Current Biology

Article

The evolutionary history of bees in time and space

Highlights

- Bees likely originated in Gondwana before the breakup of Africa and South America
- The geological history of southern continents impacted early bee diversification
- Northward range expansions in the Tertiary impacted modern bee distribution
- Southern Hemisphere origin of bees parallels the histories of many plant groups

Authors

Eduardo A.B. Almeida, Silas Bossert,
Bryan N. Danforth, ...,
Michael Kuhlmann,
Michael G. Branstetter, Marcio R. Pie

Correspondence

eduardoalmeida@usp.br (E.A.B.A.),
silas.bossert@wsu.edu (S.B.)

In brief

Almeida, Bossert, et al. combine phylogenomics and the fossil record to propose a comprehensive hypothesis of bee biogeography. Bees originated in Gondwana and later colonized all continents through a series of events that parallel the histories of numerous flowering plants, offering opportunities to interpret the evolution of pollination in space.

Almeida et al., 2023, Current Biology *33*, 3409–3422
August 21, 2023
https://doi.org/10.1016/j.cub.2023.07.005

Current Biology

CellPress

Article

The evolutionary history of bees in time and space

Eduardo A.B. Almeida,[1,18,19,*] Silas Bossert,[2,3,18,*] Bryan N. Danforth,[4] Diego S. Porto,[1,5] Felipe V. Freitas,[1,2]
Charles C. Davis,[6] Elizabeth A. Murray,[2,3] Bonnie B. Blaimer,[3,7] Tamara Spasojevic,[3,8,9] Patrícia R. Ströher,[10,11]
Michael C. Orr,[12,13] Laurence Packer,[14] Seán G. Brady,[3] Michael Kuhlmann,[15] Michael G. Branstetter,[16]
and Marcio R. Pie[10,17]
[1]Departamento de Biologia, Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo, Ribeirão Preto, São Paulo 14040-901, Brazil
[2]Department of Entomology, Washington State University, Pullman, WA 99164, USA
[3]Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA
[4]Department of Entomology, Cornell University, Comstock Hall, Ithaca, NY 14853, USA
[5]Finnish Museum of Natural History - LUOMUS, University of Helsinki, Helsinki 00014, Finland
[6]Department of Organismic and Evolutionary Biology, Harvard University Herbaria, 22 Divinity Avenue, Cambridge, MA 02138, USA
[7]Center for Integrative Biodiversity Discovery, Museum für Naturkunde, Leibniz-Institute for Evolution and Biodiversity Science, 10115 Berlin, Germany
[8]Life Sciences, Natural History Museum Basel, 4051 Basel, Switzerland
[9]Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland
[10]Departamento de Zoologia, Universidade Federal do Paraná, Curitiba, Paraná 81531-990, Brazil
[11]Department of Anthropology and Archaeology, University of Calgary, Calgary, AB T2N 1N4, Canada
[12]Entomologie, Staatliches Museum für Naturkunde Stuttgart, 70191 Stuttgart, Germany
[13]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China
[14]Department of Biology, York University, Toronto, ON M3J 1P3, Canada
[15]Zoological Museum, University of Kiel, Hegewischstr. 3, 24105 Kiel, Germany
[16]U.S. Department of Agriculture, Agricultural Research Service, Pollinating Insects Research Unit, Utah State University, Logan, UT 84322, USA
[17]Department of Biology, Edge Hill University, St Helens Rd, Ormskirk, Lancashire L39 4QP, UK
[18]These authors contributed equally
[19]Lead contact
*Correspondence: eduardoalmeida@usp.br (E.A.B.A.), silas.bossert@wsu.edu (S.B.)
https://doi.org/10.1016/j.cub.2023.07.005

SUMMARY

Bees are the most significant pollinators of flowering plants. This partnership began ca. 120 million years ago, but the uncertainty of how and when bees spread across the planet has greatly obscured investigations of this key mutualism. We present a novel analysis of bee biogeography using extensive new genomic and fossil data to demonstrate that bees originated in Western Gondwana (Africa and South America). Bees likely originated in the Early Cretaceous, shortly before the breakup of Western Gondwana, and the early evolution of any major bee lineage is associated with either the South American or African land masses. Subsequently, bees colonized northern continents via a complex history of vicariance and dispersal. The notable early absences from large landmasses, particularly in Australia and India, have important implications for understanding the assembly of local floras and diverse modes of pollination. How bees spread around the world from their hypothesized Southern Hemisphere origin parallels the histories of numerous flowering plant clades, providing an essential step to studying the evolution of angiosperm pollination syndromes in space and time.

INTRODUCTION

Bees are a vital component of terrestrial biomes worldwide,[1,2] yet their spatiotemporal origin remains uncertain. With over 20,000 described species in seven families,[1,3,4] bees are the most species-rich lineage of pollinivorous insects, inhabiting natural and agricultural ecosystems on all continents except Antarctica.[4,5] The interdependence of bees and flowering plants, which first arose in the Cretaceous,[6,7] has profoundly impacted terrestrial ecosystems globally. Although bees as a whole are most species-rich in xeric, seasonally warm temperate regions of the Southern and Northern hemispheres,[1,2,8] we lack a comprehensive understanding of the historical events that led to their present-day distributions (although advances were made for certain taxa[9–19]). These shortcomings hinder our ability to explore the role bees have played in the assembly of modern biotas, especially concerning their relevance to the evolution of the eudicot angiosperm pollination syndromes (i.e., suites of floral traits associated with pollination).

Modern historical bee biogeography builds upon Charles Michener's landmark contribution on the "Biogeography of the Bees,"[8] a major synthesis of bee and angiosperm natural history

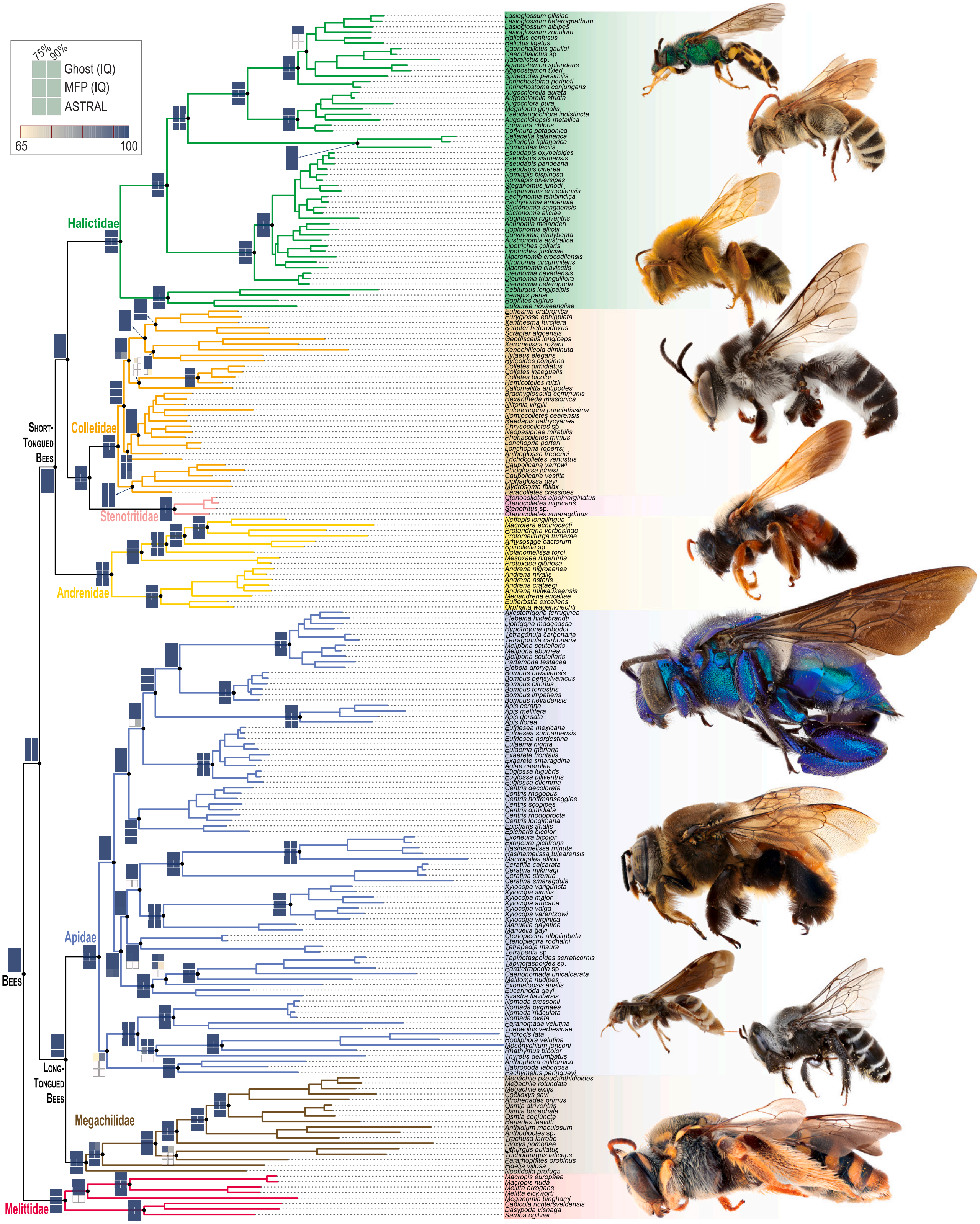

Figure 1. Bee phylogeny

Phylogenetic hypothesis of bees (Hymenoptera: Apoidea: Apiformes) inferred from 830 ultraconserved element (UCE) loci (set of loci filtered for 75% taxonomic completeness) and the general heterogeneous evolution on a single topology (GHOST) model in IQ-TREE.[30,31] All bee families and subfamilies are represented in

(legend continued on next page)

at the time, and the single most influential contribution on bee biogeography to date. Michener hypothesized that bees originated in the arid regions of Western Gondwana (modern South America and Africa), in habitats considered climatically similar to present-day bee biodiversity hotspots.[2,8]

Although Michener's *Western Gondwana hypothesis* has been widely accepted for over four decades, it was based on an understanding of bee phylogeny and evolutionary history that is dramatically different from the currently accepted view. First, the advent of model-based phylogenetic methods and the extensive use of molecular data to infer phylogenetic relationships fundamentally revised our understanding of bee evolution,[20] including major tribal, subfamily, and family-level rearrangements. For example, Michener's synthesis[8] considered the family Colletidae "most primitive among bees," while molecular-based studies consistently recover Colletidae as a relatively young family.[4,6,21–23] Several enigmatic lineages were considered families at the time (Oxaeidae, Ctenoplectridae, and Fideliidae), but recent phylogenetic evidence identified them as morphologically derived members of well-established groups (within Andrenidae, Apidae, and Megachilidae, respectively).[6,9,15,21–23] Second, Bayesian estimates of divergence times based on DNA data now allow us to establish a global time frame of bee evolution by incorporating accurate divergence times for families, subfamilies, and tribes. Finally, the past decades witnessed a significant expansion of the known bee fossil record: the number of described extinct bee species has more than doubled since 1979, now providing nearly 200 spatiotemporal data points[24–28] crucial to understanding past dispersal and extinction dynamics. These include the oldest undisputed fossil bee, *Cretotrigona prisca*, from New Jersey amber, which was recovered far outside of its lineage's present-day range.[24,26] The fossil record further suggests that some ancient bee lineages went extinct during massive extinction events, particularly the Eocene-Oligocene transition[25,28] and another event close to the K-T (Cretaceous-Tertiary) boundary.[24,29]

In light of this new evidence, it is time to revisit the longstanding question of the spatial origin of bees and explicitly test the *Western Gondwana hypothesis*. To this end, we developed the taxonomically broadest phylogenomic framework of bees to date, including species from all currently recognized families (7) and subfamilies (28). Our divergence time estimates incorporate information from nearly 200 fossil bees, providing a robust temporal backbone of bee evolution. We combine analytical results based on this detailed phylogenomic framework with a paleobiogeographic reconstruction based on the fossil record to create an integrative scenario of bee evolution in space and time. Our analysis reveals that Charles Michener's *Western Gondwana hypothesis* was correct: bees did indeed have their origins in Western Gondwana. However, our analysis reveals important new insights into both the timing and spatial expansion of bees out of their reconstructed place of origin. Based on our results, we can now interpret the current geographic distribution of bees in a well-resolved historical framework.

RESULTS AND DISCUSSION

Phylogeny

To infer a comprehensive phylogeny of bees, we sampled 216 species representing all major lineages (7/7 families; 28/28 subfamilies, 57/69 tribes) from all biogeographic regions (Data S1A). This represents a 4-fold increase in relation to the sampling of the most recent phylogenomic study to include a significant representation of bee diversity.[21] We filtered loci for completeness to generate two datasets representing two levels of taxon occupancy (percentage of taxa required to be represented for each locus): 75% and 90% filtered locus sets. These datasets included 830 loci and 364,396 bp (251,279 informative sites), and 377 loci and 181,287 bp (125,151 informative sites) of sequence data, respectively. For each dataset, we assessed the effects of the analysis type (coalescent-based species-tree approach using ASTRAL vs. maximum likelihood [ML] analyses of the concatenated data using IQ-TREE), as well as substitution models. This combination of analytical strategies and data filtering produced six trees (available for download from the project's Mendeley data repository; see data and code availability), summarized in Figure 1.

Our results recovered bees, as well as all families and subfamilies, as monophyletic based on all analyses (Figure 1). Our overall tree topology at the family level was congruent with those of recently published higher-level phylogenies,[21,32] although our taxon sampling is much more extensive. We adopted the general heterogeneous evolution on a single topology (GHOST)-ML tree inferred from the 75% complete dataset (Figure 1) as our preferred phylogeny, given its consistency with studies that investigated relationships below the family level with denser taxon sampling: Andrenidae,[9,11] Apidae,[15,33] Colletidae,[12] Halictidae,[34] Megachilidae,[14] Melittidae,[35] and Stenotritidae.[12] Although gene-tree conflicts are expected in phylogenomics, all nodes in our best tree (Figure 1) were generally favored by the loci with highest phylogenetic accuracy (Figure S1; Data S2; phylogenetic accuracy of a locus was assessed by its ability of resolving a set of well-supported nodes—refer to Table S1 and STAR Methods for details). Our results were also congruent with recent phylogenomic studies of aculeate Hymenoptera employing different kinds of high-throughput sequence data.[21,32,36] Comparing our higher-level phylogenetic results to those of previously published studies,[6,9,11–15,20,21,23,33–35,37–39] we are confident that we have sampled all major lineages and accurately captured major branching events in the early diversification of bees (a prerequisite for reliably inferring biogeographic patterns).

the tree based on the sampling of 216 species. Node support of selected clades is displayed as Navajo rugs for comparison of six strategies to analyze the data contrasting maximum likelihood-based concatenation (two models) and species-tree summary approaches, and two levels of data completeness; white rectangles indicate that the node was not recovered (see also Figure S1; Table S1; Data S1 and S2). Pictures to the right of the tree illustrate part of the morphological diversity of bees sampled in this study (from top to bottom, photos to scale)—Halictidae, *Agapostemon virescens* ♂ and *Macronomia clavisetis* ♂; Colletidae, *Mydrosoma fallax* ♂; Stenotritidae, *Ctenocolletes nigricans* ♂; Andrenidae, *Euherbstia excellens* ♀; Apidae, *Exaerete smaragdina* ♀, *Centris longimana* ♂, *Paranomada velutina* ♀; Megachilidae, *Lithurgus pullatus* ♀; and Melittidae, *Meganomia binghami* ♀.

Antiquity

To reconstruct the timeline of bee evolutionary history, we combined our phylogenetic data with 185 identified bee fossils (Data S1B) under the Bayesian fossilized birth-death process, similar to the strategy of Bossert et al.[9] but with superior sampling of fossils. Because of the computational challenges of estimating divergence times from large phylogenomic datasets, particularly with many fossil calibrations, we subsampled sets of loci from the original matrix. We contrasted three strategies of filtering loci to generate subsampled matrices, chosen either due to desirable features of the data (i.e., clocklikeness or average bootstrap values) or randomly selected sets of fragments. Such an approach can be justified as it has been demonstrated that the accuracy of divergence time estimation is not due to the amount of molecular data,[16] and it accounts for numerous potential biases.

Based on our analysis, bees originated in the Early Cretaceous (95% credibility interval: 106.6–137.9 million years ago [mya], mean: 124 mya) and all divergences among the major lineages recognized as families occurred between the mid- and the Late Cretaceous (ca. 73–124 mya: Figure 2). Although little is known about Cretaceous extinctions that may have wiped out other major bee lineages (see Engel[25]), our results confidently show that the major clades represented by modern families diverged from their closest relatives and most likely differentiated prior to the K-T boundary. Our results contrast with some recent studies that recovered younger ages (Figure 2), and we attribute that to the limited taxon sampling of certain bee lineages, methodological differences, and the placement of fossil calibration points related to the focus of those studies on the deeper nodes of Aculeata.[21,32]

Historical biogeography

Six biogeographic schemes were used to infer how the current distribution of bees occurred using a dispersal-extinction-cladogenesis (DEC) range evolution model on the timetree (Figure 3). These scenarios were based on three submodels that describe how connectivity between areas varied over time (Figure 4), each applied to two coding strategies for the geographic occurrence of bee lineages, either restricting the number of occupied areas to three or fewer or leaving the number of areas unconstrained (all reconstructions are available for download from the project's Mendeley data repository; see data and code availability). The exploration of alternative constraints to model range evolution through time was guided by geological scenarios conveying the effect of plate tectonic movements[41–47] on dispersal probabilities over time.

Early Cretaceous (120 mya)

Bees almost certainly arose in Western Gondwana (Africa + South America; Figure 3), in agreement with Michener's[8] hypothesis. Nonetheless, incorporating the significantly expanded fossil record of bees enables a much-improved resolution for understanding both the spatial and temporal aspects of range expansions and dispersal routes.

Western Gondwana was hypothesized to have been xeric,[8,42] which was one of the main reasons that led Michener to conclude that bees originated there.[8] Our results confirm the *Western Gondwana hypothesis* and highlight the incredibly long-lasting phylogenetic niche conservatism of bees and their climatic preferences: most bees have almost certainly always thrived in xeric environments and seasonal temperate climates. Thus, the climatic conditions that are believed to have favored bee origins are the same as those that characterize bee diversity hotspots today.[2]

Mid-Cretaceous (100 mya)

The separation of South America and Africa approximately 100 mya split each of the two long-tongued bee families (Apidae and Megachilidae) into African and South American lineages.

Melittidae remained restricted to Africa, based on our analysis, either because they were spatially limited to the eastern portion of Western Gondwana or because melittid lineages in South America subsequently went extinct. This renders Melittidae the only extant family of bees that likely underwent an early diversification exclusively in the Eastern Hemisphere. The close association of Melittidae with Africa is also congruent with the observation that this is the only continent to host all three subfamilies.[37]

These Cretaceous-age tectonic events also restricted the ancestors of modern Colletidae, Andrenidae, and Halictidae to South America. The Neotropical region holds an exceptionally rich flora,[48] and it is possible that the diversification of angiosperms in the Neotropics relates to their long-lasting association with bees. Some specialized associations between bees and their host plants have likely evolved in the Cretaceous.[49]

Our reconstructions suggest that bees had not yet reached Australia or the Indian landmass in the mid-Cretaceous. In fact, our results indicate that the Indian subcontinent did not host any bees for much of the period that it remained isolated from other continents (up until the Eocene). This is consistent with the observation that there are no ancient, endemic lineages of bees uniquely associated with India.

Late Cretaceous (80 mya)

Michener[8] suggested that bees originated in the Late Cretaceous. Our results render this hypothesis an underestimate; we show that bees originated earlier, diversified faster, and spread wider than he suspected. All extant bee families had likely diverged from each other and were present by the end of the Cretaceous (Figures 2 and 3), and all three southern continents harbored at least one of the extant families (Figure 3). This could be explained by the first range expansion onto the Australian continent: our biogeographic reconstructions indicate that by the Late Cretaceous a single family, Colletidae, had reached Australia from South America (Figure 3). This interchange occurred when climatic conditions and connectivity between South America, Antarctica, and Australia permitted a biogeographic integration of faunas among these continents from the Cretaceous until the Eocene,[12] although apparently only colletid bees took this southern route.[9] Currently, 50% of the Australian bee fauna comprises colletid bees,[50] paralleling the dominance of marsupials there. Other lineages of bees currently present in Australia (Apidae, Megachilidae, and Halictidae) arrived in Australia much later (see below).

Early Paleogene (60–40 mya)

According to the fossil record and DEC reconstructions, bees began colonizing the Northern Hemisphere in the Late

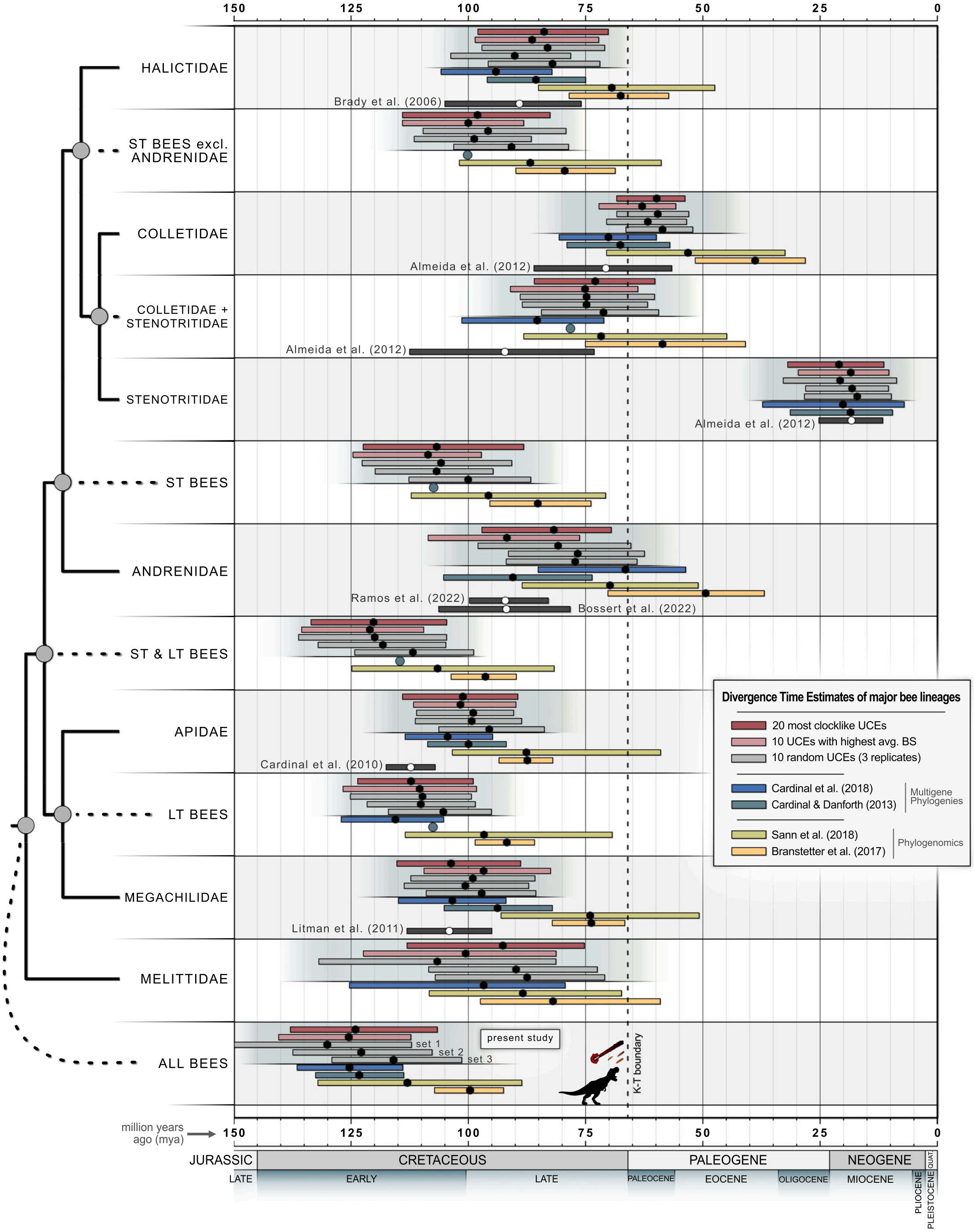

Figure 2. Divergence ages of bees and their major lineages

Comparison across datasets and studies of the estimated antiquity of 13 clades of bees (Apoidea: Apiformes) expressed by the means of their crown-group ages (dots) with 95% highest posterior density intervals (bars) when available. Five sets of divergence times estimated in this study (top bars) are contrasted with other comprehensive dating analyses of bees,[6,21,22,32,40] as well as research focusing on specific lineages.[9,11,12,14,40] The tree on the left summarizes the family-level phylogeny of bees (Figure 1); BS denotes bootstrap support; ST bees denote the short-tongued bees (Andrenidae, Colletidae, Halictidae, and Stenotritidae), and LT bees denote the long-tongued bees (Apidae and Megachilidae).

Cretaceous and early Paleogene via connections between South and North America, and between Africa and Europe, beginning at ca. 60 mya (Figure 3). The combination of different routes used by taxa expanding their ranges northward in combination with interchanges across northern land masses (Bering Land Bridge and North Atlantic Land Bridges[51]) created a complex

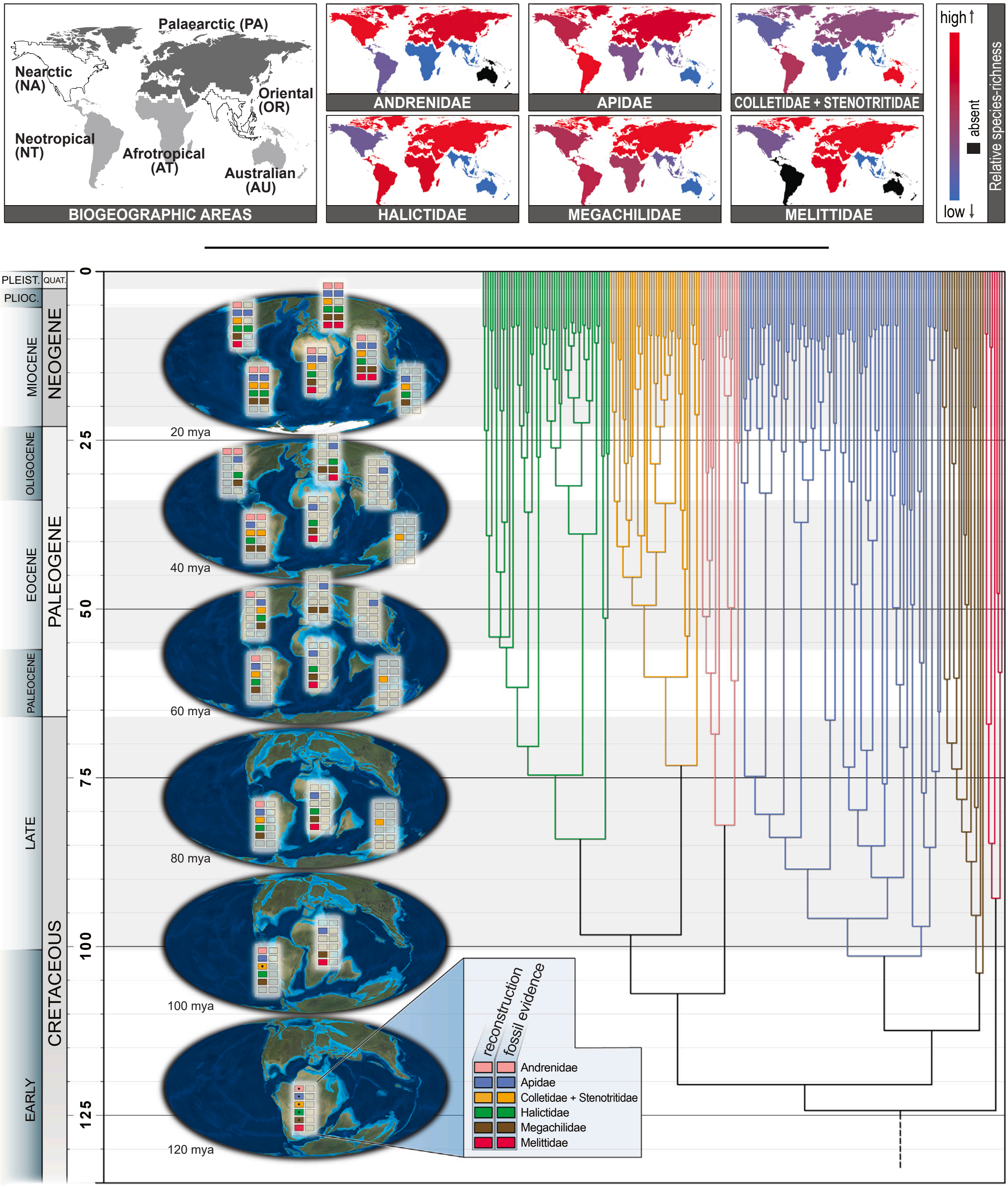

Figure 3. Biogeographic reconstruction of the history of bees based on fossil and model-based analytical evidence

The biogeographic interpretation of extant and fossil bees is displayed on paleomaps[41] representing geological reconstructions as 20 mya intervals from 120 mya to present-day using Bayesian dispersal extinction cladogenesis (DEC) and paleodistribution of 221 known fossils of the group (Data S3). DEC reconstructions synthesize the likely ancestral ranges (available for download from the project's Mendeley data repository; see data and code availability; see also Figure S3 and Data S1) condensed into family-level summaries. Colored rectangles indicate the inferred presence of a family in a region, as indicated by the DEC results, the fossil record, or both; rectangles containing a black dot indicate a stem-lineage presence for the respective family. The scale bar on the left shows the number of

(legend continued on next page)

biogeographic scenario partly captured by the fossil record from Eurasia and North America. Apidae and Megachilidae are documented by Paleocene fossils from Europe (Data S3F and S3AE), probably via range expansions of both lineages from Africa. These hypothesized northward range extensions were only detected by integrating different sources of evidence (Figure 3). Neither fossils on their own nor DEC model-based reconstructions on their own provided a complete picture independently.

The Paleocene and Eocene were marked by thermal maxima, when higher latitudes temporarily housed tropical and subtropical habitats.[52–54] These periods were climatically favorable for bee lineages restricted to warm conditions, likely enabling them to reach and temporarily expand their ranges northward. Bossert et al.[9] (their Figures 1 and 2) and Freitas et al.[16] (their Figure 5) also hypothesized northward range expansions of andrenid and apid bees in the Paleocene-Eocene periods, providing good examples of how such biogeographic events can be inferred with fine resolution. Past climatic cycles seemingly allowed the establishment of taxa previously endemic to Africa and South America to widen their geographic distributions, although some taxa must be assumed to have subsequently gone extinct from higher latitudes, likely when climates were no longer favorable (as seems likely in Anthophorinae[39]). The best example of such spatiotemporal fluctuation in range comes from the stingless bees (Apidae: Meliponini), currently restricted to tropical regions but with a fossil record in areas of North America and Europe that serve as vestiges of these past northward expansions.[18]

The concept of a band of tropical habitats throughout the Northern Hemisphere was named the ‘‘boreotropical hypothesis,’’ and it was invoked to explain range expansions of primarily tropical taxa to North America and Eurasia that were favored during warmer climatic periods,[55] particularly the thermal maxima of the Eocene.[52] Stingless bees provide one of the best examples of a lineage currently restricted to the tropics that expanded northward in the Paleogene in association with warming temperature in the Northern Hemisphere, as evidenced by *Cretotrigona prisca*. The boreotropical hypothesis has been invoked to explain the discrepancy between the geographical range of fossils and of extant lineages for several taxa of plants and animals,[52,56–60] but our study marks the first to invoke the boreotropical hypothesis to explain bee evolution on such a broad scale.

Late Paleogene: Eocene/Oligocene (40 mya) and the Neogene (20 mya)

According to our biogeographic hypothesis, different bee lineages continued migrating between South-North America and Africa-Europe in the late Paleogene, and between areas in northern regions likely across the Bering Land Bridge and North Atlantic Land Bridges. Melittidae, previously restricted to Africa, most likely expanded its range northward from Africa, as indicated above (Data S3AO and S3AP). It is worth noting that groups that diversified in southern continents after the periods of Eocene global warming and closure of a boreotropical corridor (e.g., Oxaeinae, Protandrenini, Emphorini[9,16]) apparently did not expand their distribution as much as taxa that began diversifying earlier. Important extinction events, partly observable via the fossil record of bees, also marked the late Eocene, such as the extinction of certain corbiculate apid bee lineages (e.g., Melikertini).[25]

All bee families reached southern Asia in the Neogene or earlier according to the DEC reconstructions (Figure 3). Notably, Apidae, Megachilidae, Halictidae all seem to have reached Australia via the ‘‘northern route’’ through southern Asia and New Guinea. Up to the late Oligocene-Neogene, the fauna of Australia was entirely composed of two closely related bee families (Colletidae and Stenotritidae). Extinct lineages of corbiculate apid bees were recorded from India in the Early Eocene (Data S1B),[28] indicating that faunistic exchanges with Asia began when these landmasses collided around 50 mya. India had harbored tropical forests since the Cretaceous,[28] but the novel pollinator interactions made possible by the arrival of bees in the Eocene probably influenced the local floral diversification dramatically, warranting future study.

Biogeographic associations with angiosperms

There are numerous parallels between bee biogeography outlined here and the biogeographic history of their primary mutualistic partners, the angiosperms. The first is the recognition that approximately one-quarter of angiosperm species (70,000) belong to the large and diverse rosid clade, which constitutes most tropical and temperate tree species diversity that form the terrestrial matrix for bees.[61] The major outlines of rosid phylodiversity arose exceptionally rapidly, similar to the tempo and timescale (108–91 mya[61]) in which bees diversified. Along these lines, several key angiosperm clades similarly show early vicariance, implicating the fragmentation of Western Gondwana in their biogeographic histories.[62–66] Despite these connections, however, there remain few exceptionally clear-cut cases of modern flowering plant clades (families) with obvious Western Gondwanan biogeographic affinities.[62,67] More recent northward migrations of angiosperms from their southern tropical origins to Laurasia in the north, as hypothesized here for bees, are more plentiful and involve bidirectional corridors via the proto-Caribbean or the emerging Isthmus of Panama, and interspersed land connections bridging Africa with Eurasia.[56,68–71] Similarly, once established in the north, the migration of tropical lineages involving North Atlantic land connections was likely common,[53,56,68,69,72–74] until paleoclimate and land configurations deteriorated this boreotropical superhighway, triggering plant extinction, range contraction, or adaptation as temperate environments arose. The evidence is less compelling for tropical plants using Beringian corridors during the Cenozoic owing to less suitable climates.[75,76] We hypothesize similar physiological constraints impacted the distribution and range expansions of bees.

The separation of the Indian subcontinent from Gondwana and its subsequent northward drifting greatly predates our inferred arrival time of bees in this region by tens of millions of years. At least a small number of angiosperm clades are hypothesized

million years before the present and corresponding geological periods and epochs. The depth of divergence times of the chronogram topology on the right is based on the Bayesian analysis of the 20 most clocklike UCE loci, and branch colors match the family coding of the box on the bottom center. The top section shows the six large world divisions (left) and a family-level representation of the present species richness in each area (right; Table S2).

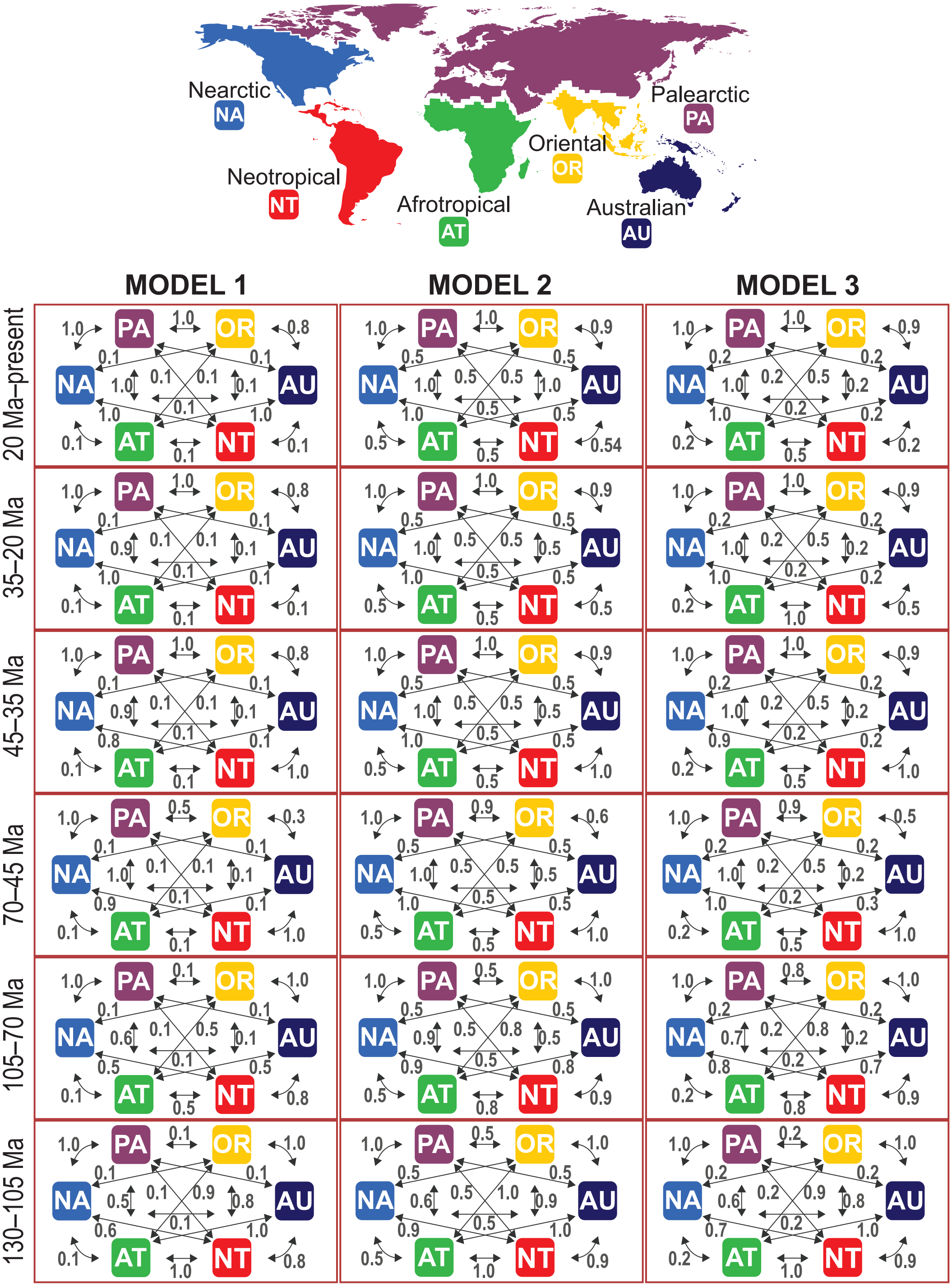

Figure 4. Biogeographic regions and models applied to the estimation of bee range evolution

Six areas defined for the biogeographic reconstruction of bee evolution include the major regions recognized in zoogeography (map on top): Afrotropical, Australian, Nearctic, Neotropical, Oriental, and Palearctic. Dispersal-extinction cladogenesis (DEC) analyses considered three scenarios of connectivity between areas, represented by models 1–3, in which six time-slices were defined and varying prior probabilities of connectivity were assigned to pairs of areas in each case. In this graphical summary of the connectivity matrices, it is shown that model 1 is most restrictive toward long-distance dispersals and overall connectivity between areas (i.e., probability of dispersals between disjunct areas close to null), model 2 is the most permissive, whereas model 3 represents an intermediary scenario among them. See STAR Methods for details of time-slice delimitation and geological hypotheses justifying the probabilities assigned to area connectivity.

to have journeyed northward on this floating land mass before it collided with Asia[66,77] in the Eocene,[43–45] suggesting that the early evolution of the angiosperm flora in India may have occurred largely in the absence of bees. The lack of early bee pioneers in India fits within our limited understanding of angiosperm diversification in this region.

Our results may also help to explain patterns of host use in bees. The Australian bee fauna consists of an exceptionally unique mix of early colonists (Colletidae) that arrived via an austral route[12] and lineages that arrived much later (Apidae, Halictidae, and Megachilidae). The dominant angiosperm families in Australia are Myrtaceae, Fabaceae, and Proteaceae.[78] Recent studies of host-plant use among endemic Australian colletid groups, such as Euryglossinae, Hylaeinae, and Neopasiphaeinae, indicate that these ancient Australian lineages retain a narrow host-association with the predominant "austral" flowering plant families.[50] The southern austral route hypothesized here for bees is well-established for plants, as determined from numerous studies based on phylogenetic[79,80] and fossil[81,82] evidence. An obvious and outstanding question along these lines is to what extent pronounced episodes of aridification in Africa and Australia during the Cenozoic (as the latter continent rifted and both land masses migrated northward) similarly affected the extinction rates and species turnover in bees as it did in plants.[56,83,84]

A critical challenge for timing the biogeography and co-diversification of bees and angiosperms is that many studies of angiosperm biogeography have focused solely at the family level or at lower taxonomic levels, which is insufficient to explore these more ancient patterns, or commonalities among groups. This is demonstrated by the rosid radiation, which occurred deeper in the angiosperm phylogeny, largely at the taxonomic rank of order. Deeper explorations of the plant phylogeny are warranted to best explore the early biogeographic history and co-diversification of angiosperms and bees. In addition, the modern distributions and diversity of bees and angiosperms are likely different from those in the deep past, confounding our efforts to understand this history, especially as this key mutualism was arising. Specifically, it appears that the most species-rich angiosperm clades that possess an essential suite of putatively adaptive floral traits for bee pollination—petals present, few stamens, and floral zygomorphy—are surprisingly recent innovations within angiosperms and only evolved within the last ca. 65 mya.[85] By extension, the next obvious step is to explore the more recent biogeographic history of bee lineages and how they correspond with these megadiverse angiosperm clades. Nevertheless, our results represent an essential first step in elucidating the geographic mosaic of the ancient angiosperm-bee coevolutionary syndrome.

Conclusions

Our thorough exploration of the bee fossil record and our reconstruction of historical biogeography via DEC-based reconstruction of range evolution provide the first high-resolution account of the evolutionary history of bees in time and space. Bees certainly arose in the western part of Gondwana, and their biogeographic history was thereafter impacted by the fragmentation of the southern continents in the Cretaceous (separating South American and African lineages) and the expansion of geographic distributions to the Northern Hemisphere. Because of the timing of the separation of India from the other southern continents, we hypothesize that India was "bee free" for a significant portion of its geological history. Likewise, Australia only hosted bees of two closely related families (Colletidae and Stenotritidae) for much of its geological history and only since the Late Cretaceous, which may help explain the recent radiations of some elements of the Australian flora during the Paleogene-Neogene.[78] Bees are the dominant pollinators in most contemporary ecosystems, and our results provide new insights into how their Cretaceous-age association with flowering plants evolved through time and around the world. By reconstructing the temporal and spatial history of bee diversification, we are closer than ever to building a truly integrative view of the evolution of plant-pollinator interactions.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
    - ○ Lead contact
    - ○ Materials availability
    - ○ Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
    - ○ Taxon sampling
    - ○ Phylogenomic data acquisition
    - ○ Processing of UCE sequence data
    - ○ Phylogenetic analyses
    - ○ Divergence time estimation
    - ○ Integration of analytical biogeography and the fossil record of bees
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2023.07.005.

ACKNOWLEDGMENTS

We would like to acknowledge the following people for their generosity in donating, exchanging, or facilitating the acquisition of specimens used in this study: Antonio Aguiar, Isabel Alves dos Santos, John Ascher, Ricardo Ayala, Airton Carvalho, Eduardo Dominguez, Connal Eardley, Terry Griswold, Terry Houston, Daercio Lucena, Hugo Medeiros, Gabriel Melo, Paulo Milet Pinheiro, José Eustáquio Santos Jr., Fernando Silveira, Michael Schwarz, and José Amilcar Tavares. All sequenced samples were obtained legally. We thank Mariela C. Castro for her guidance on the paleogeographic reconstructions with GPlates, Max C. Langer for geological/paleontological clarifications, Gabriel Melo and Brunno Rosa for discussions about apoid wasps, and two anonymous reviewers for their constructive suggestions to improving this work. We are grateful to Michael Landis for sharing his Bayesian implementation of DEC in the open software RevBayes. E.A.B.A. discloses support for the research of this work from São Paulo Research Foundation (FAPESP) (grant number 2018/09666-5) and Brazilian National Council for Scientific and Technological Development (CNPq) (grant number 310111/2019-6). S.B., B.N.D., E.A.M., S.G.B., and M.G.B. disclose support for publication of this work from the U.S. National Science Foundation (grant numbers DEB-1555905, DEB-2127744, and DEB-2127745). M.G.B. discloses support for publication of this work from the U.S. Department of Agriculture (USDA-ARS project 2080-21000-019-000-D). E.A.B.A., F.V.F., D.S.P., and P.R.S. disclose support for publication of this work from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brazil (CAPES) (Finance Code 001). Computational resources were provided by the Smithsonian High Performance Cluster (SI/HPC), Smithsonian Institution (https://doi.org/10.25572/SIHPC). USDA is an equal opportunity provider and employer.

AUTHOR CONTRIBUTIONS

Conceptualization, E.A.B.A., S.B., B.N.D., and M.R.P.; methodology and formal analysis, E.A.B.A., S.B., D.S.P., F.V.F., E.A.M., T.S., and M.R.P.; investigation, E.A.B.A., S.B., D.S.P., F.V.F., B.B.B., P.R.S., M.C.O., M.K., L.P., and M.G.B.; visualization, E.A.B.A., S.B., D.S.P., and F.V.F; data curation, E.A.B.A., S.B., D.S.P., F.V.F., and M.R.P.; funding acquisition, E.A.B.A., S.B., B.N.D., E.A.M., S.G.B., and M.G.B.; writing of original draft, E.A.B.A., S.B., B.N.D., and C.C.D.; all authors reviewed the manuscript and helped improve it.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 19, 2023
Revised: July 4, 2023
Accepted: July 4, 2023
Published: July 27, 2023

REFERENCES

1. Michener, C.D. (2007). The Bees of the World, Second Edition (John Hopkins University Press).
2. Orr, M.C., Hughes, A.C., Chesters, D., Pickering, J., Zhu, C.-D., and Ascher, J.S. (2021). Global patterns and drivers of bee distribution. Curr. Biol. *31*, 451–458.e4. https://doi.org/10.1016/j.cub.2020.10.053.
3. Ascher, J.S., and Pickering, J. (2020). Discover Life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila). Draft-55. http://www.discoverlife.org/mp/20q?guide=Apoidea_species.
4. Danforth, B.N., Minckley, R.L., and Neff, J.L. (2019). The Solitary Bees. Biology, Evolution, Conservation, First Edition (Princeton University Press).
5. Ollerton, J. (2017). Pollinator diversity: distribution, ecological function, and conservation. Annu. Rev. Ecol. Evol. Syst. *48*, 353–376. https://doi.org/10.1146/annurev-ecolsys-110316-022919.
6. Cardinal, S., and Danforth, B.N. (2013). Bees diversified in the age of eudicots. Proc. Biol. Sci. *280*, 20122686. https://doi.org/10.1098/rspb.2012.2686.
7. Murray, E.A., Bossert, S., and Danforth, B.N. (2018). Pollinivory and the diversification dynamics of bees. Biol. Lett. *14*, 20180530. https://doi.org/10.1098/rsbl.2018.0530.
8. Michener, C.D. (1979). Biogeography of the bees. Ann. Mo. Bot. Gard. *66*, 277–347. https://doi.org/10.2307/2398833.
9. Bossert, S., Wood, T.J., Patiny, S., Michez, D., Almeida, E.A.B., Minckley, R.L., Packer, L., Neff, J.L., Copeland, R.S., Straka, J., et al. (2022). Phylogeny, biogeography and diversification of the mining bee family Andrenidae. Syst. Entomol. *47*, 283–302. https://doi.org/10.1111/syen.12530.
10. Pisanty, G., Richter, R., Martin, T., Dettman, J., and Cardinal, S. (2022). Molecular phylogeny, historical biogeography and revised classification of andrenine bees (Hymenoptera: Andrenidae). Mol. Phylogenet. Evol. *170*, 107151. https://doi.org/10.1016/j.ympev.2021.107151.
11. Ramos, K.S., Martins, A.C., and Melo, G.A.R. (2022). Evolution of andrenine bees reveals a long and complex history of faunal interchanges through the Americas during the Mesozoic and Cenozoic. Mol. Phylogenet. Evol. *172*, 107484. https://doi.org/10.1016/j.ympev.2022.107484.
12. Almeida, E.A.B., Pie, M.R., Brady, S.G., and Danforth, B.N. (2012). Biogeography and diversification of colletid bees (Hymenoptera: Colletidae): emerging patterns from the southern end of the world. J. Biogeogr. *39*, 526–544. https://doi.org/10.1111/j.1365-2699.2011.02624.x.
13. Almeida, E.A.B., Packer, L., Melo, G.A.R., Danforth, B.N., Cardinal, S.C., Quinteiro, F.B., and Pie, M.R. (2019). The diversification of neopasiphaeine bees during the Cenozoic (Hymenoptera: Colletidae). Zool. Scr. *48*, 226–242. https://doi.org/10.1111/zsc.12333.
14. Litman, J.R., Danforth, B.N., Eardley, C.D., and Praz, C.J. (2011). Why do leafcutter bees cut leaves? New insights into the early evolution of bees. Proc. Biol. Sci. *278*, 3593–3600. https://doi.org/10.1098/rspb.2011.0365.
15. Bossert, S., Murray, E.A., Almeida, E.A.B., Brady, S.G., Blaimer, B.B., and Danforth, B.N. (2019). Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. Mol. Phylogenet. Evol. *130*, 121–131. https://doi.org/10.1016/j.ympev.2018.10.012.
16. Freitas, F.V., Branstetter, M.G., Casali, D.M., Aguiar, A.J.C., Griswold, T., and Almeida, E.A.B. (2022). Phylogenomic dating and Bayesian biogeography illuminate an antitropical pattern for eucerine bees. J. Biogeogr. *49*, 1034–1047. https://doi.org/10.1111/jbi.14359.
17. Hines, H.M. (2008). Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). Syst. Biol. *57*, 58–75. https://doi.org/10.1080/10635150801898912.
18. Rasmussen, C., and Cameron, S.A. (2010). Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal. Biol. J. Linn. Soc. *99*, 206–232. https://doi.org/10.1111/j.1095-8312.2009.01341.x.
19. Ramírez, S.R., Roubik, D.W., Skov, C., and Pierce, N.E. (2010). Phylogeny, diversification patterns and historical biogeography of euglossine orchid bees (Hymenoptera: Apidae). Biol. J. Linn. Soc. *100*, 552–572. https://doi.org/10.1111/j.1095-8312.2010.01440.x.
20. Danforth, B.N., Cardinal, S., Praz, C., Almeida, E.A.B., and Michez, D. (2013). The impact of molecular data on our understanding of bee phylogeny and evolution. Annu. Rev. Entomol. *58*, 57–78. https://doi.org/10.1146/annurev-ento-120811-153633.
21. Branstetter, M.G., Danforth, B.N., Pitts, J.P., Faircloth, B.C., Ward, P.S., Buffington, M.L., Gates, M.W., Kula, R.R., and Brady, S.G. (2017). Phylogenomic insights into the evolution of stinging wasps and the origins of ants and bees. Curr. Biol. *27*, 1019–1025. https://doi.org/10.1016/j.cub.2017.03.027.
22. Cardinal, S., Buchmann, S.L., and Russell, A.L. (2018). The evolution of floral sonication, a pollen foraging behavior used by bees (Anthophila). Evolution *72*, 590–600. https://doi.org/10.1111/evo.13446.
23. Hedtke, S.M., Patiny, S., and Danforth, B.N. (2013). The bee tree of life: a supermatrix approach to apoid phylogeny and biogeography. BMC Evol. Biol. *13*, 138. https://doi.org/10.1186/1471-2148-13-138.
24. Engel, M.S. (2000). A new interpretation of the oldest fossil bee (Hymenoptera, Apidae). Am. Mus. Novit. *3296*, 1–11.
25. Engel, M.S. (2001). Monophyly and extensive extinction of advanced eusocial bees: insights from an unexpected Eocene diversity. Proc. Natl. Acad. Sci. USA *98*, 1661–1664. https://doi.org/10.1073/pnas.98.4.1661.
26. Michener, C.D., and Grimaldi, D.A. (1988). A *Trigona* from Late Cretaceous amber of New Jersey (Hymenoptera: Apidae: Meliponinae). Am. Mus. Novit. *2917*, 1–10.
27. Michez, D., Vanderplanck, M., and Engel, M.S. (2012). Fossil bees and their plant associates. In Evolution of Plant-Pollinator Relationships. Systematics Association Special Volume Series, S. Patiny, ed. (Cambridge University Press), pp. 103–164. https://doi.org/10.1017/CBO9781139014113.006.
28. Rust, J., Singh, H., Rana, R.S., McCann, T., Singh, L., Anderson, K., Sarkar, N., Nascimbene, P.C., Stebner, F., Thomas, J.C., et al. (2010). Biogeographic and evolutionary implications of a diverse paleobiota in amber from the Early Eocene of India. Proc. Natl. Acad. Sci. USA *107*, 18360–18365. https://doi.org/10.1073/pnas.1007407107.

29. Rehan, S.M., Leys, R., and Schwarz, M.P. (2013). First evidence for a massive extinction event affecting bees close to the K-T boundary. PLoS One *8*, e76683. https://doi.org/10.1371/journal.pone.0076683.
30. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 1530–1534. https://doi.org/10.1093/molbev/msaa015.
31. Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. *32*, 268–274. https://doi.org/10.1093/molbev/msu300.
32. Sann, M., Niehuis, O., Peters, R.S., Mayer, C., Kozlov, A., Podsiadlowski, L., Bank, S., Meusemann, K., Misof, B., Bleidorn, C., et al. (2018). Phylogenomic analysis of Apoidea sheds new light on the sister group of bees. BMC Evol. Biol. *18*, 71. https://doi.org/10.1186/s12862-018-1155-8.
33. Cardinal, S., Straka, J., and Danforth, B.N. (2010). Comprehensive phylogeny of apid bees reveals the evolutionary origins and antiquity of cleptoparasitism. Proc. Natl. Acad. Sci. USA *107*, 16207–16211. https://doi.org/10.1073/pnas.1006299107.
34. Danforth, B.N., Brady, S.G., Sipes, S.D., and Pearson, A. (2004). Single-copy nuclear genes recover Cretaceous-age divergences in bees. Syst. Biol. *53*, 309–326. https://doi.org/10.1080/10635150490423737.
35. Michez, D., Patiny, S., and Danforth, B.N. (2009). Phylogeny of the bee family Melittidae (Hymenoptera: Anthophila) based on combined molecular and morphological data. Syst. Entomol. *34*, 574–597. https://doi.org/10.1111/j.1365-3113.2009.00479.x.
36. Peters, R.S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S., Meusemann, K., Kozlov, A., Podsiadlowski, L., Petersen, M., Lanfear, R., et al. (2017). Evolutionary history of the Hymenoptera. Curr. Biol. *27*, 1013–1018. https://doi.org/10.1016/j.cub.2017.01.027.
37. Danforth, B.N., Sipes, S., Fang, J., and Brady, S.G. (2006). The history of early bee diversification based on five genes plus morphology. Proc. Natl. Acad. Sci. USA *103*, 15118–15123. https://doi.org/10.1073/pnas.0604033103.
38. Freitas, F.V., Branstetter, M.G., Griswold, T., and Almeida, E.A.B. (2021). Partitioned gene-tree analyses and gene-based topology testing help resolve incongruence in a phylogenomic study of host-specialist bees (Apidae: Eucerinae). Mol. Biol. Evol. *38*, 1090–1100. https://doi.org/10.1093/molbev/msaa277.
39. Orr, M.C., Branstetter, M.G., Straka, J., Yuan, F., Leijs, R., Zhang, D., Zhou, Q., and Zhu, C.-D. (2022). Phylogenomic interrogation revives an overlooked hypothesis for the early evolution of the bee family Apidae (Hymenoptera: Apoidea), with a focus on the subfamily Anthophorinae. Insect Syst. Divers. *6*, ixac022. https://doi.org/10.1093/isd/ixac022.
40. Brady, S.G., Sipes, S., Pearson, A., and Danforth, B.N. (2006). Recent and simultaneous origins of eusociality in halictid bees. Proc. Biol. Sci. *273*, 1643–1649. https://doi.org/10.1098/rspb.2006.3496.
41. Blakey, R.C. (2008). Gondwana paleogeography from assembly to breakup—a 500 m.y. odyssey. In Resolving Late Paleozoic Ice Age Time Space Geological Society of America, Special Paper 441 (The Geological Society of America), pp. 1–28.
42. Scotese, C.R. (2005). Paleomap project. http://www.scotese.com.
43. Scotese, C.R. (2004). A continental drift flipbook. J. Geol. *112*, 729–741. https://doi.org/10.1086/424867.
44. McLoughlin, S. (2001). The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. Aust. J. Bot. *49*, 271–300. https://doi.org/10.1071/BT00023.
45. Seton, M., Müller, R.D., Zahirovic, S., Gaina, C., Torsvik, T., Shephard, G., Talsma, A., Gurnis, M., Turner, M., Maus, S., et al. (2012). Global continental and ocean basin reconstructions since 200Ma. Earth Sci. Rev. *113*, 212–270. https://doi.org/10.1016/j.earscirev.2012.03.002.
46. Müller, R.D., Cannon, J., Qin, X., Watson, R.J., Gurnis, M., Williams, S., Pfaffelmoser, T., Seton, M., Russell, S.H.J., and Zahirovic, S. (2018). GPlates: building a virtual Earth through deep time. Geochem. Geophys. Geosyst. *19*, 2243–2261. https://doi.org/10.1029/2018GC007584.
47. Evenick, J.C. (2021). Glimpses into Earth's history using a revised global sedimentary basin map. Earth Sci. Rev. *215*, 103564. https://doi.org/10.1016/j.earscirev.2021.103564.
48. Hughes, C.E., Pennington, R.T., and Antonelli, A. (2013). Neotropical plant evolution: assembling the big picture. Bot. J. Linn. Soc. *171*, 1–18. https://doi.org/10.1111/boj.12006.
49. Martins, A.C., Melo, G.A.R., and Renner, S.S. (2015). Gain and loss of specialization in two oil-bee lineages, *Centris* and *Epicharis* (Apidae). Evolution *69*, 1835–1844. https://doi.org/10.1111/evo.12689.
50. Slattery, P.S., Parslow, B.A., Lee, M.S.Y., Batley, M., Walker, K.L., and Schwarz, M.P. (2023). Disparate continental scale patterns in floral host breadth of Australian colletid bees (Colletidae: Hymenoptera). Apidologie *54*, 17. https://doi.org/10.1007/s13592-023-00996-z.
51. Praz, C.J., and Packer, L. (2014). Phylogenetic position of the bee genera *Ancyla* and *Tarsalia* (Hymenoptera: Apidae): a remarkable base compositional bias and an early Paleogene geodispersal from North America to the Old World. Mol. Phylogenet. Evol. *81*, 258–270. https://doi.org/10.1016/j.ympev.2014.09.003.
52. Lavin, M., and Luckow, M. (1993). Origins and relationships of tropical North America in the context of the boreotropics hypothesis. Am. J. Bot. *80*, 1–14. https://doi.org/10.1002/j.1537-2197.1993.tb13761.x.
53. Meseguer, A.S., Lobo, J.M., Cornuault, J., Beerling, D., Ruhfel, B.R., Davis, C.C., Jousselin, E., and Sanmartín, I. (2018). Reconstructing deep-time palaeoclimate legacies in the clusioid Malpighiales unveils their role in the evolution and extinction of the boreotropical flora. Glob. Ecol. Biogeogr. *27*, 616–628. https://doi.org/10.1111/geb.12724.
54. Sanmartín, I., Enghoff, H., and Ronquist, F. (2001). Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biol. J. Linn. Soc. *73*, 345–390. https://doi.org/10.1006/bijl.2001.0542.
55. Zachos, J.C., Dickens, G.R., and Zeebe, R.E. (2008). An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. Nature *451*, 279–283. https://doi.org/10.1038/nature06588.
56. Davis, C.C., Bell, C.D., Mathews, S., and Donoghue, M.J. (2002). Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. Proc. Natl. Acad. Sci. USA *99*, 6833–6837. https://doi.org/10.1073/pnas.102175899.
57. Yang, T., Lu, L.M., Wang, W., Li, J.H., Manchester, S.R., Wen, J., and Chen, Z.D. (2018). Boreotropical range expansion and long-distance dispersal explain two amphi-Pacific tropical disjunctions in Sabiaceae. Mol. Phylogenet. Evol. *124*, 181–191. https://doi.org/10.1016/j.ympev.2018.03.005.
58. Branstetter, M.G., Longino, J.T., Reyes-López, J.L., Brady, S.G., and Schultz, T.R. (2022). Out of the temperate zone: a phylogenomic test of the biogeographical conservatism hypothesis in a contrarian clade of ants. J. Biogeogr. *49*, 1640–1653. https://doi.org/10.1111/jbi.14462.
59. Ye, Z., Zhen, Y., Zhou, Y., and Bu, W. (2017). Out of Africa: biogeography and diversification of the pantropical pond skater genus *Limnogonus* Stål, 1868 (Hemiptera: Gerridae). Ecol. Evol. *7*, 793–802. https://doi.org/10.1002/ece3.2688.
60. Vicente, N., Kergoat, G.J., Dong, J., Yotoko, K., Legendre, F., Nattier, R., and Robillard, T. (2017). In and out of the Neotropics: historical biogeography of Eneopterinae crickets. J. Biogeogr. *44*, 2199–2210. https://doi.org/10.1111/jbi.13026.
61. Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R., and Soltis, D.E. (2009). Rosid radiation and the rapid rise of angiosperm-dominated forests. Proc. Natl. Acad. Sci. USA *106*, 3853–3858. https://doi.org/10.1073/pnas.0813376106.
62. Beaulieu, J.M., Tank, D.C., and Donoghue, M.J. (2013). A Southern Hemisphere origin for campanulid angiosperms, with traces of the break-up of Gondwana. BMC Evol. Biol. *13*, 80. https://doi.org/10.1186/1471-2148-13-80.

63. Dutta, S., Tripathi, S.M., Mallick, M., Mathews, R.P., Greenwood, P.F., Rao, M.R., and Summons, R.E. (2011). Eocene out-of-India dispersal of Asian dipterocarps. Rev. Palaeobot. Palynol *166*, 63–68. https://doi.org/10.1016/j.revpalbo.2011.05.002.

64. Hauenschild, F., Favre, A., Michalak, I., and Muellner-Riehl, A.N. (2018). The influence of the Gondwanan breakup on the biogeographic history of the ziziphoids (Rhamnaceae). J. Biogeogr. *45*, 2669–2677. https://doi.org/10.1111/jbi.13448.

65. Luebert, F., Couvreur, T.L.P., Gottschling, M., Hilger, H.H., Miller, J.S., and Weigend, M. (2017). Historical biogeography of Boraginales: west Gondwanan vicariance followed by long-distance dispersal? J. Biogeogr. *44*, 158–169. https://doi.org/10.1111/jbi.12841.

66. Sanil, M.S., Balakrishnan, S., Sreekumar, V.B., and Dev, S.A. (2022). Dipterocarps used India as a raft from Gondwana to Eurasia. Taxon *71*, 1214–1229. https://doi.org/10.1002/tax.12794.

67. Ruhfel, B.R., Bove, C.P., Philbrick, C.T., and Davis, C.C. (2016). Dispersal largely explains the Gondwanan distribution of the ancient tropical clusioid plant clade. Am. J. Bot. *103*, 1117–1128.

68. Huang, X., Deng, T., Moore, M.J., Wang, H., Li, Z., Lin, N., Yusupov, Z., Tojibaev, K.S., Wang, Y., and Sun, H. (2019). Tropical Asian origin, boreotropical migration and long-distance dispersal in nettles (Urticeae, Urticaceae). Mol. Phylogenet. Evol. *137*, 190–199. https://doi.org/10.1016/j.ympev.2019.05.007.

69. Thomas, D.C., Chatrou, L.W., Stull, G.W., Johnson, D.M., Harris, D.J., Thongpairoj, U., and Saunders, R.M.K. (2015). The historical origins of palaeotropical intercontinental disjunctions in the pantropical flowering plant family Annonaceae. Perspect. Plant Ecol. Evol. Syst. *17*, 1–16. https://doi.org/10.1016/j.ppees.2014.11.001.

70. Willis, C.G., Franzone, B.F., Xi, Z., and Davis, C.C. (2014). The establishment of Central American migratory corridors and the biogeographic origins of seasonally dry tropical forests in Mexico. Front. Genet. *5*, 433. https://doi.org/10.3389/fgene.2014.00433.

71. Willis, C.G., and Davis, C.C. (2015). Rethinking migration. Science *348*, 766. https://doi.org/10.1126/science.348.6236.766-a.

72. Baker, W.J., and Couvreur, T.L.P. (2013). Global biogeography and diversification of palms sheds light on the evolution of tropical lineages. I. Historical biogeography. J. Biogeogr. *40*, 274–285. https://doi.org/10.1111/j.1365-2699.2012.02795.x.

73. Mandel, J.R., Dikow, R.B., Siniscalchi, C.M., Thapa, R., Watson, L.E., and Funk, V.A. (2019). A fully resolved backbone phylogeny reveals numerous dispersals and explosive diversifications throughout the history of Asteraceae. Proc. Natl. Acad. Sci. USA *116*, 14083–14088. https://doi.org/10.1073/pnas.1903871116.

74. Yan, Y., Davis, C.C., Dimitrov, D., Wang, Z., Rahbek, C., and Borregaard, M.K. (2021). Phytogeographic history of the tea family inferred through high-resolution phylogeny and fossils. Syst. Biol. *70*, 1256–1271. https://doi.org/10.1093/sysbio/syab042.

75. Manchester, S.R., and Tiffney, B.H. (2001). Integration of paleobotanical and neobotanical data in the assessment of phytogeographic history of Holarctic angiosperm clades. Int. J. Plant Sci. *162*, S19–S27. https://doi.org/10.1086/323657.

76. Tiffney, B.H., and Manchester, S.R. (2001). The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary. Int. J. Plant Sci. *162*, S3–S17. https://doi.org/10.1086/323880.

77. Rutschmann, F., Eriksson, T., Schönenberger, J., and Conti, E. (2004). Did Crypteroniaceae really disperse out of India? Molecular dating evidence from rbcL, ndhF, and rpl16 intron sequences. Int. J. Plant Sci. *165*, S69–S83. https://doi.org/10.1086/383335.

78. Crisp, M.D., and Cook, L.G. (2013). How was the Australian flora assembled over the last 65 million years? A molecular phylogenetic perspective. Annu. Rev. Ecol. Evol. Syst. *44*, 303–324. https://doi.org/10.1146/annurev-ecolsys-110512-135910.

79. Cook, L.G., and Crisp, M.D. (2005). Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. Proc. Biol. Sci. *272*, 2535–2544. https://doi.org/10.1098/rspb.2005.3219.

80. Sanmartín, I., and Ronquist, F. (2004). Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. Syst. Biol. *53*, 216–243. https://doi.org/10.1080/10635150490423430.

81. Kooyman, R.M., Wilf, P., Barreda, V.D., Carpenter, R.J., Jordan, G.J., Sniderman, J.M.K., Allen, A., Brodribb, T.J., Crayn, D., Feild, T.S., et al. (2014). Paleo-Antarctic rainforest into the modern Old World tropics: the rich past and threatened future of the "southern wet forest survivors.". Am. J. Bot. *101*, 2121–2135. https://doi.org/10.3732/ajb.1400340.

82. Kooyman, R.M., Ivory, S.J., Benfield, A.J., and Wilf, P. (2022). Gondwanan survivor lineages and the high-risk biogeography of Anthropocene Southeast Asia. J. Syst. Evol. *60*, 715–727. https://doi.org/10.1111/jse.12853.

83. Crayn, D.M., Rossetto, M., and Maynard, D.J. (2006). Molecular phylogeny and dating reveals an Oligo-Miocene radiation of dry-adapted shrubs (former Tremandraceae) from rainforest tree progenitors (Elaeocarpaceae) in Australia. Am. J. Bot. *93*, 1328–1342. https://doi.org/10.3732/ajb.93.9.1328.

84. Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J.J., and Brunet, M. (2006). Tectonic uplift and Eastern Africa aridification. Science *313*, 1419–1423. https://doi.org/10.1126/science.1129158.

85. O'Meara, B.C., Smith, S.D., Armbruster, W.S., Harder, L.D., Hardy, C.R., Hileman, L.C., Hufford, L., Litt, A., Magallón, S., Smith, S.A., et al. (2016). Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. Proc. Biol. Sci. *283*, https://doi.org/10.1098/rspb.2015.2304.

86. Faircloth, B.C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics *32*, 786–788. https://doi.org/10.1093/bioinformatics/btv646.

87. Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat. Biotechnol. *29*, 644–652. https://doi.org/10.1038/nbt.1883.

88. Harris, R.S. (2007). Improved pairwise alignment of genomic DNA. PhD thesis (The Pennsylvania State University).

89. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

90. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. *17*, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334.

91. Mirarab, S., and Warnow, T. (2015). ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics *31*, i44–i52. https://doi.org/10.1093/bioinformatics/btv234.

92. Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., and Calcott, B. (2017). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. *34*, 772–773. https://doi.org/10.1093/molbev/msw260.

93. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. *61*, 539–542. https://doi.org/10.1093/sysbio/sys029.

94. Rambaut, A., Drummond, A.J., Xie, D., Baele, G., and Suchard, M.A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. *67*, 901–904. https://doi.org/10.1093/sysbio/syy032.

95. Höhna, S., Landis, M.J., Heath, T.A., Boussau, B., Lartillot, N., Moore, B.R., Huelsenbeck, J.P., and Ronquist, F. (2016). RevBayes: Bayesian phylogenetic inference using graphical models and an interactive

model-specification language. Syst. Biol. *65*, 726–736. https://doi.org/10.1093/sysbio/syw021.

96. Tribble, C.M., Freyman, W.A., Landis, M.J., Lim, J.Y., Barido-Sottani, J., Kopperud, B.T., Höhna, S., and May, M.R. (2022). RevGadgets: an R package for visualizing Bayesian phylogenetic analyses from RevBayes. Methods Ecol. Evol. *13*, 314–323. https://doi.org/10.1111/2041-210X.13750.
97. Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.-Y. (2017). ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. *8*, 28–36. https://doi.org/10.1111/2041-210X.12628.
98. Behrensmeyer, A., and Turner, A. (2017). Fossilworks: Gateway to the Paleobiology Database. http://www.fossilworks.org/cgi-bin/bridge.pl?action=home.
99. Behrensmeyer, A.K., and Turner, A. (2013). Taxonomic occurrences of suidae recorded in the Paleobiology Database. Fossilworks. https://web.archive.org/web/20140731140924/http://fossilworks.org/.
100. Blaimer, B.B., Mawdsley, J.R., and Brady, S.G. (2018). Multiple origins of sexual dichromatism and aposematism within large carpenter bees. Evolution *72*, 1874–1889. https://doi.org/10.1111/evo.13558.
101. Bossert, S., Murray, E.A., Pauly, A., Chernyshov, K., Brady, S.G., and Danforth, B.N. (2021). Gene tree estimation error with ultraconserved elements: an empirical study on *Pseudapis* bees. Syst. Biol. *70*, 803–821. https://doi.org/10.1093/sysbio/syaa097.
102. Faircloth, B.C., Branstetter, M.G., White, N.D., and Brady, S.G. (2015). Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. Mol. Ecol. Resour. *15*, 489–501. https://doi.org/10.1111/1755-0998.12328.
103. Grab, H., Branstetter, M.G., Amon, N., Urban-Mead, K.R., Park, M.G., Gibbs, J., Blitzer, E.J., Poveda, K., Loeb, G., and Danforth, B.N. (2019). Agriculturally dominated landscapes reduce bee phylogenetic diversity and pollination services. Science *363*, 282–284. https://doi.org/10.1126/science.aat6016.
104. Elsik, C.G., Worley, K.C., Bennett, A.K., Beye, M., Camara, F., Childers, C.P., de Graaf, D.C., Debyser, G., Deng, J., Devreese, B., et al. (2014). Finding the missing honey bee genes: lessons learned from a genome upgrade. BMC Genom. *15*, 86. https://doi.org/10.1186/1471-2164-15-86.
105. Kapheim, K.M., Pan, H., Li, C., Salzberg, S.L., Puiu, D., Magoc, T., Robertson, H.M., Hudson, M.E., Venkat, A., Fischman, B.J., et al. (2015). Social evolution. Genomic signatures of evolutionary transitions from solitary to group living. Science *348*, 1139–1143. https://doi.org/10.1126/science.aaa4788.
106. Kocher, S.D., Li, C., Yang, W., Tan, H., Yi, S.V., Yang, X., Hoekstra, H.E., Zhang, G., Pierce, N.E., and Yu, D.W. (2013). The draft genome of a socially polymorphic halictid bee, Lasioglossum albipes. Genome Biol. *14*, R142. https://doi.org/10.1186/gb-2013-14-12-r142.
107. Park, D., Jung, J.W., Choi, B.S., Jayakodi, M., Lee, J., Lim, J., Yu, Y., Choi, Y.S., Lee, M.L., Park, Y., et al. (2015). Uncovering the novel characteristics of Asian honey bee, *Apis cerana*, by whole genome sequencing. BMC Genom. *16*, 1. https://doi.org/10.1186/1471-2164-16-1.
108. Rehan, S.M., Glastad, K.M., Lawson, S.P., and Hunt, B.G. (2016). The genome and methylome of a subsocial small carpenter bee, *Ceratina calcarata*. Genome Biol. Evol. *8*, 1401–1410. https://doi.org/10.1093/gbe/evw079.
109. Sadd, B.M., Barribeau, S.M., Bloch, G., de Graaf, D.C., Dearden, P., Elsik, C.G., Gadau, J., Grimmelikhuijzen, C.J., Hasselmann, M., Lozier, J.D., et al. (2015). The genomes of two key bumblebee species with primitive eusocial organization. Genome Biol. *16*, 76. https://doi.org/10.1186/s13059-015-0623-3.
110. Branstetter, M.G., Longino, J.T., Ward, P.S., and Faircloth, B.C. (2017). Enriching the ant tree of life: enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. Methods Ecol. Evol. *8*, 768–776. https://doi.org/10.1111/2041-210X.12742.
111. Blaimer, B.B., Brady, S.G., Schultz, T.R., Lloyd, M.W., Fisher, B.L., and Ward, P.S. (2015). Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. BMC Evol. Biol. *15*, 271. https://doi.org/10.1186/s12862-015-0552-5.
112. Blaimer, B.B., Lloyd, M.W., Guillory, W.X., and Brady, S.G. (2016). Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. PLoS One *11*, e0161531. https://doi.org/10.1371/journal.pone.0161531.
113. Faircloth, B.C. (2013). Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. GitHub. https://github.com/faircloth-lab/illumiprocessor.
114. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
115. Bossert, S., and Danforth, B.N. (2018). On the universality of target-enrichment baits for phylogenomic research. Methods Ecol. Evol. *9*, 1453–1460. https://doi.org/10.1111/2041-210X.12988.
116. Rannala, B., and Yang, Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics *164*, 1645–1656. https://doi.org/10.1093/genetics/164.4.1645.
117. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods *14*, 587–589. https://doi.org/10.1038/nmeth.4285.
118. Lanfear, R., Calcott, B., Kainer, D., Mayer, C., and Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol. Biol. *14*, 82. https://doi.org/10.1186/1471-2148-14-82.
119. Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. *35*, 518–522. https://doi.org/10.1093/molbev/msx281.
120. Crotty, S.M., Minh, B.Q., Bean, N.G., Holland, B.R., Tuke, J., Jermiin, L.S., and Haeseler, A.V. (2020). GHOST: recovering historical signal from heterotachously evolved sequence alignments. Syst. Biol. *69*, 249–264. https://doi.org/10.1093/sysbio/syz051.
121. Sayyari, E., and Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. *33*, 1654–1668. https://doi.org/10.1093/molbev/msw079.
122. Heath, T.A., Huelsenbeck, J.P., and Stadler, T. (2014). The fossilized birth–death process for coherent calibration of divergence-time estimates. Proc. Natl. Acad. Sci. USA *111*, E2957–E2966. https://doi.org/10.1073/pnas.1319091111.
123. Stadler, T. (2010). Sampling-through-time in birth-death trees. J. Theor. Biol. *267*, 396–404. https://doi.org/10.1016/j.jtbi.2010.09.010.
124. Engel, M.S. (1999). The taxonomy of recent and fossil honey bees (Hymenoptera: Apidae; Apis). J. Hymenopt Res. *8*, 165–196.
125. Engel, M.S., Hinojosa-Díaz, I.A., and Rasnitsyn, A.P. (2009). A honey bee from the Miocene of Nevada and the biogeography of *Apis* (Hymenoptera: Apidae: Apini). Proc. Calif. Acad. Sci. *60*, 23–38.
126. Engel, M.S. (1996). New augochlorine bees (Hymenoptera: Halictidae) in Dominican amber, with a brief review of fossil Halictidae. J. Kans. Entomol. Soc. 334–345.
127. Engel, M.S., and Rightmyer, M.G. (2000). A new augochlorine bee species in tertiary amber from the Dominican Republic (Hymenoptera: Halictidae). Apidologie *31*, 431–436. https://doi.org/10.1051/apido:2000133.
128. Michener, C.D., and Grimaldi, D.A. (1988). The oldest fossil bee: apoid history, evolutionary stasis, and antiquity of social behavior. Proc. Natl. Acad. Sci. USA *85*, 6424–6426.
129. Lepage, T., Bryant, D., Philippe, H., and Lartillot, N. (2007). A general comparison of relaxed molecular clock models. Mol. Biol. Evol. *24*, 2669–2680. https://doi.org/10.1093/molbev/msm193.
130. Spasojevic, T., Broad, G.R., Sääksjärvi, I.E., Schwarz, M., Ito, M., Korenko, S., and Klopfstein, S. (2021). Mind the outgroup and bare

branches in total-evidence dating: a case study of pimpliform Darwin wasps (Hymenoptera, Ichneumonidae). Syst. Biol. *70*, 322–339. https://doi.org/10.1093/sysbio/syaa079.

131. Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L., and Rasnitsyn, A.P. (2012). A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. Syst. Biol. *61*, 973–999. https://doi.org/10.1093/sysbio/sys058.

132. Holt, B.G., Lessard, J.P., Borregaard, M.K., Fritz, S.A., Araújo, M.B., Dimitrov, D., Fabre, P.H., Graham, C.H., Graves, G.R., Jønsson, K.A., et al. (2013). An update of Wallace's zoogeographic regions of the world. Science *339*, 74–78. https://doi.org/10.1126/science.1228282.

133. Ree, R.H., and Smith, S.A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. *57*, 4–14. https://doi.org/10.1080/10635150701883881.

134. Ronquist, F., and Sanmartín, I. (2011). Phylogenetic methods in biogeography. Annu. Rev. Ecol. Evol. Syst. *42*, 441–464. https://doi.org/10.1146/annurev-ecolsys-102209-144710.

135. Landis, M.J. (2017). Biogeographic dating of speciation times using paleogeographically informed processes. Syst. Biol. *66*, 128–144. https://doi.org/10.1093/sysbio/syw040.

136. Landis, M.J., Freyman, W.A., and Baldwin, B.G. (2018). Retracing the Hawaiian silversword radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty. Evolution *72*, 2343–2359. https://doi.org/10.1111/evo.13594.

137. Ree, R.H., Moore, B.R., Webb, C.O., and Donoghue, M.J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution *59*, 2299–2311. https://doi.org/10.1111/j.0014-3820.2005.tb00940.x.

138. Scotese, C.R. (2016). PALEOMAP PaleoAtlas for GPlates and the PaleoData Plotter Program. PALEOMAP Project. https://doi.org/10.13140/RG2.2.34367.00166. https://www.researchgate.net/profile/Christopher-Scotese/publication/301893396_PALEOMAP_PaleoAtlas_for_GPlates_and_the_PaleoData_Plotter_Program/links/5afa0551458515c00b6baebd/PALEOMAP-PaleoAtlas-for-GPlates-and-the-PaleoData-Plotter-Program.pdf.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Samples of 57 species of Apoidea (Hymenoptera)	Various sources as identified in BioProject and Data S1A	NCBI SRA BioProject PRJNA978572: https://doi.org/10.17632/j233njx65x.1
Deposited data		
Raw Illumina sequencing reads for 57 species of Apoidea (Hymenoptera)	NCBI SRA BioProject PRJNA978572	NCBI SRA https://www.ncbi.nlm.nih.gov/bioproject/PRJNA978572
UCE assembly files for 57 samples	Present Study	Mendeley Data Repository: https://doi.org/10.17632/j233njx65x.1
DNA alignments of UCE loci	Present Study	Mendeley Data Repository: https://doi.org/10.17632/j233njx65x.1
Phylogenetic trees	Present Study	Mendeley Data Repository: https://doi.org/10.17632/j233njx65x.1
Scripts and program input files	Present Study	Mendeley Data Repository: https://doi.org/10.17632/j233njx65x.1
Table S1. List of reliable nodes of the tree of bees considered for the estimation of phylogenetic accuracy	Present Study	Supplemental Information: Table S1
Table S2. Number of bee species by families per biogeographical region	Present Study	Supplemental Information: Table S2
Data S1A. Taxon sampling of the UCE-based phylogenomic study of 216 species of bees and 25 apoid wasp outgroups (Hymenoptera: Apoidea)	Present Study	Supplemental Information: Data S1A
Data S1B. Fossil specimens included in divergence time estimation analyses and paleobiogeographic reconstructions	Present Study	Supplemental Information: Data S1B
Data S1C. Distribution of the 72 terminal taxa used for the Dispersal-Extinction-Cladogenesis (DEC) biogeographic analyses	Present Study	Supplemental Information: Data S1C
Software and algorithms		
GPlates	Müller et al.[46]	https://www.gplates.org
IQ-TREE	Minh et al.,[30] Nguyen et al.[31]	http://www.iqtree.org
Illumiprocessor	Faircloth[86]	https://illumiprocessor.readthedocs.io
Trinity	Grabherr et al.[87]	https://github.com/trinityrnaseq/trinityrnaseq/wiki
LASTZ	Harris[88]	https://www.bx.psu.edu/~rsharris/lastz/
MAFFT	Katoh and Standley[89]	https://mafft.cbrc.jp/alignment/software/
PHYLUCE	Castresana[90]	https://phyluce.readthedocs.io
ASTRAL	Mirarab and Warnow[91]	https://github.com/smirarab/ASTRAL
PartitionFinder	Lanfear et al.[92]	https://www.robertlanfear.com/partitionfinder/
MrBayes	Ronquist et al.[93]	https://nbisweden.github.io/MrBayes/
Tracer	Rambaut et al.[94]	https://github.com/beast-dev/tracer
RevBayes	Höhna et al.[95]	https://revbayes.github.io
RevGadgets	Tribble et al.[96]	https://github.com/revbayes/RevGadgets
Ggtree	Yu et al.[97]	https://github.com/YuLab-SMU/ggtree

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
Discover Life checklist	Ascher and Pickering[3]	https://www.discoverlife.org/mp/20q?guide=Apoidea_species&flags=HAS: 24 October 2020
Fossilworks	Behrensmeyer and Turner[98,99]	http://www.fossilworks.org

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Eduardo A. B. Almeida (eduardoalmeida@usp.br).

Materials availability

No materials were generated during this study.

Data and code availability

- Accession numbers and other data used in the study are listed in the key resources table. Detailed specimen data (extant and fossil taxa) are available as supplemental Excel spreadsheets, and paleoreconstructions of fossil bees are available as supplemental data. Additional information (alignments, assembly files, tree files, as well as scripts and input files used in this study) can be found online at the Mendeley Data Repository and is publicly available as of the date of publication. The tree files comprise six species of trees reconstructed using ML methods and ASTRAL, five Bayesian chronograms, and six reconstructions of the biogeographic history of bees generated in this study using the DEC model. The DOI is listed in the key resources table.
- All original scripts have been deposited at the Mendeley Data Repository associated with this article and are publicly available as of the date of publication. The DOI is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Samples of 57 species of Apoidea from Argentina, Australia, Brazil, Chile, Madagascar, Mexico, South Africa, Spain, and the USA were used to generate new DNA sequence data. Detailed information on these samples is provided in Data S1A: scientific species names, their higher-level classification, and the respective collecting localities. DNA extractions were performed from dry-pinned museum specimens or individuals stored in absolute ethanol. Voucher specimens of these samples are housed in three depositories: Cornell University Insect Collection, Ithaca, USA (CUIC); Coleção Entomológica 'Prof. J.M.F.Camargo', Universidade de São Paulo, Ribeirão Preto, Brazil (RPSP); Packer Collection at York University, Toronto, Canada (PCYU), and are specified in the BioProject information for the respective samples. Data S1A further specifies the BioProject repository identifier (PRJNA978572) for the raw sequence data associated with these samples. We combined this newly generated DNA sequence data with sequences of 184 previously published samples. This additional data can be identified by the sequence identifiers listed in Data S1A and the listed NCBI assembly identifiers in case of previously published whole genome data.

METHOD DETAILS

Taxon sampling

We compiled the taxonomically broadest phylogenomic data set of bees to date. We generated new sequence data for 57 species and combined them with ultraconserved elements (UCEs) from previously generated phylogenetic studies of Hymenoptera using UCE markers.[9,15,21,38,100–103] Furthermore, we identified and extracted UCE loci from 16 publicly available bee genomes.[104–109] Detailed information on the included samples and associated voucher information is available in Data S1A. The complete dataset includes 241 tips, comprising representatives from all seven bee families and all 28 subfamilies, collected across all continents, in addition to samples of 25 apoid wasp outgroups (Data S1A).

Phylogenomic data acquisition

After DNA extraction, we enriched dual-indexed Illumina libraries for ultraconserved elements (UCEs) of 57 species. We used two different versions of the Hymenoptera bait sets: libraries enriched with the Hym-v1 bait set were processed according to Faircloth et al.[102] and sequenced by Rapid Genomics LLC. Samples enriched with the Hym-v2 bait set[110] were processed according to Blaimer et al.[111,112] and sequenced at the Cornell Core Facilities.

Processing of UCE sequence data

Raw read data were processed with the PHYLUCE pipeline version 1.6[86] and associated programs. We integrated the 57 newly sequenced samples with UCEs from 184 previously sequenced species[9,15,21,38,100,101,103] and 16 publicly available bee genomes. After sequence alignment and trimming, we assembled 75% and 90% completeness matrices, meaning that each included locus is present for at least 75% (=183) or 90% (=220) of all examined terminals, respectively.

The reads were trimmed with Illumiprocessor version 2.0.6,[113] which employs the Trimmomatic trimming tool.[114] We then used PHYLUCE[86] to assemble the reads with Trinity[87] and used LASTZ,[88] as included in PHYLUCE, to identify and extract UCEs from the 16 included genome assemblies, including 850 bp of flanking regions. The newly generated UCE contigs were then combined with sequence data from previously published research and the UCE sequences extracted from the included genomes. PHYLUCE was also used to identify and extract sequences of each individual sample that match baits from the principal Hymenoptera v2 bait set, requiring the default matching criteria (–min-coverage = 80; –min-identity = 80) to identify orthologous loci. These parameters are stricter than in comparable studies using the same bait set, but they seem appropriate as lower thresholds increase the risk of introducing contamination.[115] Contigs containing UCE loci were extracted and parsed according to their specific UCE bait, yielding a separate unaligned file for each recovered UCE locus. The resulting loci were aligned using MAFFT version 7.407[89] and the L-INS-I option, then trimmed for ambiguous nucleotides with Gblocks as included in PHYLUCE.[90] After examining the locus representation of different nucleotide matrices, we processed 75% and 90% completeness matrices, ensuring that each included locus is present in at least 75% or 90% of all examined taxa, respectively. This corresponds to at least 183 taxa or 220 taxa, respectively.

Phylogenetic analyses

We reconstructed the phylogeny of bees using maximum likelihood (ML) analyses of concatenated datasets and methods that involve gene-tree reconciliation under the multispecies coalescent model—MSC.[116] For estimating ML phylogenies, we used IQ-TREE2 version 2.1.2[30,31] and two separate approaches to model nucleotide substitutions. First, we used ModelFinder[117](part of IQ-Tree) to find the best-fitting partitioning scheme by using hierarchical clustering of loci[118] with subsequent model search and ML tree estimation. We combined partitions (= UCE loci) for improved model-fit using the relaxed hierarchical clustering method[118] at 30% (rcluster 30), with the subsequent search for the best-fitting substitution models (MFP+MERGE option). We assessed support with 1,000 bootstrap approximations (UFBoot2[119]). We then reconstructed ML phylogenies from both matrices using the GHOST mixture model,[120] which is an edge-unlinked model accommodating heterotachous sequence evolution. Specifically, we linked parameter estimation across partitions and inferred separate base frequencies with the +FO option, using four classes (-m GTR+FO+H4). For the GHOST analyses, we used the partitions identified through the previous ModelFinder analyses for both completeness matrices.

For the analyses under the MSC, we first inferred individual gene trees for the 75% and 90% completeness matrix using IQ-TREE version 1.6.9[30,31] and automated model search (-MFP). We then summarized the individual gene trees into species tree topologies using ASTRAL version 5.6.2[91] and estimated branch support with local posterior probabilities.[121]

We analyzed two datasets representing two levels of taxon occupancy (percentage of taxa required to be present for each locus): 75% and 90% filtered locus sets. These datasets included 830 loci and 364,396 bp (251,279 parsimony informative sites), and 377 loci and 181,287 bp (125,151 informative sites) of sequence data, respectively, for the 75% and 90% filtered sets. For each dataset, we assessed the effects of the optimality criterion (MSC vs. ML), as well as of substitution models.

To further explore different sources of signal in our 75% complete dataset and to assess the robustness of our phylogenetic estimates, we used a comparative approach to infer the level of phylogenetic accuracy of each of the 830 studied loci. We began by selecting 14 nodes on the bee tree that are well-supported by morphological, biogeographical, and phylogenetic evidence independent of our dataset (a list of the chosen nodes is provided in Table S1). We then recorded how many of those nodes were recovered by each locus in their respective gene trees estimated with IQ-TREE as our index of phylogenetic accuracy. We tested five potential locus characteristics that could be associated with high phylogenetic accuracy: clocklikeness, average bootstrap support, GC content, number of parsimony informative sites, and taxonomic coverage (i.e., number of terminals on the gene tree): Figure S1. All potential correlates were included in a multiple regression analysis. Subsequently, we chose alternative topologies for unstable nodes in our analysis or previous works and estimated the phylogenetic congruence for each locus that either agreed or disagreed with the respective node. This allowed us to discriminate low overall phylogenetic signal (few loci agreeing with any of the alternative topologies) from conflicting signal (e.g., a substantial number of loci with high phylogenetic accuracy both agreeing and disagreeing with particular nodes): Data S2.

Divergence time estimation

We estimated the antiquity and divergence times of bees using a Bayesian approach under the Fossilized Birth-Death model[122,123] to combine our phylogenetic data with an extensive database of fossil specimens. This approach circumvents the need to assign prior age constraints to internal nodes (‘node dating’), thus allowing us to include information on fossil specimens that would otherwise be challenging or impossible to include, as the phylogenetic position of many bee fossils is uncertain or only reasonably certain at the family or tribal level.[27] For example, several species of fossil honeybees can clearly be associated with the genus *Apis* based on several morphological features.[124,125] In contrast, others, such as the different species of *Oligochlora*, cannot be associated with an extant genus, but with a tribe that persists into the present: Augochlorini (Halictidae).[126,127] This strategy is similar to that of Bossert et al.[9] but with a much-expanded set of fossils.

We reviewed all available bee fossil and noncontroversial ichnofossil occurrence data published until July 1, 2018. The ages of the individual records were established based on information on dated deposits available on the Fossilworks website,[98] estimates provided by Michez et al.,[27] and the original publications cited in Data S1B. We filtered this fossil data based on two criteria: first, we excluded fossils for which we deemed an association to an extant bee family impossible or highly questionable, including controversial ichnotaxa like *Cellicalichnus* spp. We then removed fossil specimens which we deemed conspecific in the literature unless the fossils were recovered from different formations or distant locations. For example, we deemed it likely that the fossil specimen of *Anthophorites mellona* from the Miocene Sarmatian age (12.7–11.6 Ma), which was recovered in Oenigen (Germany), is not conspecific with the significantly older fossil with the same name from the Chattian age of the Oligocene (28.4–23.0 Ma), found in Mediterranean France.

We incorporated the 185 fossils (Data S1B) by including them as tips with missing data in the Bayesian divergence time estimation under the Fossilized Birth-Death model. We excluded outgroup taxa and employed a set of 62 partial and hard constraints (including a root constraint) to inform the tree search of the MrBayes version 3.2.7a[93] analyses for two reasons: (a) First, we consider the smaller subsets of UCEs that we use for the dating analyses (10–20 loci) less capable of accurately resolving the difficult nodes of bee phylogeny (see above) than the full data matrix of 830 UCEs. Therefore, we constrained parts of the phylogeny to resemble the topological results of the IQ-TREE2[30,31] and ASTRAL[91] phylogenies based on our extensive sensitivity analyses. (b) Second, we used the constraints to associate the fossils with the respective clades (Data S1B). Each fossil tip was assigned a uniform age prior which corresponds to the absolute age of the geological formation where the fossil was found (Data S1B). We applied an exponential prior probability distribution on the age of the root with a minimum age of 70 million years, as this is the approximate age of the oldest undisputed fossil bee, *Cretotrigona prisca*,[24,26,128] and a mean of 155 million years. This wide distribution ensures an unconstrained estimation of the root age of bees, allowing the analyses to sample from a parameter space that includes potential root ages of as low as 70 Ma to over 700 Ma (2.5% quantile of 73.9 Ma and 97.5% quantile at 642 Ma), and any previously estimated ages of bees are permitted under this prior.

Because of the computational challenges of estimating divergence times from large phylogenomic datasets, we subsampled sets of loci from the original matrix. We contrasted three strategies of filtering loci to generate five sub-sampled matrices, chosen due to desirable features of the data (i.e., clocklikeness or average bootstrap values) and contrasted them with three randomly selected subsets of fragments. Filtering loci and creating subsets of the entire dataset do not appear to reduce the quality of the divergence time inferences obtained from such data.[16] For each of these subsets, we ran PartitionFinder version 2.1.1[92] to designate partitions for the subsequent analyses. Using PartitionFinder, we carried out greedy searches, provided the initial UCE loci as data blocks and tested the following models: JC+G, K80+G, TrNef+G, SYM+G, HKY+G, TrN+G, and GTR+G. Subsequently, we used the resulting partitions to designate subsets for nucleotide substitution modeling in MrBayes[93] under a relaxed clock model.[129]

For divergence times estimates in MrBayes, we used the independent gamma rates model (IGR) for estimating the clock and, as Spasojevic et al.,[130] informed our analyses with prior information from Ronquist et al.[131] To inform the variance of the IGR model, we used an exponential prior of 37.12[131] and applied a log normal prior for the clock rate with -7.08069 and a standard deviation of 1.0.[130] Since we are including extinct, fossilized species as well as extant taxa in our analyses, we set the prior probability distribution for branch lengths (brlensp) to ‘clock:fossilization’. Because we developed a non-random, phylogenetically broad taxon sampling, we used the diversified sampling strategy (sampleStrat = diversity) and assigned a sampling probability of 0.01. This probability is calculated by dividing the number of included extant species (220 spp.) by the approximate total number of presently described species (*ca.* 20,500) of bees. For the tree model, we applied broad priors on the speciation probability (uniform(0,1)), extinction probability (beta(1,1)), and fossilization probability (beta(1,1)).

We executed six runs with four chains for each analysis and sampled every 1000th generation for at least 125 million generations. Using a 20% burn-in, we assessed convergence in Tracer version 1.7.1[94] to be sufficient once the average standard deviation of split frequencies fell below 0.01 and all parameters, including the clock rate, reached combined effective sample size (ESS) values of $\geq$200. Topological convergence was assessed using the average standard deviation of split frequencies (ASDSF). Because of the large number of included fossils without data, convergence of the MCMC was difficult to achieve to the usually applied threshold of 0.01, even after executing 150 million generations. We therefore considered topological convergence acceptable once ASDSF remained unchanged at 0.02 with fossils and fell below 0.01 after exclusion of fossils. The MrBayes analysis was repeated for each of the five different subsets of UCEs and included fossils as tips in the phylogeny.

Integration of analytical biogeography and the fossil record of bees

We investigated broad patterns of the early evolution of bees across six major zoogeographic regions[132]: Australian, Afrotropical, Neotropical, Nearctic, Oriental, and Palearctic, using a continuous-time model to reconstruct geographic range evolution by dispersal, extinction, and cladogenesis: the DEC model.[133] This stochastic model allows the extrapolation of present ranges to the past using a dated phylogenetic hypothesis and the current distribution of species in predefined areas delimited to represent historical units (Data S1C; Figure 4). DEC is arguably one of the most realistic stochastic models available for historical biogeography.[134] In reticulate models, area configurations are allowed to change over time: an area splitting (i.e., reduced connectivity) is associated with vicariance, whereas fusion of areas in the geological time (i.e., higher connectivity or removal of a previous dispersal barrier) leads to range expansions. Dispersal is modeled along phylogenetic branches as a continuous-time Markov chain with two parameters: range expansion, where an additional area is added to the current range, and range contraction, with the removal of an area from the ancestral range.

We used the Bayesian implementation of the DEC model[135,136] in RevBayes 1.1.0[95] to infer geographical range evolution of lineages on the dated phylogeny pruned to include only one species per genus or monophyletic set of genera (Figures S2 and S3). This measure allowed for the most biogeographic information to be extracted from the actual taxon sampling in concert with the incorporation of distributional data of all known bee taxa (sampled or not for our phylogeny). Additionally, we avoided causing biases in the reconstructions either because of overrepresentation or lack of certain higher-level taxa in our species-level sampling. This is a conservative approach that results in the loss of part of the distributional and phylogenetic information of some taxa. Still, it reduces the risks of unwarranted generalizations while maintaining a safe and comprehensive representation of all deeper nodes. The phylogenetic basis for the biogeographic investigation was the chronogram based on the 20 most clocklike loci (available for download from the project's Mendeley data repository; see data and code availability), which has branch lengths in units of absolute time since divergence.

We applied stratified models[133,137] with area adjacency matrices and variable dispersal rates to consider how area connectivity changed over time using the scripts for stratified analyses available on the program website (https://revbayes.github.io/tutorials/biogeo/biogeo_epoch.html). To implement these stratified models, the bee timetree was divided into time slices and the expected baseline connectivity rate was scaled by discrete values according to a hypothesis of geological connectivity among areas over time.[41,42] Stratification was defined using intervals, instead of hard bounds among the different time slices considered,[136] reflecting the uncertainty in the biogeographic connectivity model. We began the strata at the Cretaceous, corresponding to the estimated depth of bee phylogeny[6,21,32] and in agreement with our results. We considered the major geological and global climatic changes that occurred during the Cretaceous and Tertiary[41–45,47] and implemented six time slices to express the major geological and climatic events that occurred on the planet since the Early Cretaceous, when bees are hypothesized to have originated[6,21,32,36]: older than 105 Ma, 105–70 Ma, 70–45 Ma, 45–35 Ma, 35–20 Ma, and 20 Ma to today (Figure 4). For each time slice, we constructed a matrix of scaling factors (between 0.1 and 1.0) expressing the expected connectivity between areas (directly proportional to dispersal probabilities), account for the changing distances between the areas over geological time.[41–45] These scaling parameters were set to 0.1 (low connectivity), 0.2–0.4 (low-medium), 0.5 (moderate), 0.6–0.8 (medium-high), or 0.9–1.0 (high connectivity), then converted into three models that considered the sensitivity of the resulting reconstructions to the defined connectivity probabilities through the time strata (Figure 4): Model 1 is the least permissive (i.e., dispersal probabilities between disjunct terrains especially those geographically distant can be as low as 10% of the highest dispersal probabilities); Model 2 is the most permissive (i.e., dispersal probabilities between any two areas are never lower than 50% of the highest dispersal probabilities); and Model 3 represents an intermediate position between models 1 and 2. The split between Africa and South America during the Cretaceous was followed by higher connectivity between Africa and the Palearctic region during the Paleogene (and closing of India on Asia). The Paleocene-Eocene periods were also marked by higher connectivity between South America and the Australian region due to climatic[55] and geological factors,[41,44] which was reversed by a severe temperature drop at the Eocene-Oligocene[55] boundary coupled with the separation of Antarctica, Australia, and South America. This kind of historical information on the areas was used to inform the dispersal probabilities implemented in the models shown in Figure 4.

We assigned each terminal taxon to one or more geographical areas based on the natural distribution of all its described members. If unconstrained, some taxa are treated as widespread, being present in all or most of the six biogeographic regions delimited. In addition to the coding of widespread as such, we performed an alternative coding in which the size of widespread ancestral ranges was limited to a maximum of three areas to improve the resolution of the resulting reconstruction and reduce the complexity in our analysis. These alternative coding schemes are shown in Figure S3 for the 72 terminal taxa.

To assess MCMC convergence, each analysis was run twice; the overlapping of parameters of the posterior distribution between runs was verified, and ESS values for all were calculated and certified to exceed 200 in all cases. The final reconstructions were plotted using the R packages RevGadgets version 1.1.0[96] and ggtree version 3.7.[97]

The geographic and temporal distribution of 221 bee fossils (185 fossils used for the Fossilized Birth-Death analyses plus 36 additional fossils to enrich the geographic range information) were integrated into the biogeographic interpretation. The palaeocoordinates of each fossil (Data S1B) were obtained with Fossilworks,[98,99] and plotted onto the corresponding geological layer using GPlates version 2.0[46] with the paleogeographic reconstructions and the tectonic plates rotation model of PALEOMAP.[138] Different temporal layers were assembled for all fossil data points (Data S1B and S3), then integrated with the analytical biogeographic reconstructions. We prepared summary maps superimposing the ranges inferred with the DEC reconstruction and fossil paleoreconstructions, as six time-bands: (a) 125–110 Ma, (b) 110–90 Ma, (c) 90–70 Ma, (d) 70–50 Ma, (e) 50–30 Ma, and (f) 30 Ma–present. The interpretation of presence or absence of each bee family in these time slices is based on the most likely DEC-reconstructions for each node. The placement of fossil bees (Data S3) in the modern areas examined for this study considered the approximate correspondence of the record reconstructed location and the present boundaries of the six operational areas (Figure 3).

Lastly, we contrasted the reconstructions based on the fossil record of bees and analytical biogeographic models with a summary of present-day bee diversity in the six regions considered in this study. Species richness was based on Discover Life[3] visited on 24 October 2020, as given for countries or territories, which were categorized into the six biogeographic regions as follows: Afrotropical (Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central Africa, Central African Republic, Chad, Comoros, Congo, Cote D'Ivoire, Democratic Republic of the Congo, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Reunion, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Socotra, Somalia, South Africa, South Sudan, Sudan, Swaziland, Tanzania, Togo, Uganda, Zambia, Zimbabwe), Australian (Australia, Bonin Islands, Caroline Island, Cook Islands, Fiji, French Polynesia, Guam, Marianas Islands, Marshall Islands, Micronesia, New Caledonia, New Guinea whole, New Zealand, Papua New Guinea, Samoa, Solomon Islands, Tonga, Tuvalu, Vanuatu), Nearctic (Canada, Mexico, United States), Neotropical (Anguilla, Antigua and Barbuda, Argentina, Aruba, Bahamas, Barbados, Belize, Bolivia, Brazil, British Virgin Islands, Cayman Islands, Chile, Colombia, Costa Rica, Cuba, Curacao, Dominica, Dominican Republic, Ecuador, El Salvador, French Guiana, Grenada, Guadeloupe, Guatemala, Guyana, Haiti, Hispaniola, Honduras, Jamaica, Martinique, Montserrat, Netherlands Antilles, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and Grenadines, St Bartholomew, St Croix, St John, St Thomas, Suriname, Trinidad and Tobago, Uruguay, Venezuela), Oriental (Bangladesh, Brunei, Burma, Cambodia, East Timor, India, Indonesia, Laos, Malaysia, Maldives, Philippines, Singapore, Sri Lanka, Taiwan, Thailand, Timor, Vietnam), and Palearctic (Afghanistan, Albania, Algeria, Andorra, Armenia, Austria, Azerbaijan, Bahrain, Belarus, Belgium, Bhutan, Bosnia and Herzegovina, Bulgaria, China, Croatia, Cyprus, Czech Republic, Denmark, Egypt, Estonia, Finland, Former Yugoslavia, France, Georgia, Germany, Greece, Hungary, Iceland, Iran, Iraq, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Korea, Kuwait, Kyrgyzstan, Latvia, Lebanon, Libya, Liechtenstein, Lithuania, Luxembourg, Macedonia, Malta, Moldova, Monaco, Mongolia, Montenegro, Morocco, Nepal, Netherlands, North Korea, Norway, Oman, Pakistan, Poland, Portugal, Qatar, Romania, Russia, San Marino, Saudi Arabia, Serbia, Slovakia, Slovenia, South Korea, Spain, Sweden, Switzerland, Syria, Tajikistan, Tunisia, Turkestan, Turkey, Turkmenistan, Ukraine, United Arab Emirates, United Kingdom, Uzbekistan, West Bank, Western Sahara, Yemen). The resulting distribution of richness is summarized in Table S2.

QUANTIFICATION AND STATISTICAL ANALYSIS

Details of statistical analyses can be found in the results and method details. For phylogenetic analyses using maximum likelihood (ML), we used ModelFinder,[117] as implemented in IQ-TREE,[30,31] to assess the fit of substitution models on our DNA sequence alignments. ModelFinder fits various substitution models on preliminary parsimony-based trees and minimizes the score for the Bayesian information criterion (BIC). For a separate set of phylogenetic analyses using maximum likelihood, we implemented the GHOST (General Heterogeneous evolution On a Single Topology) model.[120] Node certainty for all ML analyses in IQ-TREE and IQ-TREE2 was assessed with ultrafast bootstrap approximations (UFBoot)[119] and 1,000 replicates. Branch support for the coalescent-based species tree analyses with ASTRAL[91] was assessed with local posterior probabilities.[121] Node support of the MrBayes[93] analyses was assessed with posterior probabilities. Convergence of the Bayesian analyses was assessed using Tracer.[94] Confidence intervals of divergence time estimates of any dating analysis carried out in this study represent 95% highest posterior densities (HPD). The fit of substitution models for the divergence time estimates (MrBayes) was estimated with PartitionFinder2,[92] testing the following models: JC+G, K80+G, TrNef+G, SYM+G, HKY+G, TrN+G, and GTR+G. For the biogeographic analyses, we used the Bayesian implementation of the DEC model[135,136] in RevBayes.[95] To assess MCMC convergence, each analysis was run twice (the overlapping of parameters of the posterior distribution between runs was verified, and effective sample size [ESS] of all parameters were calculated and certified to exceed 200 in all cases).