**Title**

Global distribution and climatic preferences of C4 eudicots and how they differ from those of C4 grasses

**Authors**

Jessica A. Berasategui1,2, Anže Žerdoner Čalasan1, Alexander Zizka4 & Gudrun Kadereit1,3

1Prinzessin Therese von Bayern Lehrstuhl für Systematik, Biodiversität & Evolution der Pflanzen, Ludwig-Maximilians Universität München, Menzinger Str. 67, 80638 München, Germany

2Institute for Molecular Physiology, Johannes Gutenberg-University Mainz, 55099 Mainz, Germany

3Botanischer Garten München-Nymphenburg und Botanische Staatssammlung München, Staatliche Naturwissenschaftliche Sammlungen Bayerns, Menzinger Str. 65, 67, 80638 München, Germany

4Department of Biology, Philipps-University Marburg, Karl-von-Frisch-Straße 8, 35043 Marburg, Germany

Corresponding author: Jessica Anna Berasategui (née Los) ([J.Los@biologie.uni-muenchen.de](mailto:J.Los@biologie.uni-muenchen.de))

**Date of submission: 02.05.2023**

**No. of figures: 6 (Figures 1, 2, 3, 4 and 6 should be in colour)**

**No. of tables: 2**

**Word count: 11324**

**Supplementary material: One supplementary table and 17 supplementary figures**

**Highlights**

1. Environmental conditions in Mexico/Southern United States and in large areas of Australia likely promoted the diversification of C4 lineages in grasses and eudicots.
2. More C4 lineages originated in-situ in Mexico/Southern United States than in Australia.
3. C4 eudicots (irrespective of the presence of succulence) colonised significantly more arid regions than C4 grasses.
4. C4 eudicot lineages utilising the NAD-ME decarboxylating enzyme are found in significantly drier areas than those using the NADP-ME decarboxylating enzyme.

**Abstract**

C₄ is one of three known photosynthetic processes of carbon fixation in flowering plants. It evolved independently more than 61 times in multiple angiosperm lineages and consists of a series of anatomical and biochemical modifications to the ancestral C3 pathway increasing plant productivity under warm and light-rich conditions. The C4 lineages of eudicots belong to seven orders and 15 families, are phylogenetically less clustered than those of monocots, and entail an enormous structural and ecological diversity. Eudicot C4 lineages likely evolved the C4 syndrome along different evolutionary paths. Therefore, a better understanding of this diversity is key to understanding the evolution of this complex trait as a whole. Compiling 1,207 recognized C4 eudicots species described in the literature and presenting trait data among these species, we identify global centres of species richness and of high phylogenetic diversity. Furthermore, we discuss climatic preferences in the context of plant functional traits. We identify two hotspots of C4 eudicot diversity: arid regions of Mexico/Southern United States and Australia, where several C4 eudicot lineages diversified independently. Further eudicot C4 hotspots with many different families and genera represented are in South Africa, West Africa, Patagonia, Central Asia and the Mediterranean. In general, C4 eudicots were abundant in deserts and xeric shrublands, tropical and subtropical grasslands, savannas and shrublands. We found C4 eudicots to occur in areas with less annual precipitation than C4 grasses which can be explained by frequently associated adaptations to drought stress such as among others succulence and salt tolerance. We conclude that in most eudicot lineages C4 evolved in ancestrally drought adapted clades and enabled these to further spread in these habitats and colonise even drier areas.

**Keywords**

Biome, climatic preferences, desert, C4 photosynthesis, GBIF, salt tolerance, succulence

**1. Introduction**

By the early 1950s it was widely assumed that all plants use the same C3 carbon fixation pathway, the Calvin-Benson-Bassham cycle (CBB-cycle; Bassham et al., 1950). Shortly after a brief note on the discovery of a four-carbon CO2 fixation pathway in sugarcane - now known as C4 photosynthesis - published in the 1954 Annual Report of the Hawaiian Sugar Planters Association Experiment Station (Burr et al., 1957; Hatch, 2005), researchers set out to investigate this unexplored photosynthetic pathway. They found that the CBB-cycle and RuBisCO were restricted to bundle sheath cells (BSC) and that in the mesophyll cells (MC) an auxiliary carbon fixing pathway with phosphoenolpyruvate carboxylase (PEPC) as the key enzyme generated C4 molecules that are transported into the BSC and fuel the CBB-cycle. Since its discovery, understanding C4 photosynthesis became a vibrant research discipline, integrating the fields of biochemistry, physiology, organismic biology, ecology and evolution (Langdale, 2011). Evolving knowledge about the C4 pathways has been published in various reviews (see Sage et al., (2018), Niklaus & Kelly (2019), Schlüter & Weber (2020) and Furbank & Kelly (2021) for four recent ones on different aspects of C4 photosynthesis) and special issues (e.g., JXB special issue: C4 Photosynthesis – 50 years of discovery and innovation - Von Caemmerer et al., 2017).

C4 photosynthesis evolved in at least 18 angiosperm families and more than 60 times independently (Sage, 2017; Sage et al., 2018; Figure 1a). Around 80% of the C4 species are found in the Poales with 5,044 C4 species in Poaceae and 1,322 C4 species in Cyperaceae. With two C4 species in Hydrocharitaceae (Alismatales) this adds up to 6,368 C4 species in 339 genera in monocots, opposed to around 1,777 eudicot C4 species in 79 genera (Sage, 2017 and ref. therein) (Figure 1b). Interestingly, the eudicot C4 lineages are phylogenetically more equally distributed and occur in three rosid and four asterid families belonging to six different orders. Nevertheless, the C4 species-richest eudicot clades are restricted mostly to eight families of the Caryophyllales (Figure 1a, based on Sage, 2017: Table 3). Amaranthaceae s.l. (Caryophyllales) is by far the most species-rich C4 eudicot family, followed by Euphorbiaceae (Malpighiales, rosids), Asteraceae and Boraginaceae (Asterales and Lamiales, respectively, asterids; Figure 1a). Although C4 photosynthesis in eudicots is phylogenetically more widespread and ecologically and structurally more diverse than in monocots (e.g., Muhaidat et al., 2007; Rudov et al., 2020), the latter have received more attention (mainly in the Poaceae). This is partly due to C4 grasses, such as maize and sugarcane, being initial model species of C4 research, thus making their close relatives the focus of C4 research even today (Hatch, 2005). Furthermore, research focus on grasses can be attributed to their great economic and ecological importance (Linder et al., 2018). However, understanding the diversity of the C4 syndrome in eudicots is key to understand the evolution of this complex trait (Heyduk et al., 2019) because C4 eudicot lineages evolved the C4 syndromes along different evolutionary paths (e.g., Kadereit et al., 2012; Bohley et al., 2015; Lauterbach et al., 2019).

**(--Figure 1--)**

C4 photosynthesis, which includes an auxiliary pathway to reduce photorespiration, likely arose in hot, dry, and/or saline regions where C3 photosynthesis performance is reduced (Sage et al., 2018). This is achieved by the fixative enzyme, phosphoenolpyruvate-carboxylase (PEPC), and by generating a local high CO2 concentration around the key enzyme RuBisCO, to reduce its oxygenase activity (Sage et al., 2012). C4 photosynthesis is usually associated with warm habitats with high evapotranspiration. Yet, the distribution of C4 plants cannot be explained entirely by individual environmental factors (Christin & Osborne, 2014) because C4 species occur in a variety of habitats, for example on nutrient-poor or fertile soils, in the tropics, in deserts or in the boreal zone, on open grasslands or forest undergrowth (Collins & Jones, 1986; Mahdavi & Bergmeier, 2018; Rudov et al., 2020). This diversity results from the multifaceted evolutionary history of the C4 pathway (Christin & Osborne, 2014; Sage et al., 2011; Sage et al., 2018). However, studies linking the evolution of adaptive traits and ecological niches in C4 lineages are still insufficient. Lundgren et al. (2015) for example showed that C4 photosynthesis does not initially lead to a shift of the ancestral niche in *Alloteropsis semialata* J.Presl (Poaceae), but rather expands its niche to cover a wider range of conditions that include the ancestral ones. This improves the success of occasional long-range dispersal events and thus increases the geographical range (Lundgren et al., 2015). As C4 photosynthesis is a complex syndrome that increases the efficiency of plants to use available water and nitrogen, C4 might be advantageous under various environmental conditions, but evolved predominantly in the tropics and subtropics (Griffiths et al., 2013; Sage et al., 2018).

**1.1 Global expansion of C4 grasses**

C4 plants account for one quarter of the earth’s primary terrestrial production, and almost a quarter of the Earth’s surface is dominated by C4 grasslands and savannas (Sage et al., 2018; Barbehenn et al., 2004; Grace et al., 2006). C4 grasses likely intruded into C3 grasslands and forests from open biomes of warm regions, subsequently replaced them during the late Miocene to the Pliocene (3-8 Mya) and expanded worldwide into drier biomes (Ehleringer et al., 1997; Osborne & Freckleton, 2009; Edwards & Smith, 2010). However, molecular research suggested that C4 photosynthesis in several grass lineages evolved around 18-30 Mya earlier (mid/late-Oligocene) presumably in warm, arid locations where water limitation was the main selective force to increase photorespiration (Christin et al., 2011; Zhou et al., 2018). As the atmospheric CO2 in the late Miocene fell below ~300 ppm (Royer, 2006), a small number of hyperdominant C4 grass species that were able to outcompete C3 and C4 relatives became dominant due to their advantage of a low CO2 compensation point (Christin & Osborne, 2014, Lehmann et al., 2019). South America seems to be the major hotspot for the origin of C4 grasses (Sage et al., 2011), and today C4 grasses are confined mostly to tropical and subtropical areas (Shoko et al., 2016; Woodward & Lomas, 2004; Woodward et al., 2004). Climatic patterns and the distribution of C4 grasses in North America suggest that high minimum temperatures during the growing season favour C4 grasses at regional scale (Teeri & Stowe, 1976). Yet, at the local scale, topographic and edaphic variables may exert more influence (Yan & de Beurs, 2016).

**1.2 Evolutionary and ecological diversity of C4 photosynthesis in eudicots**

Individual C4 lineages had originated independently from the Oligocene into the Quaternary (Christin et al., 2011; Niklaus & Kelly, 2019). In Amaranthaceae s.l., which includes the largest number of C4 lineages in eudicots, the earliest assumed origins of C4 date back to the Oligocene and are roughly as old as the oldest C4 grass subfamily Chloridoideae, which evolved around 32-25 Mya. This implies that C4 eudicots are per se not younger than C4 monocots (Christin et al., 2008; Christin et al., 2011; Kadereit et al., 2012). However, many C4 lineages within eudicots (as well as in the monocots) originated in the Miocene when the climate got increasingly drier (Kadereit et al., 2003, 2010). In addition, there are many evolutionary young C4 lineages in eudicots, for instance in *Flaveria* (Asteraceae), *Sesuvium* (Aizoaceae) and *Tecticornia* (Chenopodiaceae s.s.) that arose approximately 5-1 Mya (Kadereit et al., 2012; Sage et al., 2012; Christin et al., 2011). Since the range size of C4 lineages as well as the physiological refinement of the C4 syndrome is highly dependent on time, the age of the respective C4 lineage needs to be taken into account when lineages are compared to each other (Niklaus & Kelly, 2019).

Whereas South America seems to be the hotspot of origins of the nowadays cosmopolitan C4 grasses, six geographic regions were highlighted as potential ancestral areas for C4 eudicot lineages. For most C4 eudicot lineages Central Asia, North America, South Africa, northeast Africa and Arabia count as centres of origin (Sage et al., 2011; Kadereit & Freitag, 2011; Sage, 2016), an assessment based mainly on current distribution that still awaits the review of detailed phylogenetic, biogeographical studies in the individual eudicot lineages (e.g., Lauterbach et al., 2019). Due to the diverse nature of C4 eudicots no list of the global distribution of C4 eudicots has been compiled thus far.

**1.3 Functional plant traits in C4 eudicots**

Several traits such as succulence, salt tolerance, fast seed germination and longevity seem to be associated with C4 photosynthesis in eudicots (e.g., Kadereit et al., 2012, 2017). While succulence is generally rare among grasses (only in *Spinifex littoreus* (Burm.f.) Merr.; Ho et al., 2019), this trait results in an unmatched leaf and stem anatomical diversity in C4 eudicots (Kadereit et al., 2003; Muhaidat et al., 2007; Muhaidat et al., 2018; Bohley et al., 2015; Voznesenskaya et al., 2017). While grasses usually show a classical Kranz anatomy (similar to an atriplicoid leaf anatomy), eudicots show a broad variety of succulent and non-succulent C4 leaf types in addition to the atriplicoid type (Edwards & Voznesenskaya, 2011). Among the succulent C4 leaf types, the annual *Suaeda aralocaspica* (Bunge) Freitag (= *Borszczowia aralocaspica*; Amaranthaceae s.l.), *Bienertia cycloptera* Bunge ex Boiss, *B. sinuspersici* Akhani and *B. kavirense* Akhani are particularly noteworthy because their C4 photosynthesis is carried out within a single photosynthetic cell and without the supposedly mandatory C4 Kranz anatomy (Akhani et al., 2005, 2012; Freitag & Stichler, 2000; Sharpe et al., 2020; Voznesenskaya et al., 2002, 2003). Although many different examples of anatomical diversity in C4 eudicots have been found, there are still many representatives of the other C4 eudicot lineages that are not well characterised and might contribute to the ecological and morphological diversity (Muhaidat et al., 2007).

Most succulent C4 species tolerate elevated salinity, suggesting that their succulence is primarily an evolutionary response to (physiological) drought. While in grasses a repeated gain and loss of salt tolerance throughout the history of the family prevails and halophytic grass species are isolated at the tips of the phylogeny (Bromham & Bennett, 2014), there are multiple evolutionary older halophytic lineages among eudicots that additionally acquired C4 photosynthesis. This is particularly the case for C4 lineages of Amaranthaceae (Kadereit et al., 2012; Kadereit et al., 2017; Piirainen et al., 2017) but also for Gisekiaceae (Bissinger et al., 2014), Sesuvioideae-Aizoaceae (Bohley et al., 2015) and Euphorbiaceae (Rudov et al., 2020; Ghazanfar et al., 2014). Some eudicot lineages acquired even further alternative carbon fixation pathways. The widespread succulent annual *Portulaca oleracea* L. is a halophytic C4 species that is able to conduct both C4 and CAM photosynthesis depending on the environmental conditions (Ferrari et al., 2020, 2022).

While C4 trees and large shrubs are generally rare, several eudicot C4 species can be woody and/or perennial such as *Anabasis* from Eurasian steppes and semi deserts (Lauterbach et al., 2019); the saxaul (*Haloxylon ammodendron* (C.A.Mey.) Bunge ex Fenzl), which dominates continental deserts of Asia (Pyankov et al., 1999) and the Hawaiian C4 trees *Euphorbia olowaluana* Sherff and *E. herbstii* (W.L.Wagner) Oudejans (Pearcy & Troughton, 1975; Young et al., 2020).

Despite anatomical and ecological differences, the biochemical forms that exist in C4 photosynthesis are similar in grasses and eudicots. There are three biochemical subtypes in both grasses and eudicots, which are usually constant in a C4 lineage, but may vary within and between plant families: NADP-malic enzyme (ME; e.g., in Caryophyllaceae; Sage et al., 2011), NAD-ME (e.g., C4 species in Boraginaceae, Cleomaceae; Muhaidat et al., 2007), and the third further decarboxylating enzyme, PEP-CK that is more common in C4 monocots (Wang et al., 2014). Due to the high number of fast growing and highly productive C4 grasses many of which are interesting biofuel crops (such as *Miscanthus*) one might assume that C4 grasses are more competitive than C4 eudicots given the right growing conditions. However, interestingly the species with the fastest CO2 assimilation rates of 80 µmol m−2 s−1 at 325 µmol mol−1 is not a grass species but *Amaranthus palmeri* S. Watson (Amaranthaceae, Ehleringer, 1983; Sage, 2017).

**1.4 Scope and aims**

C4 grasses entail the majority of C4 species and dominate in biomass production, yet the anatomical, physiological and ecological diversity of C4 syndromes seems larger in C4 eudicots. While shifts to C4 physiology in grasses probably represent a pre-adaptation to open and arid subtropical habitats (Osborne & Freckleton, 2009; Edwards & Smith, 2010), the evolution of the C4 pathway in eudicots, e.g., Amaranthaceae s.l. (incl. Chenopodiaceae), Nyctaginaceae and Sesuvioideae, is more likely a post adaptation to the selection pressure in dry, saline and coastal environments that enabled survival in these habitats (Kadereit et al., 2012; Bohley et al., 2015; Khoshravesh et al., 2020). Already Stowe and Teeri (1978) suggested that C4 eudicots do not follow the climate preferences that have been reported for C4 grasses and therefore might have followed a different evolutionary pathway to C4.

In this study, we aimed to characterise the global occurrence of C4 eudicots, identify diversity hotspots and climatic preferences, and assign these to functional traits specifically succulence, salt tolerance, biochemical subtype and anatomical leaf type. We hypothesized that the phylogenetic and structural diversity of C4 eudicots is reflected in their colonization of a wide range of climatic regions and environments and that the combination of C4 photosynthesis with other traits enabled C4 eudicots to invade areas not or less frequently colonised by C4 grasses. To test this hypothesis, we compare the patterns found for eudicots with those of the more species-rich C4 grasses.

**2. Materials and methods**

We compiled an initially dataset of C4 eudicots according to Sage (2017). This list consisted of 16 eudicot families with indication of lineages and the number of C4 species per lineage (Table A1). In order to list each C4 species, literature research was conducted. If trait data were available, we recorded leaf anatomy, biochemical subtypes (NAD-ME, NADP-ME), succulence, woodiness, salt tolerance and life form (perennial, annual) from floras, revisions, reports, databases and online sources (see sources in Table A1). To reduce the artificial increase in species numbers and distribution areas due to synonymization, we cross-checked for synonyms using plantsoftheworldonline.org (POWO, 2020). In grasses, there are over 60,000 published scientific names corresponding to approximately 11,313 accepted species (Clayton et al., 2002 onwards; Osborne et al., 2014). Our list of C4 eudicot species includes members of 15 families (Chenopodiaceae included in Amaranthaceae) with 1,207accepted species and a total of around 3,969 synonyms. These 1,207 species have verified information about C4 photosynthesis performance and the literature and/or online resources provided detailed information about the traits discussed above (Table A1). For comparison with C4 grasses, we compiled a list of 307 genera out of a total of 321 C4 genera (Sage, 2017) following Osborne et al. (2014) (Table A1), as all species of these genera are assumed to perform C4.

* 1. **Occurrence data**

We extracted georeferenced occurrence data of the 1,207 C4 eudicot species from the Global Biodiversity Information Facility (gbif.org) using the “rgbif” v3.2.0 package (Chamberlain et al., 2020) in R (for DOIs see GBIF-References; GBIF.org.2020a; GBIF.org.2020b; GBIF.org.2020c; GBIF.org.2020d; GBIF.org.2020e; GBIF.org.2020f; GBIF.org.2020g; GBIF.org.2020h; GBIF.org.2020i; GBIF.org.2020j; GBIF.org.2020k; GBIF.org.2020l; GBIF.org.2020m; GBIF.org.2020n; GBIF.org.2020o). We downloaded occurrence records at species level for C4 eudicots and at genus level for C4 grasses (GBIF.org.2020p; GBIF.org.2020q; GBIF.org.2021r; GBIF.org.2020s; GBIF.org.2020t). Records describing fossilised specimens, records based on literature alone and records of unknown origin were excluded (R code available at <https://docs.ropensci.org/rgbif/articles/rgbif.html>). GBIF records originated from a variety of sources, including human and machine observation (e.g., photograph), living and preserved specimens. For 208 C4 eudicot species (17%) included in our list of verified C4 species, no reliable occurrence data were available in GBIF.

* 1. **Data cleaning**

Since georeferenced occurrence records from public datasets such as gbif.org are error prone (Maldonado et al., 2015; Zizka et al., 2020), automated data cleaning of the C4 eudicot and C4 grasses coordinate datasets was performed with the “CoordinateCleaner” v2.0-18 package in R (Zizka et al., 2019) using the default options. Following the process outlined in Zizka et al. (2019), erroneous records within 1,000 m of country and/or province centroids and within 10,000 m of countries’ capitals, within urban areas, records with locations as zeros, identical values, near GBIF headquarters, near biodiversity institutions and records on an ocean surface were removed. In addition to the “CoordinateCleaner”, the dataset was manually checked for incorrect synonymization relying on plantsoftheworldonline.org as the taxonomic backbones of the GBIF are not always following the currently accepted taxonomic treatments by “The International Plant Names Index” (IPNI, 2020) and “World Checklist of Vascular Plants” (WCVP, 2020). Besides, duplicate coordinates, based on species name and coordinates, were removed. Likewise, the taxonomic reliability using the distribution information of plantsoftheworldonline.org was checked and occurrence points considered incorrect based on their distribution outside the native ranges of species were excluded. As a result of all cleaning steps, the number of coordinates was reduced extensively.

The cleaning of 2,296,101 occurrence records of C4 grasses was also carried out with the “CoordinateCleaner” package. Additionally, the occurrence points outside the native ranges and duplicate coordinates per species were excluded. Manual cleaning of the incorrectly synonymized species was not carried out here, as we focussed on the genus level only.

* 1. **Analyses**

We used 100 x 100 km grid cells to infer geographic patterns of C4 species richness, with an equal area Behrmann projection. Species richness maps for each C4 eudicot family for the uncleaned and for the cleaned dataset were generated using the package “speciesgeocodeR” v2.0-10 (Töpel et al., 2017; Supplementary Figures A1-15). Grids with species numbers were calculated using *RichnessGrid*. In addition to the individual species richness maps for each C4 eudicot family, total species richness maps for C4 eudicots and C4 grasses were created with the cleaned datasets. These maps provide information on the total distribution (showing outstanding regions of C4 species richness) of both groups. Grids showing richness above 50 species were here defined to be C4 species hotspots of high phylogenetic diversity. In order to see which grid cells are relatively more important for C4 grasses and for C4 eudicots, respectively, the difference of normalised species numbers is shown in Figure 2d. For this purpose, the species numbers for C4 eudicots and C4 grasses were first scaled separately over all grid cells between 0 and 1 (0 minimum number of species in a grid cell, 1 maximum number of species in a grid cell) and then subtracted. The normalisation was necessary due to the higher total number of grass species.

Two bioclimatic variables (Bio1 – Annual Mean Temperature (°C\*10); Bio12 – Annual Mean Precipitation (mm)) were extracted from WorldClim v.2 with a spatial resolution of 10 minutes (~340 km²) (Fick & Hijmans, 2017). We plotted annual mean temperature and precipitation values using ggplot2 to compare the distribution of C4 eudicots and C4 grasses along these two climatic variables (Wickham et al., 2016). We are aware that this approach does not integrate biogeographical history which is beyond the scope of this paper. Likewise, boxplots for each family of C4 eudicots, the C4 grasses, and all C4 eudicots together were calculated in relation to the climate variables. Statistical analyses were conducted in R v4.0.2 (R Core Team, 2020) using RStudio v1.2.5042 (RStudio, Inc. 2009-2020) and R Commander (Fox, 2005). We used the Krustal-Wallis test (between C4 eudicot families) and the Mann-Whitney-U test (between C4 eudicots and C4 grasses) to determine significance according to annual mean precipitation and annual mean temperature.

To display the distribution areas of each C4 eudicot family at the biome level (Olson et al., 2001), a table in Figure 4a was created. Only families with the highest number of species in that study according to Figure 1a were selected for this table.

**3. Results**

**3.1 The impact of data cleaning**

We included only species for which direct evidence of C4 photosynthesis (such as C4-like δ13C values or C4 leaf anatomy) is documented in the literature. This was the case for 1,207 species of the approximately 1,777 C4 eudicots according to Sage (2017). Since Sage (2017) estimates the number of C4 species per lineage, our refined C4 eudicot species list is substantially shorter (Table A1). For 208 of these 1,207 C4 eudicot species no occurrence points were documented in GBIF. Therefore, the final list of C4 eudicots analysed here included 999 species.

Performing the necessary cleaning steps reduced the raw C4 eudicot occurrence points dataset, which originally contained more than 1 million records, to less than a quarter (Table 1). After use of the “CoordinateCleaner” package, 280,935 C4 eudicot records (27.75%) of 1,012,557 were removed (Table 1). A manual check for erroneous synonymization removed 153,236 occurrence records from the remaining dataset. In the next step, 155,481 duplicate coordinates were removed. Notable is the additional reduction of 175,700 distribution points, after filtering out the outliers. After all these cleaning steps, we retained 963 species with a total of 247,205 occurrences. 620 species were represented by more than 10 records, whereas 343 species were represented by less than 10 records. Altogether, 75.59% of the occurrences for the C4 eudicots were excluded (Table 1).

**(---Table 1---)**

In the raw C4 grasses dataset, that contained 2,296,101 distribution points for 271 C4 grass genera (GBIF.org.2020p; GBIF.org.2020q; GBIF.org.2021r; GBIF.org.2020s; GBIF.org.2020t), around 382,595 points (16.66%) were removed after applying the “CoordinateCleaner” package. Excluding occurrence points outside the original distribution areas resulted in the elimination of 375,560 distribution points. Duplicate coordinates per species were removed, resulting in an additional 464,424 occurrence points being excluded. After all cleaning steps, approximately 53.25% distribution points were removed from the raw C4 grasses dataset, leaving 1,073,522 distribution points. Manual cleanup of incorrectly synonymized species was not performed, as only the genus level was considered.

Intermediate analyses with uncleaned or only partly cleaned data showed that these datasets would have led to different results (Supplementary Figures A1-15 illustrate this). The false occurrence data is prevalent to the extent that they blur any meaningful result of the clean data. The usability and consequently the sustainable success of large data repositories such as GBIF will thus in the future largely depend on the effort put into the curation of the data. Currently, these meta data should only be used with caution (Zizka et al., 2020), and a meaningful dataset can only be extracted via several filtering steps, as seen in this study.

**3.2 C4 eudicot and C4 grasses comparison**

Species richness is a commonly used measure of biodiversity (Gould, 2000; Albrecht et al., 2021). Richness maps are used to explore patterns of richness and help to investigate the processes that shape those patterns. Species richness maps of C4 grasses and C4 eudicots show the generally higher species diversity of C4 grasses (Figures 2a and b).

Overall, two regions stood out with a high C4 species richness (eudicots and grasses): Mexico and Australia (Figures 2a and b). For the C4 eudicots the hotspot of a high diversity a generic and family rank was in Mexico and extended further north into the United States, where deserts and xeric shrublands prevail, and the Australian hotspot lied in the deserts and xeric shrublands of Central Australia, but extended also into the (sub)tropical region in the north and the Mediterranean region in the west.

C4 grasses showed four diversity hotspots: 1) the tropical and subtropical open coniferous forests, as well as the adjacent deserts of thorn scrubs with fleshy plants and pastures at slightly higher elevations of Mexico, where temperate to semi-arid climate prevails; 2) the tropical and subtropical grasslands and shrublands of Queensland, Australia; 3) South Africa, dominated by montane grasslands and shrublands; and 4) West Africa in tropical and subtropical grasslands, savannas and shrublands, near the Volta-, Benue- and Niger-River. In addition, small islands such as the Galapagos and Cape Verde Islands, as well as Réunion and Mauritius, have been colonised by many C4 grass species.

The normalised difference of C4 eudicots versus C4 grass species occurrence revealed the particular importance of deserts (Namib, Caatinga) and the temperate northern hemisphere for C4 eudicots (Figure 2d).

**(--Figure 2--)**

Both, C4 grasses as well as C4 eudicots, occurred in a wide range of annual mean temperatures from 1 to 31.2 °C and 6 to 30.5 °C (95% interval), respectively (Figures 3 and 5b). The median temperature was 19.0 °C for C4 grasses and 17.4 °C for C4 eudicots. C4 grasses, in addition to increasing in the range of 15 to 18 °C, had a second steep increase in occurrence records ranging between 27 to 30 °C, dominated by the subfamilies Chloridoideae and Panicoideae. That last peak could not be observed in the C4 eudicots. Occurrence points of C4 grasses were found in a broad niche of annual mean precipitation profiles, from 0 to approx. 2,000 mm (95% interval). However, the predominant occurrence of C4 grasses tended to be in the semi-humid areas. An increase in occurrence points was seen in the range between 600 to 900 mm, with the median of 772 mm. C4 eudicots, on the other hand, occurred in distinctly less precipitation areas, with a median of 394 mm (Figures 3 and 5a). An increase of C4 eudicot occurrence points was observed in regions with approx. 300 to 600 mm precipitation/year (Figure 3). On a per continent basis the occurrence of C4 grasses and C4 eudicots differed most prominently in Europe and Africa (Supplementary Figure A16). In Africa C4 eudicots show a higher density in areas with less than 500 mm precipitation, especially in regions with cooler temperatures. In Europe, C4 eudicots show a higher density in areas with less than 400 mm precipitation and warm temperature.

Overall, the diversity and abundance of C4 plants increased with increasing annual mean temperature and dry season, and decreased with increasing cold temperatures and humidity. For C4 grasses there was a trend to more humid areas than in C4 eudicots. Cool and dry conditions are primarily colonised by C4 eudicots.

**(--Figure 3--)**

**3.3 Diversity of C4 Eudicots**

Mapping C4 occurrence points at family and genus level revealed many C4 eudicots hotspots of high taxonomic diversity at higher ranks with C4 species from greater than or equal to seven families and greater than ten genera occurring in the same area (Figures 2c and 4c). These hotspots were Mexico/Southern United States, Australia, South and West Africa, and South America (Figure 4c). In South America, the hotspot was located in the montane grasslands and shrublands of Argentinian Patagonia, whereas the Australian hotspot at the family level expanded into the tropical and subtropical grasslands, and shrublands.

At the genus richness of C4 eudicots, additional diversity hotspots were retrieved: in Asia the temperate grasslands, savannas and shrublands and the Altai-Sayan mountain range and in Europe in the Mediterranean shrublands (Figure 2c).

**(--Figure 4--)**

Amaranthaceae, Asteraceae, Euphorbiaceae, Portulacaceae and Zygophyllaceae are the five eudicot families with the highest numbers of C4 species (Figure 4a). While C4 Amaranthaceae showed high species richness in many different biomes, the major biomes of C4 eudicots were tropical and subtropical moist and dry broadleaf forests (Biome 1 & 2), tropical, subtropical and temperate grasslands, savannas and shrublands (Biome 7 & 8), and deserts and xeric shrublands (Biome 13).

We found statistically significant differences of the number of records with precipitation and temperature between most C4 eudicot families (Krustal-Wallis: for annual mean precipitation – chi-squared = 24810, df = 14, p-value < 2.2e-16; for annual mean temperature – chi-squared = 23980, df = 14, p-value < 2.2e-16) and between C4 eudicots and C4 grasses (Mann-Whitney-U Test: for annual mean precipitation – p-value < 2.2e-16; for annual mean temperature – p-value < 2.2e-16). These results point towards a wide adaptation range to diverse environmental conditions within the families and both groups (Figure 5).

Focusing on the five families (Amaranthaceae, Asteraceae, Euphorbiaceae, Portulacaceae and Zygophyllaceae) considered as species-rich per grid in Figure 4, we noticed that all five families differed significantly (p-value of <0.01) in their annual mean temperature and precipitation range. Amaranthaceae had the lowest mean value of 17.02 °C of these five families. Furthermore, the mean value of annual precipitation in Amaranthaceae was 458 mm. The mean value of the preferred annual mean temperature for Asteraceae was 19.74 °C, and annual precipitation of 702 mm, occurring in more humid and warmer areas compared to the C4 Amaranthaceae. C4 species of Euphorbiaceae preferred rather humid areas (mean = 760 mm) with a wide interquartile range and temperatures that intersect with the preferred areas of Amaranthaceae and Asteraceae (mean = 18.11 °C). A similar pattern was retrieved for C4 species of the Portulacaceae. Their preferred temperature lied between the values of Amaranthaceae and Asteraceae (mean = 18.48 °C) and in comparison to the other four families, the C4 species of Portulacaceae occurred in areas with more precipitation (mean = 790 mm). Zygophyllaceae was the family among those five families that preferred the warmest areas (mean = 20.04 °C) and their mean value of the preferred annual precipitation was second lowest at 523 mm.

Families with only one C4 lineage (genus) stood out among the others. In Polygonaceae, only one C4 genus *Calligonum* occurring in the cold deserts of Central Asia, was particularly conspicuous in climatic preferences with the lowest mean value of 14.11 °C and a precipitation preference in the very dry range (mean = 142 mm). A different picture was observed in C4 species of the family Caryophyllaceae which also contains only one C4 genus, *Polycarpaea*. These C4 species preferred comparatively warmer (mean = 26.21 °C) and more humid regions (mean = 761 mm).

**(--Figure 5--)**

**3.4 Traits**

A total of 394 (39% of total) succulent C4 species were recorded in eight eudicot families (Aizoaceae, Amaranthaceae, Caryophyllaceae, Gisekiaceae, Polygonaceae, Portulacaceae, Molluginaceae, Zygophyllaceae). These are distributed around a mean value of annual precipitation of 444 mm, which is higher than the mean for all C4 eudicots. Salt tolerance is documented in seven families with about 485 (49%) species, 333 (33%) of which are succulents from five families (Aizoaceae, Amaranthaceae, Polygonaceae, Portulacaceae, Zygophyllaceae). The biochemical subtypes were also examined. 459 (46%) investigated species within eleven families were retrieved to have the NAD-malate enzyme as the predominant decarboxylase. In contrast, 559 (56%) species within ten families exhibited the NADP-ME biochemical subtypes.

Most C4 eudicot species showed the classical atriplicoid leaf anatomy without or only little accompanied water storage tissue (Figure 6c). This anatomy with minor differences was also predominant in C4 grasses (Edwards & Voznesenskaya, 2011). In cases of succulence, there was often a deviation from atriplicoid anatomy. A high diversity in the succulent leaf anatomy of C4 eudicots was observed, with most C4 leaf types occurring only in a few species (Figure 6). However, the salsoloid leaf anatomy was clearly the most common leaf type among succulent C4 eudicots, not only in the Amaranthaceae, but also in Aizoaceae and Polygonaceae (Table A1). Among the succulent species, ca. 3% were stem succulents, while the rest were leaf succulents. The diversity of the leaf anatomy was closely linked to succulence and non-succulent leaves were usually atriplicoid. Eudicot C4 species with NAD-ME as the primary decarboxylating enzyme are distributed in areas with significantly lower annual precipitation than NADP-ME subtype eudicot C4 species (annual mean precipitation of 405.83 mm versus 698.39 mm).

**(--Figure 6--)**

**4. Discussion**

We characterised the global occurrence of C4 eudicots, identified diversity hotspots and climatic preferences, and assigned these to specific functional plant traits. We furthermore aimed at comparing patterns found for the eudicots to those of C4 grasses. The database attached to this paper (Supplementary Table A1) includes the current knowledge of physiological and morphological traits underlying large-scale patterns for C4 eudicots. Our approach combines the thus far known worldwide distribution of C4 eudicots, the evolution of photosynthesis and associated traits and climatic preferences of the individual lineages.

So called “big data”, as in our case with many distribution points of many different species from GBIF, are valuable resources for meta-analyses and can provide novel insights. However, accurate and elaborate cleaning of the data is essential to obtain meaningful results (Zizka et al., 2019). Another challenge is that in a project with the scope presented here a verification of all the identifications is unrealistic if not close to impossible. Also, we have to take into consideration the sampling density bias of Europe, North America and Australia over large areas of poorly sampled areas of Africa and Asia and thus have to interpret our findings applying to these regions with caution. These facts additionally underline the general need of well-curated data in our biodiversity repositories if we want them to be used by a broad community of researchers and to provide useful data for meta-analyses.

**4.1 Similarities and differences in the global distribution of C4 grasses and C4 eudicots and their climatic and biome preferences**

Both, C4 grasses and C4 eudicots, include multiple C4 lineages. While the C4 grasses are distinctly richer in species and all except two belong to the order Poales, the C4 eudicots belong to seven angiosperm orders (spanning across basal eudicots, rosids and asterids; Figure 1). Nevertheless, this is not indicative of differences in evolutionary age of C4 photosynthesis. In both, C4 grasses and C4 eudicots, the age of C4 origins spans from the Oligocene to the Pleistocene with most origins dating back to the Late Miocene (see Sage, 2017 for a summary), an era of global expansion of C4 vegetation due to declining atmospheric CO2 combined with global cooling and increase in climate seasonality and aridity (Cerling et al., 1997; Wen et al., 2023).

The global maps of C4 species richness (Figures 2a and b) reveal two contemporary hotspots for C4 grasses and C4 eudicots: one in Mexico/Southern United States and one in Australia. In these two regions C4 lineages seem to have diversified more intensively than in other parts of the world and in case of the C4 eudicots, this diversity has been recruited independently from multiple families and even multiple times within one family (Figures 2c and 4c; Table 2). Two additional hotspot regions for the C4 grasses are found in South and West Africa. However, since Africa is generally poorly-sampled, these two regions might appear as C4 grasses hotspots due to being proportionally more densely sampled. The distribution maps showing the diversity at the genus (Figure 2c) and family level (Figure 4c), reveal further smaller hotspots for the C4 eudicots, in South and West Africa (as in grasses), Patagonia, Central Asia and the Mediterranean. In these regions diversification within genera is less prominent, but multiple C4 eudicot lineages evidently colonised these areas as well (Figures 2c and 4c).

Whether these regions of high diversity of C4 lineages represent areas of C4 origin or were just preferably colonised by already existing C4 lineages or both needs to be evaluated for each C4 lineage and region in a phylogenetic and biogeographical context (see chapter 4.2).

Generally, both C4 grasses and C4 eudicots showed broad climatic ranges. On average C4 grasses occurred in only slightly warmer but distinctly more humid areas than C4 eudicots (Figure 3), in other words C4 grasses tend to be restricted to areas with proportionally high seasonal precipitation but are less susceptible to different temperature profiles. C4 eudicots colonise predominantly dry to very dry areas with 80% of the occurrences in areas with less than 800 mm precipitation. The occurrence of C4 grasses increases in regions with an annual rainfall of around 800 mm (Figure 3) and warm temperatures coupled with high insolation – conditions common in the southern hemisphere (Still et al., 2014). Where these climatic conditions are met C4 grasses not only tend to show a high species diversity but often also dominate the vegetation, especially when fires occur regularly. For example, the C4 grasslands in the highveld of southern Africa are dominated by *Hyparrhenia hirta* (L.) Stapf (Panicoideae) and *Sporobolus pyramidalis* P.Beauv. (Chloridoideae). These grasslands receive an annual mean precipitation between 400 - 900 mm mainly during the warm summer months (Low & Rebelo, 1996; Mills & Cowling, 2006; Bond, 2008). Another example are the species-rich C4 grasslands at the tropical Sudanian savanna near the Volta-, Benue- and Niger-River which experience a peak of summer precipitation of 600 mm in the north and 1,000 mm in the south. Here, the West African monsoons that occur between June and August result in warmer and more humid summers that support C4 vegetation (Olusegun et al., 2018). A third example are the species-rich C4 grasslands in north-east Queensland and the Northern Territory, Australia, which are associated with the tropical to subtropical climate along the coastal strip, the warm and humid summer months (December-February) and the Australian monsoon bringing up to 1,300 mm rainfall (Ondei et al., 2016).

In contrast, C4 eudicots rarely dominate the vegetation, with exception of some habitats that hamper plant growth by extreme saline and/or arid conditions and are therefore only colonised by relatively few species adapted to these conditions. These often include xerophytic (and often also halophytic) species of *Anabasis*, *Atriplex*, *Calligonum*, *Haloxylon* and *Tecticornia*. While in their diversity hotspots some of these taxa are more abundant than others, the extent does not reach the dimensions of subtropical grasslands. One prominent example are the cold deserts of Eurasia, with temperatures below the freezing point for an extensive period of time throughout the year (Winter, 1981; Johnston, 1996; Rudov et al., 2020). These are dominated by woody (sub)shrubs *Haloxylon persicum* and *H. ammodendron* as well as *Calligonum aphyllum*, *C. mongolicum* and *Anabasis brevifolia*. All five C4 species represent an integral part of the cold desert vegetation and no closely related C3 relatives are known from either of these genera (Flora of China; Kürschner, 2004; Lauterbach et al., 2019). Less prominent C4 floral elements of cold Central Asian deserts include species-poor genera such as *Horaninovia*, *Iljinia*, *Nanophyton*, *Piptoptera*, *Pyankovia*, *Turania* and *Xylosalsola*. These are together with *Anabasis* and *Haloxylon* all members of an evolutionary old C4 lineage within Salsoloideae (Amaranthaceae) that likely spread into the cold desert areas several times independently (Kadereit et al., 2012; Akhani et al., 2007). The biogeographically most comprehensively studied genus among these is *Anabasis*, which revealed the adjacent hot deserts of the Irano-Turanian Provinces as source areas for the species occurring in the cold deserts of the Mongolian Province (Lauterbach et al., 2019). Another example of vegetation dominated by C4 eudicots are the hot deserts of Central Australia where certain species of the large genus *Atriplex* such as *A. holocarpa*, *A. lindleyi* and *A. vesicaria* are highly abundant (Wilson, 1984).

In terms of preferred biomes, our analyses revealed that C4 grasses are most common in (sub)tropical grasslands, savannas and shrublands, whereas the highest species diversity of C4 eudicots was recorded from deserts and xeric shrublands (Figure 4). Both, C4 grasses and C4 eudicots, are fairly common even in (sub)tropical moist broadleaf forests. All these biomes, except the latter, are demarcated by scarcity and/or seasonality of precipitation. For example, one part of C4 grasses occur in the rainforests of the Australasian realm such as tropical and Central Range montane rainforests of Queensland. Another part occurs in the Indomalayan realm, such as Borneo lowland rainforests, Kayah-Karen and Sri Lanka montane rainforests and South Taiwan monsoon rainforests. Expectedly, both C4 groups are scarcer in water-rich and cooler biomes probably because the C4 syndrome is less advantageous and C3 species are more competitive. In regions where C3 trees dominate, the C4 syndrome might be a disadvantage due the limitations of the higher ATP-demand of this pathway in shady habitats. Within C4 eudicots, Amaranthaceae are ecologically the most diverse and are the only C4 eudicot clade found at higher latitudes in boreal forests.

**4.2 Did C4 lineages originate in the diversity hotspots or did C4 lineages colonise these areas?**

**C4 hotspot Mexico/Southern United States**

The C4 hotspot in Mexico and the Southern United States encompasses mostly deserts and xeric shrublands with different climatic regimes. This southern tip of the Nearctic realm comprises warm deserts such as Mojave Desert, Sonoran Desert, Chihuahuan Desert as well as cold deserts of the Great Basin, and large adjacent and equally diverse semi desert areas (Laity, 2008). This climate and habitat diversity probably promoted speciation, making these areas particularly species-rich. About two thirds of the flora is endemic to this region and the most common plant families represented in this flora are Cactaceae, Asteraceae and Boraginaceae (Villarreal-Quintanilla et al., 2017). The majority of endemic species and also the most widely distributed species that typify these landscapes such as *Ambrosia monogyra*, *Artemisia filifolia* and *Flourensia cernua* (Asteraceae), *Ephedra torreyana* (Ephedraceae), *Larrea tridentata* (Zygophyllaceae), *Penstemon thurberi* (Plantaginaceae), *Poliomintha incana* (Lamiaceae), *Prosopis glandulosa* and *Psorothamnus scoparius* (Fabaceae) and *Yucca elata* (Asparagaceae, Shreve, 1939), however, do not perform C4 photosynthesis. Nevertheless, here we recorded many C4 species from 17 eudicot genera belonging to eight different families (Table 2), supporting the findings of Sage et al. (2011) that Mexico and the Southern United States are a hotspot of C4 lineage diversity. For at least six of these lineages the current molecular phylogenies deliver sufficient evidence for in-situ origin of C4 photosynthesis within this area (Table 2).

One is the neotropical genus *Pectis* (Tageteae, Asteraceae) which includes approximately 90 C4 species and is represented by about 47 species in Mexico and the Southern United States (Hansen et al., 2016). The sister genus *Porophyllum* performs C3 photosynthesis and is also distributed in tropical and subtropical America. Hansen et al. (2016) show that the transition to C4 photosynthesis occurred most likely during the Late Miocene in the stem lineage of *Pectis* which was probably distributed in North/Central Mexico. Another example from Asteraceae is *Flaveria*. This taxon comprises 21 species, mainly distributed in Southern North America, with few species occurring in the Caribbean and South America (Powell, 1978). The phylogenetic tree topology suggests that the C4 pathway originated in this area, likely during the Pliocene (Morales-Briones & Kadereit, 2023). Due to the phenotypic variation of photosynthetic phenotypes, including intermediate forms, *Flaveria* became a model of C4 evolution (Monson & Moore, 1989; Sage et al., 2014). The C4 genera of Nyctaginaceae, *Boerhavia* and *Allionia*, belong to the “North American xerophytic clade” of the family (Douglas & Manos, 2007; Khoshravesh et al., 2020). This clade likely diversified in the deserts of the southwestern United States and northwestern Mexico because all genera are either confined to or represented in the area (Douglas & Manos, 2007). *Boerhavia* subsequently spread and diversified in subtropical regions worldwide. The situation in the species-rich *Euploca* (Boraginaceae) is challenging to assess as the published molecular phylogeny lacks support along the backbone (Frohlich et al., 2022). Nevertheless, an in-situ origin of the North America C4 species seems likely. Within the mega-diverse family Euphorbiaceae the evolution of carbon concentrating mechanisms led to diversification bursts (Horn et al., 2014). The C4 pathway evolved only once within the subgenus *Chamaesyce* at the stem of section *Anisophyllum* subsection *Hypericifoliae* during the Mid Miocene and gave rise to approximately 350 C4 species, which constitutes the largest eudicot C4 lineage known thus far. Current studies point to North America as the most likely area of origin for this C4 lineage (Yang & Berry, 2011; Horn et al., 2014).

Amaranthaceae s.l. are represented in Mexico and the Southern United States with eight genera containing native C4 species (Table 2), six of which belong to the Amaranthoideae. While *Tidestromia* consists entirely of C4 species all but one endemic to the region (Sánchez-del Pino & Motley, 2010), the closely related, species-rich *Alternanthera* is primarily distributed in the South American tropics from where the North American species might have originated. Unfortunately, a well resolved phylogeny of this large genus that also includes C3 species and C3-C4 intermediates is missing (Sage et al., 2007; Sánchez-del Pino et al., 2012). *Froelichia*, *Guilleminea* and *Gomphrena* belong to a species-rich, and widespread C4 clade that probably started to diversify during the Mid Miocene (Limarino & Borsch, 2020). Due to insufficient sampling, it is currently impossible to infer whether the C4 pathway originated in tropical South America or in subtropical southern North America. Insufficient phylogenetic information also prevents us from inferring the origin of North American *Amaranthus* species (Waselkov et al., 2018). However, since the entire genus exhibits C4 and probably originated in South America, *Amaranthus* seems to be a migratory C4 lineage in Mexico and the Southern United States. Other migratory C4 lineages are *Trianthema* which spread into the area from Africa (Bohley et al., 2015), *Atriplex* which arrived from South America (Žerdoner Čalasan et al., 2022), *Portulaca* (Ocampo & Columbus., 2012; Tamboli et al., 2022) and *Suaeda* (Schütze et al., 2003). The biogeography of the C4 genus *Kallstroemia* which is distributed from Central and Southern North America to tropical and subtropical South America remains unclear due to limited phylogenetic support (Lauterbach et al., 2019).

Nowadays the C4 hotspot in Mexico and the Southern United States receives a limited amount of rain, ranging from around 50 to 250 mm per year (Pearcy & Ehleringer, 1984), due to rain shadow casted by the mountain ranges of Sierra Madre Occidental and Sierra Nevada, and Sierra Madre Oriental on either side, which are of Late Mesozoic and Early Cenozoic age (Dickinson, 2004). The earliest evidence of desertification of this area dates to the Middle Miocene and corresponds with the diversification events of arid-adapted lineages (Eronen et al., 2012; Hyland et al., 2019; Vásquez‐Cruz & Sosa, 2020; Said Gutiérrez-Ortega et al., 2018). This refers also to some of the in-situ originated C4 lineages, such as *Pectis*, *Flaveria* (both Asteraceae), and *Allionia* and *Boerhavia* (both Nyctaginaceae) which have originated and spread since the Mid to Late Miocene (Table 2). We suggest that the overall high diversity of ancestral C3 lineages in the area adapted to arid conditions in addition to the high selective pressure in favour of the evolution of a carbon concentration mechanism is responsible for the exceptionally high diversity of C4 lineages that originated in Mexico and the Southern United States. In addition to these in-situ C4 lineages a high number of migratory C4 lineages occur finding suitable growing conditions in the area.

**C4 hotspot Australia**

Within Australia two regions of high C4 plant diversity with different precipitation profiles are observed. The first one constitutes deserts and xeric shrublands of the Eremaean floristic region (sensu Ebach et al., 2015), rich in C4 *Atriplex* (about 60 species) but also in C3 Camphorosmeae (about 150 species) and *Chenopodium* (about 50 species; Kadereit et al., 2005). Most other Australian C4 eudicots are restricted to the northern parts of the continent where tropical and subtropical grasslands, savannas, and shrublands prevail. Here the biggest C4 genera are *Gomphrena* (Amaranthaceae) with 30 C4 species, followed by *Euphorbia* (Euphorbiaceae) with seven, and *Portulaca* (Portulacaceae) and *Polycarpaea* (Caryophyllaceae) with five C4 species each.

While there are many different C4 plant lineages known from these areas, the majority of them did not evolve in-situ (Table 2), albeit little direct evidence is available, due to the lack of robust phylogenies. C4 *Atriplex* lineages reached Australia at least two times independently – once from the Mediterranean/Pontic region at the end of the Miocene and once from Central Asia at the end of the Pliocene (Žerdoner Čalasan et al., 2022). A similar pattern can be inferred from Australian C4 Aizoaceae (*Trianthema* and *Zaleya*) with Africa as their source area (Bohley et al., 2015). The spatial and temporal aspects of other C4 representatives of Australian flora (*Euploca, Glossocardia*, *Gomphrena*, *Polycarpaea*, *Portulaca* and *Tribulus*) remain unclear. Limited data, however, points towards ex-situ evolution of the C4 syndrome in these genera.

**(--Table 2--)**

The only clear C4 in-situ origin is currently known from *Tecticornia* (Shepherd et al., 2005; Voznesenskaya et al., 2008). This taxon is adapted to hyper-saline conditions and builds extensive vegetation stands along the edges of Australian inland salt lakes (Shepherd et al., 2004). This genus comprises about 60 species, out of which only a clade of five taxa is known to perform C4 photosynthesis. While many C3 species have rather restricted distribution areas (which may or may not be result of lack of surveys in poorly accessible Australian outback), one of the two C4 species *Tecticornia indica* exhibits a wide distribution range along saline lake shores across the whole continent (Wilson, 1984). Another, albeit less clear example is the small Australian genus *Tribulopis*. Conflicting phylogenetic signals between the nuclear and chloroplast-encoded genes point towards a complex evolutionary history of this taxon (Lauterbach et al., 2019). Contrarily to *Tecticornia*, here the C3 representatives show a wider distribution range, whereas the C4 species tend to be geographically restricted (Wilson, 1984). While reasons for this peculiar distribution remain unknown, this example clearly indicates that the factors promoting the evolution of C4 photosynthesis are multifold and that each individual C4 lineage has its own unique evolutionary history. Both taxa arrived to Australia post Miocene (Piirainen et al., 2017; Wu et al., 2018), which coincides with several geological and climate features that initiated and promoted aridification of this continent. These include the northward drift towards the equator, expansion of the Antarctic ice cap, and the formation of the circum-Antarctic Ocean current and subtropical high pressure system (Fujioka & Chappell, 2010; Kemp, 1978).

**Other C4 eudicot hotspots**

In addition to Mexico/Southern United States and Australia, evidence for in-situ origin of C4 in other diversity hotspots exists from molecular phylogenies.

Biodiversity in **Africa** has been notoriously understudied and as illustrated also from this study, large data gaps exist for its flora. Despite the scarce dataset, we were able to retrieve three regions that seem to favour C4 representatives. The first includes the tropical and subtropical grasslands, savannas, and shrublands of Africa. The second C4-rich region is located in south-east Africa. Here we should mention that the Drakensberg Mountain Centre located in this region is a known biodiversity hotspot, which may or may not influence the number of C4 species (Carbutt, 2019; Popp & Kalwij, 2021). The third region is in south-western Africa. the majority of species in the south-western region of Africa, however, perform C3 photosynthesis (Schulze & Schulze, 1976; Vogel & Seely, 1977; Fisher et al., 2015).

Most C4 eudicot species in these regions belong to Acanthaceae (*Blepharis* sect. *Acanthodium*), Aizoaceae, Gisekiaceae (genus *Gisekia*), Scrophulariaceae (genus *Anticharis*) and Zygophyllaceae (Table 2). The estimated centre of in-situ origin for C4 photosynthesis, with molecular phylogenetic evidence, in *Blepharis* (Fisher et al., 2015), *Anticharis* (Khoshravesh et al., 2012), *Gisekia* (Bissinger et al., 2014), Sesuvioideae (*Sesuvium/Trianthema/Zaleya;* Bohley et al., 2015) and the species *Zygophyllum simplex* (Zygophyllaceae; Bellstedt et al., 2012) appears to be the so called “Horn of Africa”, which defines the south of the Sahara-Sindian region, and the arid south-western Africa (Bellstedt et al., 2012 and references therein). Evidence for a south-west to north-east migration can be found in *Tetraena* (=*Zygophyllum*) *simplex* and C4 species of *Gisekia* (Bissinger et al., 2014; Bellstedt et al., 2012)*.* This migration could have occurred via the African arid corridor during the Miocene epoch in Africa (Verdcourt, 1969; Bellstedt et al., 2012).

In **South America**, two centres of C4 eudicot biodiversity are recognised, both under strong influence of arid desert or steppe climate with a pronounced dry period. These two centres of C4 diversity are most evident in *Euploca* (Boraginaceae) and *Portulaca* (Portulacaceae). However, they are also found in other eudicotyledons such as Amaranthaceae, Asteraceae and Euphorbiaceae. *Portulaca* has the highest C4 species diversity in north-eastern Brazil. C4 *Portulaca* clade started to radiate in the late Miocene/Pliocene (Ocampo & Columbus, 2012), following the expansion of C4 vegetation due to decreased CO2 levels and increased aridity (Salzmann et al., 2008; Strömberg, 2011). Nevertheless, the north-eastern portion of Brazil became consistently arid very recently in geological history – at the end of Younger Dryas (Auler et al., 2004). This leads us to believe that the high C4 diversity of *Portulaca* in that region is of refugial origin, as continuous more humid interglacial cycles prior to that diminished any advantage of C4 species over their C3 congeners. Poorly resolved phylogeny and lack of time divergence estimation preclude us from discussing potential stages in the evolutionary history of *Euploca* (Frohlich et al., 2022). For C4 *Atriplex* lineages molecular phylogenetic studies show two long distance dispersal events to reach South America – one possibly from continental Asia and one from North America (Žerdoner Čalasan et al., 2022).

It is important to mention that there are also regions, such as the **Central Asian Deserts**, that are not a main area of origin for C4 lineages, but have a high C4 lineage diversity due to a lot of migration. However, while the knowledge on the geological history of this region increased dramatically in recent years (Barbolini et al., 2020; Hurka et al., 2019), the evolutionary history of its flora remains largely unknown (Seidl et al., 2021; Žerdoner Čalasan et al., 2021).

**4.3 Functional traits lead to ecological diversity in C4 eudicots**

Precipitation and temperature preference among the C4 eudicot families differ significantly (Figure 5, Supplementary Figure A17). This points towards C4 evolution being advantageous under different environmental conditions and occurring along different spatiotemporal scales. These different evolutionary context dependencies together with the respective ancestral anatomical and physiological phenotypes are likely to have influenced the evolution of C4 and with it associated traits of the plant group. Whether C4 evolution improved fitness and evolutionary persistence of a lineage or as a key factor enhanced diversification by opening up new niches is lineage-specific and depends on C4 associated and other lineage-specific traits. Universal C4 associated traits, such as Kranz anatomy, therefore represent notable examples of phenotypic convergence across a wide range of functional, biochemical, and phylogenetic diversity. While it remains poorly understood how Kranz anatomy was initiated and how it arose in C4 plants (Schlüter & Weber, 2020), it required a variety of complex developmental changes and thus possibly represent a bottleneck to the C4 origin (Lauterbach et al., 2019). Hence, the high number of independent C4 origins and their morphological, ecological and physiological diversity in eudicot lineages is unexpected.

Kranz anatomy is an important unifying trait for almost all C4 species, and the majority of species show a similar (so called atriplicoid) C4 anatomy with Kranz cells (or bundle sheath cells) surrounding the vascular bundle and an outer ring of specialised mesophyll cells (Figure 6c). However, many eudicot C4 lineages deviate from this common anatomical type and show additional anatomical specialisations related mostly to leaf or stem succulence. All lineages that combine succulence and C4 photosynthesis seem to be derived from ancestrally succulent C3 lineages (e.g., *Tetraena simplex* in Zygophyllaceae (Lauterbach et al., 2016), C4 *Sesuvium* in Aizoaceae (Bohley et al., 2015), C4 *Tecticornia* in Amaranthaceae (Shepherd et al., 2005), and C4 lineages of Salsoleae, Camphorosmeae and Suaedeae (Kadereit et al., 2012). The annual mean precipitation for occurrence points of C4 eudicot succulents is 444.47 mm which does not indicate that the C4 succulents are confined to particularly dry areas. Non-succulent C4 eudicots such as Scrophulariaceae and Acanthaceae can be found in comparably dry (but not in saline) habitats. Sometimes succulence is lost within a C4 clade as found in *Bassia* (Kadereit et al., 2011). In the salt-tolerant C4 grass *Spinifex littoreus*, however, C4 evolution preceded that of succulence and salt-tolerance (Ho et al., 2019; Morrone et al., 2012).

Together with salt tolerance known from most of the succulent C4 eudicots (Santos et al., 2016), this combination of traits provides solutions to a broad variety of unfavourable climatic and edaphic conditions, which may explain the ecological diversity of C4 eudicots. In some succulent C4 eudicots the C4 pathway is combined with weak CAM which might provide a rescue mechanism under severe drought. The best studied example is *Portulaca oleracea* (Portulacaceae; Moreno-Villena et al., 2022), but a co-occurrence of CAM and C4 was also found in species of Sesuvioideae in Aizoaceae (Winter et al., 2021: *Trianthema* *portulacastrum*; Siadjeu and Kadereit, in prep.: *Sesuvium sesuvioides*) and might be more common than initially thought. As these studies require living specimens and a multiplexed experimental design to detect the potential co-occurrence of CAM and C4 the low number of reported cases might be due to lack of data.

According to our results, C4 eudicots with NADP-ME tend to be more prevalent in areas with significantly higher rainfall (mean value = 698.39 mm) than those utilising the NAD-ME as the decarboxylating enzyme (mean value = 405.83 mm; Figure 6d). This has been thus far supported by studies that comparatively investigated the physiology of NADP-ME and NAD-ME C4 grasses. While species with the NAD-ME subtype seem to perform better under drier condition (Ghannoum et al., 2002; Carmo-Silva et al., 2009), NADP-ME species tend to have a more effective CO2-capturing system by keeping stomata open longer under wetter conditions and can more efficiently utilise nitrogen (Schulze et al., 1996; Taub, 2000; Pinto et al., 2016). Within C4 eudicots the NAD-ME subtype seems more common than in C4 grasses where the NADP-ME pathway prevails (Sage et al., 2011). This might also explain why C4 eudicots managed to successfully occupy even the most arid regions around the globe, whereas C4 grasses are on average found in regions with significantly more rainfall (Figure 5a).

**4.4 A brief summarising family perspective of C4 evolution in eudicots**

The morphological, physiological and ecological diversity within C4 eudicots is immense. The most diverse plant family is **Amaranthaceae**, in which C4 photosynthesis developed several times independently in different environmental conditions (Kadereit et al., 2003, 2012; Sage et al., 2007). While in Salicornioideae (incl. Salsoloideae, Camphorosmoideae and Suaedoideae) and Chenopodioideae, C4 photosynthesis arose in ancestrally subtropical and temperate arid, and predominantly succulent and salt-tolerant lineages, C4 clades within Amaranthoideae more likely originated from tropical ancestors. Both NAD-ME as well as NADP-ME pathways are present in multiple lineages of this family and its C4 representatives can be found in most biomes, including the C4 less-favourable cool boreal zones. A number of rare features in the context of C4 photosynthesis are found only in Amaranthaceae. For example, the occurrence of single cell C4 in terrestrial angiosperms occurs only in the genus *Bienertia* and in *Suaeda aralocaspica* (both taxa belong to the tribe Suaedeae; Voznesenskaya et al., 2002). Furthermore, the shift from C3 photosynthesis in cotyledons to C4 in adult leaves has been so far only reported from various species of Salsoleae (Lauterbach et al., 2017). Two species of *Tecticornia* represent the only known stem-succulent C4 species with window cells in their mesophyll (Marchesini et al., 2014; Moir-Barnetson, 2014) and the *Salsola divaricata* agg. represents the first C4 lineages that was shown to have arisen from ancestral hybridization events between a C4 and a C3 lineage within Salsoleae (Morales-Briones & Kadereit, 2023).

Amaranthaceae are in terms of the C4 species richness followed by **Euphorbiaceae** and Asteraceae (Figure 1). In the former predominantly tropical plant family, the C4 syndrome evolved only once resulting in a clade of about 150 species found for the most part, but not exclusively, in seasonally dry and arid zones. In Hawaii, *Euphorbia* species occur in habitats ranging from arid coastal beaches to rainforests. One example is *Euphorbia clusiifolia* Hook. & Arn. (formerly *E. forbesii* Sherff, not accepted by Govaerts et al., 2000), a tree species that grows up to 13 m high and is endemic to the cool, mesic, subtropical forests in the geographically young Hawaiian Islands (Pearcy & Ehleringer, 1984). Further C4 trees include *Euphorbia* *olowaluana* and *E. remyi*, which occur in dry open and subalpine, and humid forests, respectively (Young et al., 2020). These examples strongly indicate that the C4 pathway is not limited to herbaceous or shrubby life forms but may also occur in trees, albeit these examples evolved woodiness secondarily from herbaceous ancestors on the island (Zizka et al., 2022). This link is particularly interesting since the secondary evolution of woodiness may in turn be linked to drought adaptation in other lineages (Dória et al., 2018, Hooft van Huysduynen et al., 2021). Many other C4 *Euphorbia* species occur in the understory rainforests of Hawaii, where the precipitation is rather high with an average range of 1,200 mm to 1,800 mm annually (Pearcy & Ehleringer, 1984) – conditions, in which C4 photosynthesis does not seem to have a physiological advantage.

In **Asteraceae** the C4 syndromeoriginated several times resulting in about 100 species that often sympatrically occur with the C3 congeners in dry and arid habitats. Due to the presence of many C3-C4 intermediate species, *Flaveria* became a model of C4 evolution (Monson & Moore, 1989; Sage et al., 2014). The species of the genus *Isostigma*, native to Bolivia, Brazil and North Argentina,show Kranz anatomy in the stems and two different types of Kranz anatomy in leaves (Peter, 2009). The *Eryngiophyllum* type, common in hot and arid conditions, has one Kranz unit per leaf and sclerenchyma tissue, whereas the *Isostigma* type, more common in places with higher humidity, shows more than one Kranz unit per leaf and no sclerenchyma tissue (Peter & Katinas, 2003). C4 representatives of Euphorbiaceae and Asteraceae rely solely on the NADP-ME pathway with the exception of *Euphorbia mongolica*, which can be found in highly seasonal temperate East Asia (Zang et al., 2021) and exhibits NAD-ME pathway.

Preliminary data suggests that C4 photosynthesis evolved only once in **Caryophyllaceae** in the genus *Polycarpaea* somewhen in the Pliocene (Kool, 2012). Apart from one widely distributed C4 species, all the other C4 species are restricted predominantly to either subtropical forests and (semi)arid zones of Australia or to the western Africa and adjacent regions. The highest species diversity of the C3 representatives are the more mesic Canary Islands. Low number of extremely xeric species might be explained by the presence of the NADP-ME mechanism, which seems less efficient under extremely arid conditions, at least in grasses (Rao & Dixon, 2016). In **Boraginaceae**, the C4 photosynthesis arose probably at least twice independently and while genera *Euploca* and *Heliotropium* are distributed world-wide, the C4 representatives of either genus are found predominantly in the seasonally dry and (semi)arid habitats around the globe. A similar pattern is observed also in C4 representatives of **Cleomaceae**, **Molluginaceae** and **Nyctaginaceae**, in all three of which the diversity of C4 species is markedly lower than that of C3 species.

While the majority of **Aizoaceae** rely on the CAM carbon fixation pathway, there are about 30 species that perform C4 photosynthesis. The details of anatomy and biochemistry vary greatly among closely related species with several origins of both NAD-ME and NADP-ME pathways (Bohley et al., 2015). The most widely distributed C4 species in this group under current taxonomic treatments is *Trianthema* *portulacastrum*, which can carry out CAM and C4 (Winter et al., 2021). While this physiological plasticity possibly facilitated its wide distribution, other widely distributed non-C4 (but also weak CAM; Winter et al., 2019) species such as *Sesuvium portulacastrum* show that C4, albeit beneficial, cannot solely explain species’ wide distribution. In monogeneric **Portulacaceae** about two thirds of currently known species from the genus *Portulaca* exhibit C4 anatomical features using NAD-ME and NADP-ME carbon fixation pathways in closely related lineages (Ocampo et al., 2013). Furthermore, many species also show reversible physiological signs of an effective CAM carbon fixation pathway under extreme drought stress (Holtum et al., 2017; Moreno-Villena et al., 2022). While this physiological plasticity may again explain the wide distribution range of *Portulaca oleracea* complex, its taxonomic uncertainty hampers our understanding of its evolutionary history and its true distribution extent (Ferrari et al., 2020).

While majority of the C4 eudicots occur in regions with a moderate drought period, **Polygonaceae**, **Scrophulariaceae** and **Zygophyllaceae** have the highest C4 species diversity in much drier arid and semi-arid regions. The shrubby species of *Calligonum* (Polygonaceae) possess only the NAD-ME mechanism for carbon fixation (more effective in drier regions), have a salsoloid Kranz type and are found only in cold and dry deserts of Eurasia and Africa (Sage et al., 2011; Sage, 2017; Pyankov et al., 2000, 2010; Muhaidat et al., 2012; Supplementary Figure A12b; Supplementary Table A1). In Africa it shares its distribution range with the only C4 taxon from Scrophulariaceae, *Anticharis*, whose four species are restricted to this warm region and exhibit NAD-ME mechanism for carbon fixation as well. In Zygophyllaceae the C4 photosynthesis developed several times independently and majority of the representatives are found in hot deserts around the world (Lauterbach et al., 2019). Interestingly, apart from one species (*Tetraena* (=*Zygophyllum*) *simplex*) all C4 species of Tribuloideae seem to possess the NADP-ME subtype. While it remains unclear why this is the case, *Tetraena simplex* is also the only species whose C4 anatomy resembles the kochioid Kranz type, while the rest of the taxa show atriplicoid Kranz type. This is not the case in other taxa, where NADP-ME and NAD-ME are not restricted to a particular Kranz anatomy.

C4 representatives of **Acanthaceae** and **Gisekiaceae** have a comparable distribution area and are found across seasonally dry to arid habitats of (sub)tropical Africa and south-western Asia. In both families the C4 syndrome arose probably only once (Fisher et al., 2015; Bissinger et al., 2014). While *Blepharis* (Acanthaceae) includes C3 representatives as well, and both NAD-ME and NADP-ME pathways are known from its C4 representatives, in *Gisekia* (monogeneric Gisekiaceae) all currently known species perform NAD-ME based C4 photosynthesis. Despite this restriction, both taxa are found along a wide range of open and often disturbed arid to mesic habitats (Bissinger et al., 2014).

**5. Conclusion/Outlook**

Does the combination of ecologically functional plant traits such as succulence or salt tolerance with C4 photosynthesis allow C4 eudicots to expand into areas not colonised by C4 grasses? Looking at the results of this research, we can say that C4 eudicots have an advantage over colonisation of other extreme habitats (e.g., deserts) due to their variety of different traits. C4 eudicots colonised significantly more arid regions than C4 grasses. Furthermore, C4 eudicot lineages utilising the NAD-ME decarboxylating enzyme are found in significantly drier areas than those using the NADP-ME decarboxylating enzyme. We identified two main hotspots of C4 eudicot diversity: arid regions of Mexico/Southern United States and Australia, where several C4 eudicot lineages diversified independently. Increasingly drier environmental conditions in Mexico/Southern United States and in large areas of Australia likely promoted diversification of C4 lineages. The evolutionary history of individual taxa in these regions, however, differs greatly. Mexico/Southern United States have a high number of C4 lineages that originated in-situ there, resulting in a higher C4 diversity a priori. Meanwhile, high C4 diversity in Australia and Central Asian deserts is predominantly a result of secondary migration of C4 lineages into these areas.

**Acknowledgements**

We thank the Global Biodiversity Information Facility for making occurrence data available, and all data collectors and contributors to GBIF for their efforts. We thank Colin Osborne for providing us with the list of C4 grass species. This work was supported by the DFG grant KA1816-7/3.

**Conflict of interest**

The authors declare no conflict of interest.

**Author contributions**

**Jessica A. Berasategui:** Conceptualization (equal); Data curation (lead); Formal analysis (lead); visualisation (lead); writing – original draft (lead); writing-review & editing (lead)

**Anže Žerdoner Čalasan:** writing – original draft (supporting); writing-review & editing (supporting)

**Alexander Zizka:** Conceptualization (supporting), Formal analysis (supporting), visualisation (equal), writing-review & editing (supporting)

**Gudrun Kadereit:** Conceptualization (equal); Formal analysis (supporting); writing – original draft (supporting); writing-review & editing (supporting)

**Data Availability Statement**

The data that support the findings of this study are available in the supplementary material of this article.

**ORCID**

**Jessica A. Berasategui (née Los):** https://orcid.org/0000-0002-0375-7397  
**Anže Žerdoner Čalasan:** https://orcid.org/0000-0003-2081-2076

**Alexander Zizka:** https://orcid.org/0000-0002-1680-9192

**Gudrun Kadereit:** https://orcid.org/0000-0003-0094-8769

**Appendix A. Supplementary data**

Supplementary data associated with this article can be found in the online version.

**Table A1:** Dataset of all C4 species occurring here in the study, arranged by family (15 eudicot families) and by C4 grasses including their occurrences, synonyms, trait data and the corresponding sources. The first sheet gives the references and the reference number.

**Figures A1-15:** Species richness maps showing the impact of data cleaning of all 15 families. (a) Map showing the uncleaned distribution points of the families. (b) Distribution points of the families after data cleaning.

**Figure A16:** Scatterplot showing annual mean temperature (°C) and annual mean precipitation (mm) data of C4 grasses (green) and C4 eudicot (blue) occurrence points per continent.

**Figure A17:** Boxplots showing the annual mean precipitation (mm) of all succulent C4 species separated by family.

**References**

Akhani, H., Barroca, J., Koteeva, N., Voznesenskaya, E., Franceschi, V., Edwards, G., Ghaffari, S.M. and Ziegler, H., 2005. *Bienertia sinuspersici* (Chenopodiaceae): a new species from Southwest Asia and discovery of a third terrestrial C4 plant without Kranz anatomy. *Systematic Botany*, *30*(2), pp.290-301. <https://doi.org/10.1600/0363644054223684>.

Akhani, H., Edwards, G. and Roalson, E.H., 2007. Diversification of the old world Salsoleae sl (Chenopodiaceae): molecular phylogenetic analysis of nuclear and chloroplast data sets and a revised classification. *International Journal of Plant Sciences, 168*(6), pp.931-956. <https://doi.org/10.1086/518263>.

Akhani, H., Chatrenoor, T., Dehghani, M., Khoshravesh, R., Mahdavi, P. and Matinzadeh, Z., 2012. A new species of *Bienertia* (Chenopodiaceae) from Iranian salt deserts: a third species of the genus and discovery of a fourth terrestrial C4 plant without Kranz anatomy. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, *146*(3), pp.550-559. <https://doi.org/10.1080/11263504.2012.662921>.

Albrecht, J., Peters, M.K., Becker, J.N., Behler, C., Classen, A., Ensslin, A., Ferger, S.W., Gebert, F., Gerschlauer, F., Helbig-Bonitz, M. and Kindeketa, W.J., 2021. Species richness is more important for ecosystem functioning than species turnover along an elevational gradient. *Nature Ecology & Evolution*, *5*(12), pp.1582-1593**.** <https://doi.org/10.6084/m9.figshare.14544207>.

Auler, A.S., Wang, X., Edwards, R.L., Cheng, H., Cristalli, P.S., Smart, P.L. and Richards, D.A., 2004. Quaternary ecological and geomorphic changes associated with rainfall events in presently semi‐arid northeastern Brazil. *Journal of Quaternary Science*, *19*(7), pp.693-701. <https://doi.org/10.1002/jqs.876>.

Barbehenn, R.V., Chen, Z., Karowe, D.N. and Spickard, A., 2004. C3 grasses have higher nutritional quality than C4 grasses under ambient and elevated atmospheric CO2. *Global Change Biology*, *10*(9), pp.1565-1575. <https://doi.org/10.1111/j.1365-2486.2004.00833.x>.

Barbolini, N., Woutersen, A., Dupont-Nivet, G., Silvestro, D., Tardif, D., Coster, P.M.C., Meijer, N., Chang, C., Zhang, H.X., Licht, A. and Rydin, C., 2020. Cenozoic evolution of the steppe-desert biome in Central Asia. *Science Advances*, *6*(41), p.eabb8227. <https://doi.org/10.1126/sciadv.abb8227>.

Bassham, J.A., Benson, A.A. and Calvin, M., 1950. The path of carbon in photosynthesis VIII. The role of malic acid. *Journal of Biological Chemistry*, *185*(2), pp.781-787. <https://doi.org/10.2172/910351.>

Bellstedt, D.U., Galley, C., Pirie, M.D. and Linder, H.P., 2012. The migration of the palaeotropical arid flora: Zygophylloideae as an example. *Systematic Botany*, *37*(4), pp.951-959. <https://doi.org/10.1600/036364412x656608>.

Bissinger, K., Khoshravesh, R., Kotrade, J.P., Oakley, J., Sage, T.L., Sage, R.F., Hartmann, H.E. and Kadereit, G., 2014. *Gisekia* (Gisekiaceae): phylogenetic relationships, biogeography, and ecophysiology of a poorly known C4 lineage in the Caryophyllales. *American Journal of Botany*, *101*(3), pp.499-509. <https://doi.org/10.3732/ajb.1300279>.

Bohley, K., Joos, O., Hartmann, H., Sage, R., Liede-Schumann, S. and Kadereit, G., 2015. Phylogeny of Sesuvioideae (Aizoaceae)–Biogeography, leaf anatomy and the evolution of C4 photosynthesis. *Perspectives in Plant Ecology, Evolution and Systematics*, *17*(2), pp.116-130. <https://doi.org/10.1016/j.ppees.2014.12.003>.

Bond, W.J., 2008. What limits trees in C4 grasslands and savannas? *Annual Review of Ecology, Evolution, and Systematics*, *39*, pp.641-659. <https://doi.org/10.1146/annurev.ecolsys.39.110707.173411>.

Bromham, L. and Bennett, T.H., 2014. Salt tolerance evolves more frequently in C4 grass lineages. *Journal of Evolutionary Biology*, *27*(3), pp.653-659. <https://doi.org/10.1111/jeb.12320>.

Burr, G.O., Hartt, C.E., Brodie, H.W., Tanimoto, T., Kortschak, H.P., Takahashi, D., Ashton, F.M. and Coleman, R.E., 1957. The sugarcane plant. *Annual Review of Plant Physiology*, *8*(1), pp.275-308. <https://doi.org/10.1146/annurev.pp.08.060157.001423.>

Carbutt, C., 2019. The Drakensberg Mountain Centre: A necessary revision of southern Africa's high-elevation centre of plant endemism. *South African Journal of Botany*, *124*, pp.508-529. <https://doi.org/10.1016/j.sajb.2019.05.032>.

Carmo-Silva, A.E., Francisco, A., Powers, S.J., Keys, A.J., Ascensão, L., Parry, M.A. and Arrabaça, M.C., 2009. Grasses of different C4 subtypes reveal leaf traits related to drought tolerance in their natural habitats: changes in structure, water potential, and amino acid content. *American Journal of Botany*, *96*(7), pp.1222-1235. <https://doi.org/10.3732/ajb.0800224>.

Cerling, T.E., Harris, J.M., MacFadden, B.J., Leakey, M.G., Quade, J., Eisenmann, V. and Ehleringer, J.R., 1997. Global vegetation change through the Miocene/Pliocene boundary. *Nature*, *389*(6647), pp.153-158. <https://doi.org/10.1038/38229>.

Chamberlain, S., Barve, V., Mcglinn, D., Oldoni, D., Desmet, P., Geffert, L. and Ram, K., rgbif: Interface to the Global Biodiversity Information Facility API. 2020. R package version 3.3.0, URL [https://CRAN.R-project.org/package=rgbif](https://cran.r-project.org/package=rgbif).

Christin, P.A., Besnard, G., Samaritani, E., Duvall, M.R., Hodkinson, T.R., Savolainen, V. and Salamin, N., 2008. Oligocene CO2 decline promoted C4 photosynthesis in grasses. *Current Biology*, *18*(1), pp.37-43. <https://doi.org/10.1016/j.cub.2007.11.058>.

Christin, P.A., Osborne, C.P., Sage, R.F., Arakaki, M. and Edwards, E.J., 2011. C4 eudicots are not younger than C4 monocots. *Journal of Experimental Botany*, *62*(9), pp.3171-3181. <https://doi.org/10.1093/jxb/err041>.

Christin, P.A. and Osborne, C.P., 2014. The evolutionary ecology of C4 plants. *New Phytologist*, *204*(4), pp.765-781. <https://doi.org/10.1111/nph.13033>.

Clayton, W.D., Vorontsova, M.S., Harman, K.T. and Williamson, H., 2002 onwards. World grass species: synonymy. *Royal Botanic Garden Kew*. URL <http://www.kew.org/data/grasses-syn.html>. [accessed 25 March 2021; 10:50 GMT]

Collins, R.P. and Jones, M.B., 1986. The influence of climatic factors on the distribution of C4 species in Europe. *Vegetatio*, *64*(2), pp.121-129. <https://doi.org/10.1007/bf00044788>.

Dickinson, W.R., 2004. Evolution of the North American cordillera. *Annual Review of Earth and Planetary Sciences*, *32*, pp.13-45. <https://doi.org/10.1146/annurev.earth.32.101802.120257>.

Dória, L.C., Podadera, D.S., del Arco, M., Chauvin, T., Smets, E., Delzon, S. and Lens, F., 2018. Insular woody daisies (*Argyranthemum*, Asteraceae) are more resistant to drought‐induced hydraulic failure than their herbaceous relatives. *Functional Ecology*, *32*(6), pp.1467-1478. <https://doi.org/10.1111/1365-2435.13085>.

Douglas, N.A. and Manos, P.S., 2007. Molecular phylogeny of Nyctaginaceae: taxonomy, biogeography, and characters associated with a radiation of xerophytic genera in North America. *American Journal of Botany*, *94*(5), pp.856-872. <https://doi.org/10.3732/ajb.94.5.856>.

Ebach, M.C., González-Orozco, C.E., Miller, J.T. and Murphy, D.J., 2015. A revised area taxonomy of phytogeographical regions within the Australian Bioregionalisation Atlas. *Phytotaxa*, *208*(4), pp.261-277. <https://doi.org/10.11646/phytotaxa.208.4.2>.

Edwards, E.J. and Smith, S.A., 2010. Phylogenetic analyses reveal the shady history of C4 grasses. *Proceedings of the National Academy of Sciences*, *107*(6), pp.2532-2537. <https://doi.org/10.1073/pnas.0909672107>.

Edwards, G.E. and Voznesenskaya, E.V., 2011. C4 Photosynthesis: Kranz Forms and Single-Cell C4 in Terrestrial Plants. Pages 29-61 in: *C4 Photosynthesis and Related CO2 Concentrating Mechanisms*. Raghavendra, A. S. and Sage, R. F. (eds.). Springer Netherlands, Dordrecht. <https://doi.org/10.1007/978-90-481-9407-0.>

Ehleringer, J., 1983. Ecophysiology of *Amaranthus palmeri*, a Sonoran Desert summer annual. *Oecologia*, *57*(1), pp.107-112. <https://doi.org/10.1007/bf00379568>.

Ehleringer, J.R., Cerling, T.E. and Helliker, B.R., 1997. C4 photosynthesis, atmospheric CO2, and climate. *Oecologia*, *112*(3), pp.285-299. <https://doi.org/10.1007/s004420050311>.

Eronen, J.T., Fortelius, M., Micheels, A., Portmann, F.T., Puolamäki, K. and Janis, C.M., 2012. Neogene aridification of the Northern Hemisphere. *Geology*, *40*(9), pp.823-826. <https://doi.org/10.1130/G33147.1>.

Ferrari, R.C., Cruz, B.C., Gastaldi, V.D., Storl, T., Ferrari, E.C., Boxall, S.F., Hartwell, J. and Freschi, L., 2020. Exploring C4-CAM plasticity within the *Portulaca oleracea* complex. *Scientific Reports*, *10*(1), pp.1-14. <https://doi.org/10.1038/s41598-020-71012-y>.

Ferrari, R.C., Kawabata, A.B., Ferreira, S.S., Hartwell, J. and Freschi, L., 2022. A matter of time: regulatory events behind the synchronisation of C4 and crassulacean acid metabolism in *Portulaca oleracea*. *Journal of Experimental Botany*, *73*(14), pp.4867-4885. <https://doi.org/10.1093/jxb/erac163>.

Fick, S.E. and Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), pp.4302-4315. <https://doi.org/10.1002/joc.5086>.

Fisher, A.E., McDade, L.A., Kiel, C.A., Khoshravesh, R., Johnson, M.A., Stata, M., Sage, T.L. and Sage, R.F., 2015. Evolutionary history of *Blepharis* (Acanthaceae) and the origin of C4 photosynthesis in section *Acanthodium*. *International Journal of Plant Sciences*, *176*(8), pp.770-790. <https://doi.org/10.1086/683011>.

Flora of China. Wu Z.Y., Raven P.H. and Hong D.Y., eds. 1994-2013. *Science Press, Beijing, China and Missouri Botanical Garden Press*, St. Louis, USA (available online at efloras.org). http://www.efloras.org/florataxon.aspx?flora\_id=2&taxon\_id=10717. [accessed 24 February 2023]

Fox, J., 2005. Getting started with the R commander: a basic-statistics graphical user interface to R. *Journal of Statistical Software*, *14*(9), pp.1-42. <https://doi.org/10.18637/jss.v014.i09>.

Freitag, H. and Stichler, W., 2000. A remarkable new leaf type with unusual photosynthetic tissue in a central Asiatic genus of Chenopodiaceae. *Plant Biology*, *2*(2), pp.154-160. <https://doi.org/10.1055/s-2000-9462>.

Frohlich, M.W., Sage, R.F., Craven, L.A., Schuster, S., Gigot, G., Hilger, H.H., Akhani, H., Mahdavi, P., Luebert, F., Weigend, M. and Thulin, M., 2022. Molecular phylogenetics of *Euploca* (Boraginaceae): homoplasy in many characters, including the C4 photosynthetic pathway. *Botanical Journal of the Linnean Society*, *199*(2), pp.497-537. <https://doi.org/10.1093/botlinnean/boab082>.

Fujioka, T. and Chappell, J., 2010. History of Australian aridity: chronology in the evolution of arid landscapes. *Geological Society, London, Special Publications*, *346*(1), pp.121-139. <https://doi.org/10.1144/sp346.8>.

Furbank, R.T. and Kelly, S., 2021. Finding the C4 sweet spot: cellular compartmentation of carbohydrate metabolism in C4 photosynthesis. *Journal of Experimental Botany, 72*(17), pp.6018-6026. <https://doi.org/10.1093/jxb/erab290.>

Ghannoum, O., Von Caemmerer, S. and Conroy, J.P., 2002. The effect of drought on plant water use efficiency of nine NAD-ME and nine NADP-ME Australian C4 grasses. *Functional Plant Biology*, *29*(11), pp.1337-1348. <https://doi.org/10.1071/fp02056>.

Ghazanfar, S.A., Altundag, E., Yaprak, A.E., Osborne, J., Tug, G.N. and Vural, M., 2014. Halophytes of southwest Asia. *Sabkha Ecosystems: Volume IV: Cash Crop Halophyte and Biodiversity Conservation,* pp.105-133. <https://doi.org/10.1007/978-94-007-7411-7_8>.

Gould, W., 2000. Remote sensing of vegetation, plant species richness, and regional biodiversity hotspots. *Ecological Applications*, *10*(6), pp.1861-1870. [https://doi.org/10.1890/1051-0761(2000)010[1861:rsovps]2.0.co;2](https://doi.org/10.1890/1051-0761(2000)010%5b1861:rsovps%5d2.0.co;2).

Govaerts, R., Frodin, D.G., Radcliffe-Smith, A. and Carter, S., 2000. World checklist and bibliography of Euphorbiaceae (with Pandaceae). *Royal Botanic Gardens, Kew*.

Grace, J., Jose, J.S., Meir, P., Miranda, H.S. and Montes, R.A., 2006. Productivity and carbon fluxes of tropical savannas. *Journal of Biogeography*, *33*(3), pp.387-400. <https://doi.org/10.1111/j.1365-2699.2005.01448.x>.

Griffiths, H., Weller, G., Toy, L.F. and Dennis, R.J., 2013. You're so vein: bundle sheath physiology, phylogeny and evolution in C3 and C4 plants. *Plant, Cell & Environment*, *36*(2), pp.249-261. <https://doi.org/10.1111/j.1365-3040.2012.02585.x>.

Hansen, D.R., Jansen, R.K., Sage, R.F., Villaseñor, J.L. and Simpson, B.B., 2016. Molecular phylogeny of *Pectis* (Tageteae, Asteraceae), a C4 genus of the Neotropics, and its sister genus Porophyllum. *Lundellia, 19*(1), pp.6-38. <https://doi.org/10.25224/1097-993x-19.1.6>.

Hatch, M.D., 2005. C4 photosynthesis: discovery and resolution. *Discoveries in Photosynthesis*, pp.875-880. <https://doi.org/10.1142/9789812817563_0013.>

Heyduk, K., Moreno-Villena, J.J., Gilman, I.S., Christin, P.A. and Edwards, E.J., 2019. The genetics of convergent evolution: insights from plant photosynthesis. *Nature Reviews Genetics*, *20*(8), pp.485-493. <https://doi.org/10.1038/s41576-019-0107-5>.

Ho, C.L., Chiang, J.M., Lin, T.C. and Martin, C.E., 2019. First report of C4/CAM-cycling photosynthetic pathway in a succulent grass, *Spinifex littoreus* (Brum. f.) Merr., in coastal regions of Taiwan. *Flora*, *254*, pp.194-202. <https://doi.org/10.1016/j.flora.2018.08.005>.

Holtum, J.A., Hancock, L.P., Edwards, E.J. and Winter, K., 2017. Optional use of CAM photosynthesis in two C4 species, *Portulaca cyclophylla* and *Portulaca digyna*. *Journal of Plant Physiology*, *214*, pp.91-96. <https://doi.org/10.1016/j.jplph.2017.01.010>.

Hooft van Huysduynen, A., Janssens, S., Merckx, V., Vos, R., Valente, L., Zizka, A., Larter, M., Karabayir, B., Maaskant, D., Witmer, Y. and Fernández‐Palacios, J.M., 2021. Temporal and palaeoclimatic context of the evolution of insular woodiness in the Canary Islands. *Ecology and Evolution*, *11*(17), pp.12220-12231. <https://doi.org/10.1002/ece3.7986>.

Horn, J.W., Xi, Z., Riina, R., Peirson, J.A., Yang, Y., Dorsey, B.L., Berry, P.E., Davis, C.C. and Wurdack, K.J., 2014. Evolutionary bursts in *Euphorbia* (Euphorbiaceae) are linked with photosynthetic pathway. *Evolution, 68*(12), pp.3485-3504. <https://doi.org/10.1111/evo.12534>.

Hurka, H., Nikolai, F., Karl-Georg, B., Barbara, N., Sergej, S., Alexander, S. and Frank, B., 2019. The Eurasian steppe belt: Status quo, origin and evolutionary history. *Turczaninowia*, *22*(3), pp.5-71. <https://doi.org/10.14258/turczaninowia.22.3.1>.

Hyland, E.G., Sheldon, N.D., Smith, S.Y. and Strömberg, C.A., 2019. Late Miocene rise and fall of C4 grasses in the western United States linked to aridification and uplift. *Bulletin*, *131*(1-2), pp.224-234. <https://doi.org/10.1130/B32009.1>.

IPNI. International Plant Names Index, 2020. https://beta.ipni.org/. (Accessed 15 August 2020).

Johnston, W.H., 1996. The place of C4 grasses in temperate pastures in Australia. *New Zealand Journal of Agricultural Research*, *39*(4), pp.527-540. <https://doi.org/10.1080/00288233.1996.9513213>.

Kadereit, G., Borsch, T., Weising, K. and Freitag, H., 2003. Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C4 photosynthesis. *International Journal of Plant Sciences*, *164*(6), pp.959-986. <https://doi.org/10.1086/378649>.

Kadereit, G., Gotzek, D. and Freitag, H., 2005. Origin and age of Australian Chenopodiaceae. *Organisms Diversity & Evolution*, *5*(1), pp.59-80. <https://doi.org/10.1016/j.ode.2004.07.002>.

Kadereit, G., Mavrodiev, E.V., Zacharias, E.H. and Sukhorukov, A.P., 2010. Molecular phylogeny of Atripliceae (Chenopodioideae, Chenopodiaceae): Implications for systematics, biogeography, flower and fruit evolution, and the origin of C4 photosynthesis. *American Journal of Botany*, *97*(10), pp.1664-1687. <https://doi.org/10.3732/ajb.1000169.>

Kadereit, G. and Freitag, H., 2011. Molecular phylogeny of Camphorosmeae (Camphorosmoideae, Chenopodiaceae): Implications for biogeography, evolution of C4‐photosynthesis and taxonomy. *Taxon*, *60*(1), pp.51-78. <https://doi.org/10.1002/tax.601006>.

Kadereit, G., Ackerly, D. and Pirie, M.D., 2012. A broader model for C4 photosynthesis evolution in plants inferred from the goosefoot family (Chenopodiaceae ss). *Proceedings of the Royal Society B: Biological Sciences*, *279*(1741), pp.3304-3311. <https://doi.org/10.1098/rspb.2012.0440>.

Kadereit, G., Newton, R.J. and Vandelook, F., 2017. Evolutionary ecology of fast seed germination - a case study in Amaranthaceae/Chenopodiaceae. *Perspectives in Plant Ecology, Evolution and Systematics*, *29*, pp.1-11. <https://doi.org/10.1016/j.ppees.2017.09.007>.

Kemp, E.M., 1978. Tertiary climatic evolution and vegetation history in the southeast Indian Ocean region. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *24*(3), pp.169-208. <https://doi.org/10.1016/0031-0182(78)90042-1>.

Khoshravesh, R., Akhani, H., Sage, T.L., Nordenstam, B. and Sage, R.F., 2012. Phylogeny and photosynthetic pathway distribution in *Anticharis* Endl.(Scrophulariaceae). *Journal of Experimental Botany, 63*(15), pp.5645-5658. <https://doi.org/10.1093/jxb/ers218>.

Khoshravesh, R., Stata, M., Adachi, S., Sage, T.L. and Sage, R.F., 2020. Evolutionary Convergence of C4 Photosynthesis: A Case Study in the Nyctaginaceae. *Frontiers in Plant Science*, *11*. <https://doi.org/10.3389/fpls.2020.578739.>

Kool, A., 2012. Desert Plants and Deserted Islands: *Systematics and Ethnobotany in Caryophyllaceae* (Doctoral dissertation, Acta Universitatis Upsaliensis). ISBN 978-91-554-8471-2.

Kürschner, H., 2004. Phytosociological studies in the Alashan Gobi-A contribution to the flora and vegetation of Inner Mongolia (NW China). *Phytocoenologia*, pp.169-224. https://doi.org/10.1127/0340-269x/2004/0034-0169.

Langdale, J.A., 2011. C4 cycles: past, present, and future research on C4 photosynthesis. *The Plant Cell*, *23*(11), pp.3879-3892. <https://doi.org/10.1105/tpc.111.092098>.

Lauterbach, M., van der Merwe, P.D.W., Keßler, L., Pirie, M.D., Bellstedt, D.U. and Kadereit, G., 2016. Evolution of leaf anatomy in arid environments-a case study in southern African *Tetraena* and *Roepera* (Zygophyllaceae). *Molecular Phylogenetics and Evolution*, *97*, pp.129-144. <https://doi.org/10.1016/j.ympev.2016.01.002>.

Lauterbach, M., Billakurthi, K., Kadereit, G., Ludwig, M., Westhoff, P. and Gowik, U., 2017. C3 cotyledons are followed by C4 leaves: intra-individual transcriptome analysis of *Salsola soda* (Chenopodiaceae). *Journal of Experimental Botany*, *68*(2), pp.161-176. <https://doi.org/10.1093/jxb/erw343>.

Lauterbach, M., Zimmer, R., Alexa, A.C., Adachi, S., Sage, R., Sage, T., MacFarlane, T., Ludwig, M. and Kadereit, G., 2019. Variation in leaf anatomical traits relates to the evolution of C4 photosynthesis in Tribuloideae (Zygophyllaceae). *Perspectives in Plant Ecology, Evolution and Systematics*, *39*, p.125463. <https://doi.org/10.1016/j.ppees.2019.125463>.

Laity, J.J., 2008. Deserts and desert environments (Vol. 3). *Wiley-Blackwell*, pp.37-43. <https://doi.org/10.1177/0959683609345799>.

Lehmann, C.E., Griffith, D.M., Simpson, K.J., Anderson, T.M., Archibald, S., Beerling, D.J., Bond, W.J., Denton, E., Edwards, E.J., Forrestel, E.J. and Fox, D.L., 2019. Functional diversification enabled grassy biomes to fill global climate space. *BioRxiv*, p.583625. <https://doi.org/10.1101/583625>.

Limarino, T.O. and Borsch, T., 2020. *Gomphrena* (Amaranthaceae, Gomphrenoideae) diversified as a C4 lineage in the New World tropics with specialisations in floral and inflorescence morphology, and an escape to Australia. *Willdenowia, 50*(3), pp.345-381. <https://doi.org/10.3372/wi.50.50301>.

Linder, H.P., Lehmann, C.E., Archibald, S., Osborne, C.P. and Richardson, D.M., 2018. Global grass (Poaceae) success underpinned by traits facilitating colonisation, persistence and habitat transformation. *Biological Reviews*, *93*(2), pp.1125-1144. <https://doi.org/10.1111/brv.12388>.

Low, A.B. and Rebelo, A.G., 1996. Vegetation of South Africa, Lesotho and Swaziland. *Department of Environmental Affairs and Tourism, Pretoria*. ISBN 0–621–17316–9.

Lundgren, M.R., Besnard, G., Ripley, B.S., Lehmann, C.E., Chatelet, D.S., Kynast, R.G., Namaganda, M., Vorontsova, M.S., Hall, R.C., Elia, J. and Osborne, C.P., 2015. Photosynthetic innovation broadens the niche within a single species. *Ecology Letters*, *18*(10), pp.1021-1029. <https://doi.org/10.1111/ele.12484>.

Mahdavi, P. and Bergmeier, E., 2018. Distribution of C4 plants in sand habitats of different climatic regions. *Folia Geobotanica*, *53*(2), pp.201-211. <https://doi.org/10.1007/s12224-018-9320-9>.

Maldonado, C., Molina, C.I., Zizka, A., Persson, C., Taylor, C.M., Albán, J., Chilquillo, E., Rønsted, N. and Antonelli, A., 2015. Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases? *Global Ecology and Biogeography*, *24*(8), pp.973-984. <https://doi.org/10.1111/geb.12326>.

Marchesini, V.A., Yin, C., Colmer, T.D. and Veneklaas, E.J., 2014. Drought tolerances of three stem-succulent halophyte species of an inland semiarid salt lake system. *Functional Plant Biology*, *41*(12), pp.1230-1238. https://doi.org/10.1071/FP14108.

Mills, A.J. and Cowling, R.M., 2006. Rate of carbon sequestration at two thicket restoration sites in the Eastern Cape, South Africa. *Restoration Ecology*, *14*(1), pp.38-49. <https://doi.org/10.1111/j.1526-100x.2006.00103.x>.

Moir-Barnetson, L., 2014. Ecophysiological responses to changes in salinity and water availability in stem-succulent halophytes(*Tecticornia spp.*) from an ephemeral salt lake (Doctoral dissertation, University of Western Australia).

Monson, R.K. and Moore, B.D., 1989. On the significance of C3-C4 intermediate photosynthesis to the evolution of C4 photosynthesis. *Plant, Cell & Environment*, *12*(7), pp.689-699. <https://doi.org/10.1111/j.1365-3040.1989.tb01629.x>.

Morales-Briones, D.F. and Kadereit, G. (in press): Exploring the possible role of hybridization in the evolution of photosynthetic pathways in *Flaveria* (Asteraceae), the prime model of C4 photosynthesis evolution. *Frontiers in Ecology and Evolution*. <https://www.biorxiv.org/content/10.1101/2022.01.31.478436v1>.

Moreno-Villena, J.J., Zhou, H., Gilman, I.S., Tausta, S.L., Cheung, C.M. and Edwards, E.J., 2022. Spatial resolution of an integrated C4+CAM photosynthetic metabolism. *Science Advances*, *8*(31), p.eabn2349. <https://doi.org/10.1126/sciadv.abn2349>.

Morrone, O., Aagesen, L., Scataglini, M.A., Salariato, D.L., Denham, S.S., Chemisquy, M.A., Sede, S.M., Giussani, L.M., Kellogg, E.A. and Zuloaga, F.O., 2012. Phylogeny of the Paniceae (Poaceae: Panicoideae): integrating plastid DNA sequences and morphology into a new classification. *Cladistics*, *28*(4), pp.333-356. <https://doi.org/10.7934/p942>.

Muhaidat, R., Sage, R.F. and Dengler, N.G., 2007. Diversity of Kranz anatomy and biochemistry in C4 eudicots. *American Journal of Botany*, *94*(3), pp.362-381. <https://doi.org/10.3732/ajb.94.3.362>.

Muhaidat, R., McKown, A.D., Khateeb, W.A., Al-Shreideh, M., Domi, Z. B., Hussein, E. and El-Oqlah, A., 2012. Full assessment of C4 photosynthesis in *Blepharis attenuata* Napper (Acanthaceae) from Jordan: evidence from leaf anatomy and key C4 photosynthetic enzymes. *Asian Journal of Plant Sciences*, *11*(5), p.206. <https://doi.org/10.3923/ajps.2012.206.216.>

Muhaidat, R., McKown, A.D., Al Zoubi, M., Bani Domi, Z. and Otoum, O., 2018. C4 photosynthesis and transition of Kranz anatomy in cotyledons and leaves of *Tetraena simplex*. *American Journal of Botany*, *105*(5), pp.822-835. <https://doi.org/10.1002/ajb2.1087>.

Niklaus, M. and Kelly, S., 2019. The molecular evolution of C4 photosynthesis: opportunities for understanding and improving the world’s most productive plants. *Journal of Experimental Botany*, *70*(3), pp.795-804. <https://doi.org/10.1093/jxb/ery416.>

Ocampo, G. and Columbus, J.T., 2012. Molecular phylogenetics, historical biogeography, and chromosome number evolution of *Portulaca* (Portulacaceae). *Molecular Phylogenetics and Evolution*, *63*(1), pp.97-112. <https://doi.org/10.1016/j.ympev.2011.12.017>.

Ocampo, G., Koteyeva, N.K., Voznesenskaya, E.V., Edwards, G.E., Sage, T.L., Sage, R.F. and Columbus, J.T., 2013. Evolution of leaf anatomy and photosynthetic pathways in Portulacaceae. *American Journal of Botany*, *100*(12), pp.2388-2402. <https://doi.org/10.3732/ajb.1300094>.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C. and Loucks, C.J., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, *51*(11), pp.933-938. [https://doi.org/10.1641/0006-3568(2001)051[0933:teotwa]2.0.co;2](https://doi.org/10.1641/0006-3568(2001)051%5b0933:teotwa%5d2.0.co;2).

Olusegun, C.F., Oguntunde, P.G. and Gbobaniyi, E.O., 2018. Simulating the impacts of tree, C3, and C4 plant functional types on the future climate of West Africa. *Climate*, *6*(2), p.35. <https://doi.org/10.3390/cli6020035>.

Ondei, S., Prior, L.D., Vigilante, T. and Bowman, D.M., 2016. Post-fire resprouting strategies of rainforest and savanna saplings along the rainforest-savanna boundary in the Australian monsoon tropics. *Plant Ecology*, *217*(6), pp.711-724. <https://doi.org/10.1007/s11258-015-0531-3>.

Osborne, C.P. and Freckleton, R.P., 2009. Ecological selection pressures for C4 photosynthesis in the grasses. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1663), pp.1753-1760. <https://doi.org/10.1098/rspb.2008.1762>.

Osborne, C.P., Salomaa, A., Kluyver, T.A., Visser, V., Kellogg, E.A., Morrone, O., Vorontsova, M.S., Clayton, W.D. and Simpson, D.A., 2014. A global database of C4 photosynthesis in grasses. *New Phytologist*, *204*(3), pp.441-446. <https://doi.org/10.1111/nph.12942>.

Pearcy, R.W. and Troughton, J., 1975. C4 photosynthesis in tree form *Euphorbia* species from Hawaiian rainforest sites. *Plant Physiology*, *55*(6), pp.1054-1056. <https://doi.org/10.1104/pp.55.6.1054>.

Pearcy, R.W. and Ehleringer, J., 1984. Comparative ecophysiology of C3 and C4 plants. *Plant, Cell & Environment*, *7*(1), pp.1-13. <https://doi.org/10.1111/j.1365-3040.1984.tb01194.x>.

Peter, G., 2009. Systematic revision of the genus *Isostigma* Less.(Asteraceae, Coreopsideae). *Candollea*, *64*(1), pp.5-30. <https://doi.org/10.3372/wi.34.34216>.

Peter, G. and Katinas, L., 2003. A new type of Kranz anatomy in Asteraceae. *Australian Journal of Botany*, *51*(2), pp.217-226. https://doi.org/10.1071/BT02080.

Piirainen, M., Liebisch, O. and Kadereit, G., 2017. Phylogeny, biogeography, systematics and taxonomy of Salicornioideae (Amaranthaceae/Chenopodiaceae) - A cosmopolitan, highly specialised hygrohalophyte lineage dating back to the Oligocene. *Taxon*, *66*(1), pp.109-132. <https://doi.org/10.12705/661.6>.

Pinto, H., Powell, J.R., Sharwood, R.E., Tissue, D.T. and Ghannoum, O., 2016. Variations in nitrogen use efficiency reflect the biochemical subtype while variations in water use efficiency reflect the evolutionary lineage of C4 grasses at inter‐glacial CO2. *Plant, Cell & Environment*, *39*(3), pp.514-526. [ttps://doi.org/10.1111/pce.12636](https://doi.org/10.1111/pce.12636).

Popp, M.R. and Kalwij, J.M., 2021. Abiotic conditions shape the relationship between indigenous and exotic species richness in a montane biodiversity hotspot. *Plant Ecology*, *222*(4), pp.421-432. <https://doi.org/10.1007/s11258-021-01116-6>.

Powell, A.M., 1978. Systematics of Flaveria (Flaveriinae--Asteraceae). *Annals of the Missouri Botanical Garden*, pp.590-636. <https://doi.org/10.2307/2398862>.

POWO, 2020. Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; http://www.plantsoftheworldonline.org/. [accessed 18 October 2021]

Pyankov, V.I., Black Jr, C.C., Artyusheva, E.G., Voznesenskaya, E.V., Ku, M.S. and Edwards, G.E., 1999. Features of photosynthesis in *Haloxylon* species of Chenopodiaceae that are dominant plants in Central Asian deserts. *Plant and Cell Physiology*, *40*(2), pp.125-134. <https://doi.org/10.1093/oxfordjournals.pcp.a029519>.

Pyankov, V.I., Gunin, P.D., Tsoog, S. and Black, C.C., 2000. C4 plants in the vegetation of Mongolia: their natural occurrence and geographical distribution in relation to climate. *Oecologia*, *123*(1), pp.15-31. <https://doi.org/10.1007/s004420050985.>

Pyankov, V.I., Ziegler, H., Akhani, H., Deigele, C. and Luettge, U., 2010. European plants with C4 photosynthesis: geographical and taxonomic distribution and relations to climate parameters. *Botanical Journal of the Linnean Society*, *163*(3), pp.283-304. <https://doi.org/10.1111/j.1095-8339.2010.01062.x.>

Rao, X. and Dixon, R.A., 2016. The differences between NAD-ME and NADP-ME subtypes of C4 photosynthesis: more than decarboxylating enzymes. *Frontiers in Plant Science*, *7*, p.1525. <https://doi.org/10.3389/fpls.2016.01525>.

Royer, D.L., 2006. CO2-forced climate thresholds during the Phanerozoic. *Geochimica et Cosmochimica Acta*, *70*(23), pp.5665-5675. <https://doi.org/10.1016/j.gca.2005.11.031>.

Rudov, A., Mashkour, M., Djamali, M. and Akhani, H., 2020. A review of C4 plants in southwest Asia: An ecological, geographical and taxonomical analysis of a region with high diversity of C4 eudicots. *Frontiers in Plant Science*, *11*, p.1374. <https://doi.org/10.3389/fpls.2020.546518>.

Sage, R.F., Sage, T.L., Pearcy, R.W. and Borsch, T., 2007. The taxonomic distribution of C4 photosynthesis in Amaranthaceae sensu stricto. *American Journal of Botany, 94*(12), pp.1992-2003. <https://doi.org/10.3732/ajb.94.12.1992>.

Sage, R.F., 2016. Tracking the evolutionary rise of C4 metabolism. *Journal of Experimental Botany*, *67*(10), pp.2919-2922. <https://doi.org/10.1093/jxb/erw137>.

Sage, R.F., 2017. A portrait of the C4 photosynthetic family on the 50th anniversary of its discovery: species number, evolutionary lineages, and Hall of Fame. *Journal of Experimental Botany*, *68*(2), pp.e11-e28. <https://doi.org/10.1093/jxb/erx005>.

Sage, R.F., Christin, P.A. and Edwards, E.J., 2011. The C4 plant lineages of planet Earth. *Journal of Experimental Botany*, *62*(9), pp.3155-3169. <https://doi.org/10.1093/jxb/err048>.

Sage, R.F., Sage, T.L. and Kocacinar, F., 2012. Photorespiration and the evolution of C4 photosynthesis. *Annual Review of Plant Biology*, *63*, pp.19-47. <https://doi.org/10.1146/annurev-arplant-042811-105511>.

Sage, R.F., Monson, R.K., Ehleringer, J.R., Adachi, S. and Pearcy, R.W., 2018. Some like it hot: the physiological ecology of C4 plant evolution. *Oecologia*, *187*(4), pp.941-966. <https://doi.org/10.1007/s00442-018-4191-6>.

Sage, R.F., Khoshravesh, R. and Sage, T.L., 2014. From proto-Kranz to C4 Kranz: building the bridge to C4 photosynthesis. *Journal of Experimental Botany*, *65*(13), pp.3341-3356. <https://doi.org/10.1093/jxb/eru180>.

Said Gutiérrez-Ortega, J., Yamamoto, T., Vovides, A.P., Angel Pérez-Farrera, M., Martinez, J.F., Molina-Freaner, F., Watano, Y. and Kajita, T., 2018. Aridification as a driver of biodiversity: a case study for the cycad genus *Dioon* (Zamiaceae). *Annals of Botany*, *121*(1), pp.47-60. <https://doi.org/10.1093/aob/mcx123>.

Salzmann, U., Haywood, A.M., Lunt, D.J., Valdes, P.J. and Hill, D.J., 2008. A new global biome reconstruction and data‐model comparison for the Middle Pliocene. *Global Ecology and Biogeography*, *17*(3), pp.432-447. <https://doi.org/10.1111/j.1466-8238.2008.00381.x>.

Sánchez-del Pino, I. and Motley, T.J., 2010. Evolution of *Tidestromia* (Amaranthaceae) in the deserts of the southwestern United States and Mexico. *Taxon*, *59*(1), pp.38-48. <https://doi.org/10.1002/tax.591005>.

Sánchez-del Pino, I., Motley, T.J. and Borsch, T., 2012. Molecular phylogenetics of *Alternanthera* (Gomphrenoideae, Amaranthaceae): resolving a complex taxonomic history caused by different interpretations of morphological characters in a lineage with C4 and C3-C4 intermediate species. *Botanical Journal of the Linnean Society*, *169*(3), pp.493-517. <https://doi.org/10.1111/j.1095-8339.2012.01248.x>.

Santos, J., Al-Azzawi, M., Aronson, J. and Flowers, T.J., 2016. eHALOPH a database of salt-tolerant plants: helping put halophytes to work. *Plant and Cell Physiology*, *57*(1), pp.e10-e10. <https://doi.org/10.1093/pcp/pcv155>.

Schlüter, U. and Weber, A.P., 2020. Regulation and evolution of C4 photosynthesis. *Annual Review of Plant Biology*, *71*, pp.183-215. <https://doi.org/10.1146/annurev-arplant-042916-040915>.

Schulze, E.D. and Schulze, I., 1976. Distribution and control of photosynthetic pathways in plants growing in the Namib Desert, with special regard to *Welwitschia mirabilis* Hook. fil. *Madoqua*, *1976*(3), pp.5-13.

Schulze, E.D., Ellis, R., Schulze, W., Trimborn, P. and Ziegler, H., 1996. Diversity, metabolic types and δ13C carbon isotope ratios in the grass flora of Namibia in relation to growth form, precipitation and habitat conditions. *Oecologia*, *106*(3), pp.352-369. <https://doi.org/10.1007/bf00334563>.

Schütze, P., Freitag, H. and Weising, K., 2003. An integrated molecular and morphological study of the subfamily Suaedoideae Ulbr.(Chenopodiaceae). *Plant Systematics and Evolution*, *239*, pp.257-286. <https://doi.org/10.1007/s00606-003-0013-2>.

Seidl, A., Tremetsberger, K., Pfanzelt, S., Blattner, F.R., Neuffer, B., Friesen, N., Hurka, H., Shmakov, A., Batlai, O., Žerdoner Čalasan, A. and Vesselova, P.V., 2021. The phylogeographic history of Krascheninnikovia reflects the development of dry steppes and semi-deserts in Eurasia. *Scientific Reports*, *11*(1), p.6645. <https://doi.org/10.1038/s41598-021-85735-z>.

Sharpe, R.M., Williamson-Benavides, B., Edwards, G.E. and Dhingra, A., 2020. Methods of analysis of chloroplast genomes of C3, Kranz type C4 and Single Cell C4 photosynthetic members of Chenopodiaceae. *Plant Methods*, *16*(1), pp.1-14. <https://doi.org/10.1186/s13007-020-00662-w>.

Shepherd, K.A., Waycott, M. and Calladine, A., 2004. Radiation of the Australian Salicornioideae (Chenopodiaceae) - based on evidence from nuclear and chloroplast DNA sequences. *American Journal of Botany*, *91*(9), pp.1387-1397. <https://doi.org/10.3732/ajb.91.9.1387>.

Shepherd, K.A., Macfarlane, T.D. and Waycott, M., 2005. Phylogenetic analysis of the Australian Salicornioideae (Chenopodiaceae) based on morphology and nuclear DNA. *Australian Systematic Botany*, *18*(1), pp.89-115. <https://doi.org/10.1071/sb04031>.

Shreve, F., 1939. Observations on the vegetation of Chihuahua. *Madroño*, *5*(1), pp.1-13.

Shoko, C., Mutanga, O. and Dube, T., 2016. Progress in the remote sensing of C3 and C4 grass species aboveground biomass over time and space. *ISPRS Journal of Photogrammetry and Remote Sensing*, *120*, pp.13-24. <https://doi.org/10.1016/j.isprsjprs.2016.08.001>.

Still, C.J., Pau, S. and Edwards, E.J., 2014. Land surface skin temperature captures thermal environments of C3 and C4 grasses. *Global Ecology and Biogeography*, *23*(3), pp.286-296. <https://doi.org/10.1111/geb.12121>.

Stowe, L.G. and Teeri, J.A., 1978. The geographic distribution of C4 species of the Dicotyledonae in relation to climate. *The American Naturalist*, *112*(985), pp.609-623. <https://doi.org/10.1086/283301>.

Strömberg, C.A., 2011. Evolution of grasses and grassland ecosystems. *Annual Review of Earth and Planetary Sciences*, *39*, pp.517-544. <https://doi.org/10.1146/annurev-earth-040809-152402>.

Tamboli, A.S., Dalavi, J.V., Kadam, S.K., Yadav, S.R., Govindwar, S.P., Choo, Y.S. and Pak, J.H., 2022. Molecular Phylogenetic Evidence and Biogeographic History of Indian Endemic *Portulaca* L.(Portulacaceae) Species. *Diversity*, *14*(6), p.443. <https://doi.org/10.3390/d14060443>.

Taub, D.R., 2000. Climate and the US distribution of C4 grass subfamilies and decarboxylation variants of C4 photosynthesis. *American Journal of Botany*, *87*(8), pp.1211-1215. <https://doi.org/10.2307/2656659>.

Team, R.C., 2020. R: 2019. *A Language and Environment for Statistical Computing version*, *3*(1).

Teeri, J.A. and Stowe, L.G., 1976. Climatic patterns and the distribution of C4 grasses in North America. *Oecologia*, *23*(1), pp.1-12. <https://doi.org/10.1007/bf00351210>.

Töpel, M., Zizka, A., Calió, M.F., Scharn, R., Silvestro, D. and Antonelli, A., 2017. SpeciesGeoCoder: fast categorization of species occurrences for analyses of biodiversity, biogeography, ecology, and evolution. *Systematic Biology*, *66*(2), pp.145-151. <https://doi.org/10.1093/sysbio/syw064>.

Vásquez‐Cruz, M. and Sosa, V., 2020. Assembly and origin of the flora of the Chihuahuan Desert: The case of sclerophyllous Rosaceae. *Journal of Biogeography*, *47*(2), pp.445-459. <https://doi.org/10.1111/jbi.13745>.

Verdcourt, B., 1969. The arid corridor between the north-east and south-west areas of Africa. *Palaeoecology of Africa*, *4*, pp.140-144.

Villarreal-Quintanilla, J.A., Bartolomé-Hernández, J.A., Estrada-Castillón, E., Ramírez-Rodríguez, H. and Martínez-Amador, S.J., 2017. The endemic element of the Chihuahuan Desert vascular flora. *Acta Botánica Mexicana*, (118), pp.65-96. <https://doi.org/10.21829/abm118.2017.1201>.

Vogel, J.C. and Seely, M.K., 1977. Occurrence of C4 plants in the Central Namib Desert. *Madoqua*, *1977*(1), pp.75-78.

Von Caemmerer, S., Ghannoum, O. and Furbank, R.T., 2017. C4 photosynthesis: 50 years of discovery and innovation. *Journal of Experimental Botany*, *68*(2), p.97. <https://doi.org/10.1093/jxb/erw491>.

Voznesenskaya, E.V., Franceschi, V.R., Kiirats, O., Artyusheva, E.G., Freitag, H. and Edwards, G.E., 2002. Proof of C4 photosynthesis without Kranz anatomy in *Bienertia cycloptera* (Chenopodiaceae). *The Plant Journal*, *31*(5), pp.649-662. <https://doi.org/10.1046/j.1365-313x.2002.01385.x>.

Voznesenskaya, E.V., Edwards, G.E., Kiirats, O., Artyusheva, E.G. and Franceschi, V.R., 2003. Development of biochemical specialisation and organelle partitioning in the single‐cell C4 system in leaves of *Borszczowia aralocaspica* (Chenopodiaceae). *American Journal of Botany*, *90*(12), pp.1669-1680. <https://doi.org/10.3732/ajb.90.12.1669>.

Voznesenskaya, E.V., Akhani, H., Koteyeva, N.K., Chuong, S.D., Roalson, E.H., Kiirats, O., Franceschi, V.R. and Edwards, G.E., 2008. Structural, biochemical, and physiological characterization of photosynthesis in two C4 subspecies of *Tecticornia indica* and the C3 species *Tecticornia pergranulata* (Chenopodiaceae). *Journal of Experimental Botany*, *59*(7), pp.1715-1734. <https://doi.org/10.1093/jxb/ern028>.

Voznesenskaya, E.V., Koteyeva, N.K., Edwards, G.E. and Ocampo, G., 2017. Unique photosynthetic phenotypes in *Portulaca* (Portulacaceae): C3-C4 intermediates and NAD-ME C4 species with Pilosoid-type Kranz anatomy. *Journal of Experimental Botany*, *68*(2), pp.225-239. <https://doi.org/10.1093/jxb/erw393>.

Wang, Y., Bräutigam, A., Weber, A.P. and Zhu, X.G., 2014. Three distinct biochemical subtypes of C4 photosynthesis? A modelling analysis. *Journal of Experimental Botany*, *65*(13), pp.3567-3578. <https://doi.org/10.1093/jxb/eru058>.

Waselkov, K.E., Boleda, A.S. and Olsen, K.M., 2018. A phylogeny of the genus *Amaranthus* (Amaranthaceae) based on several low-copy nuclear loci and chloroplast regions. *Systematic Botany*, *43*(2), pp.439-458. <https://doi.org/10.1600/036364418X697193>.

WCVP. World Checklist of Vascular Plants, version 2.0, 2020. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; Retrieved from http://wcvp.science.kew.org/. [accessed 25 January 2021]

Wen, Y., Zhang, L., Holbourn, A.E., Zhu, C., Huntington, K.W., Jin, T., Li, Y. and Wang, C., 2023. CO2-forced Late Miocene cooling and ecosystem reorganisations in East Asia. *Proceedings of the National Academy of Sciences*, *120*(5), p.e2214655120. <https://doi.org/10.1073/pnas.2214655120>.

Wickham, H., Chang, W. and Wickham, M.H., 2016. Package ‘ggplot2’. *Create Elegant Data Visualisations Using the Grammar of Graphics. Version*, *2*(1), pp.1-189. <https://doi.org/10.1111/j.1541-0420.2011.01616.x>.

Wilson, P.G. 1984. Chenopodiaceae. In George, A.S. (ed.) *Flora of Australia. Volume 4. Australian Government Publishing Service, Canberra*.

Winter, K., 1981. C4 plants of high biomass in arid regions of asia-occurrence of C4 photosynthesis in Chenopodiaceae and Polygonaceae from the Middle East and USSR. *Oecologia*, *48*(1), pp.100-106. <https://doi.org/10.1007/bf00346994>.

Winter, K., Garcia, M., Virgo, A. and Holtum, J.A., 2019. Operating at the very low end of the crassulacean acid metabolism spectrum: *Sesuvium portulacastrum* (Aizoaceae). *Journal of Experimental Botany*, *70*(22), pp.6561-6570. <https://doi.org/10.1093/jxb/ery431>.

Winter, K., Garcia, M., Virgo, A., Ceballos, J. and Holtum, J.A., 2021. Does the C4 plant *Trianthema portulacastrum* (Aizoaceae) exhibit weakly expressed crassulacean acid metabolism (CAM)?. *Functional Plant Biology*, *48*(7), pp.655-665. <https://doi.org/10.1071/fp20247>.

Woodward, F.I. and Lomas, M.R., 2004. Vegetation dynamics-simulating responses to climatic change. *Biological Reviews*, *79*(3), pp.643-670. <https://doi.org/10.1017/s1464793103006419>.

Woodward, F.I., Lomas, M.R. and Kelly, C.K., 2004. Global climate and the distribution of plant biomes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1450), pp.1465-1476. <https://doi.org/10.1098/rstb.2004.1525>.

Wu, S.D., Zhang, L.J., Lin, L., Yu, S.X., Chen, Z.D. and Wang, W., 2018. Insights into the historical assembly of global dryland floras: the diversification of Zygophyllaceae. *BMC Evolutionary Biology*, *18*, pp.1-10. <https://doi.org/10.1186/s12862-018-1277-z>.

Yan, D. and de Beurs, K.M., 2016. Mapping the distributions of C3 and C4 grasses in the mixed-grass prairies of southwest Oklahoma using the Random Forest classification algorithm. *International Journal of Applied Earth Observation and Geoinformation*, *47*, pp.125-138. <https://doi.org/10.1016/j.jag.2015.12.007>.

Yang, Y. and Berry, P.E., 2011. Phylogenetics of the Chamaesyce clade (*Euphorbia*, Euphorbiaceae): Reticulate evolution and long‐distance dispersal in a prominent C4 lineage. *American Journal of Botany, 98*(9), pp.1486-1503. <https://doi.org/10.3732/ajb.1000496>.

Young, S.N., Sack, L., Sporck-Koehler, M.J. and Lundgren, M.R., 2020. Why is C4 photosynthesis so rare in trees? *Journal of Experimental Botany*, *71*(16), pp.4629-4638. <https://doi.org/10.1093/jxb/eraa234>.

Zang, E.H., Zhang, M.X., Wang, W.L., Zhang, C.H. and Li, M.H., 2021. *Euphorbia mongoliensis* (Euphorbiaceae), a new species from Inner Mongolia, China. *Phytotaxa*, *501*(1), pp.151-161. <https://doi.org/10.11646/phytotaxa.501.1.6>.

Žerdoner Čalasan, A., Hurka, H., German, D.A., Pfanzelt, S., Blattner, F.R., Seidl, A. and Neuffer, B., 2021. Pleistocene dynamics of the Eurasian steppe as a driving force of evolution: Phylogenetic history of the genus *Capsella* (Brassicaceae). *Ecology and Evolution*, *11*(18), pp.12697-12713. <https://doi.org/10.1002/ece3.8015>.

Žerdoner Čalasan, A., Hammen, S., Sukhorukov, A.P., McDonald, J.T., Brignone, N.F., Böhnert, T. and Kadereit, G., 2022. From continental Asia into the world: Global historical biogeography of the saltbush genus *Atriplex* (Chenopodieae, Chenopodioideae, Amaranthaceae). *Perspectives in Plant Ecology, Evolution and Systematics*, *54*, p.125660. <https://doi.org/10.1016/j.ppees.2022.125660>.

Zhou, H., Helliker, B.R., Huber, M., Dicks, A. and Akçay, E., 2018. C4 photosynthesis and climate through the lens of optimality. *Proceedings of the National Academy of Sciences*, *115*(47), pp.12057-12062. <https://doi.org/10.1073/pnas.1718988115>.

Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., Farooq, H., Herdean, A., Ariza, M., Scharn, R. and Svantesson, S., 2019. CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, *10*(5), pp.744-751. <https://doi.org/10.1111/2041-210x.13152>.

Zizka, A., Carvalho, F.A., Calvente, A., Baez-Lizarazo, M.R., Cabral, A., Coelho, J.F.R., Colli-Silva, M., Fantinati, M.R., Fernandes, M.F., Ferreira-Araújo, T. and Moreira, F.G.L., 2020. No one-size-fits-all solution to clean GBIF. *PeerJ*, *8*, p.e9916. <https://doi.org/10.7717/peerj.9916>.

Zizka, A., Onstein, R.E., Rozzi, R., Weigelt, P., Kreft, H., Steinbauer, M.J., Bruelheide, H. and Lens, F., 2022. The evolution of insular woodiness. *Proceedings of the National Academy of Sciences*, *119*(37), p.e2208629119. <https://doi.org/10.1073/pnas.2208629119>.

**GBIF - References**

*C4 Eudicots*

GBIF.org.2020a. Acanthaceae (17 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.p5n2uy> (0114210-200613084148143)

GBIF.org.2020b. Aizoaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.57hjud> (0116423-200613084148143)

GBIF.org.2020c. Amaranthaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.su2qph> (0116611-200613084148143)

GBIF.org.2020d. Asteraceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.hnc7zj> (0116631-200613084148143)

GBIF.org.2020e. Boraginaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.x6654v> (0116642-200613084148143)

GBIF.org.2020f. Cleomaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.whe79v> (0116648-200613084148143)

GBIF.org.2020g. Caryophyllaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.jadxrm> (0116652-200613084148143)

GBIF.org.2020h. Euphorbiaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.kt6gaz> (0116655-200613084148143)

GBIF.org.2020i. Gisekiaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.ycp32w> (0116659-200613084148143)

GBIF.org.2020j. Molluginaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.62fqxd> (0116660-200613084148143)

GBIF.org.2020k. Nyctaginaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.3zuy37> (0116664-200613084148143)

GBIF.org.2020l. Polygonaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.2zp22x> (0116667-200613084148143)

GBIF.org.2020m. Portulacaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.r2w25u> (0116670-200613084148143)

GBIF.org.2020n. Scrophulariaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.hfe2jb> (0116676-200613084148143)

GBIF.org.2020o. Zygophyllaceae (19 Nov 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.pk453g> (0116678-200613084148143)

*C4 Grasses*

GBIF.org.2020p. Aristidoideae (17 Sep 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.7ahgpw> (0063372-200613084148143)

GBIF.org.2020q. Chloridoideae (15 Sep 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.e9k62g> (0061870-200613084148143)

GBIF.org.2021r. Hydrocharitaceae (21 Jan 2021) GBIF occurrence download: <https://doi.org/10.15468/dl.bmtypz> (0164167-200613084148143)

GBIF.org.2020s. Micrairoideae (17 Sep 2020) GBIF occurrence download: <https://doi.org/10.15468/dl.w2yqez> (0063376-200613084148143)

GBIF.org.2020t. Panicoideae (15 Sep 2020) GBIF occurrence download:

<https://doi.org/10.15468/dl.4z5c4d> (0061888-200613084148143)

**Figure Legends**

**Figure 1:**Angiosperm families in eudicots (a) and monocots (b) including C4 species. The first number represents the total number of C4 species per family according to Sage (2017) | the second number represents the number of C4 species with verified information about C4 photosynthesis performance (Supplementary Table A1) | the number in bold shows the remaining number of C4 species after various data cleaning steps (Table 1). Proportions shown in the diagrams and mentioned in the text were calculated according to species numbers in Sage (2017) but do not change much when based on the other two reduced numbers.

**Figure 2:** Global maps of total C4 grasses and C4 eudicot diversity. (a) Species richness of C4 grasses. (b) Species richness of C4 eudicots. (c) Genus richness of C4 eudicots. Scale indicates the number of families occurring in each grid (100 x 100 km). (d) Normalised difference C4 eudicots - C4 grasses. Blue grid cells indicate centres of diversity more important for C4 grasses, red grid cells indicate centres of diversity of C4 eudicots.

**Figure 3:** Scatterplot showing annual mean temperature (°C) and annual mean precipitation (mm) data of C4 grasses (green) and C4 eudicot (blue) occurrence points.

**Figure 4:** (a) Table shows the number of species of a family within a biome (Terrestrial ecoregions of the world (Olson et al., 2001)). Species are considered present in a biome if at least 5% of the distribution points are in that biome. Five C4 eudicot families with the highest number of species in this study are shown (Figure 1a), as well as C4 eudicots and C4 grasses in general. Colours correspond to the biomes shown in (b). Saturation shows the percentage of species per family or per group (C4 eudicots, C4 grasses) in each biome. (b) Terrestrial ecoregions of the world (Olson et al., 2001). (c) Global family richness map of C4 eudicots. Scale indicates the number of families occurring in each grid (100 x 100 km).

**Figure 5:** Differences in precipitation and temperature preference of the C4 species in eudicot families, and C4 eudicots, and C4 grasses in general. The Krustal-Wallis test was used over C4 eudicot families: *p* < .001 statistically significant.

**Figure 6:** (a) Distribution of Kranz types present in C4 eudicots. (b) Distribution of the traits summarised within the C4 eudicots. Separated into the number of C4 species and the number of families in which the traits occur. (c) Schematic illustration of the four common C4 leaf types (Schematic drawings for illustration of leaf anatomical types are adapted from Bohley et al., 2015). (d) Range and mean of annual precipitation of species with NADP-ME and NAD-ME as primary decarboxylating enzymes.

**Table Legends**

**Table 1:** The impact of data cleaning on the raw occurrence datasets of the various eudicots plant families containing C4 species (including the number of occurrence points at each step). Beginning with the raw record list.

**Table 2:** List of eudicots C4 genera occurring in the diversity hotspots Mexico/Southern United States and Australia according to our occurrence points and of C4 genera with molecular evidence of C4 in-situ origin within the remaining diversity hotspots. Genera that are species-rich (> 10 species) in the respective region are indicated in bold (**x**). Genera that likely originated in the respective region are marked with \*.

**Table 1:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Occurrences after downloading from GBIF\*1 | Occurrences after applying CoordinateCleaner\*2 | Occurrences after manual cleaning\*3 | Occurrences after removing duplicates | Occurrences after removing outliers\*4 |
| Acanthaceae | 1,380 | 1,335 | 1,335 | 988 | 855 |
| Aizoaceae | 4,894 | 4,009 | 4,009 | 3,135 | 2,632 |
| Amaranthaceae | 645,813 | 463,332 | 336,492 | 242,485 | 144,100 |
| Asteraceae | 14,369 | 12,772 | 12,767 | 10,889 | 9,562 |
| Boraginaceae | 5,261 | 4,782 | 4,253 | 3,156 | 3,003 |
| Caryophyllaceae | 9,338 | 8,644 | 8,606 | 6,389 | 3,224 |
| Cleomaceae | 2,093 | 1,735 | 1,563 | 1,163 | 925 |
| Euphorbiaceae | 122,123 | 85,559 | 82,052 | 61,509 | 45,285 |
| Gisekiaceae | 563 | 525 | 399 | 315 | 312 |
| Molluginaceae | 1,347 | 1,263 | 1,240 | 1,031 | 528 |
| Nyctaginaceae | 21,248 | 18,395 | 16,530 | 14,271 | 13,645 |
| Polygonaceae | 765 | 719 | 719 | 391 | 210 |
| Portulacaceae | 155,597 | 106,501 | 86,490 | 61,308 | 12,197 |
| Scrophulariaceae | 239 | 216 | 204 | 171 | 169 |
| Zygophyllaceae | 27,527 | 21,835 | 21,727 | 17,504 | 10,558 |
| **C4 eudicots** | **1,012,557** | **731,622** | **578,386** | **422,905** | **247,205** |
| \*1 see references for DOI number  \*2 CoordinateCleaner v2.0-18 (Zizka et al., 2019)  \*3 Manual cleaning: Nomenclatural and taxonomic checking including [the correction of] wrong synonymizations  \*4 Deleting outliers: after checking whether the native distribution information of plantsoftheworldonline.org matches the distribution country (CountryCode) of each distribution point from GBIF. | | | | | |

**Table 2:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Family** | **Genus** | **Mexico/South US** | **Australia** | **C4 origin** | **Reference** |
| Acanthaceae | *Blepharis* |  |  | Africa | Fisher et al. 2015 |
| Aizoaceae | *Sesuvium* |  |  | Africa | Bohley et al. 2015 |
|  | *Trianthema* | x | x | Africa | Bohley et al. 2015 |
|  | *Zaleya* |  | x | Africa | Bohley et al. 2015 |
| Amaranthaceae | *Alternanthera* | x | x | more likely origin: South American tropics | Sage et al. 2007; Sánchez-del Pino et al. 2012 |
|  | *Amaranthus* | **x** | x | origin unknown |  |
|  | *Atriplex* | **x** | **x** | continental Asia | Zerdoner Calasan et al. 2022 |
|  | *Froelichia* | x |  | origin unknown |  |
|  | *Gomphrena* | x | **x** | origin unknown |  |
|  | *Guilleminea* | x |  | origin unknown |  |
|  | *Suaeda* | x |  | origin unknown |  |
|  | *Tecticornia* |  | x\* | Australia | Shepherd et al. 2005; Voznesenskaya et al. 2008 |
|  | *Tidestromia* | x\* |  | Mexico/South US | Sánchez-del Pino & Motley, 2010 |
| Asteraceae | *Flaveria* | x\* |  | Mexico/South US | Morales-Briones & Kadereit, 2023 |
|  | *Glossocardia* |  | x | origin unknown |  |
|  | *Pectis* | **x** \* |  | Mexico/South US | Hansen et al. 2016 |
| Boraginaceae | *Euploca* | x | x | origin unknown |  |
| Caryophyllaceae | *Polycarpaea* |  | x | origin unknown |  |
| Cleomaceae | *Cleome* |  | x | origin unknown |  |
| Euphorbiaceae | *Euphorbia* | **x \*** | x | Mexico/South US (*Euphorbia* subg. *Chamaesyce* sect. *Anisophyllum* subsect. *Hypericifoliae)* | Yang & Berry, 2011; Horn et al. 2014 |
| Gisekiaceae | *Gisekia* |  |  | Africa | Bissinger et al. 2014 |
| Nyctaginaceae | *Allionia* | x\* |  | Mexico/South US | Douglas & Manos, 2007 |
|  | *Boerhavia* | x\* | x | Mexico/South US | Douglas & Manos, 2007 |
| Portulacaceae | *Portulaca* | **x** | x | origin unknown |  |
| Scrophulariaceae | *Anticharis* |  |  | Africa | Khoshravesh et al. 2012 |
| Zygophyllaceae | *Kallstroemia* | x |  | origin unknown |  |
|  | *Tribulopis* |  | x(\*) | less clear origin: Australia | Lauterbach et al. 2019 |
|  | *Tribulus* |  | x | origin unknown |  |
|  | *Zygophyllum* |  |  | Africa (*Z. simplex*) | Bellstedt et al. 2012 |