



OPEN

DATA DESCRIPTOR

NAMPHORA: a fossil and modern pollen database from Northern Africa and adjacent Mediterranean and Arabian regions

Irene Solano¹✉, Jakob Bro-Jørgensen¹, Ignacio A. Lazagabaster^{1,2}, Chris D. Thomas³ & Saúl Manzano^{4,5,6}

Northern Africa's climate and vegetation underwent significant changes throughout the Holocene, particularly in connection with the termination of the African Humid Period ca. 5500 years ago. Fossil pollen records are key to reconstructing past environments, yet current databases for this region are limited by the omission of significant unpublished data, taxonomic inconsistencies, and the lack of standardised plant trait information. To address these issues, we introduce the Northern African, Arabian, and Mediterranean Pollen Holocene Records Archive (NAMPHORA)—a comprehensive, machine-readable and taxonomically-harmonised database compiling fossil and modern pollen records alongside plant functional traits, and ecological and phytogeographical information. This database includes all of Africa to the north of 7.52° N and constitutes the most complete and comprehensive resource (836 pollen records; 853 harmonised pollen types, and 13 key standardised plant traits) to improve palaeoecological reconstructions, enhance biogeographical analyses, and refine climate models for northern Africa during the Holocene. It enables direct data retrieval via programming languages such as R, and all datasets and code are openly available via Zenodo.

Background & Summary

Northern Africa has undergone multiple climatic fluctuations that have significantly shaped its physical landscape and vegetation dynamics^{1,2}. However, research on the region's ecological and biogeographical dynamics remain challenging due to difficulties in reconstructing its paleoclimate, particularly during the African Humid Period (AHP: 14,800–5,500 years BP)^{1–3}. Studies suggest that diverse but uncertain rainfall patterns—including tropical plumes, Mediterranean winter rains, and the Afro-tropical monsoon—governed northern Africa's ecosystems during the AHP, particularly in the Sahara^{1–4}. Plant migrations from surrounding regions^{3,5–8} transformed the area into a complex, heterogeneous landscape—a “melting pot” of plant communities⁹ thought to include Mediterranean, tropical and Arabian floral affinities¹⁰. Despite extensive palaeobotanical and paleoclimatic research, an integrated palaeoecological understanding of northern Africa is still lacking due to the scattering of the data through space and time leaving significant gaps remain in our understanding of the origins, past distribution, and climatic influences on northern Africa's unique vegetation.

Fossil pollen records are among the most widely used proxies in palaeoecological studies¹¹ and provide a powerful resource for reconstructing past vegetation and climate dynamics during the AHP. However, to accurately interpret fossil records, it is essential to establish robust relationships between pollen types (the palaeopalynological identification unit) and the environmental parameters they represent (e.g., vegetation composition, climate variables)¹². This is usually achieved by correlating modern pollen rain with present-day vegetation, to

¹Department of Evolution, Ecology and Behaviour, University of Liverpool, Liverpool, UK. ²Centro Nacional de Investigación Sobre Evolución Humana CENIEH, Burgos, Spain. ³Leverhulme Centre for Anthropocene Biodiversity, Department of Biology, University of York, York, UK. ⁴Quaternary Palynology Lab. Instituto de Medio Ambiente, Recursos Naturales y Biodiversidad, Universidad de León, León, Spain. ⁵The Bolus Herbarium, Department of Biological Sciences, University of Cape Town, Cape Town, 7701, South Africa. ⁶Área de Botánica, Departamento de Biodiversidad y Gestión Ambiental, Facultad de Ciencias Biológicas y Ambientales, Universidad de León, León, Spain.

✉e-mail: irene13@liverpool.ac.uk

extrapolate relationships between fossil pollen and past vegetation using these modern analogues¹³. Modern pollen records can also be used to identify past plant communities with no modern analogues, shedding light on how vegetation assemblage has responded to global change in unexpected, unpredictable manners¹⁴. Integrating modern and fossil pollen records is crucial for achieving accurate paleoenvironmental reconstructions and understanding its implications for environmental management relevant to future non-analogue climates.

From a functional ecological point of view, fossil pollen data are increasingly being integrated with plant functional traits (PFTs) to reconstruct past ecological dynamics and understand underpinning patterns of vegetation assembly¹⁵. PFTs represent groups of morphological, physiological, and phenological traits that influence plant survival, growth, and adaptation^{16,17}. Since these traits are related to plant responses to environmental conditions, they can be used to hindcast functional changes in plant communities and to understand past vegetation shifts, and ecological processes¹⁶. However, no comprehensive database has yet compiled modern and fossil pollen records alongside detailed plant functional traits for the study region, hampering an integrated understanding of the long-term assembly of plant communities in northern Africa and adjacent regions.

Currently, two major pollen databases are available: the Neotoma Paleocology Database¹⁸ (<https://www.neotomadb.org/>), which hosts global multi-proxy records and integrates several regional databases, and the African Pollen Database¹⁹ (APD: <https://africanpollendatabase.ipsl.fr/#/home>), which focuses specifically on African pollen data. While Neotoma (established in 2009) offers an accessible web interface and an R package for data retrieval²⁰, it includes fewer pollen records for Africa and the Arabian Peninsula than are included in the APD (developed in 1996). Despite their considerable value, neither database covers the full range of available pollen records from northern Africa and the Arabian Peninsula, and both lack the inclusion of plant functional traits and a clearly documented taxonomic harmonisation suitable for integrating records from diverse sources. Additionally, accessing the APD can be slow and inconvenient for retrieving multiple datasets, as it does not support direct access via programming languages like R. These constrain large-scale, reproducible palaeoecological research on the region.

In addition to pollen databases, there are several plant functional trait (PFT) databases, such as the Botanical Information and Ecology Network²¹ (BIEN: <https://bien.nceas.ucsb.edu/bien/>) and the TRY Plant Trait Database²² (TRY: <https://www.try-db.org/TryWeb/Home.php>), which compile traits from thousands of species. However, there is no database that provides a direct linkage between pollen types and the species-level traits they represent. Pollen grains are assigned to particular morphological ‘types’, which often correspond to multiple species or higher taxa (e.g., genus or family), which can produce morphologically indistinguishable pollen grains, making trait assignment challenging. Furthermore, this process requires extensive data cleaning and processing, which is time-consuming and complicates the integration of functional traits into palaeoecological analyses.

Here we present NAMPHORA: the Northern Africa, Arabian, and Mediterranean Pollen Holocene Records Archive—a comprehensive, reproducible and open-access database that integrates both open data sources and includes a significant amount of information extracted from the grey literature. These include palaeoecological datasets, taxonomical information to harmonise palynomorph identifications (morphotypes) and a selection of key PFTs consolidated into a findable, interoperable, accessible and reusable (FAIR)²³ machine-readable format, enabling reproducible and large-scale analyses. It will serve as the critical resource for reconstructing past vegetation and climate dynamics across northern Africa during the Holocene. All datasets and code are openly available via Zenodo²⁴, and herewith we invite researchers to access, contribute to, and refine the database for future studies. All new contributions to the database will undergo a review process by our team to ensure they meet the same quality standards as the original dataset.

Methods

We created a workflow for each stage of database creation, from data collection to the final database products (Fig. 1). We used R version 4.1.1²⁵ for part of the data collection, and for the complete preparation and processing of the database (see details in the Zenodo repository²⁴).

Study area. The study area spans from 7.52° to 40.96° latitude and from −21.03° to 60.83° longitude (Fig. 2), covering northern Africa and its surrounding areas, including the Arabian Peninsula and the south of the Mediterranean region. These areas were included because of the strong floristic connections between northern Africa, southern Europe and western Asia²⁶. The flora of this region reflects the interaction of several biogeographic elements, including Saharo-Arabian, Tropical-African and Mediterranean floristic components¹⁰. Accordingly, the spatial extent was defined to include the main ecological regions adjacent and thus influencing northern Africa vegetation. The southern boundary (7.52° N) captures the Sudanian region and the regions above, including the Sahel, Sahara and Mediterranean zones. The northern boundary (40.96° N) includes sites along the southern Mediterranean margin of Europe, while limiting the expansion of the database further into Europe in order to maintain the focus on northern Africa and exclude the influence of sub-mediterranean and temperate taxa that become more prevalent in the vegetation above 40°N^{27,28}.

The study area therefore exhibits diverse vegetation types, such as deserts and xeric shrublands, Mediterranean woodlands, and tropical grasslands and savannas.

Data collection. Pollen records—both fossil (fossilised pollen grains in sedimentary deposits) and modern (modern pollen rain, collected since the 1980s)—from the Mediterranean, Arabian Peninsula, Saharan, Sahelian and Sudanian regions were extracted from Neotoma¹⁸ (<https://www.neotomadb.org/>) and the African Pollen Database¹⁹ (APD: <https://africanpollendatabase.ipsl.fr/#/home>). Additional records, either from grey literature or openly unavailable records were obtained directly from the authors. Although Neotoma integrates several regional databases, including the European Pollen Database (EPD)²⁹ and the African Pollen Database (APD)¹⁹,

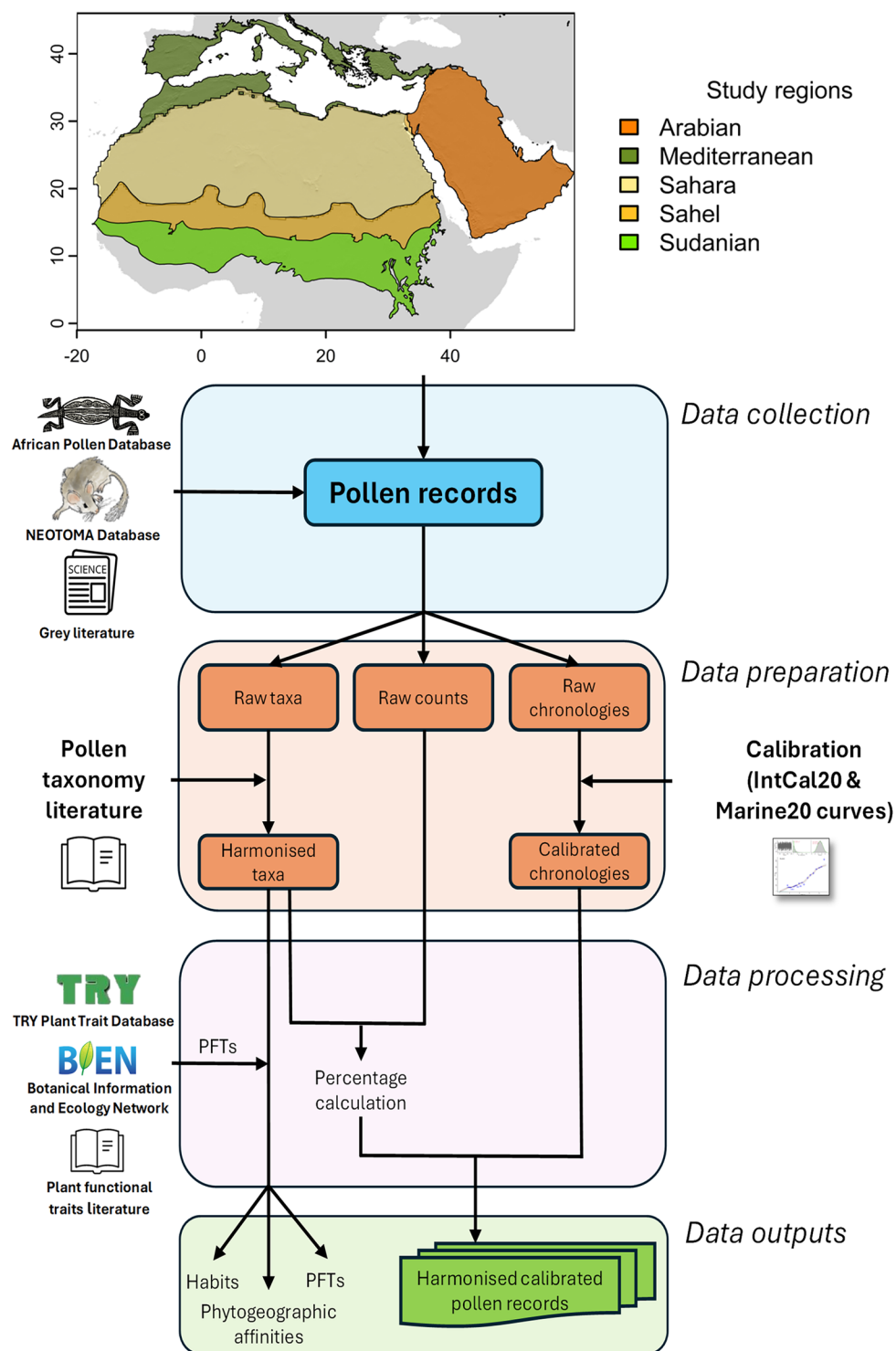


Fig. 1 Workflow of database development.

we also consulted the APD directly to ensure complete coverage of Saharan records, as some datasets in the APD have not yet been integrated into Neotoma at the time of compilation.

To extract the records from the Neotoma database, we used the *neotoma2* R package v.1.0.5²⁰, using the dataset IDs that were manually retrieved from the Neotoma website. For standardisation and comparison, the Neotoma records were modified to match the structure of the APD (see the “01_pollen_list.R” script in the Zenodo repository). In contrast, the records from the APD were manually retrieved from the APD website (<https://africanpollendatabase.ipsl.fr/#/home>). Information for each pollen record, including geographic coordinates, record type, source reference(s), and the URLs to the Neotoma and APD databases where the corresponding records are stored, is provided in the file “database.csv” located in the metadata folder of the repository.

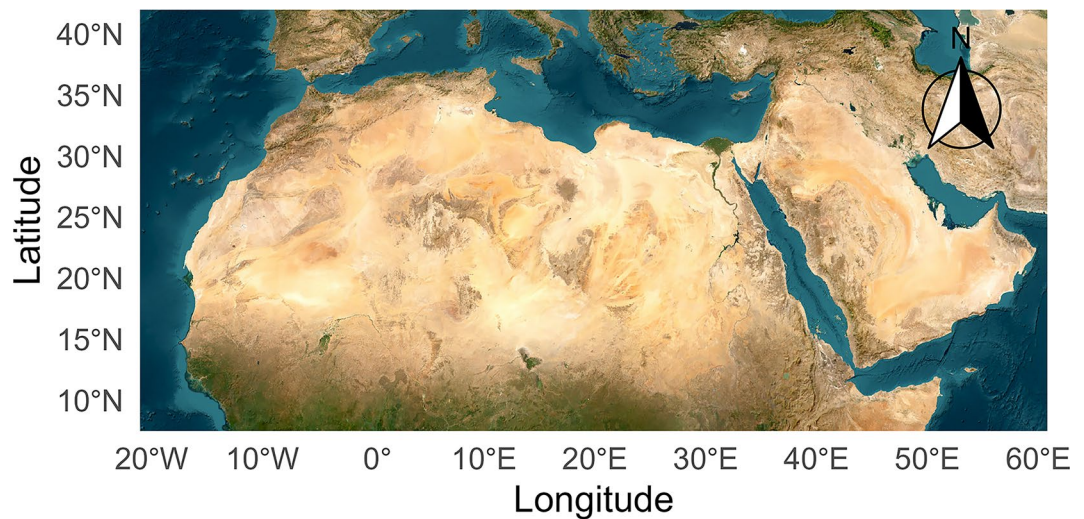


Fig. 2 Study area.

Additional records that were unavailable in openly accessible databases were obtained directly from the authors. Information about these records, including the authors' details (i.e. the name of the researcher who provided the data, the name of the organisation, and contact details) and the method of data collection, is provided in the file "*pollen_data_received_from_authors_information.csv*", located in the metadata folder of the repository.

This database also includes chronological data (uncalibrated radiocarbon dates), which were obtained alongside the pollen records in the same repositories (APD or Neotoma), retrieved from independent research articles, provided by the authors, or sourced from Giesecke *et al.*³⁰. Information about the origin of the chronological data and the calibration method is provided in the file "*calibration_information.csv*", located in the metadata folder of the repository. The current temporal coverage of data available spans the entire Holocene. Only a few records (4 in total) extend into the Middle Pleistocene (up to approximately 240,000 years BP), highlighting the limited availability of older datasets.

Additionally, plant functional traits were collected from trait databases, specifically from the Botanical Information and Ecology Network²¹ (BIEN: <https://bien.nceas.ucsb.edu/bien/>) and TRY Plant Trait Database²² (TRY: <https://www.try-db.org/TryWeb/Home.php>). Since the BIEN database has an associated R package³¹ (*Bien* R package v. 1.2.7), trait data from this source were collected and processed in R, and the results were stored in the processed data folder. In contrast, the TRY database does not provide an R package for direct data access, so the "leaf type" trait data taken from this source were manually extracted from the website (<https://www.try-db.org/TryWeb/Prop0.php>), using the "Request by traits/species" interface. The downloaded dataset was stored in the raw data folder, after which data cleaning and processing were performed.

All raw data, including the original pollen records, radiocarbon dates and pollen counts, are stored in the repository's raw data folder to ensure transparency and reproducibility. All data are legally accessible and reusable, having been obtained from public CC-BY databases or from contributors with permission.

Data preparation. *Taxonomic harmonisation.* Given the large geographic coverage of this database, the different pollen extraction protocols, and the varied taxonomical approaches to pollen identification, which depend largely on the pollen analyst and the region of study, standardisation of the pollen taxonomy and nomenclature is crucial to allow for the demarcation of comparable pollen taxonomical units (hereafter pollen types). In palynological analysis, the name given to a certain pollen grain refers to a pollen type (a morphological pollen category), rather than a plant systematic taxon, as multiple species, often related but sometimes in different genera or families, can produce morphologically indistinguishable pollen grains³². Hence, different analysts may use different names for the same pollen morphological category (synonymy of pollen types). To address this issue, we standardised the palaeopalynological records by harmonising the pollen taxonomy and nomenclature. First, we extracted the raw pollen taxa from all fossil and modern pollen records and compiled them in a list. Taxonomical harmonisation was then based on fully referenced pollen morphological information in original pollen taxonomical works (including detailed descriptions, images, keys and pollen type taxonomies), as well as our expert knowledge, to build FAIR pollen taxonomical and nomenclatural entities (pollen types) with a homogeneous nomenclature at the family, genus, species and pollen type levels. For pollen type nomenclature, we followed the nomenclatural standard proposed by Joosten and de Klerk³². A complete description of the harmonisation protocol will be provided in Manzano *et al.* (in prep.).

Additionally, we included other not as comprehensive pollen harmonisation proposals, including Lézine *et al.*¹², Mottl *et al.*³³ and the African Pollen Database taxonomy list (<https://africanpollendatabase.ipls.fr/#/taxon-dict>) to enhance interoperability of the data. The standardised list of harmonised pollen taxa and the supporting palynological literature for this standardisation can be found in the harmonised taxonomy list file.

Taxonomic harmonisation was applied to each of the raw pollen data files; hence, the original pollen types were standardised to their harmonised pollen type names using the harmonised pollen taxonomy list.

Age calibration. To ensure chronological consistency, all radiocarbon dates were recalibrated to calendar years with the most recently updated calibration curves IntCal20³⁴ and Marine20³⁵ for terrestrial and marine archives respectively. For records that reported two or more radiocarbon dates, we employed Bayesian age-depth modelling using the *rbacon* R package v.3.2.0³⁶. In cases where sites had only one dated depth or lacked depth information, we calibrated radiocarbon ages using the probability distribution method implemented in the *rice* R package v. 1.0.0³⁷. In fossil pollen records containing post-bomb dates (i.e. after 1950CE) at shallow depths, the percent modern carbon (pMC) values were converted to 14C dates using the *pMC.age* function within the *rice* package, and then calibrated with the corresponding post-bomb calibration curves³⁸ (NH zone 1, zone 2 or zone 3), determined through <http://calib.org/CALIBomb/>. When available, we accounted for reservoir effects by applying the reported delta values and incorporated hiatuses when documented in the original publication. For sites with chronologies based on tephra layers, biostratigraphic markers, or oxygen isotope dates, we retained the original ages.

Data processing. *Plant species corresponding to each pollen type.* From the harmonised pollen taxonomy list, each pollen type has been linked to all the potential plant sources extant in the study area. Since pollen types are morphological categories and not plant taxonomic units, depending on the case, a pollen type may source from any plant in a family (e.g., Annonaceae), in a group of genera (e.g., *Acacia* type, which includes the species in the genera *Acacia*, *Senegalia* and *Vachellia*), in a group of species (e.g., *Plantago major/media*, including both *P. major* and *P. media*) or even just from a single species (e.g., *Adansonia digitata*). Thus, to assign plant taxa to pollen types, we consulted the palynological literature to identify the plant species ascribable to each pollen type, considering only angiosperm and gymnosperm pollen types. We then used the function *name_lookup* from the R package *rgbif* v. 3.8.0³⁹ to retrieve all species names of plant taxa at genus and family level source to every particular pollen type. This resulted in a preliminary list of species corresponding to the taxa assigned to each pollen type.

Next, we discarded those species that were not present in the study area by checking their current distributions in the Global Biodiversity Information Facility (GBIF)⁴⁰. Because retrievals from GBIF via R are limited to a maximum of 100,000 species at a time, it was not feasible to process the full species list directly, as doing so would require excessive time and computational resources. Instead, we extracted a species list filtered by our study area (polygon) directly from the GBIF website⁴⁰. We then matched this GBIF species list with our own species list to exclude species absent from the study area. Plant taxonomy followed the GBIF Backbone Taxonomy⁴¹.

This process resulted in a final list of pollen types with their corresponding plant sources present in the study area. This list was manually inspected and cross checked against the Kew Botanic Gardens Plants of the World online website (POWO; <https://powo.science.kew.org/>), to identify introduced species and correct potential errors.

Habit categorisation. Plant habit (or growth form) is a key ecological strategy of the plant in response to the surrounding environment⁴². Plant habit categories are a classification of plants according to their morphology, mechanical architecture and development⁴³ and are strictly correlated with the abiotic and biotic conditions in which they develop⁴⁴. We assigned plant habits to pollen types by considering all source plants to each pollen type. For example, if a pollen type includes two genera (e.g., *Erythrococca*/*Micrococca* pollen type includes all the *Erythrococca* and *Micrococca* species), we assigned a plant habit to the pollen type based on the habits of the species circumscribed. When the genera exhibit different habits (e.g., shrub for *Erythrococca* and herb for *Micrococca*), both habits were included in the classification.

Therefore, the harmonised pollen types were sorted into separate groups based on their habit. Two main classifications were conducted: one based on habit diversity (such as trees, shrubs and herbs, as well as forms like epiphytes, or lianas) with 35 categories (Table 1: column b), and another based on competition for space (i.e. trees, shrubs and herbs vs aquatic species) with 17 categories (Table 1, column b). The latter classification was used to calculate summary pollen percentages afterward. To assess the most probable habit of each harmonised taxon, we consulted POWO (<https://powo.science.kew.org/>), as well as relevant research articles and regional floras, to examine the habit of the plants included in the corresponding pollen type.

Phytogeographical affinity categorisation. Phytogeographical affinity indicates the distribution of a plant species, that is, its association with a particular geographic region based on its origin, evolutionary history, and dispersal patterns⁴⁵. As with habit categorisation, each pollen type was assigned a phytogeographical affinity category based on all taxa corresponding to the pollen type. Phytogeographical categories were assigned only to pollen types corresponding to angiosperms, gymnosperms, ferns, and bryophytes.

To assign the most prevalent phytogeographical affinity to each harmonised pollen taxon (Fig. 3), we consulted GBIF⁴⁰, POWO (<https://powo.science.kew.org/>), and the works of Watrin *et al.*⁶, Hély *et al.*⁷, White⁴⁶, Quézel⁴⁷ and Olson *et al.*⁴⁸, among others, to examine the current distribution of the plants included in the pollen types. 19 phytogeographical affinity categories were established (Table 2) based on White⁴⁶ and Olson *et al.*⁴⁸.

Plant functional traits. We selected plant functional traits commonly used in palaeoecological studies and available in public trait databases (TRY and BIEN)^{3,12,16,17}. In total, we retrieved 13 plant functional traits (PFTs) (Table 3), including widely used PFTs, such as plant height, seed mass, leaf area and leaf dry mass per area (LMA), as well as less common but ecological important traits, such as flower pollination syndrome and whole plant sexual system.

The processing of plant functional traits was conducted in several steps (Fig. 4), following the methodology outlined by Brussel and Brewer¹⁶. From the final list of pollen types and their corresponding species, we extracted all available traits from all the species in the list. Continuous trait measurements were log10-transformed and

(a) Habit	(b) Habit (for percentage calculation)	Description	Examples of representative taxa
Algae	Algae	Photosynthetic organisms, mostly aquatic, occurring in unicellular, colonial, filamentous or multicellular forms ⁵⁸	<i>Spirogyra</i> , <i>Botryococcus</i>
Anthocerotophyta	Anthocerotophyta	Small and thalloid bryophytes with a horn-like sporophyte ⁵⁹	<i>Anthoceros</i>
Aquatic	Aquatic	Vascular plants adapted to life in or on water, typically with specialised floating structures ⁶⁰	<i>Myriophyllum</i> , <i>Nuphar</i>
Bryophyta	Bryophyta	Non-vascular plants forming mats or cushions, with a gametophyte-dominated life cycle, and spores produced in a single capsule ⁶¹	<i>Hedwigia</i> , <i>Sphagnum</i>
Climbers	Herbs/Shrubs	Vascular plants with flexible and vining stems that use external support structures, such as aerial roots, to grow vertically ⁶²	<i>Hedera</i> , <i>Clematis</i>
Climbing herbs	Herbs	Herbs with climbing habit.	<i>Luffa</i> , <i>Phaseolus</i>
Climbing shrubs	Shrubs	Shrubs with climbing habit.	<i>Baizea</i>
Fern	Fern	Vascular, spore-producing plants with rhizomatous stems, curled fronds, and free-living gametophytes ⁶³	<i>Adiantum</i> , <i>Equisetum</i>
Fungi	Fungi	Non-photosynthetic eukaryotes that absorb nutrients through hyphae and reproduce by spores ⁶⁴	<i>Apiosordaria</i> , <i>Sordaria</i>
Geophytes	Geophytes	Herbaceous plants that survive adverse seasons using underground storage organs such as bulbs, tubers or rhizomes ⁶⁵	<i>Allium</i> , <i>Crocus</i>
Herbs	Herbs	Non-woody plants with annual shoots and annual, biennial or perennial life cycles ⁶⁶	<i>Centaurea scabiosa</i> , <i>Plantago</i>
Herbs/Climbers	Herbs	Herbs, erect or climbing.	<i>Cucumis</i> , <i>Trifolium</i>
Herbs/Shrubs	Herbs/Shrubs	Herbs or shrubs.	<i>Acalypha</i> , <i>Artemisia</i>
Herbs/Shrubs/Climbers	Herbs/Shrubs	Herbs or shrubs, erect or climbing.	<i>Jacquemontia</i> , <i>Vigna</i>
Herbs/Shrubs/Lianas	Herbs/Shrubs	Herbs or shrubs or lianas (woody climbers).	<i>Stephania</i> , <i>Cucurbitaceae</i>
Herbs/Shrubs/Trees	Herbs/Shrubs/Trees	Herbs or shrubs or trees.	<i>Abutilon</i> , <i>Ricinus</i>
Herbs/Shrubs/Trees/Climbers	Herbs/Shrubs/Trees	Herbs or shrubs, sometimes climbing, or trees.	<i>Heliotropium</i> , <i>Solanum</i>
Herbs/Shrubs/Trees/Lianas	Herbs/Shrubs/Trees	Herbs or shrubs or trees or lianas (woody climbers).	<i>Adenia</i> , <i>Mimosa</i>
Lianas	Shrubs	Woody, perennial vines that take root in the ground and climb up other vegetation ⁶⁶	<i>Tetracera alnifolia</i> , <i>Cissampelos</i>
Mangrove	Aquatic	Woody plants adapted to saline coastal environments, with leaves that excrete salt and specialised aerial roots ⁶⁷	<i>Avicennia</i> , <i>Rhizophora</i>
Marchantiophyta	Marchantiophyta	Non-vascular bryophytes with thalloid or frondose forms that colonise moist habitats ⁶⁸	<i>Riccia</i>
Palm	Trees	Tropical monocotyledonous plants with a single, unbranched stem; usually tall and woody, without secondary thickening ⁶⁹	<i>Phoenix</i> , <i>Elaeis</i>
Parasitic	Parasitic	Angiosperms that obtain resources from host plants via haustoria (specialised organs) ⁷⁰	<i>Cytinus</i> , <i>Loranthus</i>
Selaginellales	Selaginellales	Small plants with scale-like leaves and branched stems, often mat-forming and moss-like in appearance ⁷¹	<i>Selaginella</i>
Shrubs	Shrubs	Woody perennial plants with multiple main stems emerging near ground level ⁶⁶	<i>Myrtus</i> , <i>Ephedra</i>
Shrubs/Climbers	Shrubs	Shrubs, erect or climbing.	<i>Jasminum</i> , <i>Rosa</i>
Shrubs/Lianas	Shrubs	Erect shrubs or lianas.	<i>Tinospora</i> , <i>Sherbournia</i>
Shrubs/Trees	Shrubs/Trees	Shrubs or trees.	<i>Nerium</i> , <i>Syzygium</i>
Shrubs/Trees/Climbers	Shrubs/Trees	Shrubs, erect or climbing, or trees.	<i>Allophylus</i> , <i>Zanthoxylum</i>
Shrubs/Trees/Epiphytic	Shrubs/Trees	Shrubs, sometimes epiphytic (growing on the surface of another plant), or trees.	<i>Vaccinium</i>
Shrubs/Trees/Lianas	Shrubs/Trees	Erect shrubs or lianas (a group of woody climbers) or trees	<i>Capparis</i> , <i>Periploca</i>
Shrubs/Trees/Lianas/Epiphytic	Shrubs/Trees	Shrubs, erect or climbing, sometimes epiphytic or trees.	<i>Schefflera</i>
Trees	Trees	Tall, woody perennial plants with a single main stem supporting a branched crown ⁶⁶	<i>Adansonia digitata</i> , <i>Faidherbia albida</i>
Trees/Lianas	Shrubs/Trees	Trees or lianas (a group of woody climbers).	<i>Pycnanthus</i>
Unknown/Indeterminable	Unknown/Indeterminable	Unknown or indeterminable habit due to unknown pollen grain identification.	Unknown/Indeterminable

Table 1. Habit categories.

normalised to approximate a normal distribution and reduce the influence of outliers. For species with multiple observations, we calculated the mean and standard deviation. Finally, for each pollen type, we calculated the mean value of each trait across all associated species, so that we had a final matrix of pollen types with their averaged traits.

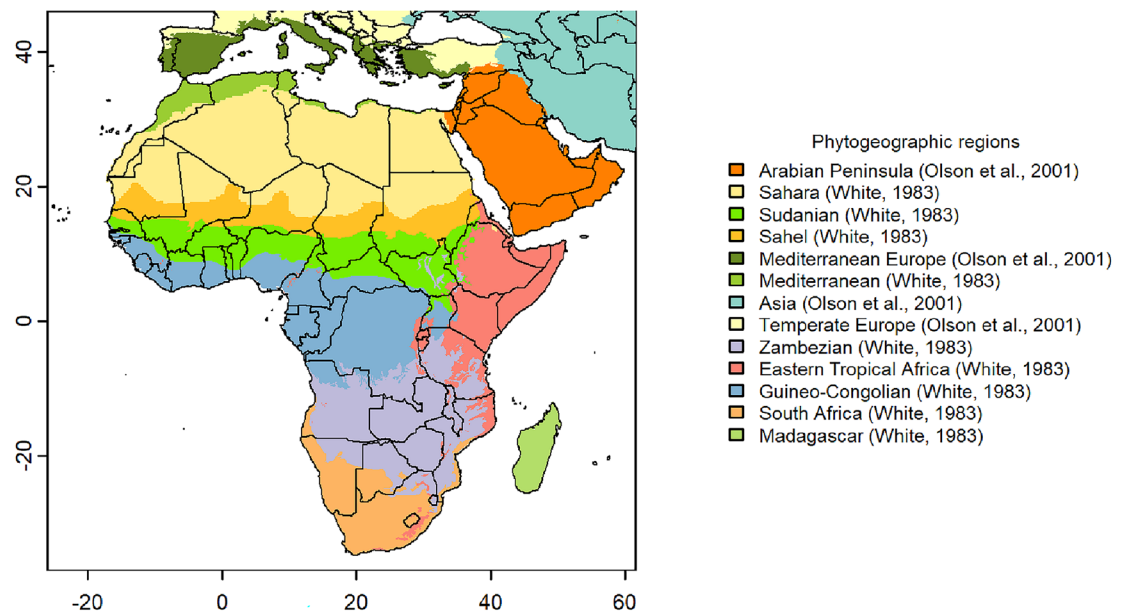


Fig. 3 Main summarised phylogeographic regions based on White⁴⁶ and Olson *et al.*⁴⁸, used to categorise pollen types into phylogeographical affinities. Regions from America, Macaronesia, Australasia, and parts of Asia and Europe are excluded from the map.

Phylogeographical affinity categories from the works of White ⁴⁶ for the African continent and Olson <i>et al.</i> ⁴⁸ for the rest of the world	Summarised phylogeographical affinity of pollen type	Description
Guineo-Congolian centre of endemism; Guineo-Congolian/Sudanian transition zone & Guineo-Congolian/Zambezian regional transition zone ⁴⁶ .	Guineo-Congolian	Parent plant occurring in the Guineo-Congolian regional centre of endemism.
Sudanian regional centre of endemism ⁴⁶ .	Sudanian	Parent plant occurring in the Sudanian regional centre of endemism.
Zambezian centre of endemism ⁴⁶ .	Zambezian	Parent plant occurring in the Zambezian regional centre of endemism.
Sahel regional transition zone ⁴⁶ .	Sahel	Parent plant occurring in the Sahel.
Mediterranean Africa, Mediterranean/Sahara transition zone ⁴⁶ & Mediterranean region (Mediterranean forests, woodlands and shrubs) ⁴⁸ .	Mediterranean	Parent plant occurring in the Mediterranean region.
Sahara regional transition zone ⁴⁶ .	Sahara	Parent plant occurring in the Sahara.
Karoo-Namib centre of endemism, Cape centre of endemism, Tongaland-Pondoland mosaic & Kalahari-Highveld transition zone ⁴⁶ .	South Africa	Parent plant occurring in South Africa.
Eastern part of the Zambezian centre of endemism, Somalia-Masai centre of endemism, Afromontane and Afroalpine archipelagos, Lake Victoria mosaic & Zanzibar-Inhambane mosaic ⁴⁶ .	Eastern Tropical Africa	Parent plant occurring in Eastern Tropical Africa (East Zambezian area, Somalia-Masai, Afromontane Archipelago, Lake Victoria or/and Zanzibar-Inhambane).
East Malagasy centre of endemism & West Malagasy centre of endemism ⁴⁶ .	Madagascar	Parent plant native to Madagascar.
Azonal vegetation: Mangrove, and fresh-water swamp vegetation ⁴⁶ .	African coast	Parent plant occurring in the coast of Tropical Africa, mangrove type vegetation.
Temperate Europe: Temperate broadleaf and mixed forests, coniferous forests, temperate grasslands, savannas and shrubs, boreal forests & tundra ⁴⁸ .	Temperate	Parent plant occurring in temperate Europe.
Palaearctic Asia & Indo-Malay: Temperate broadleaf and mixed forests, coniferous forests, temperate grasslands, savannas and shrublands, boreal forests, tundra, asian desert and xeric shrublands, montane grasslands and shrublands, tropical and subtropical moist and dry broadleaf forests ⁴⁸ .	Asia	Parent plant occurring in Asia.
Arabian Peninsula: Arabian desert and xeric shrublands, Arabian xeric grasslands, savannas and shrubs ⁴⁸ .	Arabia	Parent plant occurring in the Arabian Peninsula.
Neartic and Neotropic: from Tropical and Subtropical forests to Boreal forests and tundra ⁴⁸ .	America	Parent plant occurring in North America and/or South America.
Australasia: from Mediterranean forests, woodlands and scrubs to deserts and tropical and subtropical forests ⁴⁸ .	Australasia	Parent plant occurring in Australasia.
Canary Islands dry woodlands and forests & succulent thickets ⁴⁸ .	Macaronesia	Parent plant occurring in the Macaronesia.
N/A	Cosmopolitan	Parent plant occurring in every continent.
N/A	Exotic/Introduced	Parent plant native to America, Australia or eastern Asia and has been introduced to Africa and/or Europe.
N/A	Widespread	Parent plant occurring in every continent but not present in the Sahara.

Table 2. Phylogeographical affinity categories.

PFT	Data type	Source	Description
Dispersal syndrome	Categorical: biotic/abiotic	BIEN database ^{31,72}	Dispersal of seeds, fruits or spores by animals (zoochory), by floating in water, or wind (abiotic) ¹⁷ .
Leaf type	Categorical: broad-, needle- or scale- leaved	TRY database ^{73–87}	Overall shape and structure of leaves. Broad-leaved species adopt more active water-use strategies, while needle-leaved species often exhibit more conservative strategies that limit photosynthesis but increases resilience under extreme conditions ⁸⁸ .
Vegetative phenology	Categorical: evergreen/ deciduous	BIEN database ^{31,89}	Vegetative phenology distinguishes evergreen from deciduous leaf habits. Evergreen leaves are long-lived and require fewer nutrients, whereas deciduous leaves are shed to reduce water loss during dry periods, lowering maintenance costs and enabling higher carbon gain when conditions improve ^{90,91} .
Plant height	Continuous (m)	BIEN database ^{31,87,92–145}	Plant height is the vertical distance from the ground to the top of the photosynthetic tissue ¹⁷ . Greater height promotes seed dispersal and light absorption, but increases construction and maintenance costs ^{16,146} , thereby shaping competitive ability ¹⁴⁷ .
Seed mass	Continuous (mg)	BIEN database ^{31,103,104,107,111,114–116,130,132,148–164}	Seed mass influences dispersal, colonisation and establishment ^{146,147} . Lighter seeds are produced in greater numbers and disperse further, whereas, heavier seeds are more beneficial for surviving the early stages of recruitment ¹⁶ .
Leaf area	Continuous (mm ²)	BIEN database ^{31,93,97,99,104,106,108,111,114,115,125–128,134,156,165–183}	Leaf area influences photosynthesis, water balance and resource investment ^{146,147,184} . Larger leaves capture more light, increasing photosynthetic capacity but with higher construction costs ¹⁶ .
Habit categories (growth form)	Categorical: tree, shrub, or herb.	Literature and reference regional floras	Habit categories compiled manually for this database; used for pollen percentage calculations.
Whole plant growth form diversity	Categorical: tree, shrub, herb or aquatic.	BIEN database ^{31,185}	Plant growth form categories are a classification of plants according to their morphology, mechanical architecture and development ⁴³ .
Leaf dry mass per area (LMA)	Continuous (g/mm ²)	BIEN database (calculated from leaf dry mass and leaf area data)	Leaf dry mass per area (LMA) is the ratio of the dry mass of the leaf to its area ^{17,184} . It is related to resource allocation and leaf economy strategies, and is a key predictor of ecosystem productivity and nutrient cycling ¹⁷ .
Whole plant sexual system	Categorical: hermaphrodite/ dioecious	BIEN database ^{31,72}	The plant sexual system describes how sex is allocated within a species, with hermaphrodites and deciduous plants being the most common ¹⁸⁶ . Most flowering plants are hermaphroditic and have high reproductive success, although they increase the risk of inbreeding depression. Dioecious species have lower reproductive success but avoid inbreeding depression and can live in harsher environments ¹⁸⁷ .
Leaf life span	Continuous (months)	BIEN database ^{31,126,148,174,188–190}	Leaf lifespan is the period during which a leaf remains physiologically active. It reflects nutrient-use strategies, decomposition rates, and palatability, and influences both ecosystem productivity and the hydrological cycle. Short-lived leaves allocate nutrients towards high rates of photosynthesis, while longer-lived leaves tend to invest nutrients in leaf protection ¹⁷ .
Leaf nitrogen content per leaf dry mass	Continuous (mg/g)	BIEN database ^{31,108,111,119,120,126,127,137,148,150,167,168,170–175,178,180,183,188–210}	Leaf nitrogen concentration is calculated by dividing the total amount of nitrogen in the leaf by its dry mass. This trait is important for protein and nucleotide synthesis ¹⁷ .
Longest whole plant longevity	Continuous (years)	BIEN database ³¹	This trait reflects the environmental adaptation of plant communities, and how they respond to biotic and abiotic stresses ¹⁷ . Short-lived plants have rapid growth, while long-lived plants have slower growth and live in harsh conditions with low resources ¹⁷ .
Flower pollination syndrome	Categorical: biotic/abiotic	BIEN database ^{31,72}	Floral pollination syndrome is the set of floral traits that are linked to a specific functional group of pollinators ^{211,212} .

Table 3. Plant functional traits compiled for this database. Traits were extracted from the BIEN and TRY databases, with the original sources acknowledged alongside each trait. In addition, habit categories were compiled from the literature and regional floras specifically for this database.

12 categorical and continuous traits were obtained from the BIEN database²¹ using the *Bien* R package v. 1.2.7³¹. Since the leaf dry mass per area (LMA) trait is not directly available in the BIEN database, but its components (leaf dry mass and leaf area) are, we calculated LMA by dividing leaf dry mass by leaf area for each taxon that had data for both traits. The categorical trait “leaf type” was obtained directly from the TRY database²² website (<https://www.try-db.org/>).

Calculation of percentages in harmonised pollen records. Pollen types in the original datasets were replaced by the harmonised pollen types. Raw pollen counts from the harmonised pollen records were transformed into percentages to remove the effect of sampling and counting effort⁴⁹ on the basis of the terrestrial pollen sum, including trees, shrubs, and herbs (Table 1, column b). Next, we calculated the total sum, which includes all habit types—terrestrial, parasitic, aquatic, algae, bryophytes, ferns, fungi, and indeterminate or unknown taxa (Table 1, column b).

Percentages were then calculated in relation to the pollen sum, except for indeterminate or unknown pollen types, which were calculated relative to the total pollen sum (all palynomorphs included). The harmonised pollen records with percentage values were subsequently archived in the processed data folder providing a valuable resource for the examination of site-to-site vegetation dynamics.

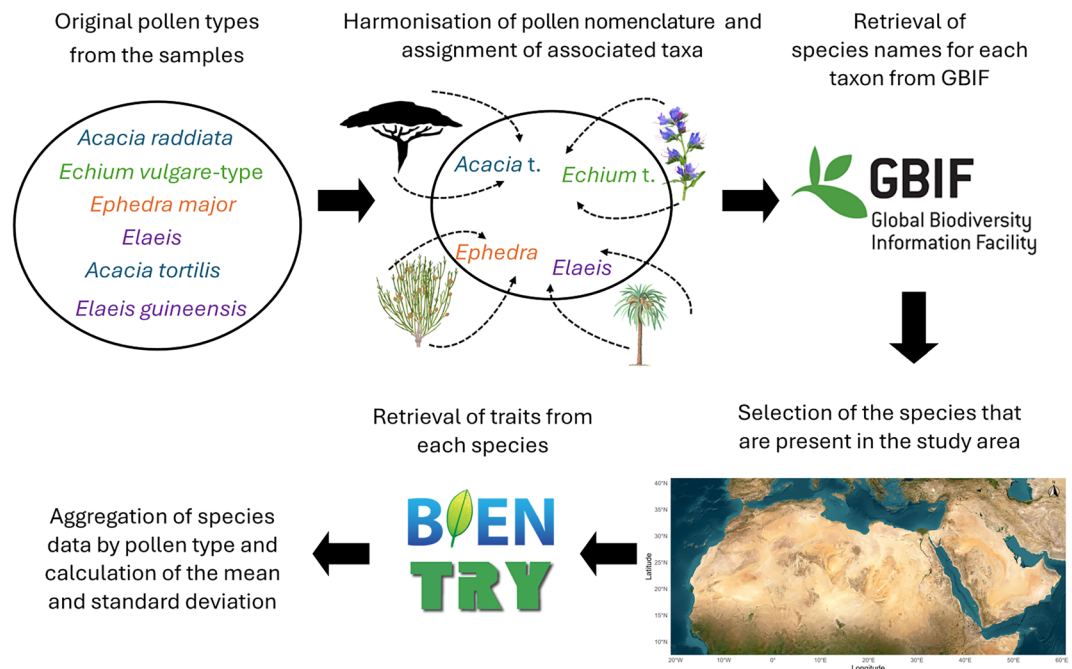


Fig. 4 Workflow for assigning plant species to pollen types and processing plant functional traits.

Addition of calibrated dates to harmonised pollen records. To ensure temporal consistency across the dataset, only the median recalibrated dates were extracted and associated with their corresponding harmonised pollen records. Consequently, each harmonised pollen record presented the median calibrated age for each depth, alongside its harmonised pollen taxa. This allows for consistent chronological comparison across sites.

Data Records

The NAMPHORA database is available via Zenodo²⁴. It includes pollen records, chronological data in raw radio-carbon dates and calibrated calendar years, a pollen taxonomic harmonisation list, habit and phytogeographical categorisation, and plant functional traits retrieved from TRY and BIEN databases, for the harmonised pollen taxonomical entities. All files, metadata, data outputs and R scripts are stored under a CC-BY license.

All datasets have associated references, which can be found alongside the processed data files in a reference column. Moreover, a complete list of references associated with each dataset is provided in the Supplementary Information file, together with an acknowledgment of all data contributors.

In addition, the PDF documents of the original studies from which the data were extracted (e.g. pollen records, chronological data and plant functional traits), as well as the references used for taxonomic harmonisation, have been compiled and made available for download in the Zenodo repository²⁴.

The database includes the following data entries:

Pollen records. Pollen data is structured into three separate main folders: raw data, processed data and metadata.

1. **Raw data:** Contains individual raw pollen records (in csv format) with original taxa names, separated into fossil and modern pollen records.
2. **Processed data:** Contains individual harmonised pollen records (in csv format), separated into fossil and modern pollen records, which are further divided into:
 - a. **Harmonised counts:** In each pollen record, the original pollen type names have been transformed into harmonised pollen type names, with counts summed according to taxa regrouping.
 - b. **Harmonised percentages:** In each pollen record, the original pollen type names have also been transformed into harmonised pollen type names, with counts converted to percentages.
3. **Metadata:** A list with geographical, temporal and other relevant information of all the pollen records has been compiled in the “database.csv” file. Specific metadata details are available in the “pollen_metadata.csv” and “pollen_metadata.html” files. The metadata includes several key information:

- a. **Site location**, records the site name (both original and modified name for machine readability), country, biogeographic area, latitude and longitude in decimal degrees, and altitude in meters.
- b. **Data source**, details whether pollen records were retrieved from Neotoma, the APD, or directly from authors if were not publicly available. The *Link to database* field provides a direct link to the data repository where each site was retrieved. The *dataset_id* field stores the Neotoma dataset identifier for downloading pollen records via the *neotoma2* R package²⁰. Finally, the *Database* field indicates whether a site is available in a public data repository or was obtained directly from the authors.
- c. **Record type** specifies whether the pollen record is modern or fossil.
- d. **Archive type** indicates the type of core, such as terrestrial, lacustrine, archaeological, etc.
- e. **Chronological information** includes details on whether a site has an associated chronology (*Dated*) and whether raw radiocarbon dates have been recalibrated using updated calibration curves (*Recalibrated*). Sites are classified by the geological time period, including the Late Holocene (Meghalayan: 0 to 4.2 ka BP), Mid Holocene (Northgrippian: 4.2 to 8.2 ka BP), Early Holocene (Greenlandian: 8.2 to 11.7 ka BP), Holocene (0 to 11.7 ka BP), Holocene-Pre-Holocene (present to 240 ka BP) and Pre-Holocene (11.7 to 240 ka BP). Additionally, for each dated fossil site, the youngest age (*Minimum mean age*), the oldest age (*Maximum mean age*) and whether the site coincides fully or partially with the African Humid Period (14.8 to 5.5 ka BP) are detailed.
- f. **Bibliographic references** include the original studies from which the data originate.

Chronologies. Chronological data is structured into three separate main folders: raw data, processed data and metadata.

1. **Raw data:** Contains radiocarbon dates for each dated fossil pollen record (in csv format) in the radiocarbon folder, and calibration information for each dated fossil pollen record in the calibration folder, which is further divided into:
 - a. **Bacon runs:** Includes files in csv format for those records that have been calibrated using a Bayesian age-depth model. The files are formatted so they can be calibrated with the *rbacon* v. 3.2.0 package³⁶.
 - b. **Single depth calibration:** Includes the plot of individual calibrated dates in pdf format for each fossil pollen record that has been calibrated with the *rice* v. 1.0.0 package³⁷.
2. **Processed data:** Contains the updated calibrated dates for all recalibrated pollen records.
3. **Metadata:** A list with details on the calibration information for each site has been compiled in the “*calibration_information.csv*” file. Details on specific metadata information are available in the “*calibration_metadata.csv*” and “*calibration_metadata.html*” files. The metadata documents information on the:
 - a. **Site location** records the site name (both original and modified name for machine readability), country and biogeographic area.
 - b. **Chronological information** includes the following fields:
 - i. *Dated*: Whether the site contains radiocarbon dates.
 - ii. *Recalibrated*: Whether the raw radiocarbon dates have been recalibrated with updated calibration curves.
 - iii. *Depths*: Whether the record reports depths along with the radiocarbon dates.
 - iv. *Material dated*: Material used for radiocarbon dating (e.g., wood, bulk sediment, etc.).
 - v. *Dating method*: Method used for radiocarbon dating (conventional or AMS).
 - vi. *Calibration curve*: Calibration curve used in the calibration process (IntCal20³⁴ or Marine20³⁵).
 - vii. *Hiatus and reservoir effect*: Whether there was a hiatus or an effect of reservoir on radiocarbon dating.
 - viii. *Post-bomb curve*: Whether there was a post-bomb date and the corresponding post-bomb curve³⁸ (NH1, NH2 or NH3) used in the calibration process.
 - ix. *Accumulation rate (yr/cm)*: Rate of accumulation in years per cm used in the calibration.
 - c. **Bibliographic references** include the original studies from which the data originate.

Taxonomy. Taxonomical data is also separated into three main folders: raw data, processed data and metadata.

1. **Raw data:** Contains three subfolders:
 - a. **Pollen taxonomy harmonisation lists:** Includes published harmonisation tables^{12,33} (in csv and xlsx formats) for cross-reference, and the APD taxonomy list retrieved directly from the APD website (<https://africanpollendatabase.ipsl.fr/#/taxon-dict>).
 - b. **Raw taxa list:** Contains the list of raw taxon names (in csv format) extracted from all pollen records (modern and fossil).
 - c. **GBIF species list:** Contains the list of plant species filtered by the study area, directly obtained from the GBIF website⁵⁰.

1. **Processed data:** Contains the harmonised taxonomy list, the phytogeographical affinity list, the habit list, the list of taxa (family, genus or at species level) corresponding to the pollen types, the list of species names corresponding to the taxa, and the final list of filtered species associated to the pollen types (all in csv format). The harmonised pollen taxonomy list includes the original pollen type names and the consequent taxonomic harmonisation. The phytogeographical affinity and habit lists include the phytogeographical affinities and habits for each harmonised pollen taxon respectively. The list of taxa corresponding to pollen types includes all plants associated with each pollen type, either at family, genus or species level, depending on the resolution of the pollen type. The list of species names includes all the unfiltered species corresponding to each taxon associated to each pollen type. Finally, the species list includes all the filtered species corresponding to the pollen types that are present in the study area.
2. **Metadata:** Specific metadata information for the harmonised pollen taxonomy list is available in the “*taxonomy_metadata.csv*” and “*taxonomy_metadata.html*” files. In addition, there is a file (“*phytogeographical_affinity_categories.csv*”) that provides a definition for each phytogeographic affinity category (Table 2) and the corresponding references. The harmonised pollen taxonomy list metadata documents information on:
 - a. **Original pollen taxon** reports the original pollen taxonomic names given by the pollen analysts to the grains they identified, these are taken from the original, raw pollen records.
 - b. **Taxonomical information** includes an assignation of the pollen taxon to a plant taxonomical entity at the order, family, genus and or species level.
 - c. **Harmonised pollen taxon** refers to the standardised name assigned to each pollen taxon, based on our literature-based interpretation of palynotaxonomy.
 - d. **Number of pollen records** (dated fossil records, undated fossil records, modern records and total records) in which the original pollen type is present.
 - e. **Number of pollen records** (dated fossil records, undated fossil records, modern records and total records) in which the harmonised pollen taxon is present.
 - f. **Pollen type harmonisations** proposals from Lézine *et al.*¹², Mottl *et al.*³³ and from the APD (<https://africanpollendatabase.ipsl.fr/#/taxon-dict>).
 - g. **Pollen taxonomical references** include the original studies used for taxonomic harmonisation and for habit and phytogeographical affinity categorisations.

Plant functional traits. Plant functional trait data is structured into raw data, processed data and metadata folders.

1. **Raw data:** Contains the leaf type trait information retrieved from the TRY database (“*leaf_type_TRY.zip*”).
2. **Processed data:** Contains the final processed plant functional traits data from both databases (BIEN and TRY) in the file “*total_pfts.csv*”.
3. **Metadata:** Specific metadata information for the final plant functional traits list is available in the “*pfts_metadata.csv*” and “*pfts_metadata.html*” files. It also contains a file (“*pfts_categories.csv*”) that provides a definition for each plant functional trait and the corresponding reference. The metadata documents information about the:
 - a. **Taxonomical information:** Family, genus and species name of the plant taxa corresponding to the harmonised pollen taxon.
 - b. **Harmonised pollen taxon:** Morphological harmonisation of the pollen taxon.
 - c. **Url source:** URL linking to the original data source of the measured trait.
 - d. **Project Pi information:** Principal investigator of the original project in which the trait was measured, and its contact information.
 - e. **The plant functional traits** for each pollen taxa.
 - f. **Bibliographic citations** include the original studies where the trait leaf-type has been retrieved (from the TRY database).

Data Overview

The database includes a total of 836 pollen records (Table 4, Fig. 5), which were sourced from the APD, Neotoma and individual researchers (Fig. 6). While the dataset covers a broad geographical extent, pollen records are unevenly distributed, with the Mediterranean region being by far the most studied (Fig. 7). In contrast, the Sahara, Sahel and Sudanian zones, contain relatively fewer sites, highlighting the gaps in pollen record collection in these regions. This scarcity is largely due to challenges in physical access, political and economic instability, and the limited availability of palaeoecological archives in extreme environments such as deserts^{19,51}.

Most of the fossil pollen records relied on radiocarbon dating, which were calibrated to calendar years before present (BP). From the total of 429 fossil pollen records compiled in the database, 320 included chronological data (Table 4). A total of 277 records were recalibrated using updated calibration curves^{34,35} (Table 5). 234 records were recalibrated using a Bayesian age-depth model with the *rbacon* v. 3.2.0 package³⁶ and 43 records were recalibrated with the *rice* v. 1.0.0 package³⁷, due to insufficient or missing depth information. The remaining 43 records could not be updated due to missing radiocarbon dates, lack of associated errors, or because their chronologies were based on tephra or biostratigraphic timescales, and thus calibration is not needed.

Biogeographic region	Total records compiled		Fossil records with chronological data
	Modern	Fossil	
Arabian Peninsula	74	51	39
Mediterranean	109	232	188
Sahara	49	76	41
Sahel	70	25	22
Sudanian	105	45	30
Total	407	429	320

Table 4. Pollen records summary table.

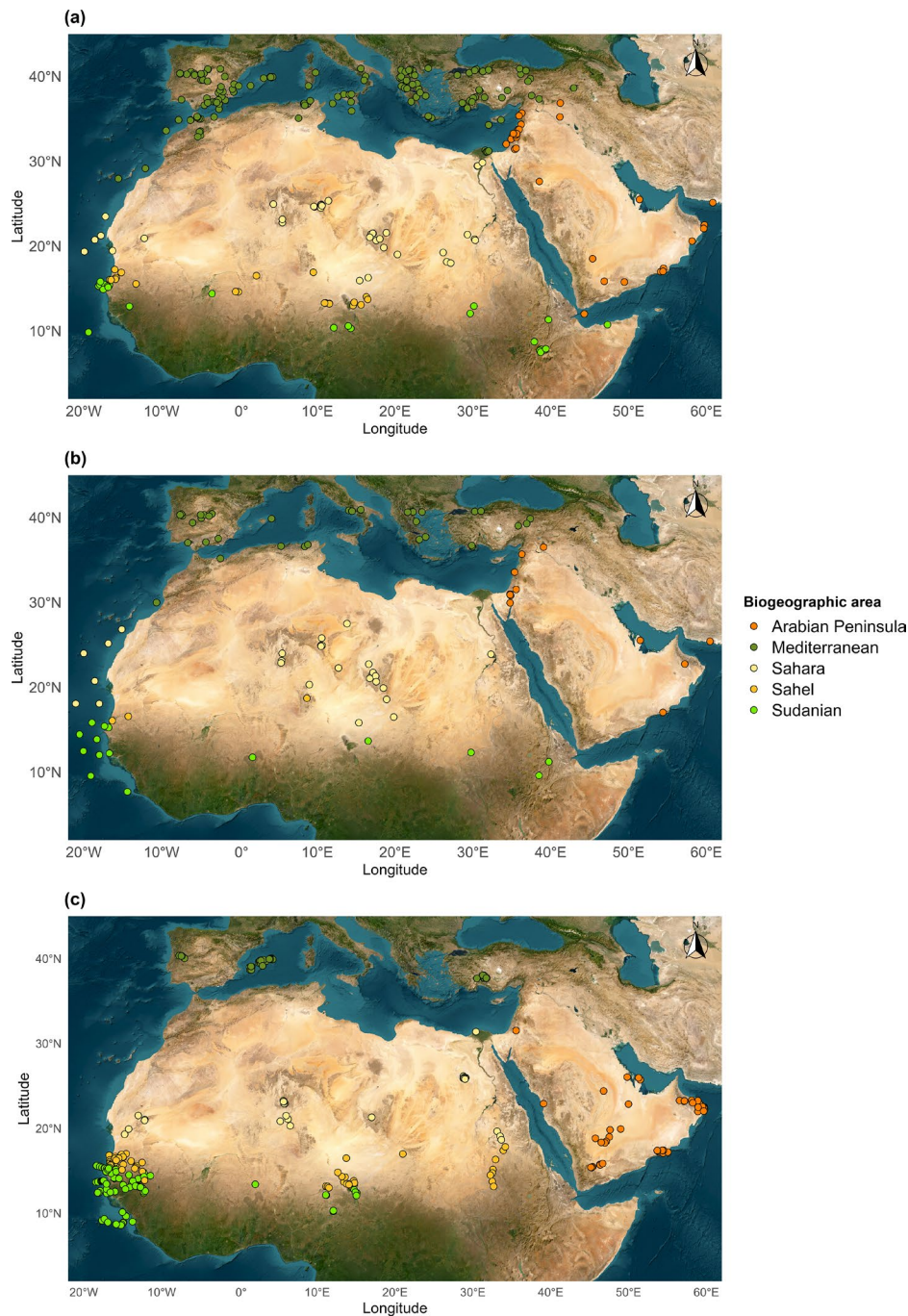


Fig. 5 The NAMPHORA database: dated (a) and undated (b) fossil pollen records, and modern pollen records (c). For a more detailed inspection of the pollen records see Fig. 1 in database website (https://irene131998.github.io/NAMPHORA_database/).

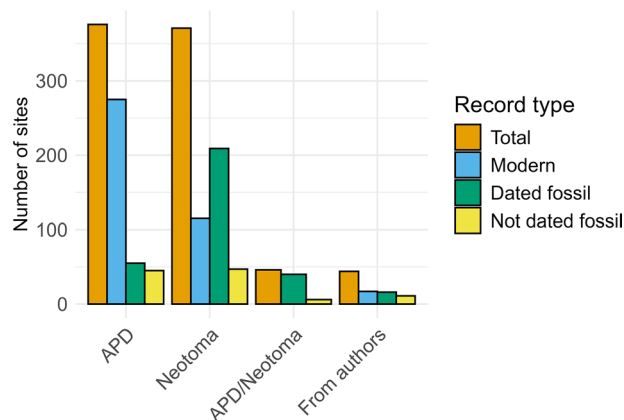


Fig. 6 Pollen records according to source database (APD, Neotoma, present in both databases (APD/Neotoma), and received directly from the authors).

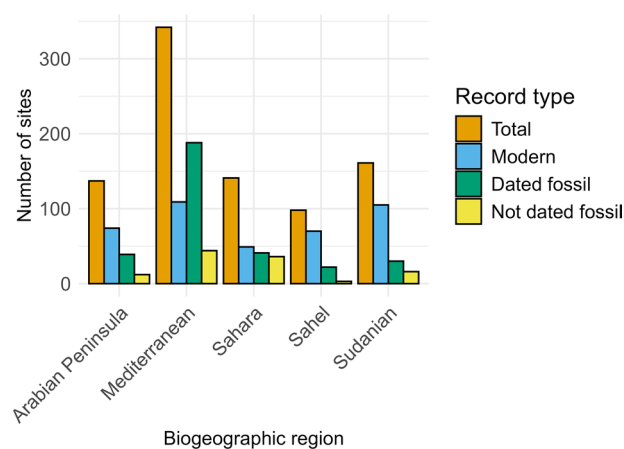


Fig. 7 Pollen records according to biogeographic region (Arabian Peninsula, Atlantic Ocean, Mediterranean, Sahara, Sahel and Sudanian regions).

Biogeographic region	Fossil records dated but not recalibrated		Re-calibrated with Rbacon v.3.2.0 ³⁶	Re-calibrated with rice v.1.0.0 ³⁷	Total number of recalibrated records (Rbacon ³⁶ & rice ³⁷)
	Could not be recalibrated	Calibration not necessary			
Arabian Peninsula	1	5	29	4	33
Mediterranean	8	19	147	14	161
Sahara	3	1	19	18	37
Sahel	2	1	14	5	19
Sudanian	2	1	25	2	27
Total	16	27	234	43	277

Table 5. Fossil pollen records chronological information.

As shown in Fig. 8, most dated fossil records belong to the Late Holocene, with some extending into the Pleistocene. Although many records, especially from the Sahara, are from the African Humid Period (AHP: 14,800 to 5,500 BP), they are still less abundant than more recent records, which highlights the need to obtain more Early and Mid-Holocene records to refine our understanding of the AHP.

After harmonisation of the pollen nomenclature, the initial dataset comprising 5,614 pollen types was reduced to 853 harmonised pollen types. After processing the species corresponding to each pollen type, we obtained a final list of 44,970 species (from 202 families and 1,103 genera) assigned to pollen types and present in the study area. These comprise 546 terrestrial pollen types (trees, shrubs, and herbs), 32 aquatic pollen types, and 148 pollen types classified as indeterminate or belonging to other groups (e.g., bryophytes, ferns, algae, fungi) (Table 6).

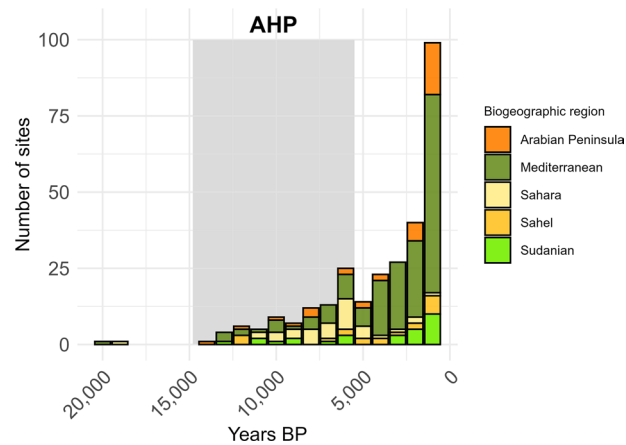


Fig. 8 Dated pollen records in 1000-year intervals. The shaded area represents the African Humid Period (AHP: 14,800 to 5,500 years BP). Dates older than 20,000 were omitted, as they were only represented by few records.

Habit (a)	Number or harmonised taxa
Algae	38
Anthocerotophyta	1
Aquatic	32
Bryophyta	4
Climbers	10
Climbing herbs	2
Climbing shrubs	2
Fern	49
Fungi	40
Geophytes	5
Herbs	110
Herbs or Climbers	3
Herbs or Shrubs	79
Herbs or Shrubs or Climbers	7
Herbs or Shrubs or Lianas	3
Herbs or Shrubs or Trees	61
Herbs or Shrubs or Trees or Climbers	11
Herbs or Shrubs or Trees or Lianas	15
Lianas	6
Mangrove	2
Marchantiophyta	2
Palm	8
Parasitic	11
Selaginellales	1
Shrubs	19
Shrubs or Climbers	6
Shrubs or Lianas	4
Shrubs or Trees	178
Shrubs or Trees or Climbers	15
Shrubs or Trees or Epiphytic	2

Table 6. Number of pollen taxa per habit category.

Illustrative examples of database use. We present two examples to introduce users to the NAMPHORA database and illustrate its potential applications, ranging from a simple use case (Example A) to a more detailed analysis incorporating trends in plant functional traits (Example B).

Example A. A researcher interested in the Holocene distribution of *Quercus evergreen* pollen in the Sahara region may extract all dated fossil pollen records for this taxon and visualise temporal changes in pollen abundance (expressed as pollen percentages), providing an overview of its temporal dynamics in the region (Fig. 9).

Illustrative example of dataset use for a specific pollen taxon

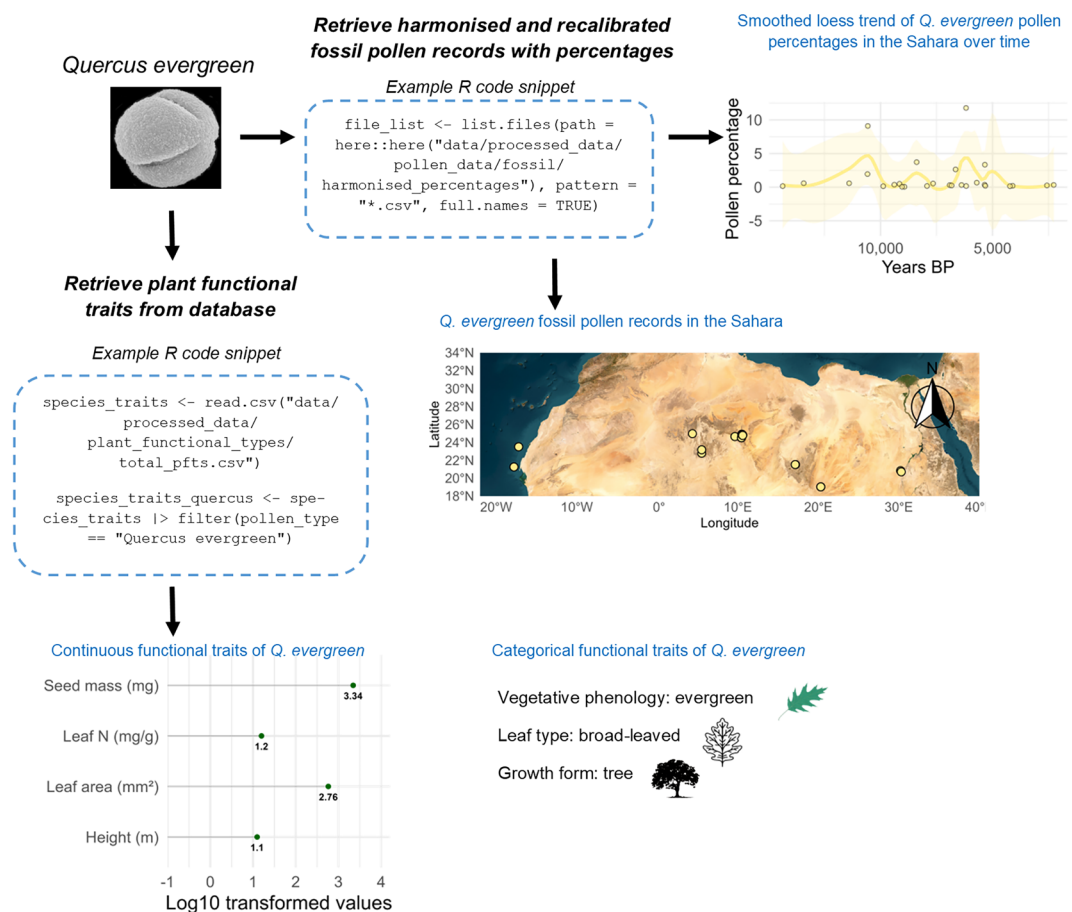


Fig. 9 Illustrative example of dataset use for a specific pollen taxon. In this example, a researcher interested in the Holocene distribution of *Quercus evergreen* pollen in the Sahara extracts harmonised and recalibrated fossil pollen records from the NAMPHORA database, selects the sites where *Q. evergreen* is present, and retrieves the corresponding pollen percentages and sample ages. These data are then used to visualise temporal changes in *Q. evergreen* pollen abundance using a LOESS smoothed trend and to map the spatial distribution of fossil occurrences, providing an initial overview of its past abundance and distribution. In addition, the database allows retrieval of plant functional traits associated with the taxon, including both continuous and categorical traits, enabling a basic ecological interpretation of the pollen record.

The researcher may also retrieve information on the plant functional traits associated with this taxon to complement the interpretation of the pollen record.

Example B. In a more advanced application, a research group investigates changes in vegetation strategies recorded in the pollen sequence of Lake Yoa (Chad)^{52,53}, which is available in NAMPHORA database and has associated trait data. The researchers link the identified pollen taxa to selected plant functional traits (e.g., mean plant height and growth forms) and calculate the community weighted mean (CWM) for each trait⁵⁴ using the *functcomp* function from *FD* R package⁵⁵. This provides the mean trait value of the fossil pollen assemblage weighted by the relative abundance of each pollen taxon. The researchers then model the CWM through time using generalised additive models (GAM) to analyse how dominant ecological strategies changed throughout the Holocene. This approach enables the inference of shifts in vegetation structure and ecological strategies in response to climatic changes associated with the AHP (Fig. 10).

However, the linkage between pollen types and plant functional traits is often incomplete or uncertain (Fig. 10). This uncertainty may arise not only from missing trait information but also because a single pollen taxon can represent species with different functional strategies⁵⁶. For example, a pollen type may include both woody and herbaceous species. In such cases, assigning a unique plant functional trait to the pollen type may oversimplify its ecological interpretation. Researchers can address this issue in several ways: (i) restricting the analysis to taxa with well-defined trait assignments, excluding taxa associated with multiple categories for the same trait (as in this example), (ii) assigning traits based on the most likely ecological affinity using additional ecological or biogeographical information, or (iii) exploring multiple scenarios by testing alternative trait assignments for ambiguous taxa (e.g., one scenario treating the taxon as herbaceous, another as woody). Comparing

Illustrative example of dataset use for a specific pollen record

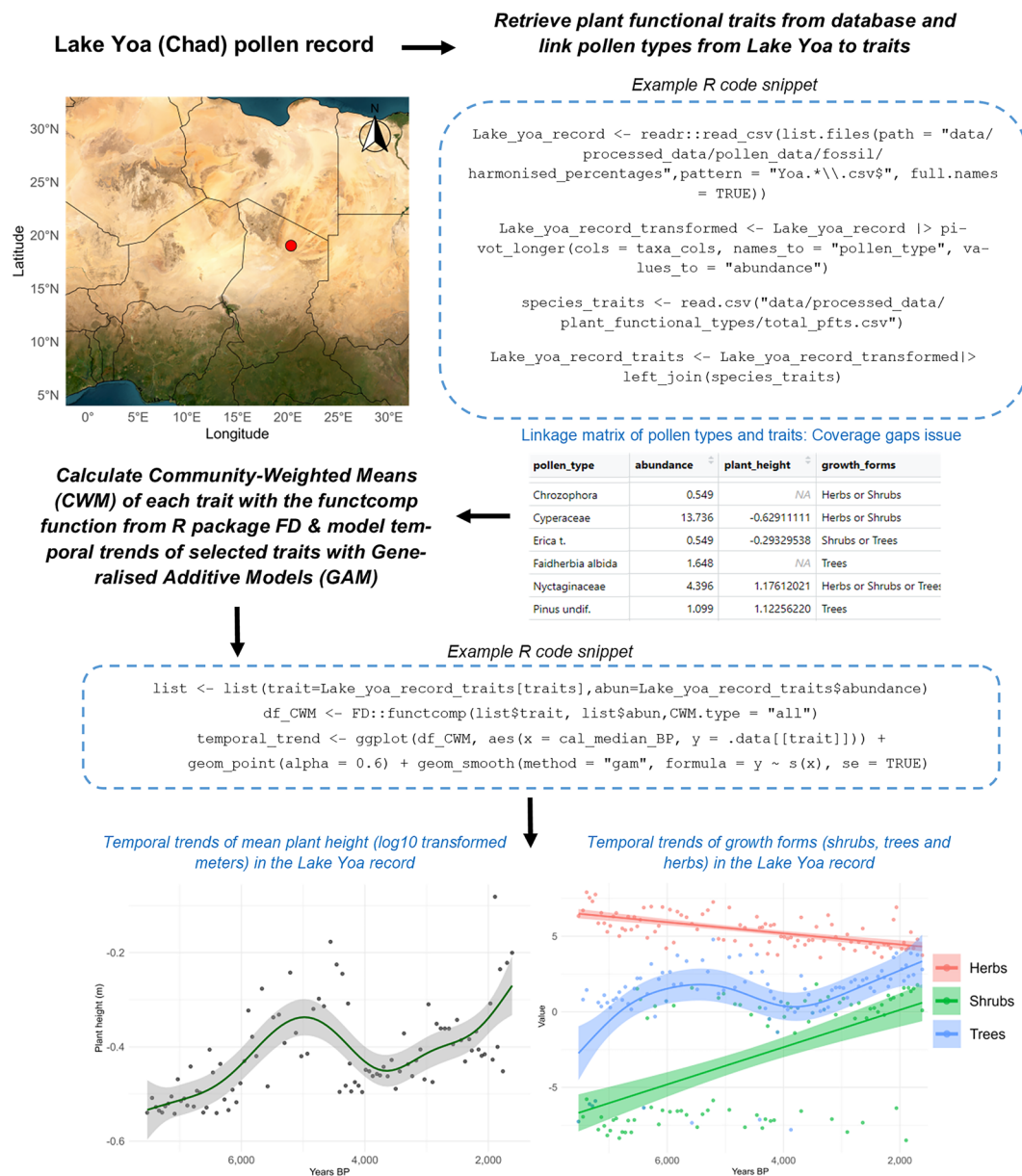


Fig. 10 Illustrative example of dataset use for a specific pollen record. In this example, researchers aim to reconstruct temporal trends in plant functional traits in the Lake Yoa pollen record (Chad). They use the harmonised and recalibrated Lake Yoa record from the NAMPHORA database and link the identified pollen types to selected plant functional traits (growth forms and mean plant height). They then calculate the community weighted mean (CWM) of each trait for every pollen sample and apply generalised additive models (GAMs) to estimate temporal trends in these traits, excluding taxa with intermediate growth forms from the analysis. This approach provides a way to estimate changes in ecological strategies through time in a pollen record, complementing traditional vegetation reconstructions.

the trends resulting from these different approaches allows researchers to evaluate the robustness of inferring functional changes through time while accounting for this source of uncertainty.

The workflow underlying these examples is fully documented in the “11_illustrative_examples_database_use.R” script, which is available in the Zenodo repository²⁴.

Technical Validation

The pollen records used in this study were sourced from multiple databases, and to ensure consistency, all files were standardised to conform to the structure of the APD. Each record was carefully reviewed, and any discrepancies were cross-checked against the original sources to confirm accuracy. Data processing was carried

out using R v. 4.1.1²⁵, where the datasets were processed and standardised. Geographic coordinates underwent additional quality control through map projection (Fig. 5).

For pollen taxonomic harmonisation, a comprehensive literature review was conducted, and the harmonised pollen list was cross-referenced with established nomenclature from key studies (Lézine *et al.*¹², Mottl *et al.*³³ and the APD dictionary) to ensure robust taxonomic consistency.

Radiocarbon dates were reviewed for reliability, and when necessary, the original sources were consulted to obtain more accurate chronological data.

The phytogeographical affinity and growth habit of each taxon were assigned based on current distribution data from authoritative sources, including GBIF, POWO, and relevant literature (monographs, research articles, theses, and books).

Usage Notes

The project structure and data pipeline are thoroughly explained in the database website (https://irene131998.github.io/NAMPHORA_database/). Additionally, the R code used for data preparation and processing is fully available and reproducible.

This database has several potential applications, including reconstructing past vegetation and climate and serving as validation for paleoclimatic simulations^{2,57}. Contributions and error reports are highly encouraged and can be submitted directly to the corresponding author or through the issues or pull requests sections in the associated GitHub repository (https://github.com/Irene131998/NAMPHORA_database).

Future efforts will focus on expanding the database as new records become available, and the possibility of interconnecting NAMPHORA with the Neotoma database will be explored with data stewards.

Data availability

All raw and processed data are available for download as CSV files from Zenodo repository²⁴.

Code availability

The code used to extract, clean, and process the data, as well as to generate each figure, and map in this manuscript, is written in R programming language. It is openly accessible via Zenodo²⁴.

Received: 19 December 2025; Accepted: 22 April 2026;

Published online: 06 May 2026

References

1. Pausata, F. S. R. *et al.* The Greening of the Sahara: Past Changes and Future Implications. *One Earth* **2**, 235–250 (2020).
2. Dallmeyer, A., Claussen, M., Lorenz, S. J. & Shanahan, T. The end of the African humid period as seen by a transient comprehensive Earth system model simulation of the last 8000 years. *Clim. Past* **16**, 117–140 (2020).
3. Phelps, L. N. *et al.* Asymmetric response of forest and grassy biomes to climate variability across the African Humid Period: influenced by anthropogenic disturbance? *Ecography* **43**, 1118–1142 (2020).
4. Cheddadi, R. *et al.* Early Holocene greening of the Sahara requires Mediterranean winter rainfall. *Proc. Natl. Acad. Sci. USA*. **118**, e2024898118 (2021).
5. Houérou, H. N. L. Outline of the biological history of the Sahara. *Journal of Arid Environments* **22**, 3–30 (1992).
6. Watrin, J., Lézine, A.-M. & Hély, C. Plant migration and plant communities at the time of the “green Sahara”. *Comptes Rendus Geoscience* **341**, 656–670 (2009).
7. Hély, C., Lézine, A.-M. & Contributors, A. Holocene changes in African vegetation: tradeoff between climate and water availability. *Clim. Past* **10**, 681–686 (2014).
8. Wright, D. K. Humans as Agents in the Termination of the African Humid Period. *Front. Earth Sci.* **5** (2017).
9. Migliore, J. *et al.* Surviving in Mountain Climate Refugia: New Insights from the Genetic Diversity and Structure of the Relict Shrub *Myrtus nivellei* (Myrtaceae) in the Sahara Desert. *PLoS ONE* **8**, e73795 (2013).
10. Médail, F. & Quézel, P. *Biogéographie de La Flore Du Sahara: Une Biodiversité En Situation Extrême*. (IRD Éditions, 2018).
11. Bhatta, K. P. *et al.* Exploring spatio-temporal patterns of palynological changes in Asia during the Holocene. *Front. Ecol. Evol.* **11**, 1115784 (2023).
12. Lézine, A.-M., Watrin, J., Vincens, A. & Hély, C. Are modern pollen data representative of west African vegetation? *Review of Palaeobotany and Palynology* **156**, 265–276 (2009).
13. Davis, B. A. S. *et al.* The European Modern Pollen Database (EMPD) project. *Veget. Hist. Archaeobot.* **22**, 521–530 (2013).
14. Williams, J. W. & Jackson, S. T. Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment* **5**, 475–482 (2007).
15. Gillison, A. N. 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 (2019).
16. Brussel, T. & Brewer, S. C. Functional Paleocology and the Pollen-Plant Functional Trait Linkage. *Front. Ecol. Evol.* **8**, 564609 (2021).
17. McElwain, J. C. *et al.* Functional traits of fossil plants. *New Phytologist* **242**, 392–423 (2024).
18. Williams, J. W. *et al.* The Neotoma Paleocology Database, a multiproxy, international, community-curated data resource. *Quat. res.* **89**, 156–177 (2018).
19. Lézine, A. M., Ivory, S. J., Gosling, W. D. & Scott, L. The African Pollen Database (APD) and tracing environmental change: State of the Art. (2021).
20. Socorro, D. & Goring, S. Working with the Neotoma Paleocology Database. (2023).
21. Enquist, B. J. *et al.* BIEN: A biodiversity informatics ecosystem advancing open and reproducible workflows for plant observation, plot and trait data. *Methods Ecol. Evol.* <https://doi.org/10.1111/2041-210X.70274> (2026).
22. Kattge, J. *et al.* TRY plant trait database – enhanced coverage and open access. *Global Change Biology* **26**, 119–188 (2020).
23. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
24. Solano, I., Bro-Jørgensen, J., Lazagabaster, I. A., Thomas, C. D. & Manzano, S. NAMPHORA database. *Zenodo* <https://doi.org/10.5281/zenodo.17914358> (2025).
25. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. (2025).
26. Fady, B. *et al.* Native Trees of the Mediterranean Region: Distribution, Diversity and Conservation Challenges. *Curr. For. Rep.* **11**, 20 (2025).

27. Blasi, C. La Vegetazione d'Italia con Carta delle Serie di Vegetazione in scala 1 : 500000. *Palombi Editori* (2010).
28. Costa, M., Morla, C. & Sainz, L. Los bosques ibéricos - una interpretación geobotánica. 497–505 (1998).
29. Fyfe, R. M. *et al.* The European Pollen Database: past efforts and current activities. *Veget Hist Archaeobot* **18**, 417–424 (2009).
30. Giesecke, T. *et al.* Towards mapping the late Quaternary vegetation change of Europe. *Veget Hist Archaeobot* **23**, 75–86 (2014).
31. Maitner, B. *et al.* The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods in Ecology and Evolution* **9**, 373–379 (2017).
32. Joosten, H. & de Klerk, P. What's in a name? Some thoughts on pollen classification, identification, and nomenclature in Quaternary palynology. *Review of Palaeobotany and Palynology* (2002).
33. Mottl, O. *et al.* Global acceleration in rates of vegetation change over the past 18,000 years. *Science* **372**, 860–864 (2021).
34. Reimer, P. J. *et al.* The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).
35. Heaton, T. J. *et al.* Marine20—The Marine Radiocarbon Age Calibration Curve (0–55,000 cal BP). *Radiocarbon* **62**, 779–820 (2020).
36. Blaauw, M., Christen, J. A. & Aquino Lopez, M. A. rbacon: Age-Depth Modelling using Bayesian Statistics. 3.3.1 (2017).
37. Blaauw, M. rice: Radiocarbon Equations. <https://CRAN.R-project.org/package=rice> (2024).
38. Reimer, P. J., Brown, T. A. & Reimer, R. W. Discussion: Reporting and Calibration of Post-Bomb¹⁴ C Data. *Radiocarbon* **46**, 1299–1304 (2004).
39. Chamberlain, S. A. & Boettiger, C. R. Python, and Ruby clients for GBIF species occurrence data. Preprint at <https://doi.org/10.7287/peerj.preprints.3304v1> (2017).
40. GBIF: The Global Biodiversity Information Facility. What is GBIF? Available from <https://www.gbif.org/what-is-gbif> (2025).
41. GBIF Secretariat. GBIF Backbone Taxonomy. Checklist dataset <https://doi.org/10.15468/39omei> (2023).
42. Taylor, A., Weigelt, P., Denelle, P., Cai, L. & Kreft, H. The contribution of plant life and growth forms to global gradients of vascular plant diversity. *New Phytologist* **240**, 1548–1560 (2023).
43. Rowe, N. & Speck, T. Plant growth forms: an ecological and evolutionary perspective. *New Phytologist* **166**, 61–72 (2005).
44. Dos Santos, P. *et al.* Plant growth forms dictate adaptations to the local climate. *Front. Plant Sci.* **13**, 1023595 (2022).
45. Björne, A. E. PALEOBOTANY: Paleophytogeography. in *Encyclopedia of Quaternary Science* 730–733 (Elsevier, 2013).
46. White, F. *The Vegetation of Africa: A Descriptive Memoir to Accompany the Unesco/AETFAT/UNSO Vegetation Map of Africa*. (Unesco, Paris, 1983).
47. Quézel, P. Analysis of the Flora of Mediterranean and Saharan Africa. *Annals of the Missouri Botanical Garden* **65**, 479 (1978).
48. Olson, D. M. *et al.* Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience* **51**, 933 (2001).
49. Correa-Metrio, A., Bush, M. B., Pérez, L., Schwalb, A. & Cabrera, K. R. Pollen distribution along climatic and biogeographic gradients in northern Central America. *The Holocene* **21**, 681–692 (2011).
50. GBIF.org. GBIF Occurrence Download. <https://doi.org/10.15468/dl.66vbhj> (2025).
51. Kaufman, D. *et al.* A global database of Holocene paleotemperature records. *Sci Data* **7**, 115 (2020).
52. Lézine, A.-M., Zheng, W., Braconnot, P. & Krinner, G. Late Holocene plant and climate evolution at Lake Yoa, northern Chad: pollen data and climate simulations. *Clim. Past* **7**, 1351–1362 (2011).
53. Kröpelin, S. *et al.* Climate-Driven Ecosystem Succession in the Sahara: The Past 6000 Years. *Science* **320**, 765–768 (2008).
54. Carvalho, F. *et al.* A method for reconstructing temporal changes in vegetation functional trait composition using Holocene pollen assemblages. *PLoS ONE* **14**, e0216698 (2019).
55. Laliberté, E., Legendre, P. & Shipley, B. FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package version 1.0-12.3. (2014).
56. Veeken, A., Santos, M. J., McGowan, S., Davies, A. L. & Schrödt, F. Pollen-based reconstruction reveals the impact of the onset of agriculture on plant functional trait composition. *Ecology Letters* **25**, 1937–1951 (2022).
57. Hopcroft, P. O., Valdes, P. J. & Ingram, W. Using the Mid-Holocene “Greening” of the Sahara to Narrow Acceptable Ranges on Climate Model Parameters. *Geophysical Research Letters* **48**, e2020GL092043 (2021).
58. Raven, J. A. & Giordano, M. Algae. *Current Biology* **24**, R590–R595 (2014).
59. Frangedakis, E. *et al.* The hornworts: morphology, evolution and development. *New Phytologist* **229**, 735–754 (2021).
60. Oyediji, A. A. & Abowei, J. F. N. The Classification, Distribution, Control and Economic Importance of Aquatic Plants. *International Journal of Fisheries and Aquatic Sciences* (2012).
61. Crandall-Stotler, B. J. & Bartholomew-Began, S. E. Morphology of Mosses (Phylum Bryophyta). *Flora of North America North of Mexico* (2007).
62. Gianoli, E. The behavioural ecology of climbing plants. *AoB PLANTS* **7** (2015).
63. Christenhusz, M. J. M. & Chase, M. W. Trends and concepts in fern classification. *Annals of Botany* **113**, 571–594 (2014).
64. Carris, L. M., Little, C. R. & Stiles, C. M. Introduction to Fungi. *The plant health instructor* (2012).
65. Howard, C. C., Folk, R. A., Beaulieu, J. M. & Cellinese, N. The monocotyledonous underground: global climatic and phylogenetic patterns of geophyte diversity. *American J of Botany* **106**, 850–863 (2019).
66. Simpson, M. G. Plant Systematics. *Academic Press* (2019).
67. Kathiresan, K. & Bingham, B. L. Biology of mangroves and mangrove Ecosystems. in *Advances in Marine Biology* vol. 40 81–251 (Elsevier, 2001).
68. Crandall-Stotler, B., Stotler, R. E. & Long, D. G. Phylogeny and classification of the Marchantiophyta. *Edinburgh J. Bot.* **66**, 155–198 (2009).
69. Tomlinson, P. B. The uniqueness of palms. *Botanical Journal of the Linnean Society* **151**, 5–14 (2006).
70. Těšitel, J. Functional biology of parasitic plants: a review. *Plecevo* **149**, 5–20 (2016).
71. Weststrand, S. & Korall, P. A subgeneric classification of *Selaginella* (Selaginellaceae). *American J of Botany* **103**, 2160–2169 (2016).
72. Bezeng, S. B., Davies, J. T., Yessoufou, K. & Maurin, O. & Van Der Bank, M. Revisiting Darwin's naturalization conundrum: explaining invasion success of non-native trees and shrubs in southern Africa. *Journal of Ecology* **103**, 871–879 (2015).
73. Chen, Y., Han, W., Tang, L., Tang, Z. & Fang, J. Leaf nitrogen and phosphorus concentrations of woody plants differ in responses to climate, soil and plant growth form. *Ecography* **36**, 178–184 (2013).
74. Cornelissen, J. H. C. An Experimental Comparison of Leaf Decomposition Rates in a Wide Range of Temperate Plant Species and Types. *The Journal of Ecology* **84**, 573 (1996).
75. Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B. & Otero-Casal, C. Hydrologic regulation of plant rooting depth. *Proc. Natl. Acad. Sci. USA*. **114**, 10572–10577 (2017).
76. Han, W. *et al.* Floral, climatic and soil pH controls on leaf ash content in China's terrestrial plants. *Global Ecology and Biogeography* **21**, 376–382 (2012).
77. He, T., Pausas, J. G., Belcher, C. M., Schwill, D. W. & Lamont, B. B. Fire-adapted traits of *Pinus* arose in the fiery Cretaceous. *New Phytologist* **194**, 751–759 (2012).
78. Iversen, C. M. *et al.* A global Fine-Root Ecology Database to address below-ground challenges in plant ecology. *New Phytologist* **215**, 15–26 (2017).
79. Kattge, J., Knorr, W., Raddatz, T. & Wirth, C. Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology* **15**, 976–991 (2009).
80. Onoda, Y. *et al.* Global patterns of leaf mechanical properties: Global patterns of leaf mechanical properties. *Ecology Letters* **14**, 301–312 (2011).

81. Onstein, R. E., Carter, R. J., Xing, Y. & Linder, H. P. Diversification rate shifts in the Cape Floristic Region: The right traits in the right place at the right time. *Perspectives in Plant Ecology, Evolution and Systematics* **16**, 331–340 (2014).
82. Prentice, I. C. *et al.* Evidence of a universal scaling relationship for leaf CO₂ drawdown along an aridity gradient. *New Phytologist* **190**, 169–180 (2011).
83. Quested, H. M. *et al.* Decomposition of sub-arctic plants with differing nitrogen economies: a functional role for hemiparasites. *Ecology* **84**, 3209–3221 (2003).
84. Reich, P. B., Oleksyn, J. & Wright, I. J. Leaf phosphorus influences the photosynthesis–nitrogen relation: a cross-biome analysis of 314 species. *Oecologia* **160**, 207–212 (2009).
85. White, M. A., Thornton, P. E., Running, S. W. & Nemani, R. R. Parameterization and Sensitivity Analysis of the BIOME–BGC Terrestrial Ecosystem Model: Net Primary Production Controls. *Earth Interact.* **4**, 1–85 (2000).
86. *Old-Growth Forests: Function, Fate and Value*. vol. **207** (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009).
87. Wright, I. J. *et al.* The worldwide leaf economics spectrum. *Nature* **428**, 821–827 (2004).
88. Ma, H. *et al.* The global biogeography of tree leaf form and habit. *Nat. Plants* **9**, 1795–1809 (2023).
89. Zanne, A. E. *et al.* Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).
90. Givnish, T. Adaptive significance of evergreen vs. deciduous leaves: solving the triple paradox. *Silva Fenn.* **36**, (2002).
91. Van Ommen Kloeke, A. E. E., Douma, J. C., Ordoñez, J. C., Reich, P. B. & Van Bodegom, P. M. Global quantification of contrasting leaf life span strategies for deciduous and evergreen species in response to environmental conditions. *Global Ecology and Biogeography* **21**, 224–235 (2012).
92. Aakala, T., Shimatani, K., Abe, T., Kubota, Y. & Kuuluvainen, T. Crown asymmetry in high latitude forests: disentangling the directional effects of tree competition and solar radiation. *Oikos* **125**, 1035–1043 (2016).
93. Abakumova, M., Zobel, K., Lepik, A. & Semchenko, M. Plasticity in plant functional traits is shaped by variability in neighbourhood species composition. *New Phytologist* **211**, 455–463 (2016).
94. Ackerly, D. D. Adaptation, niche conservatism, and convergence: comparative studies of leaf evolution in the California chaparral. *The American Naturalist* **163**, 654–671 (2004).
95. Anderson-Teixeira, K. J. *et al.* Size-related scaling of tree form and function in a mixed-age forest. *Functional Ecology* **29**, 1587–1602 (2015).
96. Beaulieu, J., Doerksen, T., Clément, S., MacKay, J. & Bousquet, J. Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity* **113**, 343–352 (2014).
97. Carmona, C. P., Rota, C., Azcárate, F. M. & Peco, B. More for less: sampling strategies of plant functional traits across local environmental gradients. *Functional Ecology* **29**, 579–588 (2015).
98. Carus, J., Paul, M. & Schröder, B. Vegetation as self-adaptive coastal protection: Reduction of current velocity and morphologic plasticity of a brackish marsh pioneer. *Ecology and Evolution* **6**, 1579–1589 (2016).
99. Cavender-Bares, J. *et al.* Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): a genomic and population genetics approach. *Molecular Ecology* **24**, 3668–3687 (2015).
100. Dalponte, M. & Coomes, D. A. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. *Methods Ecol Evol* **7**, 1236–1245 (2016).
101. DeWalt, S. J., Bourdy, G., Chavez de Michel, L. R. & Quenevo, C. Ethnobotany of the Tacana: Quantitative inventories of two permanent plots of Northwestern Bolivia. *Economic Botany* **53**, 237–260 (1999).
102. Dostál, P., Fischer, M., Chytrý, M. & Prati, D. No evidence for larger leaf trait plasticity in ecological generalists compared to specialists. *Journal of Biogeography* **44**, 511–521 (2017).
103. Easdale, T. A. & Healey, J. R. Resource-use-related traits correlate with population turnover rates, but not stem diameter growth rates, in 29 subtropical montane tree species. *Perspectives in Plant Ecology, Evolution and Systematics* **11**, 203–218 (2009).
104. Feng, Y. & Van Kleunen, M. Phylogenetic and functional mechanisms of direct and indirect interactions among alien and native plants. *Journal of Ecology* **104**, 1136–1148 (2016).
105. Gamal El-Dien, O. *et al.* Prediction accuracies for growth and wood attributes of interior spruce in space using genotyping-by-sequencing. *BMC Genomics* **16**, 370 (2015).
106. Kleyer, M. *et al.* The LEDA Traitbase: a database of life-history traits of the Northwest European flora. *Journal of Ecology* **96**, 1266–1274 (2008).
107. Letcher, S. G. *et al.* Environmental gradients and the evolution of successional habitat specialization: a test case with 14 Neotropical forest sites. *Journal of Ecology* **103**, 1276–1290 (2015).
108. Li, W., Xu, F., Zheng, S., Taube, F. & Bai, Y. Patterns and thresholds of grazing-induced changes in community structure and ecosystem functioning: species-level responses and the critical role of species traits. *Journal of Applied Ecology* **54**, 963–975 (2017).
109. Louda, S. M., Dixon, P. M. & Huntly, N. J. Herbivory in sun versus shade at a natural meadow-woodland ecotone in the Rocky Mountains. *Plant Ecology* **72**, 141–149 (1987).
110. Loughnan, D. & Gilbert, B. Trait-mediated community assembly: distinguishing the signatures of biotic and abiotic filters. *Oikos* **126**, 1112–1122 (2017).
111. Marx, H. E., Giblin, D. E., Dunwiddie, P. W. & Tank, D. C. Deconstructing Darwin's Naturalization Conundrum in the San Juan Islands using community phylogenetics and functional traits. *Diversity and Distributions* **22**, 318–331 (2016).
112. Mazer, S. J. Ecological, taxonomic, and life history correlates of seed mass among Indiana dune angiosperms. *Ecological Monographs* **59**, 153–175 (1989).
113. McHugh, N., Edmondson, J. L., Gaston, K. J., Leake, J. R. & O'Sullivan, O. S. Modelling short-rotation coppice and tree planting for urban carbon management – a citywide analysis. *Journal of Applied Ecology* **52**, 1237–1245 (2015).
114. Meers, T. L., Kasel, S., Bell, T. L. & Enright, N. J. Conversion of native forest to exotic *Pinus radiata* plantation: Response of understorey plant composition using a plant functional trait approach. *Forest Ecology and Management* **259**, 399–409 (2010).
115. Milla, R., Morente-López, J., Alonso-Rodrigo, J. M., Martín-Robles, N. & Stuart Chapin, F. Shifts and disruptions in resource-use trait syndromes during the evolution of herbaceous crops. *Proc. R. Soc. B.* **281**, 20141429 (2014).
116. Molinari, N. A. & D'Antonio, C. M. Structural, compositional and trait differences between native- and non-native-dominated grassland patches. *Functional Ecology* **28**, 745–754 (2014).
117. Mottet, M.-J., DeBlois, J. & Perron, M. High genetic variation and moderate to high values for genetic parameters of *Picea abies* resistance to *Pissodes strobi*. *Tree Genetics & Genomes* **11**, 58 (2015).
118. Mutete, P., Murepa, R. & Gapare, W. J. Genetic parameters in subtropical pine F1 hybrids: heritabilities, between-trait correlations and genotype-by-environment interactions. *Tree Genetics & Genomes* **11**, 93 (2015).
119. Neba, G. A., Newbery, D. M. & Chuyong, G. B. Limitation of seedling growth by potassium and magnesium supply for two ectomycorrhizal tree species of a Central African rain forest and its implication for their recruitment. *Ecology and Evolution* **6**, 125–142 (2016).
120. Nidzgorski, D. A. & Hobbie, S. E. Urban trees reduce nutrient leaching to groundwater. *Ecological Applications* **26**, 1566–1580 (2016).
121. Osuri, A. M. & Sankaran, M. Seed size predicts community composition and carbon storage potential of tree communities in rain forest fragments in India's Western Ghats. *Journal of Applied Ecology* **53**, 837–845 (2016).
122. Peet, R. K. *et al.* Vegetation-plot database of the Carolina Vegetation Survey. *Biodiversity and Ecology* **4**, 243–253 (2012).

123. Pérez-de-Lis, G. *et al.* Environmental conditions and vascular cambium regulate carbon allocation to xylem growth in deciduous oaks. *Functional Ecology* **31**, 592–603 (2017).
124. Ploton, P. *et al.* Closing a gap in tropical forest biomass estimation: taking crown mass variation into account in pantropical allometries. *Biogeosciences* **13**, 1571–1585 (2016).
125. Poorter, L. & Rozendaal, D. M. Leaf size and leaf display of thirty-eight tropical tree species. *Oecologia* **158**, 35–46 (2008).
126. Price, C. A. *et al.* Are leaf functional traits ‘invariant’ with plant size and what is ‘invariance’ anyway? *Functional Ecology* **28**, 1330–1343 (2014).
127. Ramírez-Valiente, J. A. *et al.* Elucidating the role of genetic drift and natural selection in cork oak differentiation regarding drought tolerance. *Molecular Ecology* **18**, 3803–3815 (2009).
128. Robinson, K. M., Hauzy, C., Loeuille, N. & Albrechtsen, B. R. Relative impacts of environmental variation and evolutionary history on the nestedness and modularity of tree–herbivore networks. *Ecology and Evolution* **5**, 2898–2915 (2015).
129. Rodríguez-Quilón, I. *et al.* Capturing neutral and adaptive genetic diversity for conservation in a highly structured tree species. *Ecological Applications* **26**, 2254–2266 (2016).
130. Roe, A. D. *et al.* Fitness dynamics within a poplar hybrid zone: II. Impact of exotic sex on native poplars in an urban jungle. *Ecology and Evolution* **4**, 1876–1889 (2014).
131. Russo, S. E. *et al.* Interspecific relationships among growth, mortality and xylem traits of woody species from New Zealand. *Functional Ecology* **24**, 253–262 (2010).
132. Sánchez-Robles, J. M. *et al.* Effects of tree architecture on pollen dispersal and mating patterns in *Abies pinsapo* Boiss. (Pinaceae). *Molecular Ecology* **23**, 6165–6178 (2014).
133. Simpson, K. J. *et al.* Determinants of flammability in savanna grass species. *Journal of Ecology* **104**, 138–148 (2016).
134. Steane, D. A. *et al.* Genome-wide scans detect adaptation to aridity in a widespread forest tree species. *Molecular Ecology* **23**, 2500–2513 (2014).
135. United States Department of Agriculture. Forest Inventory and Analysis.
136. Urrutia-Jalabert, R., Malhi, Y. & Lara, A. The Oldest, Slowest Rainforests in the World? Massive Biomass and Slow Carbon Dynamics of *Fitzroya cupressoides* Temperate Forests in Southern Chile. *PLoS ONE* **10**, e0137569 (2015).
137. Van Der Plas, F., Howison, R. A., Mpanza, N., Croomsigt, J. P. G. M. & Olff, H. Different-sized grazers have distinctive effects on plant functional composition of an African savannah. *Journal of Ecology* **104**, 864–875 (2016).
138. Wood, Z. T., Peart, D. R., Palmiotto, P. A., Kong, L. & Peart, N. V. Asymptotic allometry and transition to the canopy in *Abies balsamea*. *Journal of Ecology* **103**, 1658–1666 (2015).
139. Goodman, R. C., Phillips, O. L. & Baker, T. R. The importance of crown dimensions to improve tropical tree biomass estimates. *Ecological Applications* **24**, 680–698 (2014).
140. Grime, J. P., Hodgson, J. G. & Hunt, R. *Comparative Plant Ecology: A Functional Approach to Common British Species*. (Springer, 2014).
141. Hamilton, J. A., Lexer, C. & Aitken, S. N. Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* x *P. glauca*). *Molecular Ecology* **22**, 827–841 (2013).
142. Ishizuka, W. & Goto, S. Modeling intraspecific adaptation of *Abies sachalinensis* to local altitude and responses to global warming, based on a 36-year reciprocal transplant experiment. *Evolutionary Applications* **5**, 229–244 (2012).
143. King, D. A. Allometry and life history of tropical trees. *Journal of tropical ecology* **12**, 25–44 (1996).
144. Kleinschroth, F., Healey, J. R., Sist, P., Mortier, F. & Gourellet-Fleury, S. How persistent are the impacts of logging roads on Central African forest vegetation? *Journal of Applied Ecology* **53**, 1127–1137 (2016).
145. Shugart, H. Jr., Hopkins, M., Burgess, I. & Mortlock, A. Development of a succession model for subtropical rain forest and its application to assess the effects of timber harvest at Wiangaree State Forest, New South Wales. *J. Environ. Manage (United States)* **11** (1980).
146. Fyllas, N. M. *et al.* Functional Trait Variation Among and Within Species and Plant Functional Types in Mountainous Mediterranean Forests. *Front. Plant Sci.* **11**, 212 (2020).
147. E-Vojtkó, A., De Bello, F., Durka, W., Kühn, I. & Götzenberger, L. The neglected importance of floral traits in trait-based plant community assembly. *J Vegetation Science* **31**, 529–539 (2020).
148. Ametegui, A. *et al.* Shade tolerance and the functional trait: demography relationship in temperate and boreal forests. *Functional Ecology* **31**, 821–830 (2017).
149. Bufford, J. L., Lurie, M. H. & Daehler, C. C. Biotic resistance to tropical ornamental invasion. *Journal of Ecology* **104**, 518–530 (2016).
150. Cornwell, W. K., Schwillk, D. W. & Ackerly, D. D. A trait-based test for habitat filtering: convex hull volume. *Ecology* **87**, 1465–1471 (2006).
151. Correia, M., Montesinos, D., French, K. & Rodríguez-Echeverría, S. Evidence for enemy release and increased seed production and size for two invasive Australian acacias. *Journal of Ecology* **104**, 1391–1399 (2016).
152. Fricke, E. C. & Wright, S. J. The mechanical defence advantage of small seeds. *Ecology Letters* **19**, 987–991 (2016).
153. Kraft, T. S. *et al.* Seed size and the evolution of leaf defences. *Journal of Ecology* **103**, 1057–1068 (2015).
154. Liu, K., Eastwood, R., Flynn, S., Turner, R. & Stuppy, W. Seed information database (release 7.1, May 2008). Available at <http://www.kew.org/data/sid> (2008).
155. Manzano-Piedras, E., Marcer, A., Alonso-Blanco, C. & Picó, F. X. Deciphering the Adjustment between Environment and Life History in Annuals: Lessons from a Geographically-Explicit Approach in *Arabidopsis thaliana*. *PLoS ONE* **9**, e87836 (2014).
156. Medrano, M., Herrera, C. M. & Bazaga, P. Epigenetic variation predicts regional and local intraspecific functional diversity in a perennial herb. *Molecular Ecology* **23**, 4926–4938 (2014).
157. Murali, K. Patterns of Seed Size, Germination and Seed Viability of Tropical Tree Species in Southern India. *Biotropica* **29**, 271–279 (1997).
158. Paine, C. E. T. *et al.* Globally, functional traits are weak predictors of juvenile tree growth, and we do not know why. *Journal of Ecology* **103**, 978–989 (2015).
159. Paynter, Q., Buckley, Y. M., Peterson, P., Hugh Gourlay, A. & Fowler, S. V. Breaking and remaking a seed and seed predator interaction in the introduced range of Scotch Broom (*Cytisus scoparius*) in New Zealand. *Journal of Ecology* **104**, 182–192 (2016).
160. Pélabon, C., Hennet, L., Strimbeck, R., Johnson, H. & Armbruster, W. S. Blossom colour change after pollination provides carbon for developing seeds. *Functional Ecology* **29**, 1137–1143 (2015).
161. Schneider, G., Krauss, J., Riedinger, V., Holzschuh, A. & Steffan-Dewenter, I. Biological pest control and yields depend on spatial and temporal crop cover dynamics. *Journal of Applied Ecology* **52**, 1283–1292 (2015).
162. Yang, X. *et al.* Transgenes for insect resistance reduce herbivory and enhance fecundity in advanced generations of crop–weed hybrids of rice. *Evolutionary Applications* **4**, 672–684 (2011).
163. Zas, R., Cendán, C. & Sampedro, L. Mediation of seed provisioning in the transmission of environmental maternal effects in Maritime pine (*Pinus pinaster* Aiton). *Heredity* **111**, 248–255 (2013).
164. Zas, R. & Sampedro, L. Heritability of seed weight in Maritime pine, a relevant trait in the transmission of environmental maternal effects. *Heredity* **114**, 116–124 (2015).
165. Atkinson, R. R. L. *et al.* C4 photosynthesis boosts growth by altering physiology, allocation and size. *Nature Plants* **2**, 16038 (2016).

166. Burns, J. H., Halpern, S. L. & Winn, A. A. A test for a cost of opportunism in invasive species in the Commelinaceae. *Biol Invasions* **9**, 213–225 (2007).
167. De La Riva, E. G. *et al.* Disentangling the relative importance of species occurrence, abundance and intraspecific variability in community assembly: a trait-based approach at the whole-plant level in Mediterranean forests. *Oikos* **125**, 354–363 (2016).
168. Deraison, H., Badenhauer, L., Börger, L. & Gross, N. Herbivore effect traits and their impact on plant community biomass: an experimental test using grasshoppers. *Functional Ecology* **29**, 650–661 (2015).
169. Gavinet, J., Prévosto, B. & Fernandez, C. Introducing resprouters to enhance Mediterranean forest resilience: importance of functional traits to select species according to a gradient of pine density. *Journal of Applied Ecology* **53**, 1735–1745 (2016).
170. Grootemaat, S., Wright, I. J., Van Bodegom, P. M., Cornelissen, J. H. C. & Cornwell, W. K. Burn or rot: leaf traits explain why flammability and decomposability are decoupled across species. *Functional Ecology* **29**, 1486–1497 (2015).
171. Kraft, N. J., Valencia, R. & Ackerly, D. D. Functional traits and niche-based tree community assembly in an Amazonian forest. *Science* **322**, 580–582 (2008).
172. Leishman, M. R., Cooke, J. & Richardson, D. M. Evidence for shifts to faster growth strategies in the new ranges of invasive alien plants. *Journal of Ecology* **102**, 1451–1461 (2014).
173. Martin, A. R. *et al.* Intraspecific trait variation across multiple scales: the leaf economics spectrum in coffee. *Functional Ecology* **31**, 604–612 (2017).
174. Mason, C. M. & Donovan, L. A. Evolution of the leaf economics spectrum in herbs: Evidence from environmental divergences in leaf physiology across *Helianthus* (Asteraceae). *Evolution* **69**, 2705–2720 (2015).
175. Norghauer, J. M., Glauser, G. & Newbery, D. M. Seedling resistance, tolerance and escape from herbivores: insights from co-dominant canopy tree species in a resource-poor African rain forest. *Functional Ecology* **28**, 1426–1439 (2014).
176. Onstein, R. E. *et al.* Evolutionary radiations of Proteaceae are triggered by the interaction between traits and climates in open habitats. *Global Ecol. Biogeogr.* **25**, 1239–1251 (2016).
177. Prunier, R., Holsinger, K. E. & Carlson, J. E. The effect of historical legacy on adaptation: do closely related species respond to the environment in the same way? *J of Evolutionary Biology* **25**, 1636–1649 (2012).
178. Royer, D. L., Wilf, P., Janesko, D. A., Kowalski, E. A. & Dilcher, D. L. Correlations of climate and plant ecology to leaf size and shape: potential proxies for the fossil record. *American Journal of Botany* **92**, 1141–1151 (2005).
179. Spasojevic, M. J., Turner, B. L. & Myers, J. A. When does intraspecific trait variation contribute to functional beta-diversity? *Journal of Ecology* **104**, 487–496 (2016).
180. Wigley, B. J. *et al.* Leaf traits of African woody savanna species across climate and soil fertility gradients: evidence for conservative versus acquisitive resource-use strategies. *Journal of Ecology* **104**, 1357–1369 (2016).
181. Zanne, A. E. *et al.* A deteriorating state of affairs: How endogenous and exogenous factors determine plant decay rates. *Journal of Ecology* **103**, 1421–1431 (2015).
182. Zheng, S. *et al.* Effects of grazing on leaf traits and ecosystem functioning in Inner Mongolia grasslands: scaling from species to community. *Biogeosciences* **7**, 1117–1132 (2010).
183. Zhu, H. *et al.* Reducing soil erosion by improving community functional diversity in semi-arid grasslands. *Journal of Applied Ecology* **52**, 1063–1072 (2015).
184. Diaz, S. *et al.* The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).
185. Engemann, K. *et al.* A plant growth form dataset for the New World. *Ecology* **97**, 3243–3243 (2016).
186. Leonard, J. L. *Transitions Between Sexual Systems: Understanding the Mechanisms of, and Pathways Between, Dioecy, Hermaphroditism and Other Sexual Systems.* (Springer International Publishing, Cham, 2018).
187. Peng, D. *et al.* Plant sexual systems correlated with morphological traits: Reflecting reproductive strategies of alpine plants. *J of Systematics Evolution* **52**, 368–377 (2014).
188. Edwards, E. J., Chatelet, D. S., Sack, L. & Donoghue, M. J. Leaf life span and the leaf economic spectrum in the context of whole plant architecture. *Journal of Ecology* **102**, 328–336 (2014).
189. Mason, C. M., Goolsby, E. W., Humphreys, D. P. & Donovan, L. A. Phylogenetic structural equation modelling reveals no need for an 'origin' of the leaf economics spectrum. *Ecology Letters* **19**, 54–61 (2016).
190. Poorter, L. & Bongers, F. Leaf traits are good predictors of plant performance across 53 rain forest species. *Ecology* **87**, 1733–1743 (2006).
191. Balzotti, C. S. *et al.* Environmental controls on canopy foliar nitrogen distributions in a Neotropical lowland forest. *Ecological Applications* **26**, 2451–2464 (2016).
192. Bhaskar, R., Dawson, T. E. & Balvanera, P. Community assembly and functional diversity along succession post-management. *Functional Ecology* **28**, 1256–1265 (2014).
193. Bhattarai, G. P. *et al.* Biogeography of a plant invasion: genetic variation and plasticity in latitudinal clines for traits related to herbivory. *Ecological Monographs* **87**, 57–75 (2017).
194. Bonal, D., Sabatier, D., Montpied, P., Tremeaux, D. & Guehl, J.-M. Interspecific variability of $\delta^{13}C$ among trees in rainforests of French Guiana: functional groups and canopy integration. *Oecologia* **124**, 454–468 (2000).
195. Han, W., Fang, J., Guo, D. & Zhang, Y. Leaf nitrogen and phosphorus stoichiometry across 753 terrestrial plant species in China. *New Phytologist* **168**, 377–385 (2005).
196. Hayes, P., Turner, B. L., Lambers, H. & Laliberté, E. Foliar nutrient concentrations and resorption efficiency in plants of contrasting nutrient-acquisition strategies along a 2-million-year dune chronosequence. *Journal of Ecology* **102**, 396–410 (2014).
197. He, J.-S. *et al.* A test of the generality of leaf trait relationships on the Tibetan Plateau. *New Phytologist* **170**, 835–848 (2006).
198. Hess, L. J. T. & Austin, A. T. *Pinus ponderosa* alters nitrogen dynamics and diminishes the climate footprint in natural ecosystems of Patagonia. *Journal of Ecology* **102**, 610–621 (2014).
199. Karagatzides, J. & Ellison, A. Construction Costs of Carnivorous Plants and Non-Carnivorous Plants. *Harvard Forest Data Archive: HF112* (2008).
200. Li, X., Schmid, B., Wang, F. & Paine, C. E. T. Net Assimilation Rate Determines the Growth Rates of 14 Species of Subtropical Forest Trees. *PLoS ONE* **11**, e0150644 (2016).
201. Maire, V. *et al.* Global effects of soil and climate on leaf photosynthetic traits and rates. *Global Ecology and Biogeography* **24**, 706–717 (2015).
202. Moreno-Gutiérrez, C., Battipaglia, G., Cherubini, P., Delgado Huertas, A. & Querejeta, J. I. Pine afforestation decreases the long-term performance of understorey shrubs in a semi-arid Mediterranean ecosystem: a stable isotope approach. *Functional Ecology* **29**, 15–25 (2015).
203. Niu, K., He, J. & Lechowicz, M. J. Grazing-induced shifts in community functional composition and soil nutrient availability in Tibetan alpine meadows. *Journal of Applied Ecology* **53**, 1554–1564 (2016).
204. Shibata, R. *et al.* Relationships between resprouting ability, species traits and resource allocation patterns in woody species in a temperate forest. *Functional Ecology* **30**, 1205–1215 (2016).
205. Snell-Rood, E. C., Espeset, A., Boser, C. J., White, W. A. & Smykalski, R. Anthropogenic changes in sodium affect neural and muscle development in butterflies. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10221–10226 (2014).
206. Swenson, N. & Umana, M. Data from: Interspecific functional convergence and divergence and intraspecific negative density dependence underlie the seed-to-seedling transition in tropical trees. *The American Naturalist* <https://doi.org/10.5061/dryad.j2r53> (2015).

207. Szefer, P. *et al.* Determinants of litter decomposition rates in a tropical forest: functional traits, phylogeny and ecological succession. *Oikos* **126**, 1101–1111 (2017).
208. Vincent, J. B., Weiblen, G. D. & May, G. Host associations and beta diversity of fungal endophyte communities in New Guinea rainforest trees. *Molecular Ecology* **25**, 825–841 (2016).
209. Welsh, M. E., Cronin, J. P. & Mitchell, C. E. The role of habitat filtering in the leaf economics spectrum and plant susceptibility to pathogen infection. *Journal of Ecology* **104**, 1768–1777 (2016).
210. Weremijewicz, J. & Seto, K. Mycorrhizas influence functional traits of two tallgrass prairie species. *Ecology and Evolution* **6**, 3977–3990 (2016).
211. Rosas-Guerrero, V. *et al.* A quantitative review of pollination syndromes: do floral traits predict effective pollinators? *Ecology Letters* **17**, 388–400 (2014).
212. Ollerton, J., Rech, A. R., Waser, N. M. & Price, M. V. Using the Literature to Test Pollination Syndromes — Some Methodological Cautions. *J Poll Ecol* **16**, 119–125 (2015).

Acknowledgements

We extend our sincere gratitude to all data contributors. The full list of contributors is available in the Supplementary Information file. In particular, we thank the following colleagues for directly providing data (listed alphabetically): Ahmed Dawelbeit, Michèle Dinies, Kaitlyn E. Horisk, Sarah J. Ivory, Etienne Jaillard, Dafna Langgut, Suzzane Leroy, Leszek Marks, Anna Maria Mercuri, Krystyna Milecka, Serge D. Muller, and Neil Roberts. We are also grateful to Graciela Gil Romera, Suzanne A.G. Leroy, Anne-Marie Lézine and Anna Maria Mercuri for their valuable feedback on the project. We further thank Carmina Gómez Ferreras for donating classical pollen taxonomic literature used in the pollen taxonomy harmonisation. Additionally, we acknowledge the Neotoma Paleoecology Database, the African Pollen Database (APD), the Botanical Information and Ecology Network (BIEN) and TRY Plant Trait Database, along with their contributors, for providing access to valuable palynological and functional trait data. This is a contribution from I.S.'s PhD thesis, supervised by J.B.-J., I.A.L., S.M. and C.D.T. I.S. was supported by the PhD studentship *Adapting to the Challenges of a Changing Environment* (ACCE) Doctoral Training Partnership, funded by the Natural Environment Research Council (NERC), Grant: NE/S00713X/1. Additional support came from the *Biodiversity in Ancient Egypt during Societal Transitions* (BEAST) project, funded by Leverhulme Trust (Grant RPG-2021-104 to J.B.-J.). The literature compilation supporting the taxonomic homogenisation in this study was funded by MICIU/AEI/10.13039/501100011033 and by the ERDF/EU (Grant PID2024-161027NA-I00 DivAS-Cape to S.M.). I.A.L. acknowledges funding from Ministerio de Ciencia e Innovación de España under grant No. RYC2021-034991-I, funded by MCIN/AEI/10.13039/501100011033 and European Union NextGenerationEU/PRTR.

Author contributions

I.S. led the data collection and processing and wrote the manuscript; J.B.-J. conceived the project and revised and edited the manuscript; I.A.L. conceived the project and revised and edited the manuscript; C.D.T. revised and edited the manuscript; S.M. conceived the project, contributed to data collection, led the pollen taxonomical harmonisation, and revised and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-026-07319-8>.

Correspondence and requests for materials should be addressed to I.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026