

A fully resolved backbone phylogeny reveals numerous dispersals and explosive diversifications throughout the history of Asteraceae

Jennifer R. Mandel^{a,1}, Rebecca B. Dikow^b, Carolina M. Siniscalchi^a, Ramhari Thapa^a, Linda E. Watson^c, and Vicki A. Funk^d

^aDepartment of Biological Sciences, University of Memphis, Memphis, TN 38152; ^bData Science Lab, Office of the Chief Information Officer, Smithsonian Institution, Washington, DC 20024; ^cDepartment of Plant Biology, Ecology, and Evolution, Oklahoma State University, Stillwater, OK 74078-3013; and ^dDepartment of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012

Edited by Peter H. Raven, Missouri Botanical Garden, St. Louis, MO, and approved May 14, 2019 (received for review March 5, 2019)

The sunflower family, Asteraceae, comprises 10% of all flowering plant species and displays an incredible diversity of form. Asteraceae are clearly monophyletic, yet resolving phylogenetic relationships within the family has proven difficult, hindering our ability to understand its origin and diversification. Recent molecular clock dating has suggested a Cretaceous origin, but the lack of deep sampling of many genes and representative taxa from across the family has impeded the resolution of migration routes and diversifications that led to its global distribution and tremendous diversity. Here we use genomic data from 256 terminals to estimate evolutionary relationships, timing of diversification(s), and biogeographic patterns. Our study places the origin of Asteraceae at ~83 MYA in the late Cretaceous and reveals that the family underwent a series of explosive radiations during the Eocene which were accompanied by accelerations in diversification rates. The lineages that gave rise to nearly 95% of extant species originated and began diversifying during the middle Eocene, coincident with the ensuing marked cooling during this period. Phylogenetic and biogeographic analyses support a South American origin of the family with subsequent dispersals into North America and then to Asia and Africa, later followed by multiple worldwide dispersals in many directions. The rapid mid-Eocene diversification is aligned with the biogeographic range shift to Africa where many of the modern-day tribes appear to have originated. Our robust phylogeny provides a framework for future studies aimed at understanding the role of the macroevolutionary patterns and processes that generated the enormous species diversity of Asteraceae.

biogeography | Compositae | molecular dating | phylogenomics

Asteraceae (Compositae), commonly known as the sunflower or daisy family, is one of three mega-diverse families that together account for more than 25% of all extant angiosperm species. The family, with an estimated 25,000–35,000 species, comprises 10% of all flowering plant species, and is rivaled only by Orchidaceae (orchids) and Fabaceae (legumes). Members of the sunflower family occur on every continent including Antarctica (1) and in nearly every type of habitat on Earth with the greatest concentration of its species in deserts, prairies, steppes, montane regions, and areas with Mediterranean-like climates.

The unique floral and fruit traits of Asteraceae are thought to have contributed to the evolutionary and ecological success of the family. Specifically, the flower (floret) develops into a fruit that possesses a highly modified calyx (pappus) with dual functionality for seed dispersal and antiherbivory (2). While all of the Asteraceae florets essentially share this basic floral plan, tremendous floral diversity has arisen from marked differences in corolla size, degree of petal fusion, symmetry, and color. The most obvious floral feature of the family is its developmentally complex, head-like inflorescence, the capitulum, which has a compound receptacle to which many tightly packed florets are attached. This is exemplified by the iconic North American sunflower that mimics a single flower for pollinator attraction (Fig. 1*F*).

The sunflower family is morphologically distinct from its sister family and its monophyly has never been disputed; however, using

molecular data to resolve intrafamilial relationships to understand its migrations and diversifications has proven difficult. This may be due, at least in part, to its history of hybridization, rapid radiations, and rampant polyploidy (3). Large-scale gene and genome duplications have likely been a significant contributor to the evolutionary and ecological success of Asteraceae. Transcriptome data have been used to document ancient whole genome duplications (WGDs) at the base of several of the family's major radiations (4–6), and WGD may be associated with increased species diversification rates (7, 8). Furthermore, gene families have been identified that play a role in phenotypic evolution for floral symmetry and capitulum development (9, 10), as well as for the diversity and abundance of secondary metabolites that provide antiherbivory properties (11–13).

Early plastid phylogenies (14, 15) definitively placed Asteraceae's origin in South America, which remains supported by all subsequent phylogenies. Several hypotheses of dispersal out of South America have been proposed that include at least three migration pathways: a Pacific–Asian route, a North American–Asian route, and direct dispersal from South America to Africa (16–18). However, testing these hypotheses has been challenging because a robust backbone phylogeny with comprehensive sampling is not available, and the most recent supertree phylogeny has many areas that are unresolved (19)

Significance

Flowering plant species represent at least 95% of all vascular plants on Earth, and members of the sunflower family comprise roughly 10% of this diversity. The family is often considered taxonomically difficult primarily because it is enormous in size and cosmopolitan in distribution. Using phylogenomics, we were able to fully resolve the backbone of the sunflower family tree. We provide evidence for a late Cretaceous origin followed by explosive diversifications and dispersals during the middle Eocene—ultimately resulting in the family's 25,000+ extant species. Our results provide a framework to interpret the spatiotemporal patterns of migration out of South America and the family's explosive diversifications out of Africa that led to its global evolutionary and ecological success.

Author contributions: J.R.M., R.B.D., L.E.W., and V.A.F. designed research; J.R.M., R.B.D., C.M.S., R.T., L.E.W., and V.A.F. performed research; J.R.M., R.B.D., C.M.S., R.T., L.E.W., and V.A.F. analyzed data; and J.R.M., C.M.S., R.T., L.E.W., and V.A.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: Raw sequencing reads were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive and Genbank under BioProject PRJNA540287. The resulting alignments and phylogenetic trees were deposited in FigShare at <https://doi.org/10.6084/m9.figshare.7697834> and <https://doi.org/10.6084/m9.figshare.7695929>.

¹To whom correspondence may be addressed. Email: jmandel@memphis.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1903871116/-DCSupplemental.

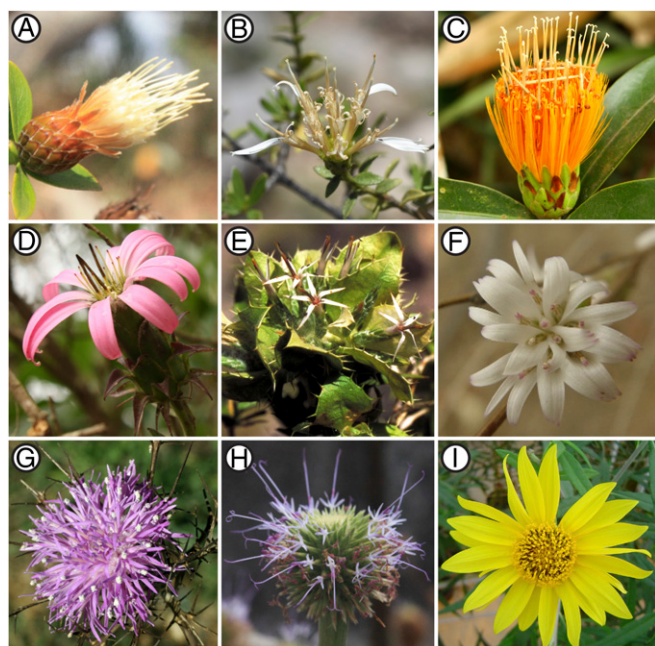


Fig. 1. Floral diversity of tribes. (A) Barnadesiidae; (B) Famatinanthae; (C) Stiffiteae; (D) Mutisieae; (E) Hecastocleidae; (F) Pertyeae; (G) Cardueae; (H) Vernoniaceae; and (I) Heliantheae. Photos of A, C, D, and H were provided by C.M.S.; B, image courtesy of J. Mauricio Bonifacio (photographer); E, by V.A.F.; F, image courtesy of Tiangang Gao (photographer); G, image courtesy of Alfonso Susanna (photographer); and I, by J.R.M.

(*SI Appendix*, Fig. S1). Until recently the family was considered to be relatively young (50–40 MYA) based on its existing fossil record (19, 20), and this time frame is consistent with molecular clock dating studies focused on the Southern Hemisphere origins of several angiosperm orders, including Asterales (21, 22). However, three recent phylogenetic studies of Asteraceae using new fossil evidence (reviewed in ref. 23) and a broader sampling of the family provided older age estimates and placed its origin sometime in the late Cretaceous, 89–69 MYA (6, 23, 24). These studies also used additional markers and increased our knowledge of the evolutionary history of the family which called into question its age and the influence of past geological and paleoclimatic events on early migrations and diversifications. However, these studies had incomplete sampling of key lineages along the backbone of the Asteraceae tree, limiting conclusions about migration routes.

In this study, we used phylogenomic data, targeting roughly 1,000 loci in the nuclear genome, to estimate the most robust and comprehensive backbone phylogeny for Asteraceae to date. We employed molecular clock dating methods and model-based biogeographic and diversification rate analyses to estimate the ages and ancestral geographic ranges for all stem nodes along the backbone of the tree and for crown nodes at the base of major radiations to interpret lineage diversifications. We aimed to gain a more precise understanding of the family's origin and major diversifications in both space and time, as well as, what role dispersals and the global climate may have played during the evolution of this tremendously successful angiosperm lineage.

Results

Phylogeny of Asteraceae. The concatenated alignment of 935 nuclear loci yielded a data matrix of 942,707 bp for 256 species (*Dataset S1*). The integration of transcriptome data was successful and is exemplified by the four pairs of duplicate taxa (four from the 1KP and four from our workflow) forming sister group relationships in our analyses (*SI Appendix*, Fig. S2). High branch support and resolution resulted, and near complete congruence between the maximum likelihood (ML) and Bayesian concatenated trees was recovered (Fig. 2 and *SI Appendix*, Figs. S2 and S3). The ASTRAL tree yielded

some topological differences compared with the concatenated phylogenies, but in nearly every case, the ASTRAL tree nodes had lower support (*SI Appendix*, Fig. S4). We found a few major differences from previous phylogenies in tribal placements. For example, in our ML tree, Hyalidae was placed sister to Stiffiteae, and Pertyeae diverged just before the Carduoideae subfamily (thistles) and immediately after the monotypic Hecastocleidae. Two subfamilies previously circumscribed (17, 25) are not supported as monophyletic: Carduoideae [here recovered as three clades: (i) Cardueae, (ii) Oldenburiaceae + Tarchonantheae, and (iii) Dicomeae, brown shading in Fig. 2]; and Cichorioideae [dandelions, here recovered as two clades: (i) tribe Cichorieae and (ii) six remaining tribes referred to herein as Vernoniaceae (26) blue shading in Fig. 2]. In the enormous subfamily Asteroideae (17,000+ species), relationships among the five tribes that comprise nearly 10,000 species, including Anthemideae (chrysanthemums), Astereae (asters), Gnaphalieae (strawflowers), Calenduleae (pot marigolds), and Senecioneae (ragworts), were resolved with high support (Fig. 2). The exception was the sister group relationship between Anthemideae and Senecioneae with bootstrap support for this relationship lower than at most other nodes at 75%. The Bayesian analysis yielded the same topology for these five tribes with posterior probabilities of 1.0 (*SI Appendix*, Fig. S3). The ASTRAL analysis resulted in high support for a different topology for these five tribes with Senecioneae as the sister group to a clade of the four remaining tribes (*SI Appendix*, Fig. S4). Within the large Heliantheae alliance, which includes sunflowers and coneflowers (supertribe Helianthodae, 13 tribes), maximal support was recovered for most intertribal relationships (Fig. 2 and *SI Appendix*, Figs. S2–S4).

Dating Analysis. The ML phylogeny, calibrated in two analyses by constraining the minimum ages of nodes with either seven or eight fossils, provided similar estimates that indicate Asteraceae likely originated during the Late Cretaceous: ~83 MYA (95% CI, 91–64 MYA, Fig. 3 and *SI Appendix*, Figs. S5 and S6 for all CIs, and *SI Appendix*, Table S1 for all scenarios tested). The earliest diverging lineage (64 MYA) is tribe Barnadesiidae followed by the monospecific Famatinanthae, which diverged from the rest of the family roughly 62 MYA near the Cretaceous–Paleogene (K–Pg) boundary. Roughly 5–10 MY passed before extant members of tribes Stiffiteae, Nassauvieae, Mutisieae, Wunderlichieae, and Gochnatieae began to diverge during the early Eocene (~53 MYA). Divergence of the large clade that includes monotypic Hecastocleidae and the rest of the family occurred in the early Eocene around 50 MYA. Explosive diversifications occurred within the remaining three subfamilies (Carduoideae, Cichorioideae, and Asteroideae) in the middle to late Eocene (42–37 MYA) that gave rise to most of the present-day tribes that harbor 95% of Asteraceae's extant species. Lastly, the radiation of the Heliantheae alliance began during the late Oligocene, roughly 25 MYA, including the origins of its tribes (*SI Appendix*, Figs. S5 and S6). Our age estimates are in general agreement with three recent publications which also used the newer fossil data (*SI Appendix*, Table S1) (6, 23, and 24).

Historical Biogeography. The AIC model selection supports the BAYAREALIKE model, including the jump (*j*) speciation parameter (Fig. 3 and *SI Appendix*, Fig. S7 and Table S2). Our ancestral range estimates support an origin of Asteraceae in southern South America with subsequent range extensions into the north and central Andes region and later into the Guiana Shield region and Brazil. From South America, Asteraceae likely migrated to North America but a possible dispersal to south central Africa or to Asia is also estimated (Fig. 3, see stem node leading to Hecastocleidae and *SI Appendix*, Fig. S7). Next, the family dispersed either to south and central Africa or to Asia. Beginning around 42 MYA, all major nodes along the backbone of the Asteraceae tree have an ancestral range of south and central Africa. The only exception is the large New World Heliantheae alliance with a broad ancestral range estimate that includes north and central Andes + Mesoamerica–Caribbean and the north and central Andes region alone. The Biogeographic Stochastic Mapping (BSM) results indicate the presence of numerous dispersal events, some requiring long-distance

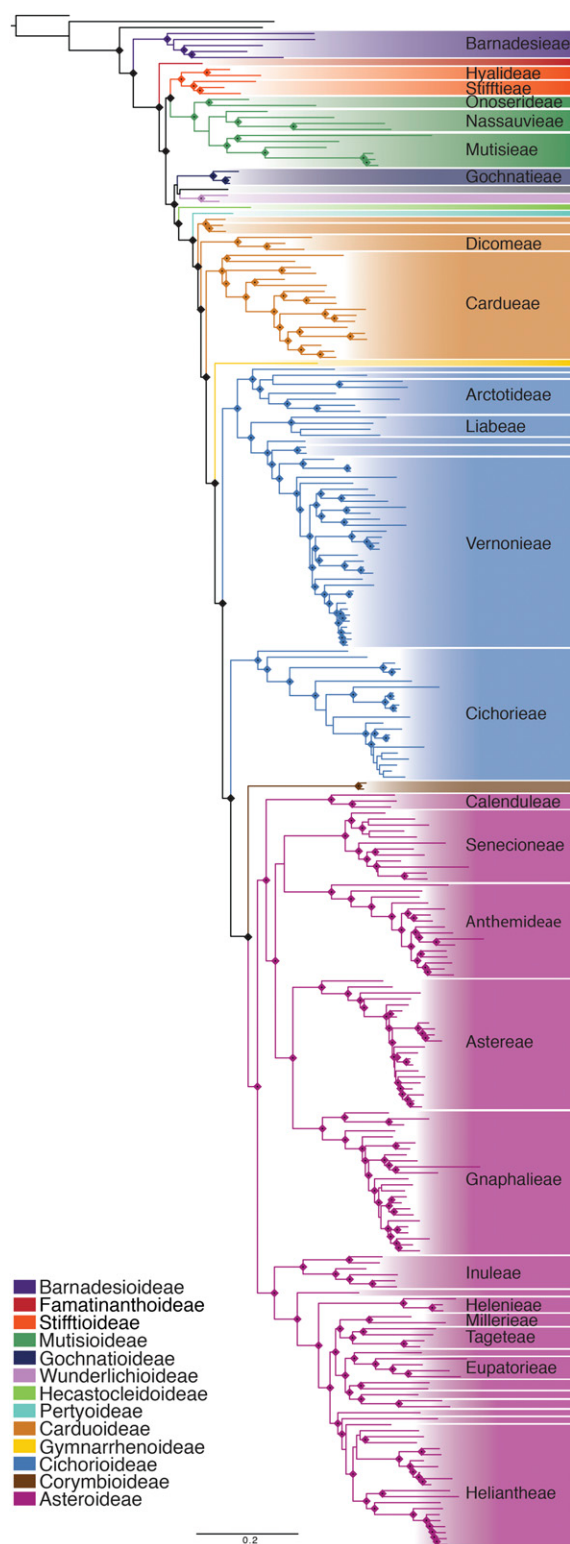


Fig. 2. Maximum likelihood tree for the Asteraceae family. Tree is color coded by subfamily with most tribes indicated in text to the right of each. Diamonds at nodes indicate bootstrap support of 94% or higher. All nodes along the backbone have maximal bootstrap support of 100%.

dispersal throughout the entirety of the Asteraceae's evolutionary history (59% of total events, *SI Appendix, Table S3*). Among these dispersals, anagenetic events (i.e., range expansions) were more prevalent than founder events (73%). The highest number of anagenetic

events occurred out of the north and central Andes region, south and central Africa, and North America, respectively. The highest number of anagenetic events was to North America and the north and central Andes region, respectively, with the fewest events to south and central Africa. By far, the greatest number of founder events occurred out of south and central Africa and coincide with the explosive diversifications during the middle Eocene. Founder events were highest to North America and mostly resulted from events out of the north and central Andes region and Mesoamerica–Caribbean regions coinciding with dispersal events during the Miocene (20–5 MYA) (Fig. 3 and *SI Appendix, Fig. S7 and Table S3*).

Diversification Analyses. The diversification analyses revealed five instances of rate acceleration (shown in red numbered boxes) and three of deceleration (shown in blue numbered boxes) (Fig. 3 and *SI Appendix, Fig. S8 and Table S4*). The rate decreases (nos. 7, 8, and 9) occur at nodes leading to tribes with only one or two species. The five accelerations coincide with major diversifications, including the core of the family following divergence from tribes Barnadesioideae and Famatinanthoideae (no. 5), the explosive diversifications in Africa during the middle Eocene (nos. 2 and 6), and on the stems leading to the Vernoniaceae (no. 4) and the Heliantheae alliance (no. 3).

Discussion

Asteraceae Phylogeny Backbone. With at least 25,000 named species and more than 1,700 genera, the backbone phylogeny of the sunflower family has been notoriously difficult to resolve. However, with increasing access to next generation sequencing technologies, resolving the phylogenies of megafamilies has become a reality. Our phylogenomics approach resulted in a highly resolved and well-supported nuclear phylogeny that shares similarities with the previously published plastid (23, 24, 27) and phylotranscriptomic (6) phylogenies despite the substantial differences in sampling of genomes, loci, and taxa. It is also important to note there is little discordance among our different analyses (ML, Bayesian, ASTRAL; Fig. 2 and *SI Appendix, Figs. S2–S4*), which may be attributed to our use of the conservative pipeline, PHYLUC, for assigning orthology (retaining only loci recovered as single copy for each taxon). Our deep sampling of genes and broad, dense sampling of select taxa representing each major lineage resulted in a fully resolved tree for each of the stem nodes along the backbone, which was previously intractable. In addition, we sampled a number of genera that have been historically difficult to place, including *Stiffia*, *Hyalis*, *Hecastocleis*, *Cavea*, *Platycarpha*, *Gundelia*, *Cyclolepis*, and *Corymbium*. The inclusion of anomalous and difficult to place genera (which also tend to be species poor) is important for biogeographic analyses of the family since, as previously noted (17), their placement nearly always anchors a large radiation.

While recent progress has been made to better understand the evolution of the Asteraceae through molecular phylogenetics using plastid markers (23, 27) and phylotranscriptomics (6), the highly resolved backbone of our nuclear phylogeny presented here further alters our perception of evolution in the family and will necessitate several changes in its classification, particularly for subfamilies Stiffioideae, Carduoideae (thistles), and Cichorioideae (dandelions). First, a few taxa (*Hyaloseris*, *Gongyolepis*, and *Leucomeris*), previously placed in Wunderlichioideae are now in Stiffioideae, which except for *Leucomeris*, are placements also supported by Panero and colleagues (23, 28). This makes more sense biogeographically and adds an interesting potential for long-distance dispersal events from South America to Asia for *Leucomeris* and its sister genus, *Nouelia* (not sampled here). Second, our data do not support the monophyly of subfamily Carduoideae (*sensu* 17). The southern African tribes Oldenburgeae + Tarchonantheae and the widespread African Dicomeae no longer share a most recent common ancestor (MRCA) with Carduoideae (core thistles) and instead form a grade paraphyletic to Carduoideae. Thus, the classification of Carduoideae needs revision to render it monophyletic. Third, the Asian Pertyeae is now shown to diverge immediately after the North American Hecastocleidae. This makes the case for an earlier arrival of the family into Asia than previously considered (see section below). Finally, our data do not support the monophyly of subfamily Cichorioideae, which necessitates a more narrowly defined

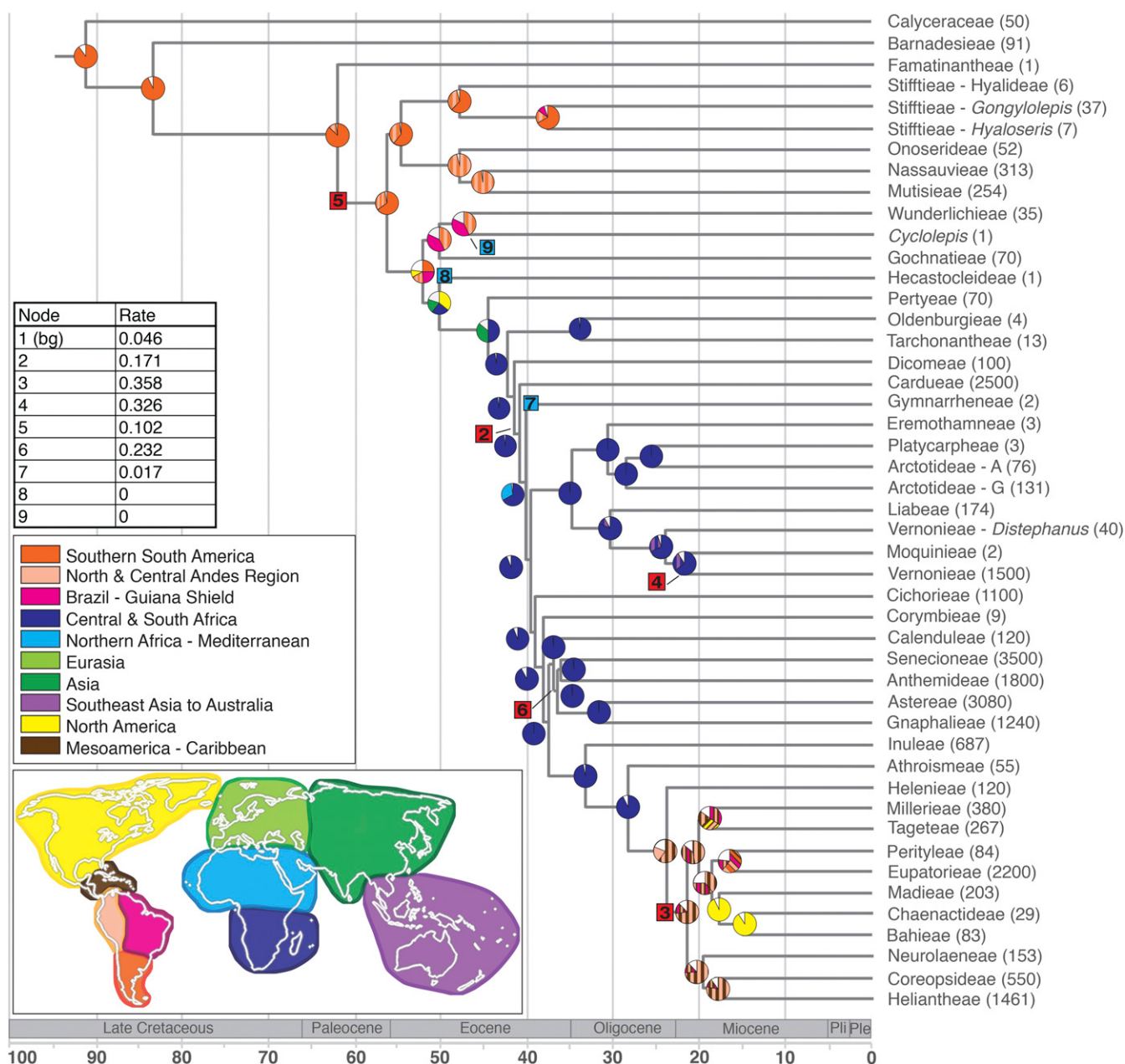


Fig. 3. Tribe-level chronogram and ancestral range estimates. Probabilities for ancestral ranges are illustrated in pie charts color coded by geographic regions on the world map. Diversification rate shifts are indicated on the phylogeny with numbered boxes corresponding to the table above the geographic legend. Rate shift increases in red and downshifts are blue boxes. Species numbers per tribe are indicated to the *Right* of tribe names.

subfamily consisting of only tribe Cichorieae. The six remaining tribes of subfamily Cichorioideae are monophyletic and need to be recognized as a separate subfamily named Vernonioidae. These vernonioid members of the currently defined Cichorioideae are atypical for Cichorieae, and understanding this new relationship may help identify the defining traits that support a newly defined Vernonioidae (already described in ref. 29).

Evolution of Asteraceae in Space and Time. Our data provide evidence that Asteraceae most likely originated during the late Cretaceous (~83 MYA, 95% CI 91–64) in southern South America and underwent several range extensions, dispersal events, and diversifications (Fig. 3). During the Cretaceous, the global climate was warmer than today with tropical forest vegetation at the poles (30). Cooling began at the end of the Cretaceous, resulting in increased seasonality (31) with subdeserts

beginning to form in central South America. The oldest lineage is the enigmatic Barnadesieae which is a small South American tribe of about 90 species with a strikingly different morphology than the rest of the family, which includes the frequent presence of prominent, paired axillary spines. Our analyses indicate that additional extant lineages of Asteraceae did not diverge until after the K–Pg. Alternatively, earlier lineages from other southern hemisphere continents may have been lost during this mass extinction event. Given the sparse fossil record for Asteraceae, and the general difficulty in assigning late Cretaceous/early Eocene pollen to extant angiosperm genera (23, 31), there are few data to more fully explore this since the only Cretaceous Asteraceae fossil evidence (*Tubuliflorides lilliei* type A, putatively attributed to Barnadesioidae) is from Antarctica and New Zealand (76–66 MYA; refs. 24 and 32), whereas the earliest diverging extant lineage is restricted to South America.

Toward the end of the Cretaceous and following the K–Pg mass extinction event (when an estimated five out of six species went extinct), the global temperature cooled while the Earth experienced several warming intervals characterized by extreme changes in climate and carbon cycling: the Paleocene–Eocene Thermal Maximum (PETM, ~56 MYA), the Early Eocene Climatic Optimum (EECO, ~53 MYA), and the Middle Eocene Climatic Optimum (MECO, ~42 MYA) (33). At the Paleocene–Eocene transition, the earliest apparent diversification from within the basal grade of Asteraceae occurred where we detected a rate acceleration on the stem lineage leading to the Mutisieae *sensu* Ortiz (34) (~55 MYA). The PETM and EECO comprise the warmest periods of the Cenozoic and are associated with dramatic terrestrial biome shifts, plant migrations, and accelerated rates of species diversification in many major plant and animal lineages (35–38). Multiple paleopolyploid events have been identified in Asteraceae after the divergence of the early lineage Barnadesieae (5) and along the stem nodes of several large tribes (6). Further, two recent studies have demonstrated a correlation between diversification accelerations and WGD in Asteraceae (7, 39). In addition, climatic instability has been associated with significant WGD events, and similar to other plant lineages (40), Asteraceae may have undergone explosive diversifications following WGDs that occurred during these intense climate upheavals.

Dispersal of Asteraceae out of South America occurred at this time of dramatic climate change roughly 50 MYA. The BioGeoBEARS analysis was ambiguous for these nodes in our phylogeny with multiple possibilities estimated for the ancestral ranges and for the dispersal out of South America (Fig. 3 and *SI Appendix*, Fig. S7). However, our phylogeny, which has Pertyeae diverging after Hecastoleideae, suggests a route from South America first through North America and then to Asia. Previous authors have hypothesized a migration route out of South America by a North American–Asian route or through a direct route via Africa (16–19, 23). A number of fossil pollen records, and the only confirmed capitulum macrofossil for the family, have been described from the early- to mid-Eocene in the Southern Hemisphere with similarities between the extant South American and African floras (e.g., refs. 18 and 24). Bremer (16) was the first to hypothesize a North American–Asian route for Asteraceae, and Panero and Funk (17) proposed Hecastoleideae as a link to Pertyeae in Asia because the two tribes share morphological similarities in floral and pollen characters. The phylogenetic placement of Pertyeae in Panero and Funk (17) was after the divergence of subfamily Carduoideae of African origin rendering the North American–Asian route less likely. Here, the link to Asia may be more plausible, given the sequential placement of Hecastoleideae followed by Pertyeae. Moreover, Eocene fossil pollen records suggest a distribution of Asteraceae in China (41), although the first reliable North American fossil evidence does not appear until the late Eocene ~37 MYA (42). Another possibility is concurrent dispersal to both North America and Africa during the mid-Eocene.

Whichever path Asteraceae took out of South America, it is clear that once the family reached Africa during the middle Eocene (~42 MYA, coincident with the MECO), an explosion of diversification occurred. By the end of the Eocene, continental interiors had begun to dry, and forests were thinning out considerably. During this time, Asteraceae experienced its greatest diversification and subsequent colonization of the globe. The diversification along the backbone was both rapid and massive. Our data support a substantial rate acceleration that ultimately resulted in nearly 95% of the family's diversity (Fig. 3, red box no. 2). From their African ancestors, these tribes diversified and colonized vast areas of the earth. These included tribes Cardueae (core thistles), Cichorieae (chicory, dandelions, and lettuce), Calenduleae (pot marigold), Senecioneae (ragwort), Anthemideae (chrysanthemums), Astereae (asters), Gnaphalieae (strawflowers), and the MRCA of the Heliantheae alliance (sunflowers and coneflowers). Another major rate acceleration and diversification occurred near the end of the Eocene around 36 MYA on the stem leading to the African-based tribes that include Senecioneae, Anthemideae, Astereae, and Gnaphalieae that today total nearly 10,000 extant species. WGDs have been detected for Gnaphalieae and Senecioneae and likely played a role in the diversification of these large tribes (6). Around 23 MYA, the MRCA of the Heliantheae

alliance diverged from the rest of the family and its descendants arrived in the New World coincident with the end of the Oligocene and another brief warming period and major climatic shift (Late Oligocene Warming Event). About 21 MYA, diversifications occurred that resulted in the segregate tribes of the Heliantheae alliance (representing more than 5,600 species), and the highest rate acceleration measured in our study was along this stem (Fig. 3, red box no. 3). Previous studies also support a WGD event at the crown node of the Heliantheae alliance (4, 6) that, along with long-distance dispersal into new habitats, likely played a major role in its diversification.

While we broadly sampled across Asteraceae, our study still represents a small portion of the total number of genera in the family (207 of the ~1,700 genera). Furthermore, extinctions that took place during the history of Asteraceae are difficult to account for in historical biogeographic analyses with a sparse pollen record. Nonetheless, our primary goal was to estimate the ancestral geographic ranges along the backbone of the tree, i.e., deep biogeographical events, thus we carried out this analysis at the level of genus to represent as much geographic diversity as possible. We coded geographic range by genus sampled, as opposed to species because species-level sampling is not practical given the size of the family and wide distribution of some sampled taxa. However, genus-level, rather than species-level coding could bias the results toward more widespread ancestral range estimates and/or lower estimates of founder effects. It is important to note that we did not see this bias in our results along the backbone of the tree, and our estimates are in general agreement with other biogeographic studies of the family (25, 43). Still we note that some interpretations of the biogeographic estimates, especially of lower taxonomic scales, will require more extensive sampling in future studies. For example, our sampling of the two very large tribes, Heliantheae and Eupatorieae, represents less than 0.005% of the ~5,600 species within the Heliantheae alliance, and therefore our limitations in sampling and/or coding could explain why the ancestral range for this lineage is estimated as more widespread than previous studies which hypothesize a North American origin (25, 43). Given that our Asteraceae-specific probe set has been successful in resolving phylogenetic relationships within both tribes and genera (ref. 44; Fig. 2), future studies testing biogeographical hypotheses at these levels may yield further insight into the migrations and dispersals of the family.

Phylogenomics coupled with biogeographic and diversification rate analyses have revealed that a series of explosive diversifications that began during the Eocene resulted in the tremendous diversity of this family. These diversifications were rapid, occurring in a few million years, and are associated with extensive dispersal events and significant climatic changes since the Cretaceous. The impact of these diversifications and dispersals on angiosperm biodiversity is substantial, with members of Asteraceae found on every continent and comprising 10% of flowering plants thus encompassing a major component of angiosperms, the world's most dominant group of terrestrial plants.

Materials and Methods

Taxon Sampling, Hyb-Seq, and Data Processing. We carried out sequence capture for 238 samples and incorporated 18 transcriptomes representing 13 subfamilies, ~45 tribes, 207 genera, and three outgroup taxa (*Dataset S2*). Sequence capture was performed and data were processed following the bioinformatic workflow and methods of ref. 45. Raw sequencing reads were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive under BioProject PRJNA540287.

Phylogenomic and Divergence Time Analyses. We generated phylogenetic trees based on a concatenated data matrix of all loci using a ML approach, Bayesian estimation, and a pseudocoalescence method to compute a consensus “species” tree based on individual gene trees. We generated a time-calibrated ML phylogeny and used fossil calibration points (*SI Appendix*, Table S1) to constrain nodes in the ML phylogeny. The resulting alignments and phylogenetic trees were deposited in FigShare DOIs 10.6084/m9.figshare.7697834 and 10.6084/m9.figshare.7695929 (46, 47).

Historical Biogeography and Diversification Analyses. Ancestral ranges were estimated using ML implemented on the dated phylogeny pruned to include only one species per genus (*SI Appendix*, Table S5). Diversification rates were estimated using the dated tree described above and pruned to the level of

tribe (*SI Appendix, Table S6*). Additional details of the methods are available in *SI Appendix*.

ACKNOWLEDGMENTS. We thank the Center for Biodiversity, High Performance Computing Cluster, and Feinstone Center at the University of Memphis; and A. Anderberg, M. Barker, R. Bayer, D. McKenna, S. Herrando-

Moraia, G. Johnson, K. Jones, C. Kellogg, C. Mason, N. Matzke, E. Pasini, E. Schilling, S. Shin, A. Susanna, and S. Wing. We also thank the many who contributed samples to the project. This project was funded by the National Science Foundation Division of Environmental Biology DEB-1745197, Undersecretary for Science Smithsonian Institution, and Fundação de Amparo à Pesquisa do Estado de São Paulo 2016/12446-1.

1. R. I. Lewis-Smith, M. Richardson, Fuegian plants in Antarctica: Natural or anthropogenically assisted immigrants? *Biol. Invasions* **13**, 1–5 (2011).
2. S. Carlquist, Tribal interrelationships and phylogeny of the Asteraceae. *Aliso* **8**, 465–492 (1976).
3. J. C. Semple, K. Watanabe, “A review of chromosome numbers in Asteraceae with hypotheses on chromosomal base number evolution” in *Systematics, Evolution, and Biogeography of Compositae*, V. A. Funk, A. Susanna, T. F. Stussey, R. J. Bayer, Eds. (IAPT, Vienna, 2009), pp. 61–72.
4. M. S. Barker *et al.*, Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. *Mol. Biol. Evol.* **25**, 2445–2455 (2008).
5. M. S. Barker *et al.*, Most Compositae (Asteraceae) are descendants of a paleohexaploid and all share a paleotetraploid ancestor with the Calyceraceae. *Am. J. Bot.* **103**, 1203–1211 (2016).
6. C. H. Huang *et al.*, Multiple polyploidization events across Asteraceae with two nested events in the early history revealed by nuclear phylogenomics. *Mol. Biol. Evol.* **33**, 2820–2835 (2016).
7. D. C. Tank *et al.*, Nested radiations and the pulse of angiosperm diversification: Increased diversification rates often follow whole genome duplications. *New Phytol.* **207**, 454–467 (2015).
8. S. Magallón, L. L. Sánchez-Reyes, S. L. Gómez-Acevedo, Thirty clues to the exceptional diversification of flowering plants. *Ann. Bot.* **123**, 491–503 (2019).
9. M. A. Chapman, J. H. Leebens-Mack, J. M. Burke, Positive selection and expression divergence following gene duplication in the sunflower CYCLOIDEA gene family. *Mol. Biol. Evol.* **25**, 1260–1273 (2008).
10. S. Tähtiharju *et al.*, Evolution and diversification of the CYC/TB1 gene family in Asteraceae—A comparative study in *Gerbera* (Mutisieae) and sunflower (Heliantheae). *Mol. Biol. Evol.* **29**, 1155–1166 (2012).
11. L. M. Calabria, V. P. Emerenciano, M. T. Scotti, T. J. Mabry, “Secondary chemistry of Compositae” in *Systematics, Evolution, and Biogeography of Compositae*, V. A. Funk, A. Susanna, T. F. Stussey, R. J. Bayer, Eds. (IAPT, Vienna, 2009), pp. 73–88.
12. L. Pazouki, H. R. Memari, A. Kännaste, R. Bichele, Ü. Niinemets, Germacrene a synthase in yarrow (*Achillea millefolium*) is an enzyme with mixed substrate specificity: Gene cloning, functional characterization and expression analysis. *Front. Plant Sci.* **6**, 111 (2015).
13. Q. Shen *et al.*, The genome of *Artemisia annua* provides insight into the evolution of Asteraceae family and artemisinin biosynthesis. *Mol. Plant* **11**, 776–788 (2018).
14. R. K. Jansen, J. D. Palmer, A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). *Proc. Natl. Acad. Sci. U.S.A.* **84**, 5818–5822 (1987).
15. R. K. Jansen, J. D. Palmer, Phylogenetic implications of chloroplast DNA restriction site variation in the Mutisieae (Asteraceae). *Am. J. Bot.* **75**, 753–766 (1988).
16. K. Bremer, *Asteraceae: Cladistics and Classification* (Timber Press, Portland, 1994).
17. J. L. Panero, V. A. Funk, The value of sampling anomalous taxa in phylogenetic studies: Major clades of the Asteraceae revealed. *Mol. Phylogenet. Evol.* **47**, 757–782 (2008).
18. L. Katinas, J. V. Crisci, P. Hoch, M. C. Telleria, M. J. Apodaca, Trans-oceanic dispersal and evolution of early composites (Asteraceae). *Perspect. Plant Ecol. Evol. Syst.* **15**, 269–280 (2013).
19. V. A. Funk *et al.*, “Compositae metatrees: The next generation” in *Systematics, Evolution, and Biogeography of Compositae*, V. A. Funk, A. Susanna, T. F. Stussey, R. J. Bayer, Eds. (IAPT, Vienna, 2009), pp. 747–777.
20. K. J. Kim, K. S. Choi, R. K. Jansen, Two chloroplast DNA inversions originated simultaneously during the early evolution of the sunflower family (Asteraceae). *Mol. Biol. Evol.* **22**, 1783–1792 (2005).
21. S. A. Smith, J. M. Beaulieu, M. J. Donoghue, An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5897–5902 (2010).
22. J. M. Beaulieu, D. C. Tank, M. J. Donoghue, A Southern Hemisphere origin for campanulid angiosperms, with traces of the break-up of Gondwana. *BMC Evol. Biol.* **13**, 80 (2013).
23. J. L. Panero, B. S. Crozier, Macroevolutionary dynamics in the early diversification of Asteraceae. *Mol. Phylogenet. Evol.* **99**, 116–132 (2016).
24. V. D. Barreda *et al.*, Early evolution of the angiosperm clade Asteraceae in the Cretaceous of Antarctica. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10989–10994 (2015).
25. V. A. Funk *et al.*, Everywhere but Antarctica: Using a supertree to understand the diversity and distribution of the Compositae. *Biol. Skr.* **55**, 343–374 (2005).
26. J. C. Loudon, *An Encyclopaedia of Plants* (Longman, Rees, Orme, Brown and Green, London, 1829).
27. Z. X. Fu *et al.*, A comprehensive generic-level phylogeny of the sunflower family: Implications for the systematics of Chinese Asteraceae. *J. Syst. Evol.* **54**, 416–437 (2016).
28. J. L. Panero *et al.*, Resolution of deep nodes yields an improved backbone phylogeny and a new basal lineage to study early evolution of Asteraceae. *Mol. Phylogenet. Evol.* **80**, 43–53 (2014).
29. B. L. Turner, “Summary of the biology of the Compositae” in *The Biology and Chemistry of the Compositae*, V. H. Heywood, J. B. Harborne, B. L. Turner, Eds. (Academic Press, New York, 1977), vol. 2, pp. 1105–1118.
30. B. L. Otto-Bliesner, G. R. Upchurch, Jr, Vegetation-induced warming of high-latitude regions during the Late Cretaceous period. *Nature* **385**, 804–807 (1997).
31. A. Graham, The age and diversification of terrestrial New World ecosystems through Cretaceous and Cenozoic time. *Am. J. Bot.* **98**, 336–351 (2011).
32. V. D. Barreda *et al.*, Reply to Panero: Robust phylogenetic placement of fossil pollen grains: The case of Asteraceae. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E412 (2016).
33. J. C. Zachos, G. R. Dickens, R. E. Zeebe, An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* **451**, 279–283 (2008).
34. S. Ortiz *et al.*, “The basal grade of Compositae: Mutisieae (Sensu Cabrera) and Carduoideae” in *Systematics, Evolution, and Biogeography of Compositae*, V. A. Funk, A. Susanna, T. F. Stussey, R. J. Bayer, Eds. (IAPT, Vienna, 2009), pp. 193–213.
35. G. J. Bowen *et al.*, Mammalian dispersal at the Paleocene/Eocene boundary. *Science* **295**, 2062–2065 (2002).
36. S. L. Wing *et al.*, Transient floral change and rapid global warming at the Paleocene-Eocene boundary. *Science* **310**, 993–996 (2005).
37. E. Schuettpelz, K. M. Pryer, Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11200–11205 (2009).
38. C. Jaramillo, M. J. Rueda, G. Mora, Cenozoic plant diversity in the neotropics. *Science* **311**, 1893–1896 (2006).
39. M. J. Landis, W. A. Freyman, B. G. Baldwin, Retracing the Hawaiian silversword radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty. *Evolution* **72**, 2343–2359 (2018).
40. L. Cai *et al.*, Widespread ancient whole-genome duplications in Malpighiales coincide with Eocene global climatic upheaval. *New Phytol.* **221**, 565–576 (2019).
41. S. Zhi-Chen, W. Wei-ming, H. Fei, Fossil pollen records of extant angiosperms in China. *Bot. Rev.* **70**, 425–458 (2004).
42. W. C. Elsik, T. E. Yancey, Palynomorph biozones in the context of changing paleoclimate, middle Eocene to lower Oligocene of the Northwest Gulf of Mexico. *Paleontology* **24**, 177–186 (2000).
43. C. McDonald-Spicer, N. J. Knerr, F. Encinas-Viso, A. N. Schmidt-Lebuhn, Big data for a large clade: Bioregionalization and ancestral range estimation in the daisy family (Asteraceae). *J. Biogeogr.* **46**, 255–267 (2019).
44. S. Herrando-Moraia; Cardueae Radiations Group, Exploring data processing strategies in NGS target enrichment to disentangle radiations in the tribe Cardueae (Compositae). *Mol. Phylogenet. Evol.* **128**, 69–87 (2018).
45. J. R. Mandel *et al.*, A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae. *Appl. Plant Sci.* **2**, 1300085 (2014).
46. R. B. Dikow *et al.*, Locus alignments and RAxML trees. FigShare. <https://doi.org/10.6084/m9.figshare.7697834.v1>. Deposited 7 May 2019.
47. R. B. Dikow *et al.*, Concatenated alignment, ML, and Bayesian tree. FigShare. <https://doi.org/10.6084/m9.figshare.7695929.v1>. Deposited 7 May 2019.