Journal of Biogeography

WILEY

Journal of Biogeography

METHOD OPEN ACCESS

Functional Signatures of Surface Pollen and Vegetation Are Broadly Similar: Good News for Past Reconstructions of Vegetation

Lucas Dugerdil[1,2] | Odile Peyron[2] | Cyrille Violle[3] | Sébastien Joannin[2] | Guillemette Ménot[1] | Pierre Denelle[4] | Helge Bruelheide[5,6] | Milan Chytrý[7] | Richard Field[8] | Mohamed Z. Hatim[9] | Hamid Gholizadeh[10] | Jiri Dolezal[11,12] | Valério D. Pillar[13] | Kamal H. Shaltout[14] | Franziska Schrodt[8] | Eric Garnier[3]

[1]ENS Lyon, UCBL, CNRS, UMR 5276 LGL-TPE, Lyon, France | [2]ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France | [3]CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France | [4]University of Göttingen, Department of Biodiversity, Macroecology & Biogeography, Göttingen, Germany | [5]Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Halle, Germany | [6]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany | [7]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic | [8]School of Geography, University of Nottingham, Nottingham, UK | [9]Plant Ecology and Nature Conservation Group, Environmental Sciences Department, Wageningen University and Research, Wageningen, the Netherlands | [10]University of Siena, Department of Life Sciences, Siena, Italy | [11]Institute of Botany of the Czech Academy of Science, Průhonice, Czech Republic | [12]Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic | [13]Department of Ecology, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil | [14]Botany Department, Faculty of Science, Tanta University, Tanta, Egypt

Correspondence: Lucas Dugerdil (lucas.dugerdil@ens-lyon.fr)

Received: 11 September 2024 | **Revised:** 21 January 2025 | **Accepted:** 22 January 2025

Funding: This research was funded, in whole or in part, by ANR (Grant ANR-22-CE27-0018-02). V.D.P. received support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil, grant no. 313315/2022-1).

Keywords: Arid Central Asia | biomes | communities | databases | functional biogeography | past climate | plant traits | pollen | trait distribution

ABSTRACT

Aim: Pollen assemblages are widely used to infer paleoenvironment features, aiming at reconstructing both past climates and biomes. However, the functional link between environmental conditions and pollen assemblages is not straightforward and requires thorough testing to be used confidently. Here, we use a trait-based approach to assess the consistency of functional signatures between pollen and plant assemblages.

Location: Arid Central Asia (ACA).

Taxon: Spermatophytes (pollen-producing plants).

Methods: We assess whether trait values and trait distributions are consistent for surface pollen samples and extant vegetation in the Arid Central Asia biogeographic region. A working plant checklist was compiled for ACA in order to assign trait values to pollen types and vegetation taxa. This was done for two widely used methods of pollen aggregation schemes (coarse and fine pollen type depend on the level of pollen identification). The functional signatures of pollen and vegetation samples were compared both at the taxon and community levels, using large-scale trait and vegetation databases, for the six traits of the global spectrum of plant form and function (i.e., plant height, seed mass, leaf area, specific leaf area, nitrogen content per leaf mass, and stem-specific density).

Results: Trait distributions and bivariate trait relationships were broadly similar for pollen types and vegetation taxa, which is also the case for the multivariate spaces of the global spectrum of plant form and function. At the community scale, the trait values weighted by taxon abundance significantly differed among biomes, and these differences were consistent for both pollen aggregation schemes and extant.

Journal of Biogeography, 2025; 52:e15100
https://doi.org/10.1111/jbi.15100

Main Conclusions: The pollen aggregation scheme does not impact the organisation of the functional space of the global spectrum of plant form and function, which compares well with that based on species actually present in the vegetation plots. This is also true at the community scale. These findings are very promising for improving climate and biome reconstructions from pollen assemblages and pave the way to a "pollen functional biogeography".

1 | Introduction

Accurate reconstructions of past vegetation and climate are increasingly required by policy-makers in conservation biology (Barnosky et al. 2017; Carter et al. 2018), researchers in biodiversity predictions and mapping (Cai et al. 2023), and climate change evaluations since it offers potential analogues for the future (Tierney et al. 2020). Reconstruction of plant functional diversity through time has also raised interests since it permits understanding the plant-climate relationships on longer time scales (Gaüzère et al. 2020; Carvalho et al. 2019; van der Sande et al. 2019; Adeleye et al. 2023). Different methods exist to quantitatively reconstruct past climate and vegetation based on pollen assemblages from marine cores or from terrestrial cores collected in lakes and wetlands. The question of the interactions between vegetation composition, as assessed by the relative abundances of pollen types, and climate, is key in palaeoenvironmental sciences (Marquer et al. 2017). Sediment accumulation within lakes and wetlands is a powerful environmental archive to record past pollen rain, mainly reflecting the watershed vegetation dynamics through time (Last et al. 2001).

However, pollen-based past climate and vegetation reconstructions suffer from various biases: sediment types, sampling methods, taphonomy (i.e., pollen fossilisation processes), pollen production, dispersion, and identification (ter Braak et al. 1993; Salonen et al. 2014; Gillison 2019; Cleal et al. 2021). To reduce these uncertainties, several mathematical methods have been developed and applied to reconstruct past vegetation (Prentice et al. 1996; Marquer et al. 2017; Sun et al. 2020; Cruz-Silva et al. 2022) or past climate (Chevalier et al. 2020), and approaches combining multiple methods have been developed to improve the reliability of the reconstructions (Peyron et al. 2013; Salonen et al. 2019). These methods directly map pollen abundances to climatic parameters without considering the intermediary link—the plants producing the pollen—resulting in uncertainties and potential biases (Chevalier et al. 2020). A first attempt to tackle this issue used Plant Functional Types (PFTs) to reconstruct climate from pollen (Prentice et al. 1992; Box 1996; Peyron et al. 1998; Tarasov et al. 1999; Kumke et al. 2004). Still, the results obtained for the Last Glacial Maximum (21,000 years BP) were not satisfactory, mainly because the assignment of non-arboreal taxa to PFT was too broad (Peyron et al. 1998).

To improve pollen-inferred climate and biome reconstructions, Harrison et al. (2010) suggested using continuous plant traits, defined as measurable and heritable morphological, physiological, and phenological characteristics of individuals usually representing a taxon, rather than discrete PFTs (Violle et al. 2007; Van Bodegom et al. 2012; Garnier et al. 2016). There is a growing consensus that a trait-based approach has strong potential to address several pending issues in ecology (for syntheses, see Garnier et al. 2016; Shipley et al. 2016), including the functional bases of how organisms relate to the environment and how the functioning of organisms scale to the community level (Enquist et al. 2015; Garnier et al. 2016). Especially, a trait-based approach instead of PFTs has already been proposed for vegetation modelling (Van Bodegom et al. 2012; Scheiter et al. 2013). Tracking past functional changes of vegetation also permits to revealing evolution in vegetation's ecological strategy in adaptation to climate change (Gaüzère et al. 2020; Adeleye et al. 2023).

In this present study, we compare the trait values and distributions of surface pollen assemblages and vegetation plots to test (1) whether relationships between traits (univariate distributions of trait values and pairwise trait–trait relationships) are conserved between pollen and vegetation. We also test (2) whether the functional space (i.e., the organisation of plant height, seed mass, leaf area, specific leaf area, nitrogen content per leaf mass, and stem-specific density trait values in multivariate space) derived from surface pollen samples matches that of extant vegetation (in vegetation plot records). This is necessary since whole phenotypes rather than individual traits respond to the environment (Anderegg 2023). Once validated, it will be possible to (3) combine sediment pollen with traits (Birks 2020) to assess the functional structure of past vegetation, from which biomes could be reconstructed (Harrison et al. 2010).

A central issue with using this approach is the taxonomic level at which the pollen is identified. There is no unified methodology to do so (De Klerk and Joosten 2007). Some authors retain the original and *finest* level of pollen identification (i.e., pollen type *fine*; Birks et al. 2016; Brussel and Brewer 2021), while others use a *coarser* and standardised identification level (i.e., pollen type *coarse*; Goring et al. 2013; Giesecke et al. 2019). For the aggregation of pollen traits, it is necessary to determine which plant taxa will be selected for aggregation. This is especially the case for family-based pollen types (e.g., Poaceae). In previous studies of pollen-based traits, various approaches have been applied, from global aggregation of families using the average value (Brussel and Brewer 2021) or Bayesian modelling (Veeken et al. 2022) to well-covered regional families (van der Sande et al. 2019) to selection of a single representative species for the entire family (Connor et al. 2018). We evaluate the influence of these two broad methods (*fine* and *coarse* pollen types) on the aggregation of traits based on pollen.

The main objective of this study is to test how comparable the functional space of surface pollen samples and vegetation plots is combined with plant traits. Since the functional space for plant species (Díaz et al. 2016) and plant community (Bruelheide et al. 2019) depicts a universal strategy, we are following the same "bulk" approach based on the biogeographic repartition of the pollen and vegetation sites. The aim is to enable assessment of the functional space of plant communities in the past using past pollen sequences. More specifically, we address the following questions:

1. What is the more reliable aggregation scheme (*coarse* or *fine*) to assign trait values to pollen types to reduce the biases between pollen and vegetation-based functional spaces?
2. Between plant species and pollen types, do we observe consistent trait distributions, univariate relations between traits, and multivariate phenotypic spaces?
3. Are the functional spaces derived from pollen and extant vegetation comparable at the level of communities?

In this present study, these issues are addressed in Arid Central Asia (ACA hereafter Chen et al. 2015; Figure 1). Due to large variations in altitude and climate, this area encompasses a wide range of biomes from warm deserts to taiga, with a predominance of drylands (Dinerstein et al. 2017). These drylands are expected to be rapidly and drastically affected by the spread of deserts into the surrounding steppes due to temperature increase and aridification (Huang et al. 2016).

2 | Materials and Methods

2.1 | Approach and Objectives

Due to extreme temperature seasonality and aridity in this region, ACA's climate forecasts and vegetation models are among the least reliable worldwide (Sherwood and Fu 2014; Huang et al. 2014). That is why the different forecasting *scenarios* should be evaluated and selected in light of pollen-based past climate reconstructions (Lioubimtseva et al. 2005). However, pollen type diversity in the ACA is reduced. A few pollen types, mostly identified at the family level, dominate drylands from Iran to Mongolia (e.g., Amaranthaceae, *Artemisia* spp., Cyperaceae, *Pinus* spp., and Poaceae) and cannot reflect temperature changes (Tarasov et al. 1998). It is especially the case since warm-dry desert-steppes (Iran, Uzbekistan…) and cold-dry desert-steppes (Mongolia, Tibetan Plateau…) present important climate differences (mainly in temperature and seasonality), although the pollen rain is very similar (Tarasov et al. 1998; Lioubimtseva 2003; Zheng et al. 2014). A trait-based approach could overcome this poor taxonomic resolution using the functional link between vegetation and climate in order to enhance past vegetation and climate reconstructions in ACA.

This study uses surface pollen and vegetation data from ACA (Chen et al. 2015; Figure 1) combined with trait data taken from worldwide databases. The full methodological workflow followed in this study is illustrated in Figure 2. Pollen surface samples and vegetation plots for ACA (Figure 2, step 1) and the major plant trait values (Figure 2, step 2) were extracted from global databases. In step 3, an ACA plant checklist was built to assign plant traits to species and pollen types (following two aggregation schemes, step 4). Finally, to compare community-weighted functional space, the trait values were gap-filled (Figure 2, step 5), and the community-weighted mean (CWM) traits were calculated. All the abbreviations used in the article are detailed in Table 1.

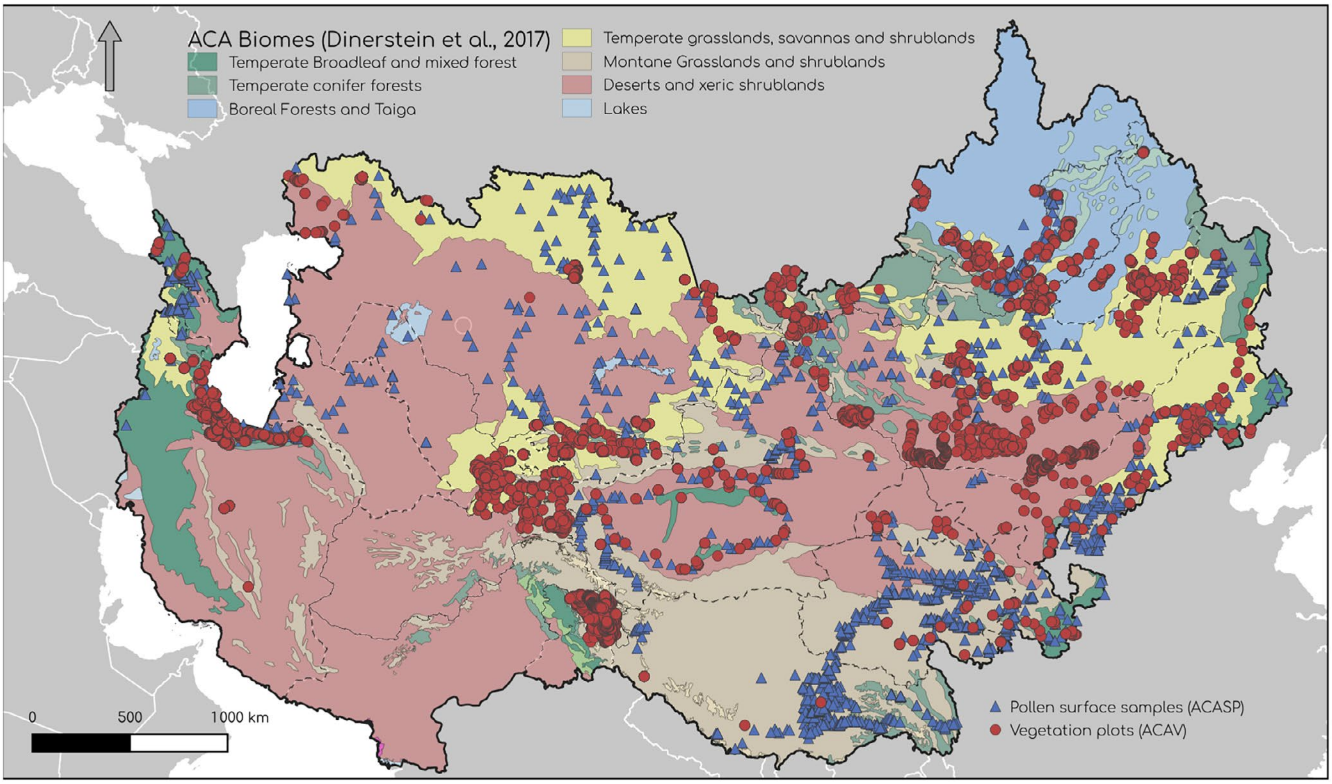

FIGURE 1 | Map of the Arid Central Asian (ACA) biomes (Dinerstein et al. 2017) with pollen surface samples (ACASP, blue triangles) and vegetation plots (ACAV, red dots).

2.2 | Arid Central Asian Geography

Central Asia covers Uzbekistan, Turkmenistan, Tajikistan, Kyrgyzstan, Kazakhstan, and Afghanistan (Figure 1). The Touran lowlands (Kyzylkum and Karakum deserts) and piedmont mountains (south-Siberian steppes, Alay, Altai) in the central part of the area are surrounded by the Caspian Sea to the east, the Pamir-Alay and Tian Shan ranges to the west, the Kopet Dagh and Hindu Kush ranges to the south, and Kazakh hills to the north (Figure S1A; Lioubimtseva et al. 2005). A climatically consistent region described by Chen et al. (2008) as the Arid Central Asia (ACA area in Figure 1) is currently used by the paleoclimate community (e.g., Chen et al. 2008, 2010; Rao et al. 2020). The ACA area encompasses central Asia as well as all the surrounding Asian drylands: Lesser Caucasus and Iranian deserts; mainland China and Mongolia (Gobi and Taklamakan deserts, Tibetan and Mongolian plateaus); and the Indo-Pakistan Thar desert. This area is large, spanning 80° longitude (40°–120° E) and 25° latitude (30°–55° N), and is located at the crossroads of the alpine and Hercynian range belts (Figure S1A); it also controls atmospheric regimes and biomes (Lioubimtseva et al. 2005).

2.3 | Pollen

Since there is currently no modern pollen dataset covering the whole ACA, we compiled pollen assemblages from available databases to build an Arid Central Asian Surface Pollen dataset (ACASP, $n_{\mathrm{site}} = 2393$, Figure 2), as part of the contribution of this study. To the Eurasian Pollen Database compiled by Peyron et al. (2017), modified by Dugerdil et al. (2021) and d'Oliveira et al. (2023), were appended open-access surface pollen sites from the European Modern Pollen Database v.2 (Davis et al. 2020), the East Asian Pollen Database (Zheng et al. 2014), and several open-access Tibetan pollen datasets from Li, Xie et al. (2020). The full list of references for pollen contributors is in Table S5. Poaceae and *Cerealia*-type were distinguished from the majority of pollen studies used in the database mainly from pollen morphological considerations (size of pollen and their pore). Surface samples dominated by human-induced pollen types such as *Plantago* spp., *Rumex* spp., and *Cerealia*-type or pollen types favoured by overgrazing were removed from the database (Peyron et al. 1998) since disturbed pollen samples do not record climate and environmental drivers (van der Sande et al. 2019; Dugerdil et al. 2021). Long-spreading pollen types (e.g., *Pinus* spp.) and exotic species (ornamental or invasive species) were not removed from the database since it is difficult to infer this species status at the biogeographical scale. Pollen counts were standardised to the pollen total (analogous fractional abundance calculation following Equation (1) below). Since the main issue in the pollen-base trait approach is the taxonomic matching between pollen type and plant species, we tested two aggregation schemes (Figure 2, panel 4):

1. ACA Surface Pollen *Fine* (ACASP-*fine*): the finest level of pollen-type identification is retained for aggregation. If the pollen type is identified at a clade level (genus, family...), the average trait values for all species in the same clade in ACA were used. We obtained $N_{\mathrm{ACASP}\text{-}fine} = 462$ different pollen types in ACA.
2. ACA Surface Pollen *Coarse* (ACASP-*Coarse*): Pollen-type identification has been down-scaled to commonly used and reliable pollen types (generally, pollen types are presented at the family level for non-arboreal pollen and genera for arboreal pollen). We obtain $\mathrm{N}_{\mathrm{ACASP}\text{-}coarse} = 163$ pollen types.

TABLE 1 | Abbreviations used in this study and abbreviations commonly used in palaeoecology or plant functional ecology. The abbreviations used for the database names and the trait are displayed, respectively, in Table 2 and in Tables S5 and S6.

Full name	Abbreviations	Commentary
ACA vegetation dataset	ACAV	This study
ACA surface pollen-*fine*	ACASP-*fine*	This study
ACA surface pollen-*coarse*	ACASP-*coarse*	This study
Angiosperm phylogeny group IV	APGIV	The Angiosperm Phylogeny Group (2016)
Arboreal pollen	AP	Usual abbreviation in paleopalynology
Community-weighted mean	CWM	Usual abbreviation in trait-based ecology
Non-arboreal pollen	NAP	Usual abbreviation in paleopalynology
Pearson correlation coefficient	*r*	
Plant functional types	PFT	
Principal component analysis	PCA	

2.4 | Vegetation

The vegetation plots used in this study were taken from the sPlot vegetation database extracted for Asia (Figure 2; Bruelheide et al. 2019), combined with some open-access vegetation plots from central Mongolia (Jamsranjav et al. 2018). The resulting vegetation database for the entire ACA region (ACAV, $n_{\mathrm{plots}} = 21{,}347$; $n_{\mathrm{species}} = 7723$) is displayed in Figure 1. The majority of site-present species abundances are in the Braun-Blanquet abundance scale, and all of them have been converted to percentage cover. Then, each vegetation plot k is expressed as the sum of fractional abundance $p_{i,k}$ defined for each species i as

$$p_{i,k} = \frac{\mathrm{SC}_{i,k}}{\sum_{j=1}^{N_{tot,k}} \mathrm{SC}_{j,k}} \tag{1}$$

with SC, the surface covered by each plant species, and $p_{i,k}$ the species fractional abundances. Although ferns, lycophytes, and bryophytes were recorded in some vegetation plots from ACA,

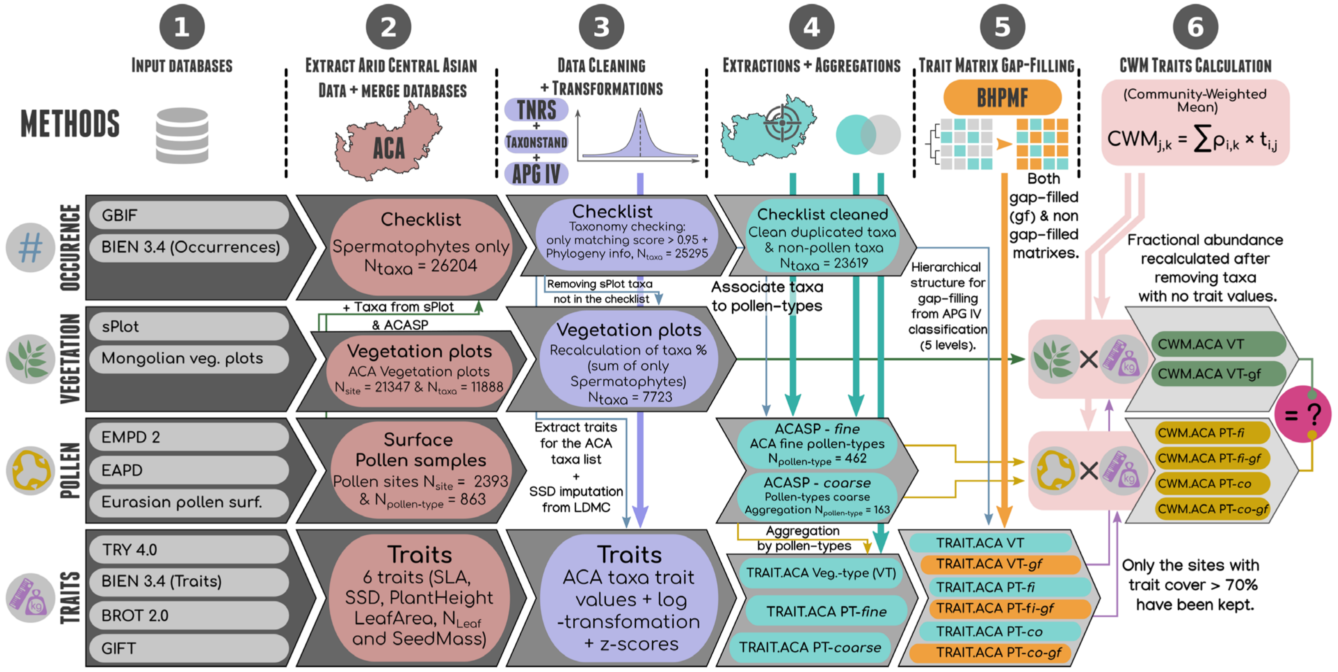

FIGURE 2 | Workflow followed in this study: From the first methodological step (input methods and databases–left-hand side) to the final correlation results. Each column of boxes represents a methodological step carried out in our study.

only the spermatophytes were selected to permit consistency among vegetation and pollen plots. As for pollen, the grass-dominated plots from forested landscapes (human clearing, crops…) and areas locally dominated by trees among open biomes (such as orchards or hedgerows) were removed.

2.5 | Plant Checklist and Taxonomy

In order to catch all the ACA functional diversity (and not only the one recorded in our vegetation plots), an ACA checklist is needed (Meyer et al. 2016). As there is no checklist available for the Central Asian area studied, we compiled an original checklist of ACA species that is as exhaustive as possible. Plant occurrences within the boundaries of ACA were extracted from GBIF (Robertson et al. 2014) and BIEN (Maitner et al. 2018). The occurrence density used to form this checklist is shown in Figure S1B. This list was then merged with the plant list extracted from the vegetation plot and the pollen sample matrices. We obtained 26,204 taxa (Figure 2, panel 2).

This checklist was then homogenised, and the taxonomy was verified using the Taxonomic Name Resolution Service (TNRS; Boyle et al. 2013) package on R. This online application uses name parsing and fuzzy matching between the input plant checklist and taxonomic databases (Boyle et al. 2013). We selected taxa with a matching score above 0.95 (1 for full matching). In parallel, we applied a second taxonomic referential checking with the Taxonstand R package from Cayuela et al. (2012). We use both algorithms since databases and approaches are not the same. Due to the heterogeneity of the pollen identification level (De Klerk and Joosten 2007), some pollen types are expressed in particular taxonomic levels such as sub-family (Asteroideae-t, Cichorioideae-t and Cardueae-t) or sub-genus (*Pinus*-Haploxylon and *P.*-Diploxylon). This taxonomic-level information was taken from the National Centre for Biotechnology Information's Taxonomy database (https://www.ncbi.nlm.nih.gov; accessed January 2021). In order to feat with the pollen database, the *Cerealia* pollen type was associated with a list of botanical taxa (e.g., *Sorghum* spp., *Avena sativa, Hordeum* spp., *Zea mays*…) using the USDA Plants Database. For oaks, two pollen types are identified: the *Quercus*-deciduous and *Q.*-evergreen types. *Quercus* species were associated with these two pollen types using the leaf phenology trait information from TRY. The upper classification levels needed for trait assignment (families, orders, kingdoms, and sub-reigons) were extracted from the APGIV classifications (The Angiosperm Phylogeny Group 2016). Then, each taxon was allocated to one of the growth forms (herb, shrub, tree, and other) extracted from TRY (Kattge et al. 2020). Each taxon was also associated with a pollen type: Arboreal Pollen (AP), Non-Arboreal Pollen (NAP, i.e., herb and shrubs), and variable (when pollen types contain herbs, shrubs, and trees). The clean ACA checklist contains 23,619 species, 2772 genera and 233 families associated with the 462 *fine* and 163 *coarse* pollen types (Figure 2, panel 3).

2.6 | Plant Trait Selection

We selected six continuous traits (Table 2; Bruelheide et al. 2019; Kattge et al. 2020) widely used in trait-based ecological studies (Garnier et al. 2016; Díaz et al. 2016; Bruelheide et al. 2018). These traits are also the most commonly available in trait databases (Kattge et al. 2020; Weigelt et al. 2020). Details on the functional significance of the selected plant trait (with trait description and measurement) can be found in Díaz et al. (2016) and Garnier et al. (2016) and are summarised in

TABLE 2 | List of the six traits used in this study with their main distribution statistical values on the ACA taxa checklist.

Trait	Full name	Unit	Mean	SD	Min	Max	$N_{species}$	Functional significances
Height	Adult plant height (vegetative)	(m)	3.02	6.39	0.004	74.7	7423	Above-ground competition, light capture, dispersal distance
LeafArea	Leaf area	(mm^2)	3580	11,900	1	272,000	3260	Light interception, energy balance
N_{Leaf}	N content per unit leaf mass	($mg.g^{-1}$)	23.9	9.18	0.97	99.6	2356	Light capture, photosynthetic rate
SeedMass	Diaspore mass	(mg)	142	5820	0.0005	404,000	4861	Dispersal and regeneration strategy, seedling competition
SLA	Specific leaf area	($mm^2.mg^{-1}$)	0.02	0.01	0.0005	0.22	3619	Photosynthetic rate, leaf longevity
SSD	Stem specific density	($mg.mm^{-3}$)	0.32	0.19	0.07	1.55	2465	Trade-off between growth potential and mortality risk

Table 2. These are: (1) plant height, the maximum plant height at plant maturity; (2) seed mass (the oven-dry mass of an average set of seeds); (3) leaf area, the average surface of one side leaf; (4) specific leaf area (SLA), the ratio between the area and the oven-dry mass of the leaf, which reflects resource use by the leaf, (5) nitrogen content per leaf mass (N_{leaf}), and (6) stem specific density (SSD), the ratio between stem dry mass and fresh volume of a stem section.

2.7 | Plant Trait Databases

Trait values were taken from TRY (Kattge et al. 2020), BROT (Tavşanoğlu and Pausas 2018), and BIEN (Maitner et al. 2018) databases and the recent Global Inventory of Floras and Traits (Weigelt et al. 2020) using the GIFT R package (Denelle et al. 2023). References of datasets from TRY used in this study are detailed in Table S6. The six selected traits (Table 2) were extracted for the species in the ACA checklist. The missing SSD values for herbaceous taxa were imputed from the leaf dry matter content following Díaz et al. (2022). Plotting the frequency distribution for each trait (Figure S2), errors in the databases (negative values, unit errors) were corrected. Then, the trait values were $\log_{10}$-transformed in order to fulfil normal requirements and reduce the influence of extreme trait values. Afterward, each trait j for each species i ($t_{i,j}$) was z-scaled with

$$z - \text{score}_{t(i,j)} = \frac{t_{i,j} - \mu_{t_{i,j}}}{\sigma_{t_{i,j}}} \tag{2}$$

to ease the comparison among traits with different units.

The trait dataset was gap-filled using the Bayesian Hierarchical Probabilistic Matrix Factorisation (Schrodt et al. 2015) machine learning technique. The R package BHPMF applied to the trait matrix and using six taxonomic levels of the Angiosperm Phylogeny Group IV (APG) classification produces two output matrices: a mean (μ_{gf}) and a standard deviation (σ_{gf}) gap-filled trait matrix. The total RMSE of the z-score traits gap-filled matrix is 0.70 ± 0.02 (RMSE average and standard deviation are calculated among all folds of the cross-validation process). Cross-validation permits removing the badly gap-filled values following the method presented by Fazayeli et al. (2014). We calculate predicted trait cross-validated values for each trait as

$$\text{CrossValidated}_{gf} = \frac{\sigma_{gf}}{\mu_{gf}} \tag{3}$$

For each $\text{CrossValidated}_{gf}$ predicted value, the BHPMF package calculated the mean (μ_{gf}) and the standard deviation (σ_{gf}) of the prediction. Then, only the gap-filled values with standard deviation vs. mean above one were retained; in other cases, the predicted value was not used and replaced by a missing value (Figure 3, panel 5; Fazayeli et al. 2014; Bruelheide et al. 2018).

2.8 | Trait Aggregation at the Community Level

The community-weighted mean (CWM) value for each trait was calculated for both pollen surface and vegetation plots following Garnier et al. (2004) as

$$\text{CWM}_{j,k} = \sum\nolimits_i^{n_k} p_{i,k} \times t_{i,j} \tag{4}$$

with k each vegetation/pollen plot, j each trait, and n the total number of taxa in plot k. Since trait coverage is not complete (Table S1) and following the recommendations of Pakeman (2014) and Borgy et al. (2017), the P_{cover} for each sample (proportion of taxa with available trait value) should be considered when using CWM trait. That is why only plots with $P_{cover} > 80\%$ were used for analysis (Borgy et al. 2017). A biome type was assigned to each vegetation plot and pollen surface sample by extraction from the biome mapping of Dinerstein et al. (2017).

Then, we analysed the relation among traits and among CWM traits using a correlation matrix, linear relations (Pearson's correlation coefficient, *r*), and principal component analysis (PCA) using R. The linear correlations among traits and among CWM traits were all performed on 10,000 permutations using the perm.cor.test from the jmuOutlier package (Garren 2017). The CWM trait functional spaces described by the two different pollen aggregation schemes (fine and coarse) were compared using a Procrustes rotation of two configurations and a PROTEST significance test between the two PCAs using the R package vegan version 2,6-8 (Dixon 2003). Finally, analysis of variance (two-way ANOVAs) and Tukey post hoc comparisons were performed using the R functions aov() and TukeyHSD().

3 | Results

3.1 | ACA Traits

3.1.1 | Traits Coverage and Distribution

Trait coverage varies for each trait and each pollen-type scheme. Mainly, the *fine* or *coarse* aggregation induced a change in trait coverage. Large differences occur according to the aggregation scheme used (Table S1): only 14% of the traits are available for the total ACA taxa checklist (on average for the six traits). After gap-filling this matrix, it reaches about 32%. Then, the percentage of coverage increases as the number of plant types decreases: there are more plant types in the vegetation ($P_{cover} = 34\%$) than in the pollen samples. Also, the cover is better for *coarse* (94% and 98% for gap-filled values) than fine pollen types (82%). The traits with the best coverage are plant height, seed mass and SLA, with 73%, 68% and 63%, respectively. By contrast, N_{leaf} is less well represented with only an average coverage of around 58%, followed by SSD (62%).

3.1.2 | Pollen Type Aggregation

The trait distribution of twelve of the major pollen types for ACA is studied as ACA pollen types representative (Figure 3 and Table S2), most common pollen types in ACA and high pollen producers (Wieczorek and Herzschuh 2020). First, trait values are missing for some pollen types (mainly SSD and N_{leaf}), and others only have values for fine or *coarse*: for example, the pollen of *Nitraria*-t is identified as Nitrariaceae for coarse and *Nitraria* spp. for *fine*. Trait variance is shown compared to the variance of the global spectrum of plant form and function (Figure 3; Díaz et al. 2022). Trait variance (Figure 3, lower panels) depends on (1) the level of aggregation (from family to species level) and (2) the type of trait. This results from the fact that some aggregated trait values are obtained from a few species (for *Pinus* types), while other types are derived from a large amount of data (families and large genera such as *Artemisia* spp. and *Ephedra* spp.). Logically, the aggregation for pollen types based on only one species (*Betula pubescens* and *Carpinus betulus*) has the smallest variance.

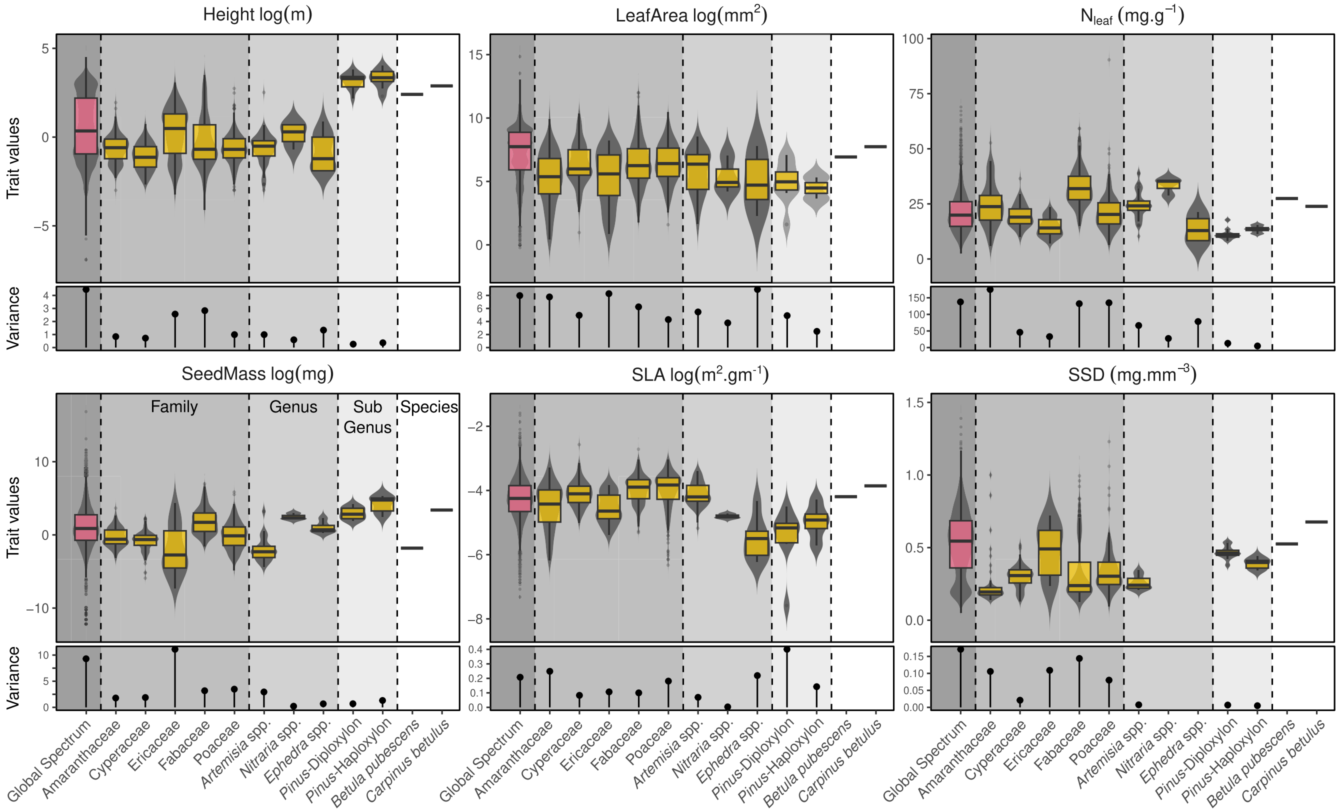

FIGURE 3 | Distribution of trait values after aggregation by pollen type for 12 selected major ACA pollen types. Yellow boxplots show all retained trait values for ACA and are compared to the global spectrum values (in pink) from Díaz et al. (2022). Lower panels show the trait variance. The selected plant traits are displayed in Table 2.

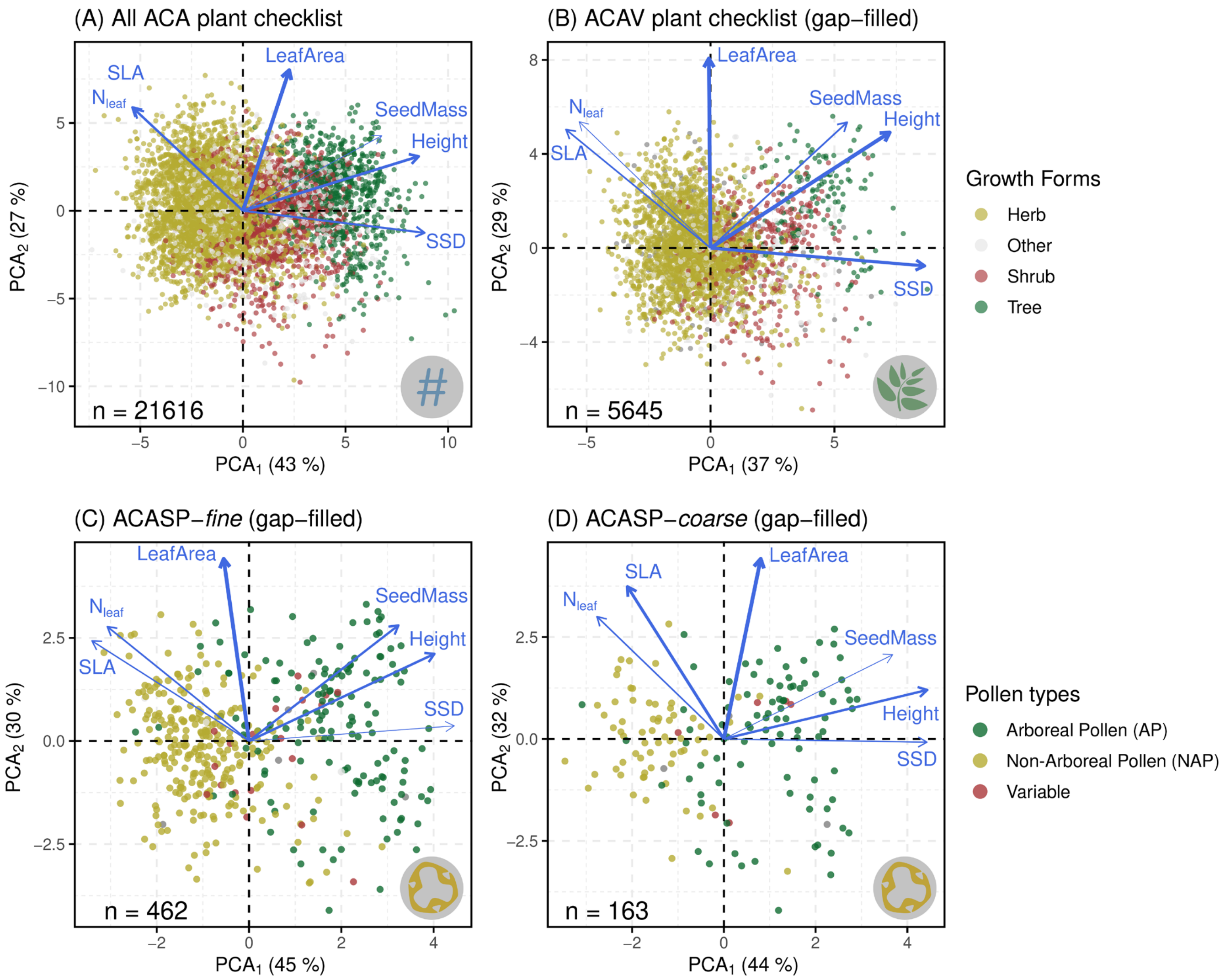

FIGURE 4 | PCA of the global spectrum of plant form based on six core traits and tested on (A) ACA plant checklist; (B) plants from the vegetation plot dataset (ACAV); (C) traits aggregated by pollen types *fine* (ACASP-*fine*); and finally (D) traits aggregated by pollen types *coarse* (ACASP-*coarse*). Dot colours (A, B) are for different growth forms and pollen types (C, D). The thickness of the arrows indicates relative contribution to explaining the variance. Details of trait correlations can be visualised in Figure S3.

Variance also depends on the homogeneity of plant growth form (tree, herbs…) among the species used for aggregation: *Pinus*-Haploxylon and *P.*-Diploxylon subgenus regroup species with similar plant growth form, resulting in small variance for each trait. By contrast, plant height in Fabaceae is very different between the tiny crawling herb *Astragalus tribuloides* (height <0.02 m) and the large tree *Albizia chinensis* (height >30 m). The same is observable for seed mass in Ericaceae, with an important amplitude between *Pyrola* spp. small seeds (mass <0.002 g) and bigger seeds of *Arbutus* spp. (mass >30 g). N_{leaf} and SSD have smaller variances, which is the consequence of a lower number of trait values available in the databases (Kattge et al. 2020).

3.1.3 | Trait Covariation

Plant height and seed mass are positively correlated (Pearson's correlation coefficient r from 0.50 to 0.64; find details in Figure S3), as well as the SLA and the N_{leaf} (r between 0.46 and 0.52). SSD is positively correlated to plant height and seed mass (same r range). The SLA–leaf area positive correlation was higher for pollen aggregation ($r_{ACASP\text{-}fine}=0.44$ and $r_{ACASP\text{-}coarse}=0.52$) than for vegetation ($r_{ACAV}=0.29$). After considering the bivariate relationships, we performed a multivariate analysis using Principal Component Analysis (PCA, Figure 4).

When all taxa from the ACA checklist were considered, the first two axes of PCA can explain 70% of the traits variance (Figure 4A). SSD, seed mass and plant height have high positive loadings on PCA_1; SLA and N_{leaf} have negative loadings, while PCA_2 is driven by leaf area. In the PCA of the whole ACA checklist (Figure 4A), taxa are clustered into three pools corresponding to the different growth forms: herbs cover mainly the left part of the plane (low plant height, SSD and small seeds), trees cover the upper-right corner (tall plants with heavy seeds), while shrubs are in between these two. When taxa from vegetation plots are used (Figure 4B), the overall relationships among traits appear similar to those from the global PCA_{ACA}. For taxa from surface pollen samples, the variance explained by the first two axes is even higher than for ACAV and the overall organisation of traits is comparable to that of the two previous PCA. Then, even if the pool of pollen types is smaller ($n_{\text{fine}} = 462$ and $n_{\text{coarse}} = 163$), the PCA space is overall similar (Figure 4C,D) to Figure 4A.

3.2 | CWM Trait Coverage and Distribution

Trait coverage for CWM traits is similar to individual traits coverage (Tables S1 and S3). However, the cover is almost complete for ACASP-*fine* and *coarse* (98% for all ACASP CWM traits). Conversely, the P_{cover} for vegetation plots is lower than for surface pollen samples. SSD and N_{leaf} have the lowest coverage (respectively 18% and 19%). The strongest correlations were found for plant height–seed mass, the plant height–SSD (also SSD–seed mass), and the SLA–N_{leaf} relationships for vegetation plots and surface pollen samples *fine* and *coarse* (Figure S3). However, some correlations were found to be significant for pollen, while they were not for vegetation plots: the plant height–N_{leaf}, the plant height–SLA, and the SLA–N_{leaf} (the difference between vegetation and pollen type *r* values is from 0.26 to 0.56, Figure S4).

3.3 | Variance in CWM Traits Among ACA Biomes

Differences in CWM traits among biomes and sample types (i.e., vegetation plots, surface pollen samples fine and coarse) were tested using two-way ANOVAs (Table 3 and Figure 5). Sample type has a significant effect (i.e., $p < 0.001$) for all CWM traits

TABLE 3 | Results of ANOVA performed to assess the impact of aggregation schemes (vegetation taxa, pollen types fine and coarse) and biome distribution on the variance of CWM traits Details of each relation are inferred by the ad-hoc Tukey significant test in Table S4.

	Sample type			Biomes		
	df	***F***	***p***	**df**	***F***	***p***
LeafN	2	359.39	< 0.001	5	195.38	< 0.001
SSD	2	2.11	0.1215	5	1050.23	< 0.001
Height	2	1376.17	< 0.001	5	2811.26	< 0.001
SeedMass	2	14.19	< 0.001	5	1350.92	< 0.001
LeafArea	2	5.3	0.005	5	948.6	< 0.001
SLA	2	187.61	< 0.001	5	509.21	< 0.001

Abbreviations: df = degree of freedom, *F* = *F*-statistic, and probability value.

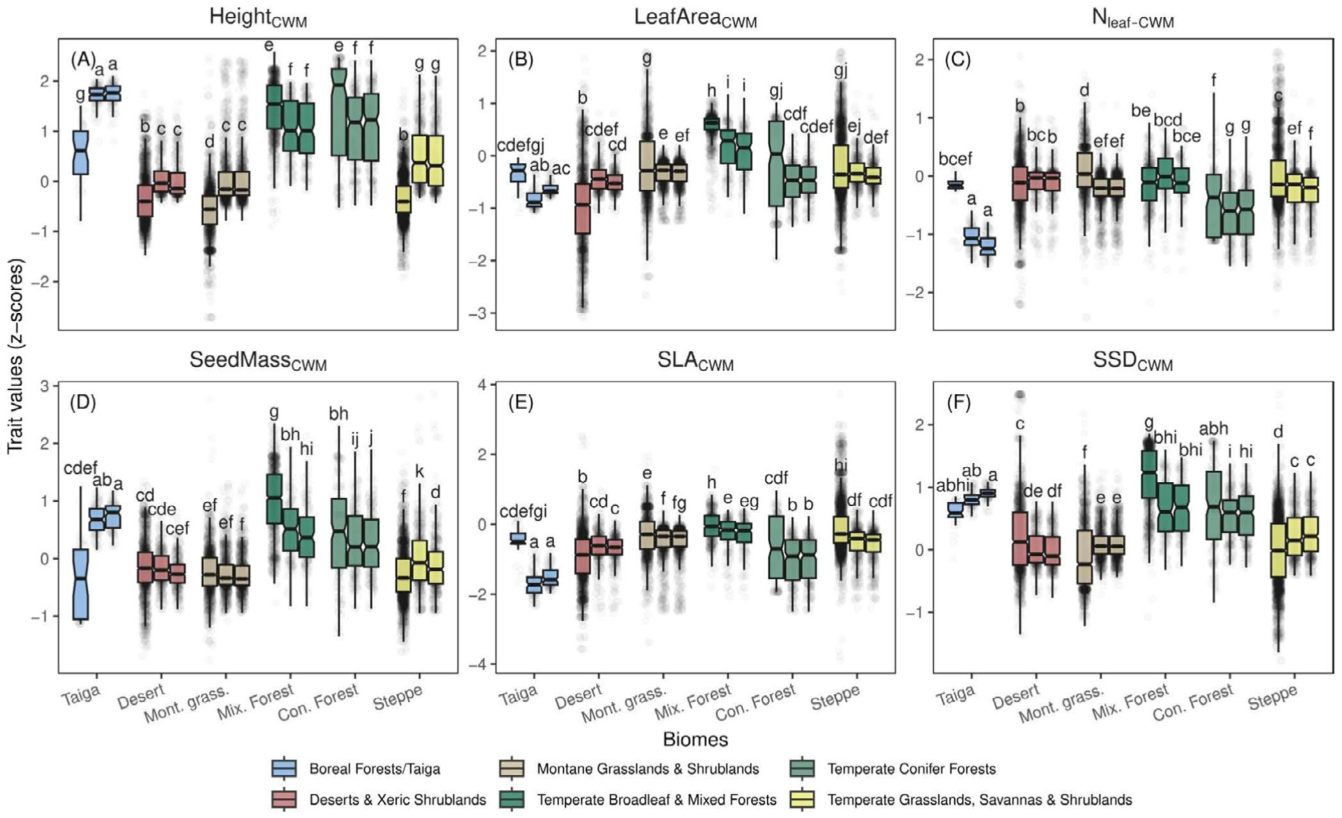

FIGURE 5 | Boxplot for the six traits grouped by biome and CWM aggregation scheme. For each trait (A–F) and biome (highlighted by boxplot colour), the three columns are, from left to right, vegetation-based CWM traits (first histogram); pollen-based CWM traits *fine* (middle one) and *coarse* (last one). Letters above boxplots highlight the two-way ANOVA and Tukey's test significance clustering (*p*-values for each test in Table S4).

except for the leaf area and the SSD due to low trait coverage for these two traits. However, the biome effect on the CWM trait distributions was always stronger than the sample type effect.

3.4 | CWM Traits and Functional Space

The first two PCA axes account for 69%, 87% and 89% of the total variance of the CWM trait values for ACAV, ACASP-*fine* and *coarse*, respectively (Figure 6). Points are clustered by biome assignment (Figure 6D): taiga sites are grouped in the upper left corner, temperate forests are in the upper right one and mountain grasslands are in the lower left one. Points are generally spread out for steppes and deserts, with negative and positive PCA$_1$, respectively. Desert sites are associated with high SSD$_{\text{CWM}}$, low SLA$_{\text{CWM}}$, and N$_{\text{leaf-CWM}}$; mountain grasslands with low plant height and seed mass as opposed to temperate forests. For pollen surface samples, the sites follow the same organisation for both ACASP-*fine* and ACASP-*coarse* (Figure 6B,C). Mountain grassland sites are clustered in the lower right corner linked with low seed mass and PlantHeight$_{\text{CWM}}$; by contrast, the temperate forests are characterised by high seed mass and PlantHeight$_{\text{CWM}}$ (upper left corner). Deserts and steppes are not well clustered and distinguishable, and unlike vegetation CWM, the taiga sites cluster around low values of SLA$_{\text{CWM}}$ and N$_{\text{leaf-CWM}}$. Comparisons between the PCA results for fine and coarse pollen types using Procrustes analysis and PROTEST (Figure S5) show similar functional spaces with a high correlation between PCA-*fine* and PCA-*coarse* (PROTEST correlation = 0.97).

To assess the consistency of pairwise relationships observed between pollen and vegetation-based CWM values, we compared the correlation coefficients (*r*) for each pairwise relationship, both for the CWM pollen types *fine* (Figure 7A) and *coarse* (Figure 7B). For some relationships, *r* values are very similar (points close to the 1:1 line). Still, the absolute *r* values tend to be higher for pollen than for vegetation-based relationships, as shown by the regression line drawn through the data points. Importantly, the direction and strength of most relationships are conserved between pollen and vegetation. Only 20% of the relations are not conserved in direction (i.e., LeafArea$_{\text{CWM}}$ vs. Height$_{\text{CWM}}$ and SSD$_{\text{CWM}}$ and N$_{\text{leaf-CWM}}$ vs. SeedMass$_{\text{CWM}}$, which are weak relationships overall identified as triangles in Figure 7).

4 | Discussion

At the biogeographic scale of Arid Central Asia, we show that (1) the functional signatures of pollen and vegetation are comparable, for key traits structuring the plant phenotype, both at the levels of taxa and communities, and (2) the pollen aggregation scheme (*Fine* or *Coarse*) has little impact on this functional signature. These points are the first step to answer the following question: Can a trait-based approach, as applied to pollen, be used to help improve past vegetation and climate reconstructions? These points are further discussed below.

4.1 | Functional Signatures of Pollen and Vegetation in ACA

For taxa found in vegetation plots of ACA, we found functional spaces consistent with the global plant spectrum of plant form and function (Díaz et al. 2016): a first axis of variation describing a contrast between tall, heavy-seeded plants and short plants with small seeds, and a second one between soft leaves (high SLA) rich in nitrogen and thicker, leathery leaves poor in nitrogen (leaf economic spectrum; cf. Wright

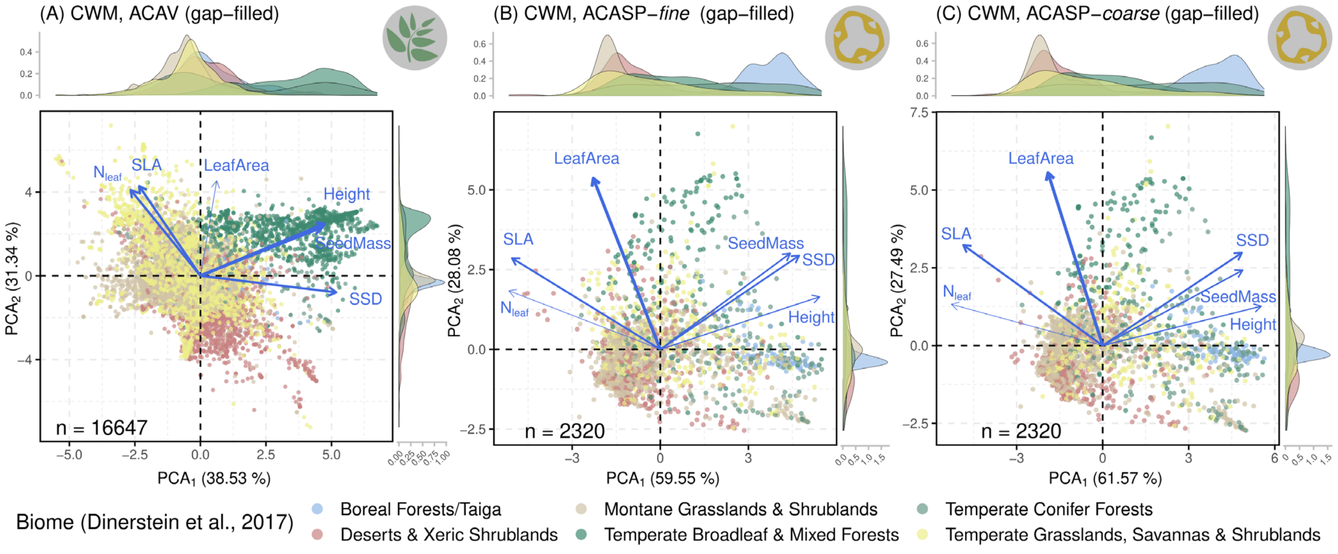

FIGURE 6 | Principal Component Analysis of the variance of the six main traits applied on CWM trait values calculated on vegetation plots (A) and surface pollen samples with pollen type fine (B) and coarse (C) aggregation scheme. The number of sites (vegetation plots or pollen surface samples) is indicated on the lower left corner of each panel. The thickness of the arrows varies with the relative contribution of each CWM trait. The distribution of the sites is shown as marginal densities above and on the right-hand side of each panel.

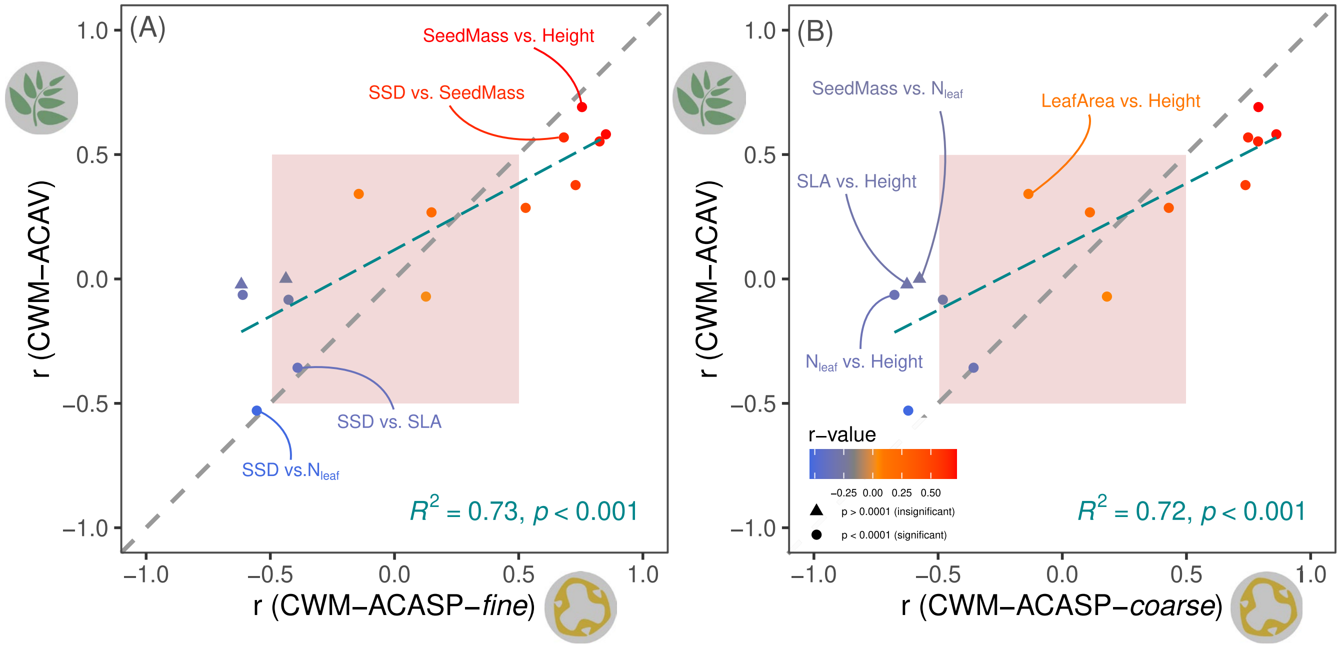

FIGURE 7 | Comparison of Pearson correlation coefficients (r) for CWM traits correlation between CWM vegetation-based and *fine* (A) or *coarse* (B) pollen aggregation scheme. The four most consistent (high r values) and least consistent (low r values) relationships are labelled on each panel. The central red squares denote relationships with low r values ($r < 0.5$). If at least one of the r-values were not significant (i.e., p-value > 0.001, the relation is shown as a triangle.

et al. 2004). These two axes of variation were also found at the community level. In this area, grasslands are well represented, confirming the analysis conducted worldwide by Bruelheide et al. (2018).

Despite the methodological limitations that are discussed in the next section, the organisation of traits in the pollen-based phenotypic space was broadly similar to that obtained for extant vegetation. Overall, it appears CWM trait correlations become tighter as coarser descriptions of the vegetation are used: the values of correlation coefficients increase from vegetation to *fine* and *coarse*, enhancing the impact of dominant pollen types in the functional space (trait values of rare taxa weightless).

However, not all CWM traits can be considered with equal reliability when applied to past pollen climate and vegetation reconstructions since we found substantial differences between pollen and plant variance for several CWM traits. In Table 3, the sample type has a limited impact on SSD_{CWM}, $\text{LeafArea}_{\text{CWM}}$, $\text{SeedMass}_{\text{CWM}}$ ($F < 15$), a slightly stronger impact on SLA_{CWM} and $\text{N}_{\text{leaf-CWM}}$ ($F < 400$), and a strong impact on $\text{PlantHeight}_{\text{CWM}}$. SSD and seed mass traits present a little variance when aggregated at fine or coarse pollen types (Figure 3), even at family-dominant species (Fabaceae, Amaranthaceae, Poaceae and Cyperaceae). In contrast, $\text{PlantHeight}_{\text{CWM}}$ and $\text{N}_{\text{leaf-CWM}}$ have to be carefully interpreted whenever applied in past pollen sequences (particularly $\text{N}_{\text{leaf-CWM}}$ with $\text{F}_{\text{biome}} < \text{F}_{\text{aggregation}}$, Table 3). For $\text{N}_{\text{leaf-CWM}}$, discrepancies between pollen and plant values may be the consequence of the limited number of values for this particular trait. For $\text{PlantHeight}_{\text{CWM}}$, discrepancies cannot be explained by weak trait coverage, and this trait has been used for European Holocene reconstructions, showing a dramatic decrease associated with the onset of agriculture and climate cooling (Veeken et al. 2022). The major difference between Europe and ACA $\text{PlantHeight}_{\text{CWM}}$ comes from the pollen types. In ACA, pollen samples are widely dominated by herbaceous pollen, leading to a higher CWM traits bias since herbaceous pollen are mostly identified at the family level rather than the species or genus level for arboreal pollen.

At the regional scale of ACA, we observe a structured trend in the distribution of pollen-based CWM trait values, comparable to that found in North America, for pollen-based (not aggregated at the level of assemblages) trait distribution (Brussel and Brewer 2021). As expected, CWM trait values were substantially different among biomes. This is especially true for $\text{PlantHeight}_{\text{CWM}}$, $\text{SeedMass}_{\text{CWM}}$ and SSD_{CWM}, corresponding to differences between forested and open land ecosystems. This agrees with observed plant trait distributions within local landscapes (Lavorel et al. 2011) and biomes (Mucina 2019; Boonman et al. 2022). $\text{LeafArea}_{\text{CWM}}$, $\text{N}_{\text{leaf-CWM}}$ and SLA_{CWM} differed (1) within forest types, between cold forests on the one hand (taiga and conifer forests) and broadleaf forests on the other hand, and (2) within open ecosystems, between deserts on the one hand and grasslands on the other hand. These observations are consistent with the worldwide distributions of competitor, stress-tolerant, and ruderal species across biomes (Pierce et al. 2017).

4.2 | Methodological Issues

4.2.1 | Taxonomic Checklist and Trait Coverage

Working with a reliable plant checklist is the very first step in plant geography and ecological studies (König et al. 2017). Such a checklist did not yet exist for the ACA (Li, Tojibaev et al. 2020) because (1) floras for this geographic area are published in local languages, making it difficult to collect and clean up synonymy (Li, Tojibaev et al. 2020), and (2) the actual floristic diversity in ACA, which presents an important level of endemism (Nowak et al. 2020), is currently understudied. This results in a lack of data for deserts (especially Iranian deserts, Karakum, Kyzylkum, and Taklamakan, which represent 43% of the total ACA surface) and an over-representation of plants within the borders of Russia (mainly the southern Siberian area and the Caucasus) and the mountain ranges (Elburz, Tian Shan, Pamir, Kashmir, Himalaya…) mainly covered by forest (19% of the ACA surface is covered by forested biomes).

To overcome this knowledge gap, it appears that the utilisation of occurrences combining GBIF, BIEN, and sPlot here is the best way to obtain an available checklist (Robertson et al. 2014), and although there are clearly gaps in databases for the ACA area (Meyer et al. 2016). This issue has to be considered for many, if not all, biogeographical studies, as plant checklists are always estimated and strongly depend on the geographical boundaries set for each case study. The checklist combined in this study will be used as a basis for further studies on ACA.

The accuracy with which the functional structure of communities is assessed depends on the coverage in trait data (Pakeman and Quested 2007; Borgy et al. 2017). Although we used a gap-filled version of the trait database (Schrodt et al. 2015), this coverage remained relatively low for some traits and/or taxa lists. It is difficult to assess how much this affected our results. However, an important finding is that over the whole plant list used, the organisation of variables in the functional spaces obtained using the “raw” or the “gap-filled” versions remained relatively stable.

4.2.2 | Spatial Scales of Pollen and Vegetation Samples

The pollen composition at a given site depends on pollen production and dispersion from an area extending beyond a local vegetation plot (Prentice 1985; ter Braak et al. 1993; Sugita 2007; Salonen et al. 2014; Gillison 2019). While a vegetation plot is generally very local (less than 50 m^2 in Bruelheide et al. 2018), a pollen surface sample captures pollen production from a 12 to 1000 m diameter circle (distance inferred by stochastic dispersal models, Theuerkauf et al. 2016). In airborne pollen studies, the dispersal distance is estimated to be several kilometres (Maya-Manzano et al. 2017). Furthermore, the amount of pollen produced and its dispersion in the environment depend on the taxa (Theuerkauf et al. 2016; Marquer et al. 2017; Maya-Manzano et al. 2017; Chevalier et al. 2020), the landscape openness (Pelánková and Chytrý 2009), and the surrounding land cover homogeneity (Maya-Manzano et al. 2017). The representativeness differences between wind- and insect-pollinated pollen or local to long-spreading pollen are considered the main limitation of the pollen-based functional approach (Carvalho et al. 2019). However, in this study, the correction of this bias using relative pollen productivity estimates does not improve the method.

The human impact on vegetation distribution, including grazing, introduction of ornamental species, and invasive species, also mitigates the pollen sample-vegetation plot match (van der Sande et al. 2019; Dugerdil et al. 2021). Such differences might explain some differences observed between vegetation-based and pollen-based CWM trait values (Carvalho et al. 2019; Cleal et al. 2021).

One limitation of our study relates to sampling differences between vegetation and pollen: the two datasets are unbalanced in size (much more vegetation plots than pollen assemblages) and in spatial organisation of the data (most pollen and vegetation plots are not in the same location). This heterogeneity restricts, to some extent, the possibility of statistical quantitative comparisons of our data, especially in the case of multivariate analyses, based on matrices of heterogeneous sizes. However, the pollen-plant functional trait linkage remains permitted at a biogeographic scale (Brussel and Brewer 2021) since the economic spectrum of traits is universal for plant taxa (Díaz et al. 2016) and communities (Bruelheide et al. 2018) and since pollen is a good proxy of vegetation communities (Sun et al. 2020) permitting reliable quantitative reconstruction of vegetation (Prentice 1985; Sugita 2007).

4.2.3 | Pollen Aggregation Scheme

Most pollen databases provide data in both an *ad hoc* pollen-type referential designed by the author of a particular study (generally *fine*; highest level of pollen identification) and a homogenised pollen-type referential (generally *coarse*, common pollen types; Davis et al. 2020). On the one hand, pollen trait values will likely be closer to those obtained for corresponding vegetation plots if the level of determination is closer for both assessments (i.e., *fine* pollen types and taxa in plots) and does not derive from an aggregation of several pollen types (Finkelstein et al. 2006). On the other hand, pollen fine aggregation is generally less reliable than coarse aggregation since the identifiable pollen structures are not distinct enough to reach the level of taxa identification and also because it heavily relies on the skills of the person who identifies the pollen (Goring et al. 2013). Here, we show that the ACASP-*fine* and ACASP-*coarse* phenotypic spaces are very similar (variables in the first plane of PCAs and percentage of total variance explained are equivalent for fine and coarse without the impact of the different trait coverage). The correlations between traits (with enough available trait coverage) are not affected by the aggregation scheme either: both the leaf economic (Wright et al. 2004) and overall plant *spectra* (Díaz et al. 2016) appear well conserved for both schemes.

At the level of pollen assemblages and vegetation plots, bivariate relationships and multivariate analyses involving CWM traits show little difference between the *fine* and *coarse* schemes. Only the SSD_{CWM} shows a difference between *fine* and coarse, especially with *fine*, which is closer to the PCA performed on the vegetation plots than *coarse*.

We thus conclude that the fine or coarse scheme describes equally well the plant functional space in the pollen sample.

This contrasts with the conclusion put forward by Finkelstein et al. (2006), stating that a *fine* aggregation should preferably be used in plant functional biogeography. This study was conducted in the boreal forest, however, where the pollen identification of a few trees was easier than in the species-rich ACA.

4.2.4 | Taxonomic Level of Pollen Aggregation

In pollen studies, data from different taxonomic levels can be retained (De Klerk and Joosten 2007). It ranges from conserving all plant observations (Brussel and Brewer 2021), only well-covered families and genera (trait coverage > 30%; van der Sande et al. 2019), to single representative species (Connor et al. 2018). This raises issues for trait assignment (Barboni et al. 2004) at higher taxonomic levels. For example, in a study in the Andes-Amazon area, van der Sande et al. (2019) found no significant difference between family and genus average trait values, which could be due to the selection of traits only measured from Neotropic plants in a forested environment (and thus for very homogeneous plant forms; these authors also considered that a pollen-aggregated trait value was reliable only if derived from at least three different species). This is not the case in ACA, where a mosaic of open and closed ecosystems is found. Defining narrower geographic boundaries for ACA would have certainly excluded the majority of tall trees of the Fabaceae family, which are mainly localised in the eastern Chinese lowlands. This would have led to a substantial reduction of plant height value assessed for this family since in ACA steppes and montane grasslands, Fabaceae are mainly crawling plants. Using a Bayesian modelling approach of pollen-type distributions instead of average values could help address this issue (Veeken et al. 2022).

4.3 | Perspectives

The past pollen signal is the most powerful tool to track vegetation changes through time, as it is the only temporally continuous, taxonomically accurate, and quantitative proxy that exists, at least for Quaternary studies (Prentice 1988; Fyfe et al. 2009; Birks 2020). It also allows one to record local and regional vegetation signatures (Carter et al. 2018). At a biogeographic scale, we have shown that the taxonomic discrepancies between pollen types and vegetation taxa are overcome by the use of plant traits, a conclusion also reached by Brussel and Brewer (2021) at the continental scale and van der Sande et al. (2021) at the local scale. This conclusion supports the proposal of Goring et al. (2013) to use functional traits to reconcile discrepancies between pollen and plant taxa (De Klerk and Joosten 2007; Birks et al. 2016).

This new approach is very promising and allows one to better understand why and how vegetation functional structures have changed through time (Adeleye et al. 2023). In Adeleye et al. (2023), the functional adaptations of vegetation were highlighted by comparing functional diversity reconstructions with independent climate reconstructions. At the local scale of a peatland, the same conclusions on the main drivers of long-term population changes have been reached using functional traits (Gaüzère et al. 2020; Connor et al. 2018). Community traits also allow the detection of human pressure thresholds when the functional response decorrelates with climate reconstructions (van der Sande et al. 2019). This promising approach could also be applied to airborne pollen data, which also reflects regional vegetation (Smith et al. 2014; Myszkowska et al. 2023).

We have shown that functional signatures of surface pollen and vegetation are broadly similar in the ACA as in other study areas of the world (Brussel and Brewer, 2021; van der Sande et al. 2021). Climate control on key community traits is also well established (Bruelheide et al. 2018). The next step will thus be to open-up the relationships between pollen-based CWM traits and climate parameters. Provided that these tests are convincing, this can lead to (1) powerful trait-based biome reconstructions (Boonman et al. 2022) improving the PFT-based classical biomization approach (Harrison et al. 2010; Sun et al. 2020) and (2) tracking past changes in ecosystem services (Lavorel et al. 2011) to improve ecosystem conservation (Barnosky et al. 2017). It will also improve (3) climate reconstructions based on PFTs, such as the inverse modelling method (Chevalier et al. 2020) using more accurate CWM traits. Finally, it will permit verification of the accuracy and reliability of pollen-based climate reconstructions (Harrison et al. 2010) and facilitate the selection of the best model (Peyron et al. 2017; Salonen et al. 2019; Dugerdil et al. 2021) by checking the vegetation strategies changes in parallel with climate reconstructions.

Author Contributions

Lucas Dugerdil, Eric Garnier, Odile Peyron, Sébastien Joannin, and Guillemette Ménot conceived the ideas and followed a methodological framework designed by Lucas Dugerdil and Eric Garnier. Lucas Dugerdil, Eric Garnier, Odile Peyron, and Pierre Denelle gathered and compiled existing pollen and trait datasets and Helge Bruelheide, Milan Chytrý, Richard Field, Mohamed Z. Hatim, Hamid Gholizadeh, Jiri Dolezal, Valério D. Pillar, Kamal H. Shaltout, and Franziska Schrodt contributed to the vegetation plot dataset. Lucas Dugerdil analysed the data with the help of Eric Garnier, Cyrille Violle, Pierre Denelle, Helge Bruelheide, Milan Chytrý, and Richard Field. Lucas Dugerdil and Eric Garnier led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Acknowledgements

This research was funded, in whole or in part, by (ANR, Grant ANR-22-CE27-0018-02). A CC-BY public copyright licence has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission, in accordance with the grant's open access conditions. Thanks to *La Tendresse*, *Le BIB* and *Team Pey Jouan* for their support. This study has been supported by the TRY initiative on plant traits (http://www.try-db.org); all the datasets from TRY used in this study are listed in Appendix S1—Data Sources, Table S5. This is an ISEM contribution ISEM 2025-006. J.D. was supported by the Czech Science Foundation (24-11954S). V.D.P. received support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil, grant no. 313315/2022-1). No fieldwork was carried out for this study since all the data analysed here were previously published in global pollen, vegetation plot, and trait databases.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

This study does not present new data. All data come from already published databases referenced in the 2. Materials and Methods section of the manuscript. The full R script and figures are available on Github. The repository link is https://github.com/LucasDugerdil/TraitPollen.

References

Adeleye, M. A., S. G. Haberle, R. Gallagher, S. C. Andrew, and A. Herbert. 2023. "Changing Plant Functional Diversity Over the Last 12,000 Years Provides Perspectives for Tracking Future Changes in Vegetation Communities." *Nature Ecology & Evolution* 7, no. 2: 224–235. https://doi.org/10.1038/s41559-022-01943-4.

Anderegg, L. D. L. 2023. "Why Can't We Predict Traits From the Environment?" *New Phytologist* 237, no. 6: 1998–2004. https://doi.org/10.1111/nph.18586.

Barboni, D., S. P. Harrison, P. J. Bartlein, et al. 2004. "Relationships Between Plant Traits and Climate in the Mediterranean Region: A Pollen Data Analysis." *Journal of Vegetation Science* 15, no. 5: 635–646. https://doi.org/10.1111/j.1654-1103.2004.tb02305.x.

Barnosky, A. D., E. A. Hadly, P. Gonzalez, et al. 2017. "Merging Paleobiology With Conservation Biology to Guide the Future of Terrestrial Ecosystems." *Science* 355: 4787. https://doi.org/10.1126/science.aah4787.

Birks, H. J. B. 2020. "Reflections on the Use of Ecological Attributes and Traits in Quaternary Botany." *Frontiers in Ecology and Evolution* 8: 166. https://doi.org/10.3389/fevo.2020.00166.

Birks, H. J. B., V. A. Felde, A. E. Bjune, J.-A. Grytnes, H. Seppä, and T. Giesecke. 2016. "Does Pollen-Assemblage Richness Reflect Floristic Richness? A Review of Recent Developments and Future Challenges." *Review of Palaeobotany and Palynology* 228: 1–25. https://doi.org/10.1016/j.revpalbo.2015.12.011.

Boonman, C. C. F., M. A. J. Huijbregts, A. Benítez-López, A. M. Schipper, W. Thuiller, and L. Santini. 2022. "Trait-Based Projections of Climate Change Effects on Global Biome Distributions." *Diversity and Distributions* 28, no. 1: 25–37. https://doi.org/10.1111/ddi.13431.

Borgy, B., C. Violle, P. Choler, et al. 2017. "Sensitivity of Community-Level Trait–Environment Relationships to Data Representativeness: A Test for Functional Biogeography." *Global Ecology and Biogeography* 26, no. 6: 729–739. https://doi.org/10.1111/geb.12573.

Box, E. O. 1996. "Plant Functional Types and Climate at the Global Scale." *Journal of Vegetation Science* 7, no. 3: 309–320. https://doi.org/10.2307/3236274.

Boyle, B., N. Hopkins, Z. Lu, et al. 2013. "The Taxonomic Name Resolution Service: An Online Tool for Automated Standardization of Plant Names." *BMC Bioinformatics* 14, no. 1: 16. https://doi.org/10.1186/1471-2105-14-16.

Bruelheide, H., J. Dengler, B. Jiménez-Alfaro, et al. 2019. "sPlot—A New Tool for Global Vegetation Analyses." *Journal of Vegetation Science* 30, no. 2: 161–186. https://doi.org/10.1111/jvs.12710.

Bruelheide, H., J. Dengler, O. Purschke, et al. 2018. "Global Trait–Environment Relationships of Plant Communities." *Nature Ecology & Evolution* 2, no. 12: 1906–1917. https://doi.org/10.1038/s41559-018-0699-8.

Brussel, T., and S. C. Brewer. 2021. "Functional Paleoecology and the Pollen-Plant Functional Trait Linkage." *Frontiers in Ecology and Evolution* 8: 564609. https://doi.org/10.3389/fevo.2020.564609.

Cai, L., H. Kreft, A. Taylor, et al. 2023. "Global Models and Predictions of Plant Diversity Based on Advanced Machine Learning Techniques." *New Phytologist* 237, no. 4: 1432–1445. https://doi.org/10.1111/nph.18533.

Carter, V. A., R. C. Chiverrell, J. L. Clear, et al. 2018. "Quantitative Palynology Informing Conservation Ecology in the Bohemian/Bavarian Forests of Central Europe." *Frontiers in Plant Science* 8: 2268. https://doi.org/10.3389/fpls.2017.02268.

Carvalho, F., K. A. Brown, M. P. Waller, M. J. Bunting, A. Boom, and M. J. Leng. 2019. "A Method for Reconstructing Temporal Changes in Vegetation Functional Trait Composition Using Holocene Pollen Assemblages." *PLoS One* 14, no. 5: e0216698. https://doi.org/10.1371/journal.pone.0216698.

Cayuela, L., Í. Granzow-de la Cerda, F. S. Albuquerque, and D. J. Golicher. 2012. "taxonstand: An r Package for Species Names Standardisation in Vegetation Databases." *Methods in Ecology and Evolution* 3, no. 6: 1078–1083. https://doi.org/10.1111/j.2041-210X.2012.00232.x.

Chen, F., Z. Yu, M. Yang, et al. 2008. "Holocene Moisture Evolution in Arid Central Asia and Its Out-of-Phase Relationship With Asian Monsoon History." *Quaternary Science Reviews* 27: 351. https://doi.org/10.1016/j.quascirev.2007.10.017.

Chen, F.-H., J.-H. Chen, J. Holmes, et al. 2010. "Moisture Changes Over the Last Millennium in Arid Central Asia: A Review, Synthesis and Comparison With Monsoon Region." *Quaternary Science Reviews* 29, no. 7–8: 1055–1068. https://doi.org/10.1016/j.quascirev.2010.01.005.

Chen, J., F. Chen, S. Feng, W. Huang, J. Liu, and A. Zhou. 2015. "Hydroclimatic Changes in China and Surroundings During the Medieval Climate Anomaly and Little Ice Age: Spatial Patterns and Possible Mechanisms." *Quaternary Science Reviews* 107: 98–111. https://doi.org/10.1016/j.quascirev.2014.10.012.

Chevalier, M., B. A. S. Davis, O. Heiri, et al. 2020. "Pollen-Based Climate Reconstruction Techniques for Late Quaternary Studies." *Earth-Science Reviews* 210: 103384. https://doi.org/10.1016/j.earscirev.2020.103384.

Cleal, C., H. S. Pardoe, C. M. Berry, et al. 2021. "Palaeobotanical Experiences of Plant Diversity in Deep Time. 1: How Well Can We Identify Past Plant Diversity in the Fossil Record?" *Palaeogeography, Palaeoclimatology, Palaeoecology* 576: 110481. https://doi.org/10.1016/j.palaeo.2021.110481.

Connor, S. E., D. Colombaroli, F. Confortini, et al. 2018. "Long-Term Population Dynamics: Theory and Reality in a Peatland Ecosystem." *Journal of Ecology* 106, no. 1: 333–346. https://doi.org/10.1111/1365-2745.12865.

Cruz-Silva, E., S. P. Harrison, E. Marinova, and I. C. Prentice. 2022. "A New Method Based on Surface-Sample Pollen Data for Reconstructing Palaeovegetation Patterns." *Journal of Biogeography* 49, no. 7: 1381–1396. https://doi.org/10.1111/jbi.14448.

Davis, B. A. S., M. Chevalier, P. Sommer, et al. 2020. "The Eurasian Modern Pollen Database (EMPD), Version 2." *Earth System Science Data* 12, no. 4: 2423–2445. https://doi.org/10.5194/essd-12-2423-2020.

De Klerk, P., and H. Joosten. 2007. "The Difference Between Pollen Types and Plant Taxa: A Plea for Clarity and Scientific Freedom." *Quaternary Science Journal* 56, no. 3: 162–171. https://doi.org/10.3285/eg.56.3.02.

Denelle, P., P. Weigelt, and H. Kreft. 2023. "GIFT—An R Package to Access the Global Inventory of Floras and Traits." *Methods in Ecology and Evolution* 14, no. 11: 2738–2748. https://doi.org/10.1111/2041-210X.14213.

Díaz, S., J. Kattge, J. H. C. Cornelissen, et al. 2016. "The Global Spectrum of Plant Form and Function." *Nature* 529, no. 7585: 167–171. https://doi.org/10.1038/nature16489.

Díaz, S., J. Kattge, J. H. C. Cornelissen, et al. 2022. "The Global Spectrum of Plant Form and Function: Enhanced Species-Level Trait Dataset." *Scientific Data* 9, no. 1: 755. https://doi.org/10.1038/s41597-022-01774-9.

Dinerstein, E., D. Olson, A. Joshi, et al. 2017. "An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm." *Bioscience* 67, no. 6: 534–545. https://doi.org/10.1093/biosci/bix014.

Dixon, P. 2003. "VEGAN, a Package of R Functions for Community Ecology." *Journal of Vegetation Science* 14, no. 6: 927–930.

d'Oliveira, L., L. Dugerdil, G. Ménot, et al. 2023. "Reconstructing 15,000 Years of Southern France Temperatures From Coupled Pollen and Molecular (brGDGT) Markers (Canroute, Massif Central).” *Climate of the Past Discussions* 2023: 1–44.

Dugerdil, L., S. Joannin, O. Peyron, et al. 2021. “Climate Reconstructions Based on GDGT and Pollen Surface Datasets From Mongolia and Baikal Area: Calibrations and Applicability to Extremely Cold–Dry Environments Over the Late Holocene.” *Climate of the Past* 17, no. 3: 1199–1226. https://doi.org/10.5194/cp-17-1199-2021.

Enquist, B. J., J. Norberg, S. P. Bonser, et al. 2015. “Scaling From Traits to Ecosystems: Developing a General Trait Driver Theory via Integrating Trait-Based and Metabolic Scaling Theories.” In *Advances in Ecological Research*, vol. 52, 249–318. Elsevier. https://www.sciencedirect.com/science/article/pii/S0065250415000070.

Fazayeli, F., A. Banerjee, J. Kattge, F. Schrodt, and P. B. Reich. 2014. “Uncertainty Quantified Matrix Completion Using Bayesian Hierarchical Matrix Factorization.” 13th International Conference on Machine Learning and Applications, 312–317. Detroit, MI, USA: IEEE. https://doi.org/10.1109/ICMLA.2014.56.

Finkelstein, S. A., K. Gajewski, and A. E. Viau. 2006. “Improved Resolution of Pollen Taxonomy Allows Better Biogeographical Interpretation of Post-Glacial Forest Development: Analyses From the North American Pollen Database.” *Journal of Ecology* 94, no. 2: 415–430. https://doi.org/10.1111/j.1365-2745.2005.01087.x.

Fyfe, R. M., J.-L. de Beaulieu, H. Binney, et al. 2009. “The European Pollen Database: Past Efforts and Current Activities.” *Vegetation History and Archaeobotany* 18, no. 5: 417–424. https://doi.org/10.1007/s00334-009-0215-9.

Garnier, E., J. Cortez, G. Billès, et al. 2004. “Plant Functional Markers Capture Ecosystem Properties During Secondary Succession.” *Ecology* 85, no. 9: 2630–2637. https://doi.org/10.1890/03-0799.

Garnier, E., M.-L. Navas, and K. Grigulis. 2016. *Plant Functional Diversity: Organism Traits, Community Structure, and Ecosystem Properties*. Oxford University Press.

Garren, S. T. 2017. “Permutation Tests for Nonparametric Statistics Using R.” *Asian Research Journal of Mathematics* 5, no. 2: 1–8. https://doi.org/10.9734/ARJOM/2017/34694.

Gaüzère, P., L. L. Iversen, A. W. R. Seddon, C. Violle, and B. Blonder. 2020. “Equilibrium in Plant Functional Trait Responses to Warming Is Stronger Under Higher Climate Variability During the Holocene.” *Global Ecology and Biogeography* 29, no. 11: 2052–2066. https://doi.org/10.1111/geb.13176.

Giesecke, T., S. Wolters, J. F. N. Van Leeuwen, P. W. O. Van Der Knaap, M. Leydet, and S. Brewer. 2019. “Postglacial Change of the Floristic Diversity Gradient in Europe.” *Nature Communications* 10, no. 1: 5422. https://doi.org/10.1038/s41467-019-13233-y.

Gillison, A. N. 2019. “Plant Functional Indicators of Vegetation Response to Climate Change, Past Present and Future: I. Trends, Emerging Hypotheses and Plant Functional Modality.” *Flora* 254: 12–30.

Goring, S., T. Lacourse, M. G. Pellatt, and R. W. Mathewes. 2013. “Pollen Assemblage Richness Does Not Reflect Regional Plant Species Richness: A Cautionary Tale.” *Journal of Ecology* 101, no. 5: 1137–1145. https://doi.org/10.1111/1365-2745.12135.

Harrison, S. P., I. C. Prentice, D. Barboni, K. E. Kohfeld, J. Ni, and J.-P. Sutra. 2010. “Ecophysiological and Bioclimatic Foundations for a Global Plant Functional Classification.” *Journal of Vegetation Science* 21, no. 2: 300–317. https://doi.org/10.1111/j.1654-1103.2009.01144.x.

Huang, J., H. Yu, X. Guan, G. Wang, and R. Guo. 2016. “Accelerated Dryland Expansion Under Climate Change.” *Nature Climate Change* 6, no. 2: 166–171. https://doi.org/10.1038/nclimate2837.

Huang, A., Y. Zhou, Y. Zhang, D. Huang, Y. Zhao, and H. Wu. 2014. “Changes of the Annual Precipitation Over Central Asia in the Twenty-First Century Projected by Multimodels of CMIP5.” *Journal of Climate* 27, no. 17: 6627–6646. https://doi.org/10.1175/JCLI-D-14-00070.1.

Jamsranjav, C., R. S. Reid, M. E. Fernández-Giménez, A. Tsevlee, B. Yadamsuren, and M. Heiner. 2018. “Applying a Dryland Degradation Framework for Rangelands: The Case of Mongolia.” *Ecological Applications* 28, no. 3: 622–642. https://doi.org/10.1002/eap.1684.

Kattge, J., G. Bönisch, S. Díaz, S. Lavorel, I. C. Prentice, and P. Leadley. 2020. “TRY Plant Trait Database–Enhanced Coverage and Open Access.” *Global Change Biology* 26, no. 1: 119–188. https://doi.org/10.1111/gcb.14904.

König, C., P. Weigelt, and H. Kreft. 2017. “Dissecting Global Turnover in Vascular Plants.” *Global Ecology and Biogeography* 26, no. 2: 228–242. https://doi.org/10.1111/geb.12536.

Kumke, T., A. Hense, and C. Schölzel. 2004. “Transfer Functions for Paleoclimate Reconstructions—Applications.” In *The Climate in Historical Times*, 245–262. Springer. https://doi.org/10.1007/978-3-662-10313-5_14.

Last, W. M., J. P. Smol, and H. J. B. Birks. 2001. *Tracking Environmental Change Using Lake Sediments. Vol. 3: Terrestrial, Algal, and Siliceous Indicators*. Kluwer Academic.

Lavorel, S., K. Grigulis, P. Lamarque, et al. 2011. “Using Plant Functional Traits to Understand the Landscape Distribution of Multiple Ecosystem Services.” *Journal of Ecology* 99, no. 1: 135–147. https://doi.org/10.1111/j.1365-2745.2010.01753.x.

Li, J.-F., G. Xie, J. Yang, et al. 2020. “Asian Summer Monsoon Changes the Pollen Flow on the Tibetan Plateau.” *Earth-Science Reviews* 202: 103114. https://doi.org/10.1016/j.earscirev.2020.103114.

Li, W., K. S. Tojibaev, H. Hisoriev, et al. 2020. “Mapping Asia Plants: Current Status of Floristic Information for Central Asian Flora.” *Global Ecology and Conservation* 24: e01220. https://doi.org/10.1016/j.gecco.2020.e01220.

Lioubimtseva, E., R. Cole, J. M. Adams, and G. Kapustin. 2005. “Impacts of Climate and Land-Cover Changes in Arid Lands of Central Asia.” *Journal of Arid Environments* 62, no. 2: 285–308. https://doi.org/10.1016/j.jaridenv.2004.11.005.

Lioubimtseva, E. 2003. “12 Arid Environments.” In *The Physical Geography of Northern Eurasia*, vol. 3, 267. Oxford University Press.

Maitner, B. S., B. Boyle, N. Casler, et al. 2018. “The Bien r Package: A Tool to Access the Botanical Information and Ecology Network (BIEN) Database.” *Methods in Ecology and Evolution* 9, no. 2: 373–379. https://doi.org/10.1111/2041-210X.12861.

Marquer, L., M.-J. Gaillard, S. Sugita, et al. 2017. “Quantifying the Effects of Land Use and Climate on Holocene Vegetation in Europe.” *Quaternary Science Reviews* 171: 20–37. https://doi.org/10.1016/j.quascirev.2017.07.001.

Maya-Manzano, J. M., M. Sadyś, R. Tormo-Molina, et al. 2017. “Relationships Between Airborne Pollen Grains, Wind Direction and Land Cover Using GIS and Circular Statistics.” *Science of the Total Environment* 584: 603–613. https://doi.org/10.1016/j.scitotenv.2017.01.085.

Meyer, C., P. Weigelt, and H. Kreft. 2016. “Multidimensional Biases, Gaps and Uncertainties in Global Plant Occurrence Information.” *Ecology Letters* 19, no. 8: 992–1006. https://doi.org/10.1111/ele.12624.

Myszkowska, D., P. Bogawski, K. Piotrowicz, et al. 2023. “Co-Exposure to Highly Allergenic Airborne Pollen and Fungal Spores in Europe.” *Science of the Total Environment* 905: 167285. https://doi.org/10.1016/j.scitotenv.2023.167285.

Mucina, L. 2019. “Biome: Evolution of a Crucial Ecological and Biogeographical Concept.” *New Phytologist* 222, no. 1: 97–114. https://doi.org/10.1111/nph.15609.

Nowak, A., S. Nowakowska, A. Nobis, et al. 2020. *Illustrated Flora of Tajikistan and Adjacent Areas*. PAN, Polish Academy of Sciences.

Pakeman, R. J. 2014. “Functional Trait Metrics Are Sensitive to the Completeness of the Species' Trait Data?” *Methods in Ecology and Evolution* 5, no. 1: 9–15. https://doi.org/10.1111/2041-210X.12136.

Pakeman, R. J., and H. M. Quested. 2007. “Sampling Plant Functional Traits: What Proportion of the Species Need to Be Measured?” *Applied Vegetation Science* 10, no. 1: 91–96. https://doi.org/10.1111/j.1654-109X.2007.tb00507.x.

Pelánková, B., and M. Chytrý. 2009. “Surface Pollen–Vegetation Relationships in the Forest-Steppe, Taiga and Tundra Landscapes of the Russian Altai Mountains.” *Review of Palaeobotany and Palynology* 157, no. 3–4: 253–265. https://doi.org/10.1016/j.revpalbo.2009.05.005.

Peyron, O., M. Magny, S. Goring, et al. 2013. “Contrasting Patterns of Climatic Changes During the Holocene Across the Italian Peninsula Reconstructed From Pollen Data.” *Climate of the Past* 9, no. 3: 1233–1252. https://doi.org/10.5194/cp-9-1233-2013.

Peyron, O., N. Combourieu-Nebout, D. Brayshaw, et al. 2017. “Precipitation Changes in the Mediterranean Basin During the Holocene From Terrestrial and Marine Pollen Records: A Model–Data Comparison.” *Climate of the Past* 13, no. 3: 249–265. https://doi.org/10.5194/cp-13-249-2017.

Peyron, O., J. Guiot, R. Cheddadi, et al. 1998. “Climatic Reconstruction in Europe for 18,000 YR B.P. From Pollen Data.” *Quaternary Research* 49, no. 2: 183–196. https://doi.org/10.1006/qres.1997.1961.

Pierce, S., D. Negreiros, B. E. L. Cerabolini, et al. 2017. “A Global Method for Calculating Plant CSR Ecological Strategies Applied Across Biomes World-Wide.” *Functional Ecology* 31, no. 2: 444–457. https://doi.org/10.1111/1365-2435.12722.

Prentice, C. 1988. “Records of Vegetation in Time and Space: The Principles of Pollen Analysis.” In *Vegetation History*, 17–42. Springer.

Prentice, C., J. Guiot, B. Huntley, D. Jolly, and R. Cheddadi. 1996. “Reconstructing Biomes From Palaeoecological Data: A General Method and Its Application to European Pollen Data at 0 and 6 Ka.” *Climate Dynamics* 12, no. 3: 185–194. https://doi.org/10.1007/BF00211617.

Prentice, I. C. 1985. “Pollen Representation, Source Area, and Basin Size: Toward a Unified Theory of Pollen Analysis.” *Quaternary Research* 23, no. 1: 76–86. https://doi.org/10.1016/0033-5894(85)90073-0.

Prentice, I. C., W. Cramer, S. P. Harrison, R. Leemans, R. A. Monserud, and A. M. Solomon. 1992. “Special Paper: A Global Biome Model Based on Plant Physiology and Dominance, Soil Properties and Climate.” *Journal of Biogeography* 19, no. 2: 117. https://doi.org/10.2307/2845499.

Rao, Z., H. Guo, J. Cao, et al. 2020. “Consistent Long-Term Holocene Warming Trend at Different Elevations in the Altai Mountains in Arid Central Asia.” *Journal of Quaternary Science* 35, no. 8: 1036–1045. https://doi.org/10.1002/jqs.3254.

Robertson, T., M. Döring, R. Guralnick, et al. 2014. “The GBIF Integrated Publishing Toolkit: Facilitating the Efficient Publishing of Biodiversity Data on the Internet.” *PLoS One* 9, no. 8: e102623. https://doi.org/10.1371/journal.pone.0102623.

Salonen, J. S., M. Korpela, J. W. Williams, and M. Luoto. 2019. “Machine-Learning Based Reconstructions of Primary and Secondary Climate Variables From North American and European Fossil Pollen Data.” *Scientific Reports* 9, no. 1: 15805. https://doi.org/10.1038/s41598-019-52293-4.

Salonen, J. S., M. Luoto, T. Alenius, et al. 2014. “Reconstructing Palaeoclimatic Variables From Fossil Pollen Using Boosted Regression Trees: Comparison and Synthesis With Other Quantitative Reconstruction Methods.” *Quaternary Science Reviews* 88: 69–81. https://doi.org/10.1016/j.quascirev.2014.01.011.

Scheiter, S., L. Langan, and S. I. Higgins. 2013. “Next-Generation Dynamic Global Vegetation Models: Learning From Community Ecology.” *New Phytologist* 198, no. 3: 957–969. https://doi.org/10.1111/nph.12210.

Sherwood, S., and Q. Fu. 2014. “A Drier Future?” *Science* 343, no. 6172: 737–739. https://doi.org/10.1126/science.1247620.

Shipley, B., F. De Bello, J. H. C. Cornelissen, E. Laliberté, D. C. Laughlin, and P. B. Reich. 2016. “Reinforcing Loose Foundation Stones in Trait-Based Plant Ecology.” *Oecologia* 180, no. 4: 923–931. https://doi.org/10.1007/s00442-016-3549-x.

van der Sande, M. T., W. Gosling, and A. Correa-Metrio. 2019. “A 7000-Year History of Changing Plant Trait Composition in an Amazonian Landscape; the Role of Humans and Climate.” *Ecology Letters* 22: 925–935. https://doi.org/10.1111/ele.13251.

van der Sande, M. T., M. B. Bush, D. H. Urrego, et al. 2021. “Modern Pollen Rain Predicts Shifts in Plant Trait Composition but not Plant Diversity Along the Andes–Amazon Elevational Gradient.” *Journal of Vegetation Science* 32, no. 1: e12925. https://doi.org/10.1111/jvs.12925.

Schrodt, F., J. Kattge, H. Shan, et al. 2015. “BHPMF—a Hierarchical Bayesian Approach to Gap-Filling and Trait Prediction for Macroecology and Functional Biogeography: Gap-Filling in Trait Databases.” *Global Ecology and Biogeography* 24, no. 12: 1510–1521. https://doi.org/10.1111/geb.12335.

Smith, M., S. Jäger, U. Berger, et al. 2014. “Geographic and Temporal Variations in Pollen Exposure Across Europe.” *Allergy* 69, no. 7: 913–923. https://doi.org/10.1111/all.12419.

Sugita, S. 2007. “Theory of Quantitative Reconstruction of Vegetation I: Pollen From Large Sites REVEALS Regional Vegetation Composition.” *Holocene* 17, no. 2: 229–241. https://doi.org/10.1177/0959683607075837.

Sun, A., Y. Luo, H. Wu, et al. 2020. “An Updated Biomization Scheme and Vegetation Reconstruction Based on a Synthesis of Modern and Mid-Holocene Pollen Data in China.” *Global and Planetary Change* 192: 103178. https://doi.org/10.1016/j.gloplacha.2020.103178.

Tarasov, P. E., R. Cheddadi, J. Guiot, et al. 1998. “A Method to Determine Warm and Cool Steppe Biomes From Pollen Data; Application to the Mediterranean and Kazakhstan Regions.” *Journal of Quaternary Science* 13, no. 4: 335–344.

Tarasov, P. E., J. Guiot, R. Cheddadi, et al. 1999. “Climate in Northern Eurasia 6000 Years Ago Reconstructed From Pollen Data.” *Earth and Planetary Science Letters* 171, no. 4: 635–645. https://doi.org/10.1016/S0012-821X(99)00171-5.

Tavşanoğlu, Ç., and J. G. Pausas. 2018. “A Functional Trait Database for Mediterranean Basin Plants.” *Scientific Data* 5, no. 1: 180135. https://doi.org/10.1038/sdata.2018.135.

ter Braak, C. J. F., S. Juggins, H. J. B. Birks, and H. Van der Voet. 1993. “Weighted Averaging Partial Least Squares Regression (WA-PLS): Definition and Comparison With Other Methods for Species-Environment Calibration.” In *Multivariate Environmental Statistics*, 525–560. Elsevier.

The Angiosperm Phylogeny Group, M. W. Chase, M. J. M. Christenhusz, et al. 2016. “An Update of the Angiosperm Phylogeny Group Classification for the Orders and Families of Flowering Plants: APG IV.” *Botanical Journal of the Linnean Society* 181, no. 1: 20. https://doi.org/10.1111/boj.12385.

Theuerkauf, M., J. Couwenberg, A. Kuparinen, and V. Liebscher. 2016. “A Matter of Dispersal: REVEALSinR Introduces State-of-the-Art Dispersal Models to Quantitative Vegetation Reconstruction.” *Vegetation History and Archaeobotany* 25, no. 6: 541–553. https://doi.org/10.1007/s00334-016-0572-0.

Tierney, J. E., C. J. Poulsen, I. P. Montañez, et al. 2020. “Past Climates Inform Our Future.” *Science* 370, no. 6517: eaay3701. https://doi.org/10.1126/science.aay3701.

Van Bodegom, P. M., J. C. Douma, J. P. M. Witte, J. C. Ordoñez, R. P. Bartholomeus, and R. Aerts. 2012. “Going Beyond Limitations of Plant Functional Types When Predicting Global Ecosystem–Atmosphere Fluxes: Exploring the Merits of Traits-Based Approaches.” *Global Ecology and Biogeography* 21, no. 6: 625–636. https://doi.org/10.1111/j.1466-8238.2011.00717.x.

Veeken, A., M. J. Santos, S. McGowan, A. L. Davies, and F. Schrodt. 2022. “Pollen-Based Reconstruction Reveals the Impact of the Onset of

Agriculture on Plant Functional Trait Composition." *Ecology Letters* 25, no. 9: 1937–1951. https://doi.org/10.1111/ele.14063.

Violle, C., M. Navas, D. Vile, et al. 2007. "Let the Concept of Trait Be Functional!" *Oikos* 116, no. 5: 882–892. https://doi.org/10.1111/j.0030-1299.2007.15559.x.

Weigelt, P., C. König, and H. Kreft. 2020. "GIFT—A Global Inventory of Floras and Traits for Macroecology and Biogeography." *Journal of Biogeography* 47, no. 1: 16–43. https://doi.org/10.1111/jbi.13623.

Wieczorek, M., and U. Herzschuh. 2020. "Compilation of Relative Pollen Productivity (RPP) Estimates and Taxonomically Harmonised RPP Datasets for Single Continents and Northern Hemisphere Extratropics." *Earth System Science Data* 12, no. 4: 3515–3528. https://doi.org/10.5194/essd-12-3515-2020.

Wright, I. J., P. B. Reich, M. Westoby, et al. 2004. "The Worldwide Leaf Economics Spectrum." *Nature* 428, no. 6985: 821–827. https://doi.org/10.1038/nature02403.

Zheng, Z., J. Wei, K. Huang, et al. 2014. "East Asian Pollen Database: Modern Pollen Distribution and Its Quantitative Relationship With Vegetation and Climate." *Journal of Biogeography* 41, no. 10: 1819–1832. https://doi.org/10.1111/jbi.12361.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.