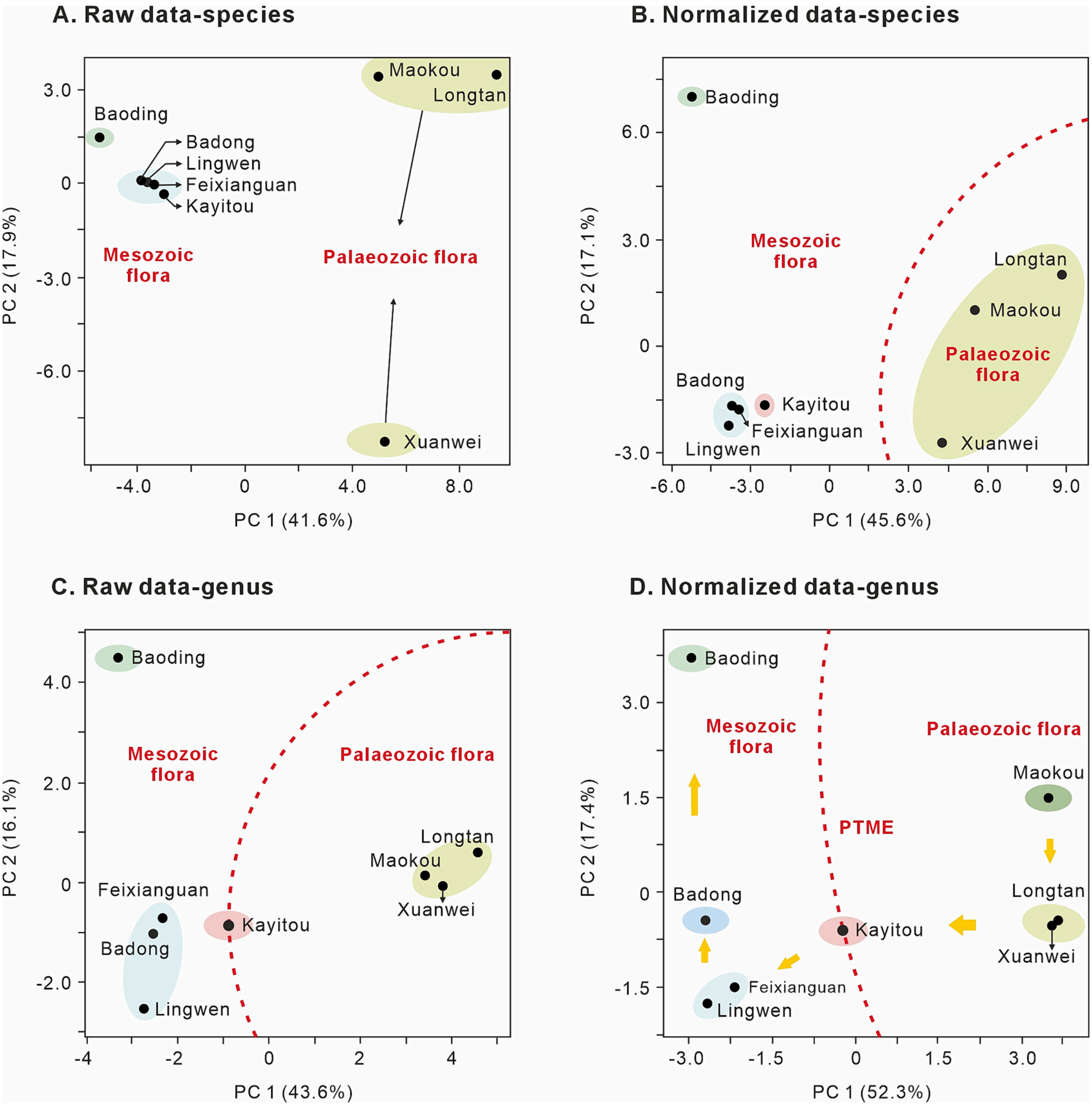


Fig. 7. Two-dimensional principal component analysis showing relationships between middle Permian to Triassic floras from the South China Plate using: species raw data **(A)**, normalized species data **(B)**, raw genus data **(C)** and normalized genus data **(D)**. Yellow arrows show the inferred direction of floral evolution through time. PTME = Permian Triassic Mass Extinction. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Wang et al., 2021b), *Scolecopteris* synangia (Li et al., 2025; Sun et al., 2024), and stems of *Botryopteris/Botryopteridium* (Zhou et al., 2023).

As a final point, we consider it important to recognize that normalizing palaeobotanical data becomes more important further back in geological time related to the "pull of the recent" and greater understanding of geologically recent floras with better whole-plant templates.

6. Conclusions

- Normalization is an essential step in evaluating fossil plant data when trying to investigate a primary floristic signal freed from taphonomic and taxonomic noise. It is especially important for assessing changes in taxon richness, floristic composition, and biogeographic patterns through time.
- Understanding whole-plant templates is key to deciphering signals hidden in fossil plant macrofossil raw taxonomic diversity and

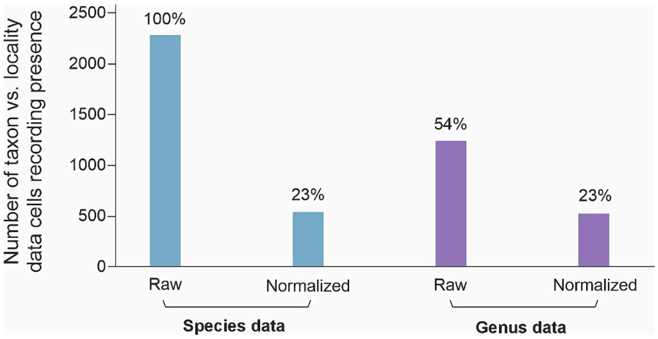


Fig. 8. Bar chart showing number of presence records (locality-taxon matrix cells with presence = 1) used to assess data redundancy.

richness counts and presence/absence palaeobiogeographic data, to recover primary signals and to minimize taxonomic artefacts of the fossil record.

- Normalizing taxonomic data aims to remove duplications; it is to some extent subjective and therefore may be subject to errors, but these are likely to be minimised compared to raw data analyses. Normalization requires the integration of more data to further approach the true whole-plant diversity.
- In our South China Permian-Triassic floral analysis, normalization enabled the evolutionary magnitude and patterns through the PTME to be examined more precisely, helping to develop a more comprehensive understanding of the co-evolution of plants and the Earth system during mass extinction events.
- In our coal-ball study, the genus level analysis reveals a relative uniformity in the make-up of the coal swamp vegetation through Pennsylvanian and Permian times across palaeotropical Euramerica

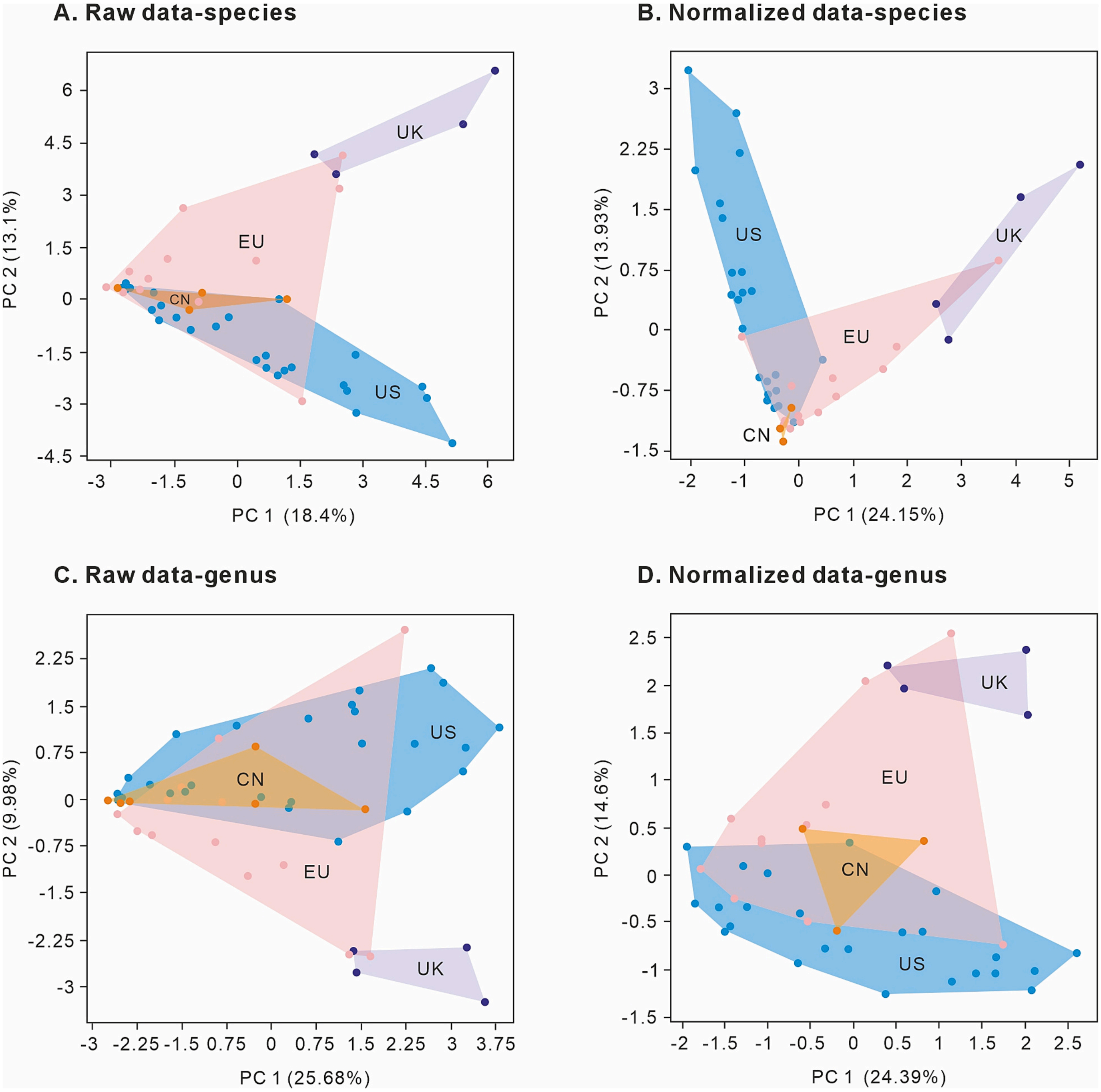


Fig. 9. Two-dimensional principal component analysis from the coal-ball data using raw species data **(A)**, normalized species data **(B)**, raw genus data **(C)**, and normalized genus data **(D)**. Name of each flora can be found in supplementary data file 2. EU = Europe, US = North America, CN = China, UK = United Kingdom.

and Cathaysia. In contrast, species level data show clearer floristic trends indicating microevolutionary evolution of the biome through time, including the greater distinction of the stratigraphically younger Chinese swamp vegetation. Normalization here allows for a clearer identification of Cathaysian endemism.
- Overall, using normalized fossil plant macrofossil data alongside raw datasets is recommended to allow for a deeper understanding of the data being analysed.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.palaeo.2025.113236.

CRediT authorship contribution statement

Zhen Xu: Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Writing – review & editing, Writing – original draft. **Emma Bateson:** Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Christopher J. Cleal:** Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Conceptualization, Writing – review & editing. **Reece Hutton:** Investigation, Data curation. **Jianxin Yu:** Supervision, Resources, Project administration, Investigation, Funding acquisition, Data curation. **Shi-Jun Wang:** Investigation, Funding acquisition, Data curation. **Andrew H. Knoll:** Validation, Methodology, Investigation, Writing – review & editing. **Benjamin J.W. Mills:** Validation, Supervision, Resources, Investigation, Funding acquisition, Writing – review & editing. **Jason Hilton:** Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization, Writing – review & editing, Writing – original draft.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgements

Case study one was originally developed as part of a PhD investigation for Zhen Xu at China University of Geosciences, Wuhan, supervised by Jianxin Yu, Hongfu Yin, and Jason Hilton, and completed during a post-doctoral position at the University of Leeds supervised by Benjamin Mills. This work is funded by the Natural Science Foundation of China (grant 42430209). Zhen Xu and Benjamin Mills are funded by the UKRI project EP/Y008790/1. Case study two was developed as BSc dissertations by Emma Bateson and Reece Hutton at the University of Birmingham supervised by Jason Hilton and Christopher Cleal. We thank Prof. Shao Longyi (China University of Mining and Technology, Beijing) for inviting Jason Hilton to present on fossil plants in palaeobiogeography at the 6th International Conference of Palaeogeography (Nanjing, 2024), leading to the framework for this contribution. Jason Hilton acknowledges financial support from the research funds of the School of Geography, Earth and Environmental Sciences at the University of Birmingham.

Data availability

All the data used in this study are provided in the supplementary files.

References

Bambach, R.K., 2006. Phanerozic biodiversity mass extinctions. Annu. Rev. Earth Planet. Sci. 34, 127–155. https://doi.org/10.1146/annurev.earth.33.092203.122654.

Bambach, R.K., Knoll, A.H., Wang, S.C., 2004. Origination, extinction, and mass depletions of marine diversity. Paleobiology 30 (4), 522–542. https://doi.org/10.1666/0094-8373(2004)030<0522:OEAMDO>2.0.CO;2.

Bateman, R.M., 1991. Palaeoecology. In: Cleal, C.J. (Ed.), Plant Fossils in Geological Investigations: The Palaeozoic. Horwood, Chichester, UK, pp. 34–116.

Bateman, R.M., Hilton, J., 2009. Palaeobotanical systematics for the phylogenetic age: applying organ-species, form-species and phylogenetic species concepts in a framework of reconstructed fossil and extant whole-plants. Taxon 58, 1254–1280. https://doi.org/10.1002/tax.584016.

Batten, D.J., 1991. Reworking of plant microfossils and sedimentary provenance. In: Morton, A.C., Todd, S.P., Haughton, P.D.W. (Eds.), Developments in Sedimentary Provenance Studies. Geol. Soc. Special Publication no, vol. 57, pp. 79–90.

Behrensmeyer, A.K., Kidwell, S.M., Gastaldo, R.A., 2000. Taphonomy and paleobiology. Paleobiology 26, 103–147. https://doi.org/10.1017/S0094837300026907.

Blomenkemper, P., Kerp, H., Abu Hamad, A., DiMichele, W.A., Bomfleur, B., 2018. A hidden cradle of plant evolution in Permian tropical lowlands. Science 362 (6421), 1414–1416. https://doi.org/10.1126/science.aau4061.

Boltovskoy, D., 1988. The Range-through Method and first-last appearance data in paleontological surveys. J. Paleontol. 62, 157–159. https://www.jstor.org/stable/1305273.

Boulter, M.C., Spicer, R.A., Thomas, B.A., 1986. Patterns of plant extinction from some palaeobotanical evidence. In: Spicer, R.A., Thomas, B.A. (Eds.), Systematic and Taxonomic Approaches in Palaeobotany, Systematics Association, Special Volume. Systematics Association, London, UK, pp. 1–36.

Brongniart, A., 1822. Sur la classification et la distribution des végétaux fossiles en général, et sur ceux des terrains de sédiment supérieur en particulier. Mém. Mus. natl. hist. nat. 8 (203–240), 297–348.

Broutin, J., Yu, J.X., Shi, X., Shu, W.C., Xue, Q., 2020. Terrestrial palaeofloral succession across the Permian–Triassic boundary in the north and South China blocks: a brief review. Paläontol. Z. 94 (1), 1–12. https://doi.org/10.1007/s12542-020-00511-0.

Cai, L., Zhu, Y., 2015. The challenges of data quality and data quality assessment in the big Data Era. Data Sci. J. 14, 1–10. https://doi.org/10.5334/dsj-2015-002.

Cascales-Miñana, B., Cleal, C.J., 2014. The plant fossil record reflects just two great extinction events. Terra Nova 26, 195–200. https://doi.org/10.1111/ter.12086.

Chaloner, W.G., 1986. Reassembling the whole fossil plant, and naming it. In: Spicer, R. A., Thomas, B.A. (Eds.), Systematic and Taxonomic Approaches in Palaeobotany, Systematics Association, Special Volume. Systematics Association, London, UK, pp. 67–78.

Chen, J.H., Yu, J.X., Huang, Q.S., Broutin, J., Song, Q.Q., Chen, B., 2011. New research progress on the paleoflora in the earliest Triassic of western Guizhou and eastern Yunnan, South China. Earth Sci. 36 (3), 500–510 (In Chinese with English abstract).

Cleal, C.J., 1986. Identifying plant fragments. In: Spicer, R.A., Thomas, B.A. (Eds.), Systematic and Taxonomic Approaches in Palaeobotany, Systematics Association, Special Volume. Systematics Association, London, UK, pp. 53–65.

Cleal, C.J., 2005. The Westphalian macrofloral record from the cratonic central Pennines Basin, UK. Z. Dtsch. Ges. Geowiss. 156, 387–402. https://doi.org/10.1127/1860-1804/2005/0156-0387.

Cleal, C.J., 2007. The Westphalian–Stephanian macrofloral record from the South Wales Coalfield, UK. Geol. Mag. 144, 465–486. https://doi.org/10.1017/S0016756807003305.

Cleal, C.J., 2008. Westphalian-Stephanian macrofloras of the Southern Pennines Basin, UK. Stud. Geol. Pol. 129, 25–41.

Cleal, C.J., Cascales-Miñana, B., 2014. Composition and dynamics of the great Phanerozoic Evolutionary Floras. Lethaia 47, 469–484. https://doi.org/10.1111/let.12070.

Cleal, C.J., Cascales-Miñana, B., 2021. Evolutionary Floras – revealing large-scale patterns in Palaeozoic vegetation history. J. Palaeosci. 70, 31–42. https://doi.org/10.54991/jop.2021.9.

Cleal, C.J., Thomas, B.A., 1994. Plant Fossils of the British Coal Measures, Palaeontological Association field guides to fossils. Palaeontological Association, London.

Cleal, C.J., Thomas, B.A., 2010. Botanical nomenclature and plant fossils. Taxon 59, 261–268. https://doi.org/10.1002/tax.591024.

Cleal, C.J., Thomas, B.A., 2019. Introduction to Plant Fossils. Cambridge University Press.

Cleal, C.J., Thomas, B.A., 2021. Naming of parts: the use of fossil-taxa in palaeobotany. Foss. Imp. 77, 166–186. https://doi.org/10.37520/fi.2021.013.

Cleal, C.J., Wang, Z.Q., 2002. A new and diverse plant fossil assemblage from the upper Westphalian Benxi Formation, Shanxi, China, and its palaeofloristic significance. Geol. Mag. 139 (2), 107–130. https://doi.org/10.1017/S0016756801006240.

Cleal, C.J., Uhl, D., Cascales-Miñana, B., Thomas, B.A., Bashforth, A.R., King, S.C., Zodrow, E.L., 2012. Plant biodiversity changes in Carboniferous tropical wetlands. Earth Sci. Rev. 114, 124–155. https://doi.org/10.1016/j.earscirev.2012.05.004.

Cleal, C.J., Pardoe, H.S., Berry, C.M., Cascales-Miñana, B., Davis, B.A.S., Diez, J.B., Filipova-Marinova, M.V., Giesecke, T., Hilton, J., Ivanov, D.A., Kustatscher, E., Leroy, S.A.G., McElwain, J.C., Opluštil, S., Popa, M.E., Seyfullah, L.J., Stolle, E., Thomas, B.A., Uhl, D., 2021. Plant diversity in deep time: 1. How well can we identify past plant diversity in the fossil record? Palaeogeogr. Palaeoecol. Palaeoclimatol. 576, 110481. https://doi.org/10.1016/j.palaeo.2021.110481.

Cohen, K., Harper, D., Gibbard, P., Car, N., 2025. The ICS international chronostratigraphic chart this decade. Episodes 48, 105–115. https://doi.org/10.18814/epiiugs/2025/025001.

DiMichele, W.A., Stein, W.E., Bateman, R.M., 2001. Ecological sorting of Vascular Plant classes during the Paleozoic Evolutionary Radiation. In: Allmon, W.D., Bottjer, D.J. (Eds.), Evolutionary Paleoecology. Columbia University Press, New York, pp. 285–336. https://doi.org/10.7312/allm10994-013.

DiMichele, W.A., Gastaldo, R.A., Pfefferkorn, H.W., 2005. Plant biodiversity partitioning in the late Carboniferous and early Permian and its implications for ecosystem assembly. Proc. Calif. Acad. Sci. Series 4 56 (Supplement 1), 32–49.

DiMichele, W.A., Bashforth, A.R., Falcon-Lang, H.J., 2020. Uplands, lowlands, and climate: Taphonomic megabiases and the apparent rise of a xeromorphic, drought-tolerant flora during the Pennsylvanian-Permian transition. Palaeogeogr. Palaeoecol. Palaeoclimatol. 559, 109965. https://doi.org/10.1016/j.palaeo.2020.109965.

Feng, Z., Wei, H.B., Guo, Y., He, X.Y., Sui, Q., Zhou, Y., Liu, H.Y., Gou, X.D., Lv, Y., 2020. From rainforest to herbland: New insights into land plant responses to the End-Permian Mass Extinction. Earth Sci. Rev. 204, 103153. https://doi.org/10.1016/j.earscirev.2020.103153.

Ferguson, D.K., 2005. Plant taphonomy: Ruminations on the past, the present, and the future. Palaios 20, 418–428. https://doi.org/10.2110/palo.2005.P05-25p.

Galtier, J., 1997. Coal-ball floras of the Namurian-Westphalian of Europe. Rev. Palaeobot. Palynol. 95, 51–72. https://doi.org/10.1016/S0034-6667(96)00027-9.

Gastaldo, R.A., 1992. Taphonomic considerations for plant evolutionary investigations. Palaeobotanist 41, 211–223. https://doi.org/10.54991/jop.1992.1124.

Hilton, J., Cleal, C.J., 2007. The relationship between Euramerican and Cathaysian tropical floras in the late Palaeozoic: Palaeobiogeographical and palaeogeographical implications. Earth Sci. Rev. 85, 85–116. https://doi.org/10.1016/j.earscirev.2007.07.003.

Isaac, N.J.B., Pocock, M.J.O., 2015. Bias and information in biological records. Biol. J. Linn. Soc. 115, 522–531. https://doi.org/10.1111/bij.12532.

Knoll, A.H., 1986. Patterns of change in plant communities through time. In: Diamond, J., Case, T. (Eds.), Community Ecology. Harper Row, New York, pp. 126–141.

Knoll, A.H., Niklas, K.J., Tiffney, B.H., 1979. Phanerozoic land-plant diversity in North America. Science 206, 1400–1402. https://doi.org/10.1126/science.206.4425.1400.

Knoll, A.H., Niklas, K.J., Gensel, P.G., Tiffney, B.H., 1984. Character diversification and patterns of evolution in early vascular plants. Paleobiology 10 (1), 34–47. https://doi.org/10.1017/S0094837300008009.

Li, X.X., Zhou, Z.Y., Cai, C.Y., Sun, G., Ouyang, S., Deng, L.H., 1995. Fossil Floras in China through the Geological Ages, English edition. Guangdong Sci. Tech. Press, Guangzhou, pp. 1–695. http://ir.nigpas.ac.cn/handle/332004/7973.

Li, H., Yu, J.X., McElwain, J.C., Yiotis, C., Chen, Z.Q., 2019. Reconstruction of atmospheric CO_2 concentration during the late Changhsingian based on fossil conifers from the Dalong Formation in South China. Palaeogeogr. Palaeoclimatol. Palaeoecol. 519, 37–48. https://doi.org/10.1016/j.palaeo.2018.09.006.

Li, D.D., Hilton, J., Sun, W.J., Zhou, W.M., Wan, M.L., Bek, J., Pšenička, J., Wang, J., 2025. Whole-plant reconstruction of the Paleozoic marattialean fern *Scolecopteris libera*. Int. J. Plant Sci. 186. https://doi.org/10.1086/733820.

Liu, D.D., Yang, Z.R., Yang, Y.D., Bao, Y.Y., Liu, B., 2009. Characteristic of the flora in the Zhenzhuchong Formation and the Jurassic–Triassic boundary in the Sichuan Basin. J. Earth Sci. Environ. 31 (3), 254–259 (In Chinese with English abstract).

Liu, L., Pšenička, J., Bek, J., Wan, M., Pfefferkorn, H.W., Wang, J., 2021. A whole calamitacean plant *Palaeostachya guanglongii* from the Asselian (Permian) Taiyuan Formation in the Wuda Coalfield, Inner Mongolia, China. Rev. Palaeobot. Palynol. 294, 104245. https://doi.org/10.1016/j.revpalbo.2020.104245.

Looy, C., Kerp, H., Duijnstee, I., DiMichele, B., 2014. The late Paleozoic ecological-evolutionary laboratory, a land-plant fossil record perspective. Sediment. Rec. 12 (4), 4–18. https://doi.org/10.2110/sedred.2014.4.

Marshall, C.R., 2023. Forty years later: the status of the “big five” mass extinctions. Camb. Prism Extinct. 1, e5. https://doi.org/10.1017/ext.2022.4.

McElwain, J.C., Punyasena, S.W., 2007. Mass extinction events and the plant fossil record. Trends Ecol. Evol. 22, 548–557. https://doi.org/10.1016/j.tree.2007.09.003.

Meng, F.S., 1982. Some Fossil Plants of *Thinnfeldia* from the Jiuligang Formation in Jingmen-Dangyang Basin, Western Hubei. J. Integr. Plant Biol. 24 (6). https://www.jipb.net/EN/Y1982/V24/I6/.

Meng, F.S., Xu, A.W., Zhang, Z.L., Lin, J.M., Yao, H.Z., 1995. Nonmarine Biota and Sedimentary Facies of the Badong Formation in the Yangtze and its Neighbouring Areas. China University of Geosciences Press, Wuhan, pp. 1–76 (In Chinese with English abstract).

Molina-Solís, A., Cleal, C.J., Capel, E., Cascales-Miñana, B., 2023. Vegetation dynamics in the Early–Middle Pennsylvanian coal swamps of the Nord-Pas-de-Calais Coalfield, France. Palaeoworld S1871174X2300046X. https://doi.org/10.1016/j.palwor.2023.04.005.

Montañez, I.P., McElwain, J.C., Poulsen, C.J., White, J.D., DiMichele, W.A., Wilson, J.P., Griggs, G., Hren, M.T., 2016. Climate, $p\mathrm{CO}_2$ and terrestrial carbon cycle linkages during late Palaeozoic glacial–interglacial cycles. Nat. Geosci. 9 (11), 824–828. https://doi.org/10.1038/ngeo2822.

Newell, N.D., 1963. Crises in the history of life. Sci. Am. 208 (2), 76–95. https://www.jstor.org/stable/24936467.

Niklas, K.J., 1988. Patterns of vascular plant diversification in the fossil record: Proof and conjecture. Ann. Mo. Bot. Gard. 75, 35. https://doi.org/10.2307/2399465.

Niklas, K.J., Tiffney, B.H., 1994. The quantification of plant biodiversity through time. Philos. Trans. R. Soc. Lond. B Biol. Sci. 345, 35–44. https://doi.org/10.1098/rstb.1994.0084.

Niklas, K.J., Tiffney, B.H., Knoll, A.H., 1986. Patterns in vascular land plant diversification: An analysis at the species level. In: Valentine, J. (Ed.), Phanerozoic Diversity Patterns. Princeton University Press, Princeton NJ, pp. 97–128. https://doi.org/10.1515/9781400855056.97.

Opluštil, S., Cleal, C.J., 2007. A comparative analysis of some late Carboniferous basins of Variscan Europe. Geol. Mag. 144, 417–448. https://doi.org/10.1017/S0016756807003330.

Opluštil, S., Šimůnek, Z., Pšenička, J., Bek, J., Libertín, M., 2017. A 25 million year macrofloral record (Carboniferous–Permian) in the Czech part of the Intra-Sudetic Basin; biostratigraphy, plant diversity and vegetation patterns. Rev. Palaeobot. Palynol. 244, 241–273. https://doi.org/10.1016/j.revpalbo.2016.11.011.

Pardoe, H.S., Cleal, C.J., Berry, C.M., Cascales-Miñana, B., Davis, B.A.S., Diez, J.B., Filipova-Marinova, M.V., Giesecke, T., Hilton, J., Ivanov, D., Kustatscher, E., Leroy, S.A.G., McElwain, J.C., Opluštil, S., Popa, M.E., Seyfullah, L.J., Stolle, E., Thomas, B.A., Uhl, D., 2021. Plant diversity in deep time: Experiences from the fossil record. 2. Assessing and measuring past biodiversity. Palaeogeogr. Palaeoclimatol. Palaeoecol. https://doi.org/10.1016/j.palaeo.2021.110618.

Phillips, T.L., 1980. Stratigraphic and geographic occurrences of permineralized coal-swamp plants – Upper Carboniferous of North America and Europe. In: Dilcher, D.L., Taylor, T.N. (Eds.), Biostratigraphy of Fossil Plants. Hutchinson and Ross, Stroudburg, Pennsylvania, UA, pp. 26–92.

Ran, W.J., 2021. Taxonomy and Morphological Evolution of the Herbaceous Lycopsids during Induan and Anisian in Hunan, Guizhou and Yunnan, South China. Master thesis,. China University of Geosciences (Wuhan), China.

Raup, D.M., Sepkoski, J.J., 1982. Mass extinctions in the marine fossil record. Science 215, 1501–1503. https://doi.org/10.1126/science.215.4539.1501.

Rees, P.M., 2002. Land-plant diversity and the end-Permian mass extinction. Geology 30 (9), 827–830. https://doi.org/10.1130/0091-7613(2002)030<0827:LPDATE>2.0.CO;2.

Ribeiro, B.R., Velazco, S.J.E., Guidoni-Martins, K., Tessarolo, G., Jardim, L., Bachman, S.P., Loyola, R., 2022. Bdc: a toolkit for standardizing, integrating and cleaning biodiversity data. Methods Ecol. Evol. 13, 1421–1428. https://doi.org/10.1111/2041-210X.13868.

Rothwell, G.W., 1981. The Callistophytales (pteridospermopsida): reproductively sophisticated Paleozoic gymnosperms. Rev. Palaeobot. Palynol. 32 (1), 103–121. https://doi.org/10.1016/0034-6667(81)90076-2.

Rothwell, G.W., 1988. Upper Pennsylvanian Stubenville coal-ball flora. Ohio J. Sci. 88 (1), 61–65. http://hdl.handle.net/1811/23241.

Sepkoski, J.J., 1986. Phanerozoic overview of mass extinction. In: Raup, D.M., Jablonski, D. (Eds.), Patterns and Processes in the History of Life. Springer, Berlin, Heidelberg, pp. 277–295. https://doi.org/10.1007/978-3-642-70831-2_15.

Šimůnek, Z., 2007. New classification of the genus *Cordaites* from the Carboniferous and Permian of the Bohemian Massif, based on cuticle micromorphology. Acta Mus. Nat. Pragae. Ser. B Hist. Nat. 62, 97–210.

Spicer, R.A., 1991. Plant taphonomic processes. In: Allison, P.A., Briggs, D.E.G. (Eds.), Taphonomy: releasing the data locked in the fossil record. Topics in Geobiology, pp. 72–113.

Spicer, R.A., Greer, A.G., 1986. Plant taphonomy in fluvial and lacustrine systems. In: Series in Geology, Notes for Short Course, vol. 15. University of Tennessee Knoxville, pp. 10–26. Land plants.

Sun, W.J., Zhou, W.M., Wu, Y.F., Pšenička, Hilton, J., Wang, J., 2024. *Scolecopteris oxydonta* sp. nov., a new marattialean fern from the early Permian Wuda Tuff Flora. Rev. Palaeobot. Palynol. 323, 105231. https://doi.org/10.1016/j.revpalbo.2024.105231.

Taylor, T.N., Taylor, E.L., Krings, M., 2009. Paleobotany, the Biology and Evolution of Fossil Plants, 2nd ed. Acad. Press, Amsterdam.

Thomas, B.A., 1970. Epidermal studies in the interpretation of *Lepidodendron* species. Palaeontology 13, 145–173.

Thomas, B.A., 1977. Epidermal studies in the interpretation of *Lepidophloios* species. Palaeontology 20, 273–293.

Thomas, B.A., 1989. Disarticulated plant fossils. In: Briggs, D.E.G., Crowther, P. (Eds.), *Palaeobiology*, a Synthesis. Blackwell, Oxford, pp. 421–423.

Thomas, B.A., Cleal, C.J., 1999. Abscission in the fossil record. In: Kurmann, M.H., Hemsley, A.R. (Eds.), The Evolution of Plant Architecture. Royal Botanic Gardens, Kew, pp. 183–203.

Thomas, B.A., Cleal, C.J., 2020. The nomenclature of fossil-taxa representing different preservational states: *Lepidodendron* as a case-study. Taxon 69 (5), 1052–1061. https://doi.org/10.1002/tax.12291.

Tian, B.L., Wang, S.J., Guo, Y.T., Li, H.Q., Chen, G.R., Hong, Z., 1996. Flora of Palaeozoic coal-balls of China. Palaeobotanist 45, 247–254.

Traverse, A., 1990. Studies of pollen and spores in rivers and other bodies of water, in terms of source-vegetation and sedimentation, with special reference to Trinity River and Bay, Texas. Rev. Palaeobot. Palynol. 64, 297–303.

Truswell, E.M., Drewry, D.J., 1984. Distribution and provenance of recycles palynomorphs in surficial sediments of the Ross Sea, Antarctica. Mar. Geol. 59, 187–214.

Turland, N.J., Wiersema, J.H., Barrie, F.R., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S., Kusber, W.H., Li, D.Z., Marhold, K., May, T.W., McNeill, J., Monro, A.M., Prado, J., Price, M.J., Smith, G.F., 2018. International Code of Nomenclature for Algae, fungi, and Plants (Shenzhen Code) Adopted by the Nineteenth International Botanical Congress. Regnum Vegetabile. Koeltz Botanical Books, Glashütten.

Uhl, D., Cleal, C.J., 2010. Late Carboniferous vegetation change in lowland and intramontane basins in Germany. Int. J. Coal Geol. 83, 318–328. https://doi.org/10.1016/j.coal.2009.07.007.

von Sternberg, K.M., 1820. Versuch einer geognostisch-botanischen Darstellung der Flora der Vorwelt. F. Fleischer, Leipzig.

Wang, S.J., Sun, K.Q., Cui, J.Z., Ma, S.M., 2009. Fossil Plants from the Coal-Balls in China. High Education Press, Beijing (in Chinese).

Wang, J., Pfefferkorn, H.W., Opluštil, S., Kerp, H., 2021a. Permian “vegetational Pompeii”: A peat-forming in situ preserved forest from the Wuda Coalfield, Inner Mongolia, China–Introduction to a volume of detailed studies. Rev. Palaeobot.

Palynol. 294, 104502. https://www.sciencedirect.com/science/article/abs/pii/S0034666721001263.

Wang, S.J., Wang, J., Liu, L., Hilton, J., 2021b. Stem diversity in the marattialean tree fern family Psaroniaceae from the earliest Permian Wuda tuff flora. Rev. Palaeobot. Palynol. 294, 104378. https://doi.org/10.1016/j.revpalbo.2021.104378.

Xu, R., Zhu, J.R., Chen, Y., Duan, S.Y., Hu, Y.F., Zhu, W.Q., 1979. China Late Triassic Baoding Flora. Science Press, Beijing.

Xu, Z., Hilton, J., Yu, J.X., Wignall, P.B., Yin, H.F., Xue, Q., Ran, W.J., Hui, L., Shen, J., Meng, F.S., 2022. Mid-Permian to late Triassic plant species richness and abundance patterns in South China: Co-evolution of plants and the environment through the Permian-Triassic transition. Earth Sci. Rev. 232, 104136. https://doi.org/10.1016/j.earscirev.2022.104136.

Xu, Z., Yu, J.X., Meredith, A.S., Yin, H., Hilton, J., Allen, B., Wignall, P.B., Dunhill, A., Gurung, K., Shen, J., Schwartzman, D., Goddéris, Y., Donnadieu, Y., Wang, Y.X., Zhang, Y.G., Poulton, S., Mills, B.J.W., 2025. Early Triassic super-greenhouse climate driven by vegetation collapse. Nat. Commun. 16, 5400. https://doi.org/10.1038/s41467-025-60396-y.

Xue, Q., 2019. The Evolution of Fossil Plants during Late Permian and Early Triassic in Western Guizhou and Eastern Yunnan, South China. Master thesis,. China University of Geosciences (Wuhan), China.

Yao, Z.Q., Xu, J.T., Zhen, Z.G., Mo, Z.G., 1980. Late Permian biostratigraphy and the Permian–Triassic boundary in Western Guizhou and Eastern Yunnan. In: Nanjing Institute of Geology and Palaeontology (Ed.), Stratigraphy and palaeontology of Late Permian coal-bearing formations in Western Guizhou and Eastern Yunnan. Science Press, Beijing, pp. 1–69 (In Chinese).

Yu, J.X., 2008. Floras (macro- and microfloras) and evolutionary dynamics across the Permian–Triassic boundary along Guizhou and Yunnan border, South China. Doctoral thesis. China University of Geosciences (Wuhan) and Universitie Pierre et Marie Curie, Paris 6.

Yu, J.X., Broutin, J., Huang, Q.S., Grauvogel-Stamm, L., 2010. *Annalepis*, a pioneering lycopsid genus in the recovery of the Triassic land flora in South China. Comptes Rendus Palevol 9, 479–486. https://doi.org/10.1016/j.crpv.2010.09.004.

Yu, J.X., Broutin, J., Chen, Z.Q., Shi, X., Li, H., Chu, D.L., Huang, Q.S., 2015. Vegetation changeover across the Permian–Triassic boundary in Southwest China: Extinction, survival, recovery and palaeoclimate: a critical review. Earth Sci. Rev. 149, 203–224. https://doi.org/10.1016/j.earscirev.2015.04.005.

Yu, J.X., Broutin, J., Lu, Z.S., 2022. Plants and Palynomorphs around the Permian-Triassic Boundary of South China. Springer Nature Singapore, Singapore. https://doi.org/10.1007/978-981-19-1492-8.

Zhang, Z.L., Meng, F.S., Sheng, X.C., 1992. Triassic. In: Wang, X.F., Ma, D.S., Jiang, D.H. (Eds.), Hainan Island Geology (1) Stratigraphic Paleontology. Geological Publishing House, Beijing, pp. 161–199.

Zheng, R.H., 2011. China Pre-Mesozoic Structural Sequence and Lithofacies Paleogeography Atlas. Geological Publishing House, Beijing (In Chinese).

Zhou, Z.Y., 1989. Late Triassic plants from Shaqiao, Hengyang, Hunan Province. Palaeontol. Cathayana 4, 131–197. http://ir.nigpas.ac.cn/handle/332004/5891.

Zhou, Z.Y., Li, B.X., 1979. A preliminary study of the early Triassic plants from the Qionghai District, Hainan Island. Acta Palaeontol. Sin. 18 (5), 444–466 (In Chinese with English abstract).

Zhou, W.M., Pšenička, J., Bek, J., Libertín, M., Wang, S.J., Wang, J., 2023. A new species of *Botryopteridium* Doweld from the early Permian Wuda Tuff Flora and its evolutionary significance. Rev. Palaeobot. Palynol. 1 (311), 104849. https://doi.org/10.1016/j.revpalbo.2023.104849.