Report

Current Biology

The radiation of Hymenoptera illuminated by Bayesian inferences from the fossil record

Highlights

- Fossil-based estimates of divergence times for extinct and extant Hymenoptera
- Hymenoptera alternated between periods of high diversification and extinction
- Successive turnovers for Hymenoptera during the MMPR
- Bayesian estimates are robust under different assumptions of diversity and root age

Authors

Corentin Jouault, Nozomu Oyama, Sergio Álvarez-Parra, Diying Huang, Vincent Perrichot, Fabien L. Condamine, Frédéric Legendre

Correspondence

jouaultc0@gmail.com

In brief

Jouault et al. infer the evolutionary history of Hymenoptera from their fossil record, estimate the ages of origin and extinction of all fossil-recorded families and superfamilies, and clarify the periods of diversification and extinction that punctuated the evolution of these insects.

Jouault et al., 2025, Current Biology *35*, 2164–2174
May 5, 2025
https://doi.org/10.1016/j.cub.2025.03.002

CellPress

CellPress
OPEN ACCESS

Current Biology

Report

The radiation of Hymenoptera illuminated by Bayesian inferences from the fossil record

Corentin Jouault,[1,2,3,4,9,*] Nozomu Oyama,[5,6] Sergio Álvarez-Parra,[7] Diying Huang,[7] Vincent Perrichot,[4,8] Fabien L. Condamine,[3,8] and Frédéric Legendre[2,8]

[1]Oxford University Museum of Natural History, University of Oxford, Parks Road, Oxford OX1 3PW, UK
[2]Institut de Systématique, Évolution, Biodiversité (ISYEB), UMR 7205, Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE-PSL, Université des Antilles, CP50, 50 rue Cuvier, 75005 Paris, France
[3]Institut des Sciences de l'Évolution de Montpellier (UMR 5554), Université de Montpellier, CNRS, Place Eugène Bataillon, 34095 Montpellier, France
[4]Géosciences Rennes (UMR 6118), Université de Rennes, CNRS, 35000 Rennes, France
[5]Centre de Recherche en Paléontologie – Paris (CR2P), MNHN – CNRS – Sorbonne Université, 75005 Paris, France
[6]The Kyushu University Museum, 812-8581 Fukuoka, Japan
[7]State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing 210008, China
[8]These authors contributed equally
[9]Lead contact
*Correspondence: jouaultc0@gmail.com
https://doi.org/10.1016/j.cub.2025.03.002

SUMMARY

Determining when lineages originated provides fundamental insights into the timing and pace of their diversification, improving our understanding of transformative paleoevents such as the Angiosperm Terrestrial Revolution (ATR)[1] and Mid-Mesozoic Parasitoid Revolution (MMPR).[2] As the MMPR overlaps with the ATR, improved age estimates help to disentangle the dynamics and temporal succession of these events that shaped modern ecosystems. Hymenoptera (ants, bees, and wasps) played an important role in the MMPR and ATR through their parasitoid and pollinating lineages. Parasitoids impact trophic networks, whereas pollinators interact with flowering plants.[3,4] However, our understanding of Hymenoptera diversification remains limited by a lack of fossil-based studies and uncertainties in phylogenetic reconstructions. Combining fossil occurrences and macroevolutionary models, we estimated the origin and diversification of Hymenoptera lineages, considering changes in preservation over time and across taxa.[5–7] Our results indicate that Hymenoptera diversification is multifaceted and lineage-specific. Sawflies diversified during the Paleozoic and Mesozoic in four episodes (middle Permian, Late Triassic to Middle Jurassic, Early Cretaceous, and the beginning of the Cenozoic) and experienced three extinction episodes (Middle Triassic, Late Jurassic, and mid-Cretaceous). The superfamily Xyeloidea originated during the middle Permian. Apocrita and parasitoid superfamilies emerged during the Early to Middle Triassic, diversified during the Late Jurassic and Early Cretaceous, and declined during the Late Cretaceous. We demonstrate that Hymenoptera experienced successive replacements during the MMPR—likely beginning in the Triassic—and synchronously with changes in floral assemblages of the ATR. We conclude with future directions to refine dating estimates from the fossil record.

RESULTS

Time frame of origination and extinction of Hymenoptera

Using the fossil record and the Bayesian Brownian bridge (BBB) model,[6] we estimated the ages of 159 families (Figure 1A; Table S1) and 32 superfamilies (Table 1) of extant and extinct Hymenoptera. All analyses were performed on total groups (i.e., stem *plus* crown group). The BBB model is known to provide robust estimates of origin and extinction ages, even for poorly sampled clades, although families with only a few fossil species often have larger 95% credibility intervals (CIs). More precise estimates and narrower CIs can be obtained for superfamilies, combining data from different families.[7] Our analyses, performed with a max boundary of uniform prior on the root age set to 281 Ma, suggest that Hymenoptera began to diversify before the Permian, with the oldest lineages being the Xyeloidea, which arose more than 270 mya (CI: 263.01–280.23 Ma; Table 1) and include the extant Xyelidae (median 257.82 Ma, CI: 242.58–276.51 Ma; Table S1).

The first major radiation of Hymenoptera occurred during the Triassic and Jurassic, after the Permo-Triassic mass extinction, in ecosystems dominated by gymnosperms (Figures 1A, 1B and S1).[9,10] It mainly concerned the "Symphyta," except for the superfamily Tenthredinoidea (Figures 1A and 1B; median 192.26 Ma, CI: 179.32–237.11 Ma), whose families mostly

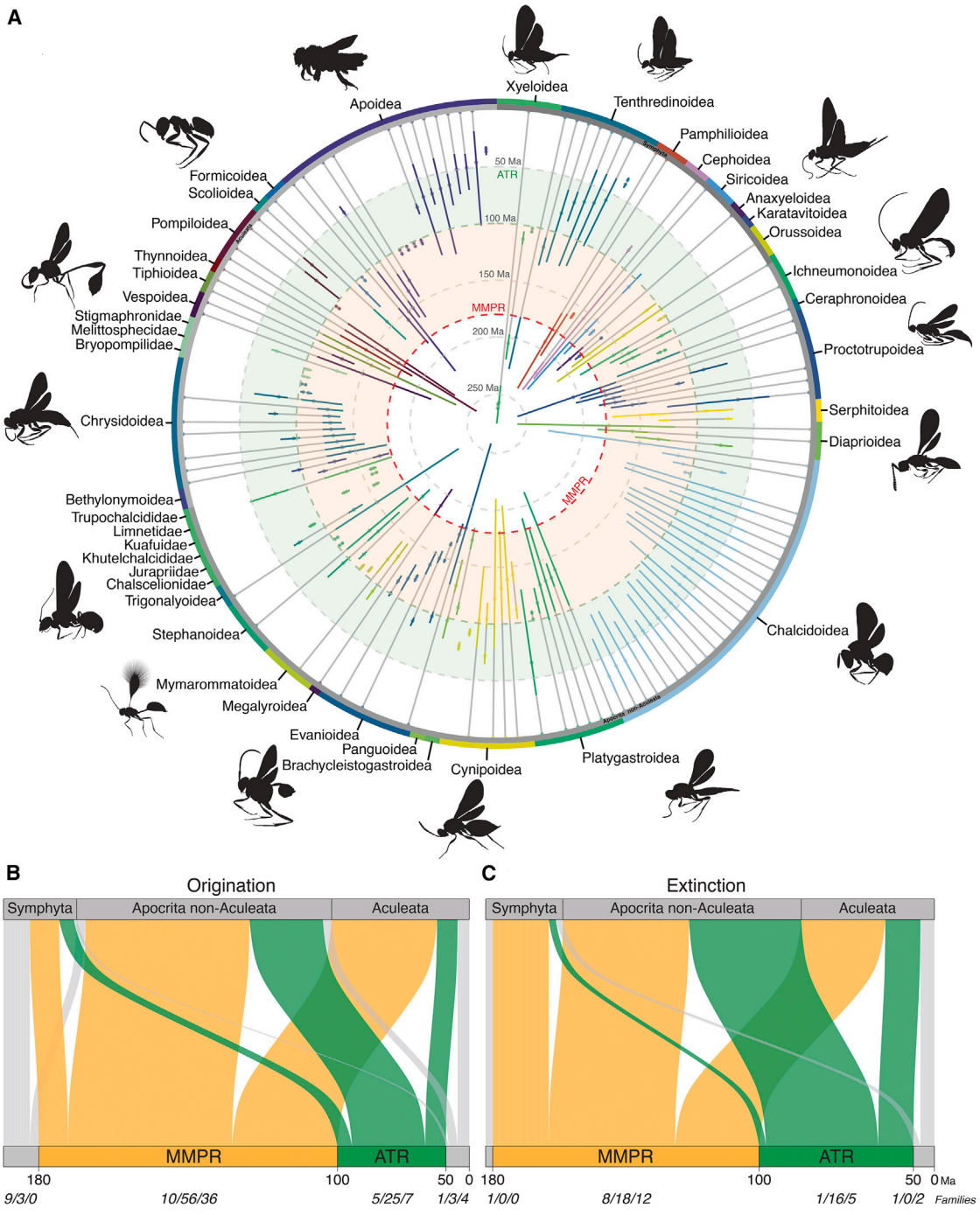


Figure 1. Estimates of clade ages and extinction times for Hymenoptera families and their distribution in key paleoevents

(A) Clade age and extinction time estimates for Hymenoptera families, inferred using the Bayesian Brownian Bridge model (root set to 281 Ma). Each line represents a family (arranged by clade but without further phylogenetic information), with 95% credible intervals in colors at the root estimates and extinction estimates (where applicable). Gray lines fill in the lineage.

(B and C) Estimated distribution of family origination (B) and extinction (C) during the different key periods. The bottom line shows the number of families in each period as follows: Symphyta/Apocrita non-Aculeata/Aculeata. Silhouettes after Goulet and Huber,[8] except Megachilidae from http://phylopic.org/ (license at https://creativecommons.org/publicdomain/zero/1.0/). ATR, Angiosperm Terrestrial Revolution; MMPR, Mid-Mesozoic Parasitoid Revolution.

See also Figure S1, Table S1, and Data S2 and S3.

Table 1. Comparison of clade age and extinction time estimates from fossil-based versus molecular approaches

	Extinction		Origination		Blaimer et al.[9]'s origination ages from topology C1		
Superfamily	Median (Ma)	95% HPD interval (Ma)	Median (Ma)	95% HPD interval (Ma)	Median (Ma)	95% HPD interval (Ma)	Age difference
Anaxyeloidea	N/A	N/A	191.55	[174.62–209.85]	N/A	N/A	N/A
Apoidea	N/A	N/A	148.22	[143.01–162.36]	102.1	[98.5–120.9]	46.12
Bethylonymoidea	97.11	[90.08–101.47]	171.91	[164.01–183.99]	N/A	N/A	N/A
Brachycleistogastroidea	111.9	[103.77–115.93]	170.52	[163.69–182.11]	N/A	N/A	N/A
Cephoidea	N/A	N/A	215.84	[208.73–233.78]	243.37	[228.4–256.2]	−27.53
Ceraphronoidea	N/A	N/A	118.15	[104.29–147.55]	136.00	[113.8–162.9]	−17.85
Chalcidoidea	N/A	N/A	114.11	[106.97–125.85]	168.6	[154.9–183.3]	−54.49
Cynipoidea	N/A	N/A	145.26	[136.04–163.95]	129.6	[109.3–154.4]	15.66
Diaprioidea	N/A	N/A	149.24	[142.73–166.68]	141.9	[116.2–173.2]	7.34
Evanioidea	N/A	N/A	177.27	[166.08–196.57]	155.00	[130.1–185.2]	22.27
Formicoidea	N/A	N/A	111.69	[104.06–129.07]	108.62	[98.5–120.9]	3.07
Ichneumonoidea	N/A	N/A	148.62	[142.01–173.98]	206.6	[192.4–221.3]	−57.98
Karatavitoidea	157.83	[156.11–158.97]	181.89	[179.02–189.01]	N/A	N/A	N/A
Megalyroidea	N/A	N/A	187.27	[178.05–206.75]	128.2	[85.1–172.9]	59.07
Mymarommatoidea	N/A	N/A	131.05	[123.16–149.83]	179.34	[165.3–193.9]	−48.29
Orussoidea	N/A	N/A	189.14	[165.09–229.99]	233.7	[220.2–247.8]	−44.56
Pamphilioidea	N/A	N/A	205.09	[197.44–224.64]	264.5	[249.4–276.6]	−59.41
Panguoidea	97.65	[84.84–101.99]	129.05	[123.02–152.78]	N/A	N/A	N/A
Platygastroidea	N/A	N/A	128.92	[124.03–141.93]	145.9	[131.1–166.5]	−16.98
Pompiloidea	N/A	N/A	126.74	[118.17–145.66]	99.1	[87.0–112.0]	27.64
Proctotrupoidea	N/A	N/A	177.54	[165.14–195.88]	174.7	[151.5–195.4]	2.84
Scolioidea	N/A	N/A	143.23	[127.76–167.58]	97.49	[76.9–11.4]	45.74
Serphitoidea	73.46	[68.59–75.99]	127.08	[124.01–135.29]	N/A	N/A	N/A
Siricoidea	N/A	N/A	190.55	[180.012–222.01]	236.7	[212.7–253.8]	−46.15
Stephanoidea	N/A	N/A	192.25	[180.14–213.69]	190.37	[167.8–210.2]	1.88
Tenthredinoidea	N/A	N/A	192.26	[179.32–237.11]	169.9	[132.9–208.6]	22.36
Thynnoidea	N/A	N/A	138.55	[117.12–201.57]	96.6	[81.8–110.8]	41.95
Tiphioidea	N/A	N/A	156.35	[104.03–250.06]	96.6	[81.8–110.8]	59.75
Trigonalyoidea	N/A	N/A	136.99	[125.3–158.61]	128.2	[85.1–172.9]	8.79
Vespoidea	N/A	N/A	127.08	[118.04–139.77]	108.6	[92.7–123.9]	18.48
Xyeloidea	N/A	N/A	270.3	[263.01–280.23]	275.67	[261.1–285.3]	−5.37

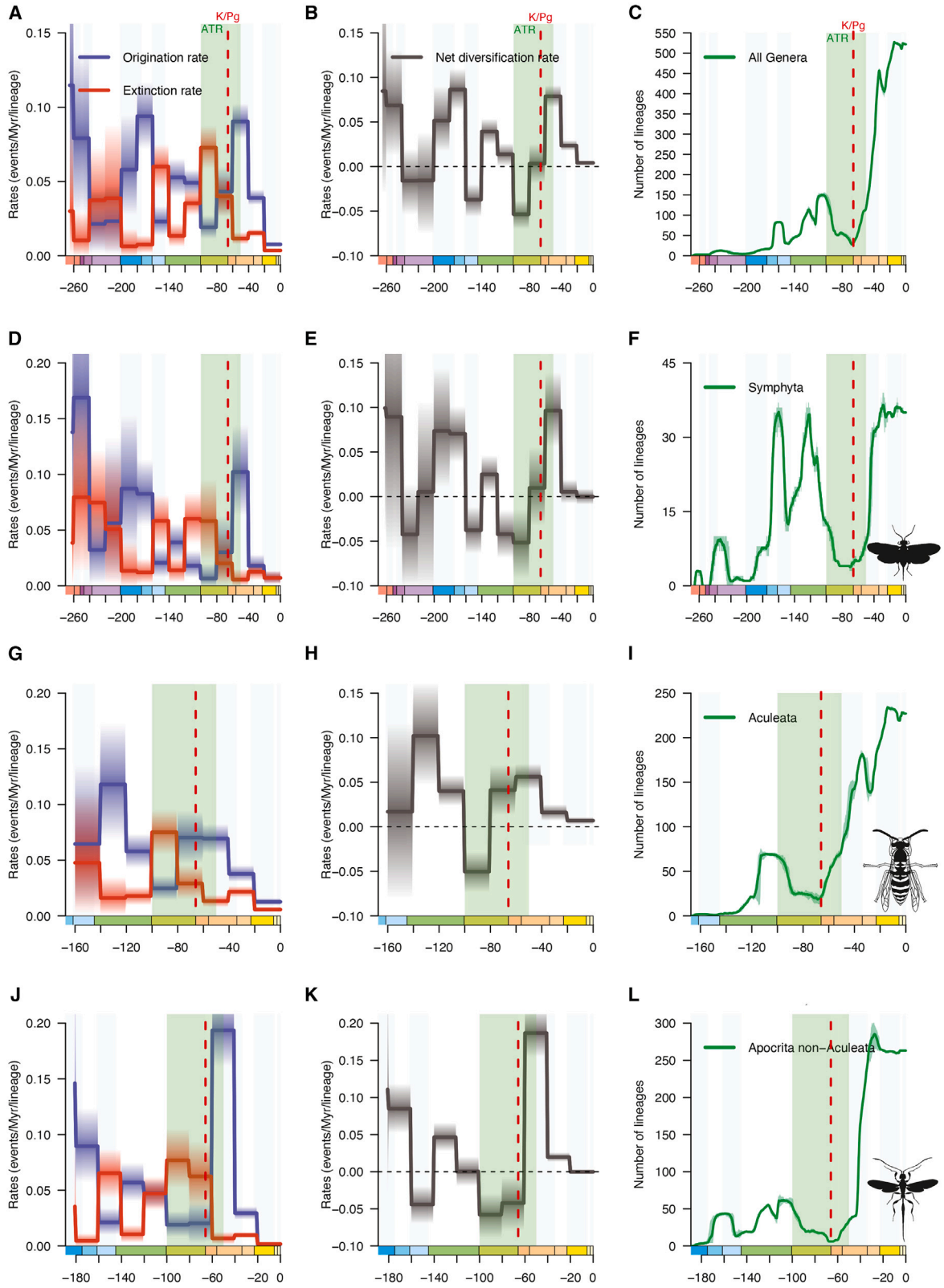
Clade age and extinction time estimates for Hymenoptera superfamilies and comparison with a recent study based on ultra-conserved elements (446 UCE loci) and 12 fossil calibrations.[9]

originated later, during the Angiosperm Terrestrial Revolution (ATR, 100–50 Ma). The ATR is characterized by a shift in ecosystem dominance, with gymnosperms being replaced by angiosperms,[1,11] affecting insect diversification.[12,13] Most parasitoid wasps are estimated to have originated during the Jurassic-Cretaceous interval (e.g., Evanioidea, median 177.27 Ma, CI: 166.08–196.57 Ma; Ichneumonidae, median 148.62 Ma, CI: 142.01–173.98 Ma), which is part of the Mid-Mesozoic Parasitoid Revolution (MMPR) (Figure 1B).[2] In addition, 20 of the 27 families of Chalcidoidea included in our analyses originated during the ATR (Figure 1).

Stinger wasps (Aculeata) are estimated to have originated during the Jurassic (e.g., Bethylonymoidea: median 171.91 Ma, CI: 164.01–183.99 Ma), but some of their constitutive clades either survived until the ATR or declined during this event, as their extinction dates fall within the ATR (Figures 1B and 1C). Many extant lineages of Apoidea, often closely associated with flowering plants,[14,15] arose during the ATR (Apidae: median 73.33 Ma, CI: 70–79.84 Ma). Apparently, many fossil families went extinct just before or during the ATR (Figures 1 and S1). However, this pattern might be skewed by the short lifetimes of many extinct families that are only known from mid-Cretaceous ambers (e.g., Ohlhoffiidae and Burmasphecidae).

Diversification dynamics of Hymenoptera lineages

We estimated the periods of diversification and extinction of Hymenoptera lineages at both genus (Figure 2) and family (Figures S2–S4) levels, using the birth-death model with constrained shifts (BDCSs).[16] Because singletons—a common sampling bias—can distort inferred patterns, we excluded them from



(legend on next page)

all our analyses[12] and cross-validated our results with the reversible-jump Markov chain Monte Carlo model (RJMCMC).[17] The genus-level analyses revealed four major periods of diversification (i.e., positive net diversification rates with origination exceeding extinction) as well as three significant declines in the diversity dynamics of hymenopteran genera (negative net diversification rates with extinction exceeding origination) (Figures 2B and S3). In contrast, family-level analyses showed no major declines in the diversification of hymenopteran families (Figures S2 and S4). At this level, the only minor extinction event occurred during the early Late Cretaceous (Figures S2–S4), whereas Hymenoptera diversified during four pulses, roughly mirroring the periods of diversification observed at the genus level.

At the genus level, the first extinction event occurred during the early Late Triassic (Figures 2A and 2B; ≈1.4-fold general background extinction rate) and impacted symphytan lineages (i.e., Xyelidoidea; Figures 2D, 2E, and S3). The second extinction period took place during the Late Jurassic (≈2.2-fold general background extinction rate) and affected both Symphyta and early diverging Apocrita (non-Aculeata) lineages (≈2-fold Apocrita background extinction rate; Figures 2E and 2K). The third extinction period occurred during the mid-Cretaceous, uniformly affecting the Symphyta, Apocrita (non-Aculeata), and Aculeata (≈2.7-fold general background extinction rate; Figures 2E, 2H, and 2K). At the family level, an early Late Cretaceous extinction was also observed (≈1.9-fold general background extinction rate; Figures S2–S4), with a similarly uniform impact (Figures S3E, S3H, and S3K). Additionally, the Symphyta experienced another, less pronounced extinction event during the Late Jurassic (Figures S2 and S4).

At the genus level, the first period of diversification occurred between the Permian and Early Triassic, primarily involving Symphyta (Figures 2A, 2B, 2D, and 2E; ≈1.7-fold general background origination rate). During the Jurassic, a second diversification event witnessed the rise of both Symphyta and early-diverging Apocrita (non-Aculeata) lineages (Figures 2A, 2B, 2D, 2E, 2J, and 2K; ≈3.2-fold general background origination rate). The third diversification period, centered on the Early Cretaceous, was uniformly recorded in Symphyta, Apocrita (non-Aculeata), and Aculeata (Figures 2D–2L; ≈2.15-fold general background diversification rate). The last episode of diversification began during the ATR and extended nearly to the present, with a similarly uniform impact across Symphyta, Apocrita (non-Aculeata), and Aculeata (Figures 2A–2L; ≈1.9-fold general background diversification rate). At the family level, Hymenoptera diversified from the Permian to the mid-Cretaceous, with a drop in net diversification rates during the Early and Middle Triassic and late Early Jurassic (Figures S3A–S3C). Additionally, a small increase in the net diversification rate was observed during the Eocene (Figures S2–S4).

Our global lineage-through-time analysis indicates that the number of genera increased during the Middle-Late Triassic, the Middle-Late Jurassic, the Early and mid-Cretaceous, and sharply after the Cretaceous-Paleogene boundary, rising from approximately 30 to over 520 genera near the present (Figure 2C). We also record a peak in the number of genera during the Eocene and Oligocene transition (EOT), uniformly recorded across our analyses (Figures 2C, 2F, 2I, and 2L). At the family level, a different pattern emerges, with an exponential increase in the number of families from the middle Permian to the Late Cretaceous, peaking at around 120 families during the mid- and Late Cretaceous (Figure S3C). Following this peak, diversity declined slightly and stabilized at around 100 families throughout the Cenozoic (Figure S3C). We also find a peak of Symphyta families around the EOT (Figure S3F).

BBB model sensitivity: Uncertainties in the root age and in the number of extant species

Across our analyses with varying maximum boundaries of the uniform prior on the root age, the estimated ages of each clade remain strikingly consistent (supplemental information). On average, across 159 families, the age discrepancy between analyses using a prior set to either 281 or 345 Ma is approximately 4.29 Ma. Differences range from as little as ≈0.027 Ma (Burmasphecidae) to a maximum of 9.95 Ma (Melittidae). At the superfamily level, the age discrepancy decreases to an average of 0.26 Ma. Notably, there is no consistent trend where analyses with the oldest/youngest root age uniformly would produce the oldest/youngest estimates (https://doi.org/10.6084/m9.figshare.28474061.v1).

In our sensitivity analyses performed on three parasitoid superfamilies (Ceraphronoidea, Chalcidoidea, and Ichneumonoidea), a 2-fold increase in the number of extant species had no effect on the estimated age of origin (maximum difference of 2.2 Ma for Ceraphronoidea). A 5-fold increase resulted in slightly older age estimates for Ceraphronoidea and Chalcidoidea but not for Ichneumonoidea (supplemental information). Using extreme increases in their extant diversity (≈33-fold increase for Ceraphronoidea and ≈22-fold increase for Chalcidoidea) resulted in older age estimates for the origin of these clades, with differences of about 26 and 10 Ma, respectively.

DISCUSSION

Origin and main diversification phases of Hymenoptera

When it comes to analyzing the diversification of organisms, dating estimates are pivotal. The BBB model[6] proves useful

Figure 2. Diversification and diversity dynamics of major Hymenoptera clades, inferred from genus-level analyses using the BDCSs, without singleton

(A, D, G, and J) Bayesian estimates of origination (blue) and extinction (red) rates for Hymenoptera.

(B, E, H, and K) Net diversification (origination minus extinction) rates for Hymenoptera.

(C, F, I, and L) Diversity (number of genera) of Hymenoptera. For each plot, solid lines indicate mean posterior rates; shaded areas show 95% CI. Reconstructions of diversity trajectories replicated to incorporate uncertainties around the ages of the fossil occurrences. ATR, Angiosperm Terrestrial Revolution; K-Pg, Cretaceous-Paleogene boundary. Time is in millions of years. The color of each geological period in the chronostratigraphic scale follows that of the International Chronostratigraphic Chart (v2023/09). Insect silhouettes from http://phylopic.org/; licenses at https://creativecommons.org/publicdomain/zero/1.0/. See also Figures S2–S4 and Data S1.

because it eliminates a complex step that has major implications for estimating divergence times: fossil calibration.[18–20] To calibrate nodes, fossils are chosen based on morphological characters, distinguishing between plesiomorphic and apomorphic states, to provide a minimum age for the most recent common ancestor of a group.[21] This step is challenging because only a subset of these characters is generally preserved in fossils, whereas extant species are delimited by many morphological characters. In addition, morphological convergence can hinder the taxonomic affiliation of fossils. This is why fossils are traditionally excluded from time-calibrated phylogenies, except for the temporal information of some of them that is used for calibrations.[20] Recent methodological developments, such as the total-evidence dating approach, have partially alleviated these limitations,[22–25] but they often require extensive morphological data, hindering their large-scale application.[26] The BBB model, however, can be used to assess which results obtained from molecular-based studies most closely align with the fossil record. It also allows for estimating the origins and extinctions of clades while fully leveraging the evidence from the fossil record (Figures 1 and S1).

Uncertainties in phylogenetic relationships also have a significant impact on divergence time estimates, often leading to discrepancies of millions of years in the origin of clades.[27] Insects are no exception to this issue. Historical conundrums in the understanding of Hymenoptera evolution include the origin and placement of the Xyelidae and Vespina (=Orussidae plus Apocrita), and the composition of Vespina.[9,10,28] Molecular phylogenetic studies have proposed three alternative topologies[29]: either Xyelidae are nested within the first radiation of Symphyta (i.e., Eusymphyta),[10] are closer to the root of Hymenoptera and represent the second earliest diverging lineage after the Tenthredinoidea,[9] or are the earliest diverging Hymenoptera lineage.[29,30] These alternatives lead to a 90-Ma age discrepancy for the superfamily, approximately from the Early Jurassic (185 Ma)[10] to the late Permian (275 Ma).[9,30] Our results favor the hypothesis of a Permian origin, with the superfamily Xyeloidea originating more than 270 mya and the Xyelidae shortly after, around 260 mya (Figures 1 and S1; Tables 1 and S1). This ancient origin is supported by fossil evidence, with extinct and extant xyelid subfamilies documented worldwide from the Triassic (e.g., ~230 Ma; Australia, Ipswich, Norian[31]; Japan, Momonoki Formation, Carnian[32]; Kyrgyzstan, Madygen, Ladinian-Carnian[33]; South Africa, Molteno, Carnian[34]; Argentina, Potrerillos Formation, Carnian[35]). Consistently, our diversification analyses indicate a peak in diversification rates during the Permian (Figures 2 and S2–S4).

Impacts of the ATR and the MMPR

The ATR is marked by a shift from gymnosperm- to angiosperm-dominated ecosystems, an upheaval that has created numerous opportunities for the evolution of new ecological relationships, such as pollination.[1,11,13] Among insects, the Hymenoptera, particularly the Apoidea—the clade encompassing the bee lineages or Antophila (e.g., Apidae, Colletidae, Halictidae, and Megachilidae)—are thought to have benefited greatly from this event.[14,15] However, uncertainties surrounding the age of the Apoidea and its constituent lineages obscure our understanding of how they actually benefited from the ATR. Recent phylogenetic analyses have estimated that the origin of the clade could as well be the Late Jurassic (~155 Ma)[10] as the mid-Cretaceous (~100 Ma).[9] Our results support a Late Jurassic origin (median: 148.22 Ma, CI = 143.01–162.36 Ma), following the recent discovery of many Apoidea species in mid-Cretaceous Kachin amber.[36,37] Most Apoidea lineages, such as the Megachilidae (median: 66.12 Ma, CI = 61–75.99 Ma), have diversified during the Cretaceous and Paleogene, particularly during the ATR and the third phase of the MMPR (Figure 1; Tables 1 and S1).[9,15] Our results suggest that the origin of the Antophila lineages coincided with the radiation of angiosperms[6,14,15,38] and occurred in the middle of the ATR.[1] Their diversification likely played a significant role in the diversification peak observed during the Late Cretaceous, a pattern also identified in molecular-based analyses.[9] Our analyses indicate that the third phase of the MMPR encompasses major radiations of Chalcidoidea and other Apocrita lineages (particularly the Aculeata) (Figure 1), with a peak in net diversification rate recorded for Apocrita (non-Aculeata) and Aculeata at the genus level (Figure 2) and for Apocrita (non-Aculeata) at the family level (Figures S2 and S4). For Chalcidoidea, our results align with a recent molecular-based study indicating that most extant chalcidoid families originated during the Cretaceous.[39] However, our analyses appear to underestimate the age of the superfamily, which molecular studies place in the Middle Jurassic, whereas our findings suggest an Early Cretaceous origin. We attribute this discrepancy to the exclusion from our analyses of the recently described chalcidoid fossil family Protoitidae. This family was not included because its description was published after the cutoff date for our data compilation.[40] Incorporating Protoitidae in future analyses is likely to push back the estimated origin of the crown-group Chalcidoidea by several million years.

One of the most critical events in the evolutionary history of Hymenoptera is the transition to a parasitoid lifestyle, a major innovation characteristic of the Vespina.[9,10] This shift is a cornerstone of the MMPR, which began in the Early Jurassic and marked the transition from bottom-up regulation of terrestrial food webs, dominated by less efficient predators, to top-down regulation by highly efficient parasitoid clades.[2] The origin of the Vespina has been dated to the Late Triassic (~200 Ma),[41] the Early to Middle Triassic (234 or 247 Ma),[9,10] or even the middle Permian (270 Ma).[30] Our analysis of the fossil record indicates that some of these estimates, particularly the Late Triassic origin of Vespina, are unrealistic when considering the fossil record and can be disregarded. Among the parasitoid Apocrita (non-Aculeata), the oldest clades—the Stephanoidea (median: 192.25 Ma, CI = 180.14–213.69 Ma), the Roproniidae (Proctotrupoidea; median: 205.27 Ma, CI = 165.30–257.16 Ma), and the Megalyridae (Megalyroidea; median: 189.13 Ma, CI = 178.09–206.73 Ma)—have their origins dated to the Late Triassic or Early Jurassic (Figure 1; Table 1). This supports an Early to Middle Triassic origin for Apocrita and Parasitoida.[9,10] Our diversification analyses further suggest a burst in the net diversification rate of Apocrita (non-Aculeata) families during the Early Jurassic (Figures S3 and S4). Therefore, the MMPR likely began as early as the Triassic, before the earliest fossil records of apocritan families (mostly Parasitoida). This hypothesis is consistent with recent analyses of the Evanioidea, an ancestrally parasitoid apocritan superfamily that originated in the Late Triassic (stem

age = 203 Ma).[42] This superfamily is a key actor of the MMPR, with an important Jurassic radiation (Praeaulacidae) showing morphological specialization for parasitoidism (e.g., a two-segmented petiole facilitating movements of the gaster).[43] Incidentally, Evanioidea is a textbook example of a clade whose familial diversity in the fossil record is higher than it is today, with five out of eight families extinct, further highlighting the importance of not overlooking fossils.[42]

The first stage of the MMPR is characterized by the emergence and diversification of extinct parasitoid families, often belonging to stem groups of modern superfamilies[2]—a pattern we found in our analyses of the diversity dynamics of Hymenoptera (Figures 1, S2, and S4). These families often exhibit a mixture of plesiomorphic (e.g., complex wing venation and absence of a wasp waist) and apomorphic traits (e.g., fore wing lacking the sub-basal loop of vein 2A or high articulation on the propodeum or not widely sclerotized ventrally at the metasomal attachment). They went extinct during the Cretaceous (extinction of Ephialtitidae, median: 112.47 Ma, CI = 106.26–115.99 Ma; Praeaulacidae, median: 97.12 Ma, CI = 91.24–101.97 Ma) before the second phase of the MMPR, marked by the rise of extant families of parasitoid wasps.[2] This turnover is confirmed here with a peak of extinction rate during the Late Jurassic and a subsequent episode of diversification during the Early Cretaceous (Figure 2). The characteristic clades of this second phase, which radiated during the Early Cretaceous, are the Chrysidoidea,[9] Chalcidoidea,[10,39] Pimpliformes (Ichneumonidae),[44] and Braconidae.[45] Our BBB analyses converge with molecular clock analyses in estimating their origins in the Jurassic or Early Cretaceous (Table 1). Additionally, our PyRate results align with the diversification patterns derived from phylogenetic analyses, particularly for the Apocrita and Aculeata.[9]

Interestingly, our results indicate a peak in hymenopteran diversity, uniformly recorded across our analyses, during the Early Cretaceous. This peak persists when singleton taxa are removed from the analyses (Figures 2 and S2). A "similar" signal was previously found, with an Aptian peak in family diversity.[46] This peak can be attributed to taphonomic conditions, such as the mass production of resin during the Cretaceous Resinous Interval,[47] which facilitated the entombment of many insects. Additionally, it coincides with the simultaneous presence of the last representatives of Mesozoic lineages, which went extinct during the Late Cretaceous, and the early representatives of Cenozoic lineages (Figure 1). This transitional period saw a proliferation of new families exhibiting a variety of morphologies, many of which are no longer common today. The coexistence of these diverse faunas may have generated competitive interactions, eventually resulting in the extinction of some clades during the Late Cretaceous (diversity dependence), which should be explored in future studies along with other factors.

To conclude, our analyses of the fossil record reveal that the evolutionary history of Hymenoptera is characterized initially by the radiation of Symphyta during the Paleozoic and Mesozoic, followed by the emergence of Apocrita, particularly the Parasitoida, during the Mesozoic and Cenozoic (Figures 1 and S2–S4). During the MMPR, Parasitoida lineages underwent successive replacements. The first burst of diversification for Parasitoida included extinct lineages with no modern representatives or stem lineages of modern groups. Most of these early lineages survived until the Cretaceous, after which they were replaced by new families and the crown representatives of their own groups. These replacements occurred synchronously with changes in floral assemblages, notably the emergence of flowering plants during the ATR (100–50 Ma). Clade competition, but also angiosperm diversification, likely favored predatory and parasitoid lineages of wasps by triggering the diversification of their phytophagous preys and hosts.[48]

Applicability and robustness of the BBB model

Estimating ages in the deep past is inherently fraught with uncertainty. Recently developed probabilistic approaches allow us to investigate how some of these uncertainties might affect these estimates. But it is still relatively rarely done. The impact of uncertainty in the "true" number of extant species, for instance, is often overlooked in macroevolutionary analyses, despite its direct implication in dating models through parameters that define the sampling probability of extant lineages.[24,49] Recent simulations suggest that increasing taxon sampling density does not necessarily improve divergence time estimates.[49] However, empirical tests of these results, particularly for insects for which diversity is much higher, remain limited.

Our sensitivity analyses indicate that age estimates derived exclusively from the hymenopteran fossil record using the BBB model are minimally affected by uncertainty in extant species diversity (up to 5-fold increase). Significant differences in age estimates occurred only for drastic increases in diversity. Likewise, age estimates were marginally affected by changes in the root age constraints. This shows that the BBB model appears to be robust for estimating clade ages with incomplete extant diversity documentation and unclear origin (see https://doi.org/10.6084/m9.figshare.28271135.v1 and https://doi.org/10.6084/m9.figshare.28474061.v1).

This robustness suggests that the BBB model is broadly applicable beyond traditionally studied groups with thoroughly documented extant diversity (e.g., plants[6] and mammals[7]), making it a valuable tool for studying insect macroevolution. The BBB model is also particularly valuable for studying the evolution of clades that have limited molecular data but a relatively robust fossil record. Nonetheless, rather than relying solely on currently documented diversity, future studies using this or other models would benefit, in particular, from including estimated species numbers and conducting sensitivity analyses in general.

Overall, our study demonstrates that age estimates derived solely from analysis of the fossil record using the BBB model are often compatible with age estimates obtained from molecular dating studies that incorporate extensive molecular and morphological datasets (Table 1). This highlights the BBB model as a highly valuable complementary or alternative approach for estimating the origin of lineages.

Perspectives and future directions

The development of new methods, such as PyRate or BBB models, represents an important step toward a better understanding of lineage evolution while accounting for the limitations of the fossil record. However, these models rely on assumptions that may not be fully consistent with real evolutionary processes. For example, PyRate estimates diversification rates under the assumption that they remain homogeneous within a clade but

can vary over time, despite evidence suggesting that diversification is heterogeneous across the Hymenoptera tree of life and through different time periods.[9]

To mitigate this limitation, particularly within a group as diverse as Hymenoptera, we generated sub-datasets to account for this heterogeneity. At the same time, refinement of the fossil record remains essential. Critical steps include clarifying the placement of key fossil taxa, such as *Avioxyela gallica* Nel et al.,[50] resolving uncertain attributions in various parasitoid lineages, and filling existing gaps in the fossil record.

For Hymenoptera, special attention should be given to the study of Permian and Campanian or Maastrichtian deposits, which remain notably scarce for fossil insects. In addition, it is essential to revisit other deposits, such as the Oise amber, which have been largely overshadowed by the recent influx of fossil material from the Burmese amber. Furthermore, the diversity of certain groups remains poorly understood, both in the fossil record and in modern ecosystems, despite their considerable extant species richness. In many cases, this is due to ecological factors, such as habitats unfavorable for fossilization (e.g., Mutillidae), a lack of taxonomic expertise,[51] or the Linnean shortfall, where numerous specimens exist but have yet to be formally studied and described (e.g., Chalcidoidea and Platygastridae).[52] Improving the fossil record would greatly benefit both fossil-based evolutionary inferences and phylogeny-based studies, leading to a more comprehensive understanding of Hymenoptera diversification.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to, and will be fulfilled by, the lead contact, Corentin Jouault (jouaultc0@gmail.com).

Materials availability

See the key resources table for materials used in this study

Data and code availability

Output files for each BBB and PyRate analysis can be found at https://doi.org/10.6084/m9.figshare.28271135.v1. All command lines have been deposited on Figshare at https://doi.org/10.6084/m9.figshare.27058510.v1, additional figures at https://doi.org/10.6084/m9.figshare.28474070.v1, and additional tables at https://doi.org/10.6084/m9.figshare.28474061.v1 and are publicly available.

ACKNOWLEDGMENTS

We thank three anonymous reviewers for their constructive comments. C.J. is grateful to Emily Carlisle for her advice. We are grateful to colleagues who provided access to institutional collections: André Nel (MNHN), Victor Shegelski (University of Alberta), Monica Solórzano-Kraemer (Senckenberg Museum), Jacek Szwedo (Gdansk University), and Bo Wang (NIGPAS). We are also grateful to the many collectors who provided access to their private collections and shared their data: Michele Baldi, Carsten Gröhn, Christel and Hans-Werner Hoffeins, Patrick Müller, Axel Niggeloh, Rainer Ohlhoff, Jean-Marc Pouillon, and Jörg Wunderlich. We are grateful to the PBDB team and contributors for their constant effort in maintaining their database and for making it freely available. S.Á.-P. was supported by the project CREI, PID2022-137316NB-C21, of the Spanish MICIU (AEI) with FEDER funds.

AUTHOR CONTRIBUTIONS

C.J., V.P., F.L.C., and F.L. designed and conceived the study. C.J. compiled the data, with discussions with N.O., S.Á.-P., D.H., and V.P. C.J. analyzed the data. C.J. wrote the first draft of the manuscript, to which all authors contributed. C.J., V.P., F.L.C., and F.L. reviewed the manuscript. All authors approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

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

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Fossil record of Hymenoptera
  - Data preparation for estimating clades' origin and extinction times
  - Extant species diversity in Hymenoptera families
  - The Bayesian Brownian Bridge model
  - Applying the Bayesian Brownian Bridge to the Hymenoptera fossil record
  - Dynamics of origination and extinction through time
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2025.03.002.

Received: September 20, 2024
Revised: January 24, 2025
Accepted: March 4, 2025
Published: March 26, 2025

REFERENCES

1. Benton, M.J., Wilf, P., and Sauquet, H. (2022). The Angiosperm Terrestrial Revolution and the origins of modern biodiversity. New Phytol. *233*, 2017–2035. https://doi.org/10.1111/nph.17822.
2. Labandeira, C.C., and Li, L. (2021). The History of Insect Parasitism and the Mid-Mesozoic Parasitoid Revolution. In The Evolution and Fossil Record of Parasitism, K. De Baets, and J.W. Huntley, eds. (Springer), pp. 377–533. https://doi.org/10.1007/978-3-030-42484-8_11.
3. Eggleton, P., and Belshaw, R. (1992). Insect parasitoids: an evolutionary overview. Phil. Trans. R. Soc. Lond. B *337*, 1–20. https://doi.org/10.1098/rstb.1992.0079.
4. Pennacchio, F., and Strand, M.R. (2006). Evolution of developmental strategies in parasitic Hymenoptera. Annu. Rev. Entomol. *51*, 233–258. https://doi.org/10.1146/annurev.ento.51.110104.151029.
5. Silvestro, D., Salamin, N., and Schnitzler, J. (2014). PyRate: A new program to estimate speciation and extinction rates from incomplete fossil data. Methods Ecol. Evol. *5*, 1126–1131. https://doi.org/10.1111/2041-210X.12263.
6. Silvestro, D., Bacon, C.D., Ding, W., Zhang, Q., Donoghue, P.C.J., Antonelli, A., and Xing, Y. (2021). Fossil data support a pre-Cretaceous origin of flowering plants. Nat. Ecol. Evol. *5*, 449–457. https://doi.org/10.1038/s41559-020-01387-8.
7. Carlisle, E., Janis, C.M., Pisani, D., Donoghue, P.C.J., and Silvestro, D. (2023). A timescale for placental mammal diversification based on Bayesian modeling of the fossil record. Curr. Biol. *33*, 3073–3082.e3. https://doi.org/10.1016/j.cub.2023.06.016.
8. Goulet, H., and Huber, J.T. (1993). Hymenoptera of the World: An Identification Guide to Families (Agriculture Canada).

9. Blaimer, B.B., Santos, B.F., Cruaud, A., Gates, M.W., Kula, R.R., Mikó, I., Rasplus, J.-Y., Smith, D.R., Talamas, E.J., Brady, S.G., et al. (2023). Key innovations and the diversification of Hymenoptera. Nat. Commun. *14*, 1212. https://doi.org/10.1038/s41467-023-36868-4.

10. 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.

11. Condamine, F.L., Silvestro, D., Koppelhus, E.B., and Antonelli, A. (2020). The rise of angiosperms pushed conifers to decline during global cooling. Proc. Natl. Acad. Sci. USA *117*, 28867–28875. https://doi.org/10.1073/pnas.2005571117.

12. Jouault, C., Condamine, F.L., Legendre, F., and Perrichot, V. (2024). The Angiosperm Terrestrial Revolution buffered ants against extinction. Proc. Natl. Acad. Sci. USA *121*, e2317795121. https://doi.org/10.1073/pnas.2317795121.

13. Peris, D., and Condamine, F.L. (2024). The angiosperm radiation played a dual role in the diversification of insects and insect pollinators. Nat. Commun. *15*, 552. https://doi.org/10.1038/s41467-024-44784-4.

14. Cardinal, S., and Danforth, B.N. (2013). Bees diversified in the age of eudicots. Proc. R. Sci. B. *280*, 20122686. https://doi.org/10.1098/rspb.2012.2686.

15. Almeida, E.A.B., Bossert, S., Danforth, B.N., Porto, D.S., Freitas, F.V., Davis, C.C., Murray, E.A., Blaimer, B.B., Spasojevic, T., Ströher, P.R., et al. (2023). The evolutionary history of bees in time and space. Curr. Biol. *33*, 3409–3422.e6. https://doi.org/10.1016/j.cub.2023.07.005.

16. Silvestro, D., Cascales-Miñana, B., Bacon, C.D., and Antonelli, A. (2015). Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. New Phytol. *207*, 425–436. https://doi.org/10.1111/nph.13247.

17. Silvestro, D., Salamin, N., Antonelli, A., and Meyer, X. (2019). Improved estimation of macroevolutionary rates from fossil data using a Bayesian framework. Paleobiology *45*, 546–570. https://doi.org/10.1017/pab.2019.23.

18. O'Reilly, J.E., dos Reis, M., and Donoghue, P.C.J. (2015). Dating tips for divergence-time estimation. Trends Genet. *31*, 637–650. https://doi.org/10.1016/j.tig.2015.08.001.

19. Rannala, B. (2016). Conceptual issues in Bayesian divergence time estimation. Philos. Trans. R. Soc. Lond. B Biol. Sci. *371*, 20150134. https://doi.org/10.1098/rstb.2015.0134.

20. dos Reis, M., Donoghue, P.C.J., and Yang, Z.H. (2016). Bayesian molecular clock dating of species divergences in the genomics era. Nat. Rev. Genet. *17*, 71–80. https://doi.org/10.1038/nrg.2015.8.

21. Parham, J.F., Donoghue, P.C.J., Bell, C.J., Calway, T.D., Head, J.J., Holroyd, P.A., Inoue, J.G., Irmis, R.B., Joyce, W.G., Ksepka, D.T., et al. (2012). Best practices for justifying fossil calibrations. Syst. Biol. *61*, 346–359. https://doi.org/10.1093/sysbio/syr107.

22. 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.

23. Ronquist, F., Lartillot, N., and Phillips, M.J. (2016). Closing the gap between rocks and clocks using total-evidence dating. Philos. Trans. R. Soc. Lond. B Biol. Sci. *371*, 20150136. https://doi.org/10.1098/rstb.2015.0136.

24. Zhang, C., Stadler, T., Klopfstein, S., Heath, T.A., and Ronquist, F. (2016). Total-evidence dating under the fossilized birth-death process. Syst. Biol. *65*, 228–249. https://doi.org/10.1093/sysbio/syv080.

25. Keating, J.N., Sansom, R.S., Sutton, M.D., Knight, C.G., and Garwood, R.J. (2020). Morphological phylogenetics evaluated using novel evolutionary simulations. Syst. Biol. *69*, 897–912. https://doi.org/10.1093/sysbio/syaa012.

26. Klopfstein, S. (2021). The age of insects and the revival of the minimum age tree. Aust. Entomol. *60*, 138–146. https://doi.org/10.1111/aen.12478.

27. Giribet, G., and Edgecombe, G.D. (2019). The Phylogeny and Evolutionary History of Arthropods. Curr. Biol. *29*, R592–R602. https://doi.org/10.1016/j.cub.2019.04.057.

28. Wutke, S., Blank, S.M., Boevé, J.-L., Faircloth, B.C., Koch, F., Linnen, C.R., Malm, T., Niu, G., Prous, M., Schiff, N.M., et al. (2024). Phylogenomics and biogeography of sawflies and woodwasps (Hymenoptera, Symphyta). Mol. Phylogenet. Evol. *199*, 108144. https://doi.org/10.1016/j.ympev.2024.108144.

29. Zhang, Y.M., Bossert, S., and Spasojevic, T. (2025). Evolving perspectives in Hymenoptera systematics: Bridging fossils and genomes across time. Syst. Entomol. *50*, 1–31. https://doi.org/10.1111/syen.12645.

30. 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.

31. Riek, E.F. (1955). Fossil insects from the Triassic beds at Mt. Crosby, Queensland. Aust. J. Zool. *3*, 654–691. https://doi.org/10.1071/ZO9550654.

32. Oyama, N., and Maeda, H. (2020). *Madygella humioi* sp. nov. from the Upper Triassic Mine Group, southwest Japan: the oldest record of a sawfly (Hymenoptera: Symphyta) in East Asia. Paleontol. Res. *24*, 64–71. https://doi.org/10.2517/2019PR005.

33. Rasnitsyn, A.P. (1969). The origin and evolution of Lower Hymenoptera. Transactions of the Paleontological Institute, Academy of Sciences of the USSR *123*, 1–196.

34. Schlüter, S. (2000). *Moltenia rieki* n. gen., n. sp. (Hymenoptera: Xyelidae?), a tentative sawfly from the Molteno Formation (Upper Triassic), South Africa. Paläont. Z. *74*, 75–78. https://doi.org/10.1007/BF02987953.

35. Lara, M.B., Rasnitsyn, A.P., and Zavattieri, A.M. (2014). *Potrerilloxyela menendezi* gen. et sp. nov. from the Late Triassic of Argentina: the oldest representative of Xyelidae (Hymenoptera: Symphyta) for Americas. Paleontol. J. *48*, 182–190. https://doi.org/10.1134/S0031030114020075.

36. Rosa, B.B., and Melo, G.A.R. (2021). Apoid wasps (Hymenoptera: Apoidea) from mid-Cretaceous amber of northern Myanmar. Cretac Res. *122*, 104770. https://doi.org/10.1016/j.cretres.2021.104770.

37. Rosa, B.B., and Melo, G.A.R. (2023). A new fossil family of aculeate wasp sheds light on early evolution of Apoidea (Hymenoptera). Syst. Entomol. *48*, 402–421. https://doi.org/10.1111/syen.12584.

38. Dimitrov, D., Xu, X., Su, X., Shrestha, N., Liu, Y., Kennedy, J.D., Lyu, L., Nogués-Bravo, D., Rosindell, J., Yang, Y., et al. (2023). Diversification of flowering plants in space and time. Nat. Commun. *14*, 7609. https://doi.org/10.1038/s41467-023-43396-8.

39. Cruaud, A., Rasplus, J.-Y., Zhang, J., Burks, R., Delvare, G., Fusu, L., Gumovsky, A., Huber, J.T., Janšta, P., Mitroiu, M.-D., et al. (2024). The Chalcidoidea bush of life: evolutionary history of a massive radiation of minute wasps. Cladistics *40*, 34–63. https://doi.org/10.1111/cla.12561.

40. Ulmer, J.M., Janšta, P., Azar, D., and Krogmann, L. (2023). At the dawn of megadiversity – Protoitidae, a new family of Chalcidoidea (Hymenoptera) from Lower Cretaceous Lebanese amber. J. Hymenopt. Res. *96*, 879–924. https://doi.org/10.3897/jhr.96.105494.

41. 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.

42. Jouault, C., Maréchal, A., Condamine, F.L., Wang, B., Nel, A., Legendre, F., and Perrichot, V. (2022). Including fossils in phylogeny: a glimpse into the evolution of the superfamily Evanioidea (Hymenoptera: Apocrita) under tip-dating and the fossilized birth–death process. Zool. J. Linn. Soc. *194*, 1396–1423. https://doi.org/10.1093/zoolinnean/zlab034.

43. Rasnitsyn, A., and Zhang, H. (2007). Nevaniinae subfam. n., a new fossil taxon (Insecta: Hymenoptera: Evanioidea: Praeaulacidae) from the

Middle Jurassic of Daohugou in Inner Mongolia, China. Insect Syst. Evol. *38*, 149–166. https://doi.org/10.1163/187631207788783987.

44. 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.

45. Kittel, R.N., Austin, A.D., and Klopfstein, S. (2016). Molecular and morphological phylogenetics of chelonine parasitoid wasps (Hymenoptera: Braconidae), with a critical assessment of divergence time estimations. Mol. Phylogenet. Evol. *101*, 224–241. https://doi.org/10.1016/j.ympev.2016.05.016.

46. Schachat, S.R., Labandeira, C.C., Clapham, M.E., and Payne, J.L. (2019). A Cretaceous peak in family-level insect diversity estimated with mark–recapture methodology. Proc. R. Sci. B. *286*, 20192054. https://doi.org/10.1098/rspb.2019.2054.

47. Delclòs, X., Peñalver, E., Barrón, E., Peris, D., Grimaldi, D.A., Holz, M., Labandeira, C.C., Saupe, E.E., Scotese, C.R., Solórzano-Kraemer, M.M., et al. (2023). Amber and the Cretaceous Resinous Interval. Earth Sci. Rev. *243*, 104486. https://doi.org/10.1016/j.earscirev.2023.104486.

48. Boderau, M., Nel, A., and Jouault, C. (2025). Diversification and extinction of Hemiptera in deep time. Commun. Biol. *8*, 352. https://doi.org/10.1038/s42003-025-07773-x.

49. Luo, A., Zhang, C., Zhou, Q.-S., Ho, S.Y.W., and Zhu, C.-D. (2023). Impacts of taxon-sampling schemes on Bayesian tip dating under the fossilized birth-death process. Syst. Biol. *72*, 781–801. https://doi.org/10.1093/sysbio/syad011.

50. Nel, A., Roques, P., Nel, P., Prokin, A.A., Bourgoin, T., Prokop, J., Szwedo, J., Azar, D., Desutter-Grandcolas, L., Wappler, T., et al. (2013). The earliest known holometabolous insects. Nature *503*, 257–261. https://doi.org/10.1038/nature12629.

51. Engel, M.S., Ceríaco, L.M.P., Daniel, G.M., Dellapé, P.M., Löbl, I., Marinov, M., Reis, R.E., Young, M.T., Dubois, A., Agarwal, I., et al. (2021). The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. Zool. J. Linn. Soc. *193*, 381–387. https://doi.org/10.1093/zoolinnean/zlab072.

52. Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M., and Ladle, R.J. (2015). Seven Shortfalls that Beset Large-Scale Knowledge of Biodiversity. Annu. Rev. Ecol. Evol. Syst. *46*, 523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400.

53. 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.

54. Engel, M.S. (2022). Establishment of some clade names for Amphiesmenoptera (Insecta: Holometabola). Novit. Paleoentomol. *24*, 1–7. https://doi.org/10.17161/np.24.18498.

55. Hong, Y.C. (2002). Amber Insects of China (Huayu Nature Book Trade).

56. Pilgrim, E.M., von Dohlen, C.D., and Pitts, J.P. (2008). Molecular phylogenetics of Vespoidea indicate paraphyly of the superfamily and novel relationships of its component families and subfamilies. Zool. Scr. *37*, 539–560. https://doi.org/10.1111/j.1463-6409.2008.00340.x.

57. Chen, H., Lahey, Z., Talamas, E.J., Valerio, A.A., Popovici, O.A., Musetti, L., Klompen, H., Polaszek, A., Masner, L., Austin, A.D., et al. (2021). An integrated phylogenetic reassessment of the parasitoid superfamily Platygastroidea (Hymenoptera: Proctotrupomorpha) results in a revised familial classification. Syst. Entomol. *46*, 1088–1113. https://doi.org/10.1111/syen.12511.

58. Sann, M., Meusemann, K., Niehuis, O., Escalona, H.E., Mokrousov, M., Ohl, M., Pauli, T., and Schmid-Egger, C. (2021). Reanalysis of the apoid wasp phylogeny with additional taxa and sequence data confirms the placement of Ammoplanidae as sister to bees. Syst. Entomol. *46*, 558–569. https://doi.org/10.1111/syen.12475.

59. Burks, R., Mitroiu, M.-D., Fusu, L., Heraty, J.M., Janšta, P., Heydon, S., Papilloud, N.D.-S., Peters, R.S., Tselikh, E.V., Woolley, J.B., et al. (2022). From hell's heart I stab at thee! A determined approach towards a monophyletic Pteromalidae and reclassification of Chalcidoidea (Hymenoptera). J. Hymenopt Res. *94*, 13–88. https://doi.org/10.3897/jhr.94.94263.

60. Niu, G.Y., Budak, M., Korkmaz, E.M., Doğan, Ö., Nel, A., Wan, S.Y., Cai, C.Y., Jouault, C., Li, M., and Wei, M.C. (2022). Phylogenomic analyses of the Tenthredinoidea support the familial rank of Athaliidae (Insecta, Tenthredinoidea). Insects *13*, 858. https://doi.org/10.3390/insects13100858.

61. Noyes, J.S. (2019). Universal Chalcidoidea Database. World Wide Web electronic publication. http://www.nhm.ac.uk/chalcidoids.

62. Nyman, T., Onstein, R.E., Silvestro, D., Wutke, S., Taeger, A., Wahlberg, N., Blank, S.M., and Malm, T. (2019). The early wasp plucks the flower: disparate extant diversity of sawfly superfamilies (Hymenoptera: "Symphyta") may reflect asynchronous switching to angiosperm hosts. Biol. J. Linn. Soc. *128*, 1–19. https://doi.org/10.1093/biolinnean/blz071.

63. Aguiar, A.P., Deans, A.R., Engel, M.S., Forshage, M., Huber, J.T., Jennings, J.T., Johnson, N.F., Lelej, A.S., Longino, J.T., Lohrmann, V., et al. (2013). Order Hymenoptera. Zootaxa *3703*, 51–62. https://doi.org/10.11646/zootaxa.3703.1.12.

64. Ascher, J.S., and Pickering, J. (2023). Discover Life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila). http://www.discoverlife.org/mp/20q?guide=Apoidea_species.

65. Agnoli, G.L., and Rosa, P. (2023). Systematics of Chrysididae. https://www.chrysis.net/chrysididae/systematics-of-chrysididae/.

66. Olmi, M., Marletta, A., and Guglielmino, A. (2020). The first record of the family Embolemidae (Hymenoptera: Chrysidoidea) in Réunion, with description of a new species of Embolemus Westwood. Isr. J. Entomol. *50*, 40–48. https://doi.org/10.5281/zenodo.3928171.

67. Azevedo, C.O., Alencar, I.D.C.C., Ramos, M.S., Barbosa, D.N., Colombo, W.D., Vargas, J.M.R., and Lim, J. (2018). Global guide of the flat wasps (Hymenoptera, Bethylidae). Zootaxa *4489*, 1–294. https://doi.org/10.11646/Zootaxa.4489.1.1.

68. Branstetter, M.G., Childers, A.K., Cox-Foster, D., Hopper, K.R., Kapheim, K.M., Toth, A.L., and Worley, K.C. (2018). Genomes of the Hymenoptera. Curr. Opin. Insect Sci. *25*, 65–75. https://doi.org/10.1016/j.cois.2017.11.008.

69. Salden, T., and Peters, R.S. (2023). Afrotropical Ceraphronoidea (Insecta: Hymenoptera) put back on the map with the description of 88 new species. Eur. J. Taxon. *884*, 1–386. https://doi.org/10.5852/ejt.2023.884.2181.

70. Heraty, J.M. (2017). Parasitoid biodiversity and insect pest management. In Insect Biodiversity: Science and Society, Second Edition, R.G. Foottit, and P.H. Adler, eds. (Springer-Verlag), pp. 603–625. https://doi.org/10.1002/9781118945568.ch19.

71. Marshall, C.R. (1990). Confidence intervals on stratigraphic ranges. Paleobiology *16*, 1–10. https://doi.org/10.1017/S0094837300009672.

72. Marshall, C.R. (1990). The fossil record and estimating divergence times between lineages: maximum divergence times and the importance of reliable phylogenies. J. Mol. Evol. *30*, 400–408. https://doi.org/10.1007/BF02101112.

73. Holland, S.M. (2016). The non-uniformity of fossil preservation. Philos. Trans. R. Soc. Lond. B Biol. Sci. *371*, 20150130. https://doi.org/10.1098/rstb.2015.0130.

74. Marshall, C.R. (2019). Using the fossil record to evaluate timetree timescales. Front. Genet. *10*, 1049. https://doi.org/10.3389/fgene.2019.01049.

75. Brocklehurst, N., Upchurch, P., Mannion, P.D., and O'Connor, J. (2012). The completeness of the fossil record of Mesozoic birds: implications

for early avian evolution. PLoS One *7*, e39056. https://doi.org/10.1371/journal.pone.0039056.

76. Montagna, M., Tong, K.J., Magoga, G., Strada, L., Tintori, A., Ho, S.Y.W., and Lo, N. (2019). Recalibration of the insect evolutionary time scale using Monte San Giorgio fossils suggests survival of key lineages through the End-Permian Extinction. Proc. R. Sci. B. *286*, 20191854. https://doi.org/10.1098/rspb.2019.1854.

77. Černý, D., Madzia, D., and Slater, G.J. (2021). Empirical and methodological challenges to the model- based inference of diversification rates in extinct clades. Syst. Biol. *71*, 153–171. https://doi.org/10.1093/sysbio/syab045.

78. Silvestro, D., Schnitzler, J., Liow, L.H., Antonelli, A., and Salamin, N. (2014). Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. Syst. Biol. *63*, 349–367. https://doi.org/10.1093/sysbio/syu006.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Fossil data	This study	Data S1
BBB data	This study	Data S2
Modern species data	This study	Data S3
Output log files	This study	https://doi.org/10.6084/m9.figshare.28271135.v1
Analytical code	This study	https://doi.org/10.6084/m9.figshare.27058510.v1
Additional Tables	This study	https://doi.org/10.6084/m9.figshare.28474061.v1
Additional Figures	This study	https://doi.org/10.6084/m9.figshare.28474070.v1
Software and algorithms		
rootBBB model	Carlisle et al.[7]	https://github.com/dsilvestro/rootBBB
PyRate	Silvestro et al.[5,17]	https://github.com/dsilvestro/PyRate
Tracer 1.7.2	Rambaut et al.[53]	https://github.com/beast-dev/tracer/releases/tag/v1.7.2

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The most recent version of the BBB model can be found at https://github.com/dsilvestro/rootBBB; the most recent version of PyRate can be found at https://github.com/dsilvestro/PyRate. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Fossil record of Hymenoptera

We compiled all species-level fossil occurrences of Hymenoptera from the literature using the Paleobiology Database (PBDB, available at: https://paleobiodb.org/; downloaded the 12 October 2021) and completed or corrected this dataset with occurrences from the Palaeoentomology database of the Laboratory of Arthropods, Palaeontological Institute, Russian Academy of Sciences, Moscow (available at: http://palaeoentomolog.ru/english.html; accessed the 26 July 2023). The dataset obtained from PBDB initially contained 5,103 occurrences for 3,607 species placed in 1,691 genera. The dataset was cleaned of trace fossils, synonyms, outdated combinations, *nomina dubia*, and other erroneous and doubtful records, based on revisions provided in the literature and/or on the expertise of the authors. After correction and data augmentation (*i.e.*, the addition of recently described taxa until September 2023), our first database resulted in an unprecedented dataset (Data S1) composed of 4,330 'species' resulting from an in-depth study and curation of the bibliography on fossil Hymenoptera, and spanning from the Westphalian (Carboniferous) to the Holocene with Copal and Defaunation resin. We consider that the species *Avioxyela gallica* Nel et al. belongs to a clade different from the Hymenoptera.[50,54] After our correction and data augmentation, our dataset of fossil occurrences (*i.e.*, fossil specimens originating from a given stratigraphic horizon assigned to a given clade) encompasses 34,569 occurrences.

Although the vast majority of the taxa included in the datasets are nominal taxa (published and named), we included unnamed taxa (genera or species) that are considered separate from others, although not formally named in the literature yet. These unpublished taxa are identifiable by the notation 'nov.' following their names in the datasets. Some fossil families are known to be poly- or paraphyletic (*e.g.*, Angarosphecidae)[37] and their systematics are not yet resolved. We refrained from splitting these families until comprehensive studies of their limits and constitutive taxa, and we used their classical delineations.

Our approach acknowledges that many occurrences are likely correctly attributed, even though some refinements are necessary. This includes older specimens that have undergone significant revision in recent decades. However, we are also mindful of work known to contain significant errors, such as those stemming from the belief that fossils cannot belong to extant clades—a recurring issue in some palaeontological studies. In our study, we actively avoided such biases. For instance, we excluded most taxa from the Fushun amber described by Hong[55] due to their uncertain placement. This author assigned nearly every fossil insect from Fushun amber to a new genus and species, often without adequately considering their affinities with extant taxa. Furthermore, these classifications have not undergone subsequent revision, further justifying their exclusion. By implementing this mixed approach, we aim to balance the inclusion of valuable data with the exclusion of potentially misleading occurrences.

Fossil species originate from a given stratigraphic horizon, obtained from PBDB, the latter is sometimes corrected for a more precise age (generally stage sometimes substage), and the age of each time bin boundaries relies mostly on the stratigraphic framework proposed in the International Chronostratigraphic Chart v2023/06 (https://stratigraphy.org/chart). Tentative species identifications originally placed with uncertainty (reported as "aff." or "?") were most of the time excluded from our dataset.

Data preparation for estimating clades' origin and extinction times

The Data S2 was constructed by retaining only one occurrence by species, which allows representing the species diversity without biasing the analyses with overrepresented fossil species (*e.g.*, *Ctenobethylus goepperti* (Mayr, 1868)). When a species is found in two deposits of different ages (*i.e.*, not in the same stratigraphic stage), we kept the species in each stage. We then computed the 'mean age of species occurences' by taking into account both the lower and upper boundaries of the life span of every species. This data was then employed to compile the number of species into one-million-year time intervals (Data S2). This approach was adopted to avoid artificially reducing or inflating the age of species when using the lower or upper boundary. This risk is particularly relevant for species classified within a period spanning two geological stages, such as the Langhian-Burdigalian for the Dominican amber, or the Albian-Cenomanian for the Kachin amber. The complete datasets used for each analysis are available in Data S2, for family level analyses (Data S2A), and superfamily level (Data S2B).

Extant species diversity in Hymenoptera families

The Bayesian Brownian Bridge (BBB) model is informed with the present-day diversity of extant groups.[6] To ensure accuracy, we compiled the diversity of extant families and superfamilies while incorporating recent taxonomic revisions. Indeed, many subfamilies of extant apoid, chalcidoid, platygastroid, tenthredinoid lineages have been elevated at a familial rank following recent molecular-based phylogenies.[39,56–60] To account for these changes, we have subtracted the diversity of these former subfamilies from the total diversity of their former family. For example, the Crabronidae *s.l.* (= including Astatidae, Bembicidae, Dinetinae, Eremiaspheciidae, Mellinidae, Pemphredonidae, and Philanthidae) encompass 9,133 species (https://www.calacademy.org/scientists/projects/catalog-of-sphecidae). After the subtraction of the diversity of the subfamilies recently elevated at a familial rank the Crabronidae diversity is now of 4,909 species. For the Chalcidoidea we follow the species count of Noyes[61] even if we are aware that the diversity of this lineage is largely underestimated.[39] For the diversity of extant symphytan lineages, we mainly followed Nyman et al.[62] and compared it with the previous compilation of Aguiar et al.[63] Similarly, we based our Anthophila species number on Ascher and Pickering[64] and compared it with the previous compilation of Aguiar et al.[63] The diversity of Chrysidoidea lineages is based on Agnoli and Rosa[65] for the Chrysididae, Olmi et al.[66] for the Embolemidae, Azevedo et al.[67] for the Bethylidae, and Aguiar et al.[63] for other lineages. We provide a detailed list of the references used to estimate the diversity of extant Hymenoptera families in Data S3. For the superfamily level, we mainly based our diversity on Branstetter et al.[68] Note that the number of species for most extant Hymenoptera lineages is likely underestimated as new discoveries are published on a daily basis.

To assess the impact of this uncertainty, we focused on three superfamilies—Ceraphronoidea, Chalcidoidea, and Ichneumonoidea—whose estimated extant diversity significantly exceeds the currently documented species counts. We performed three sets of analyses: 1) An artificial twofold increase in their extant diversity, using the three variations of the maximum boundary of the uniform prior on the root age (as detailed in *Applying the Bayesian Brownian Bridge to the Hymenoptera Fossil Record*); 2) An artificial fivefold increase in their extant diversity, using the three variations of the maximum boundary of the uniform prior on the root age; 3) Estimates of their 'true' diversity based on published literature, with values of 20,000 species for Ceraphronoidea,[69] 500,000 species for Chalcidoidea,[70] and 100,000 species for Ichneumonoidea[4] (results can be found at https://doi.org/10.6084/m9.figshare.28271135.v1).

The Bayesian Brownian Bridge model

When applied to clades with living descendants, the Bayesian Brownian Bridge (BBB) model[6] operates under the assumption that a clade's diversity evolves akin to a random walk. This walk is represented through a Brownian bridge, anchored at both the starting point (where diversity stands at one) and the extant diversity at present. In essence, the BBB model generates diversity trajectories that emerge from an unknown underlying process of speciation and extinction. These simulated trajectories align with both the fossil record and the present diversity of the clade. Unlike striving to pinpoint a single best-fitting trajectory, this framework encompasses a broad spectrum of plausible diversification histories that stem from random walks. This inclusive approach accommodates various scenarios including linear or exponential diversity growth, fluctuations in species richness, and patterns of waxing and waning.[6]

The BBB model relies on two input datasets: a vector depicting the sampled diversity of the clade through time, which draws from the number of occurrences documented in the fossil record, and the present-day species diversity. The model estimates the clade's origination time (*root_est*) and lower and upper 95% credible intervals (*root_lower*, *root_upper*), the trend parameter (*a_est*, *a_lower*, *a_upper*), the Brownian bridge rate (*sig2_est*, *sig2_lower*, sig2_upper) (which describes how rapidly diversity can change through time), and the sampling over time (*q_est*, *q_lower*, *q_upper*).

The diversity data, presented as a vector, is distributed into predefined time intervals, which, in this study, are fixed at one-million-year spans. This particular duration is chosen to ensure the necessary precision for scrutinizing subfamilies in proximity to the present era. Within our datasets, many of these intervals contained zero sampled species due to the incomplete nature of the fossil record. To manage this, the model employs data augmentation, generating vectors that represent the concealed actual diversity between the estimated point of origin and the present-day diversity.[6] This process is further governed by two conditions: (1) the clade cannot go extinct before the present, even when no fossils are found in the intermediate time intervals, and (2) the assessed true diversity cannot

fall below the recorded diversity in any given time bin. The model is designed to account for increasing sampling rates within the fossil record as it approaches the present,[71–74] as well as the resulting diminished sampling rates close to a clade's origin.[75] This is achieved through the inclusion of a parameter that models an exponential growth in the sampling rate over time, with the extent of this exponential rise being determined by the model itself (*-q_var 1*). While more intricate models of variability in sampling rates are feasible, they might not be applicable to clades with limited fossil records.[6]

Recently, the capacity of the BBB model was broadened to encompass the study of extinct clades.[7] This expansion involved the inclusion of a novel parameter that estimates the unknown time of extinction (*ext_est*) and lower and upper 95% credible intervals (*ext_lower*, *ext_upper*). The estimation of this parameter occurs within the time interval stretching from the latest known fossil occurrence to the present day. Consequently, the model undertakes a simultaneous estimation of various factors, including the times of clade origination and extinction, the variance of the Brownian bridge, and the parameters that gauge the rate of sampling and its fluctuations over time.

Applying the Bayesian Brownian Bridge to the Hymenoptera fossil record

For each clade, we performed three different analyses using alternatively a maximum boundary of uniform prior on the root age set to 281 Ma (*-max_age 281*), 309 Ma (*-max_age 309*), or 345 Ma (*-max_age 345*). The first setting allows for testing an age commonly found with molecular clock approaches[9,10,76] or a total-evidence dating approach.[30] The second is derived from the species *Avioxyela gallica* Nel et al., which belongs to a group (Apatohymenoptera) that may (or may not) be closely related to the Hymenoptera.[50,54] Although this requires further investigation, if the total-group Hymenoptera and the order Apatohymenoptera are determined to be closely related (e.g., as sister groups), the fossil of *Avioxyela* could provide a minimum age for the divergence between these groups. This would suggest that the total-group Hymenoptera may have existed as far back as the early Pennsylvanian. However, the phylogenetic relationships between Apatohymenoptera and other Holometabola must first be rigorously analyzed within a robust phylogenetic framework before drawing definitive conclusions. The third is based on the upper limit of the median age estimates for crown-group Hymenoptera as reported in certain analyses by Montagna et al.,[76] which suggests that the total group Hymenoptera may have been present since the early Mississippian. We set the preservation models with a time-increasing rate (*-q_var 1*), which reflects the increase in species number towards present of most extant clades.[73,74]

Each total group (i.e., stem *plus* crown group) of the 159 families and 31 superfamilies (excluding Chrysidoidea because of the non-monophyly) were analyzed independently using the BBB model run for 1,000,000 Markov chain Monte Carlo (MCMC) iterations and a sampling frequency every 1,000. We examined the results using Tracer 1.7.2[53] to assess the convergence of the runs, quantified through an effective sampling size (ESS) greater than 200 following the procedure in Silvestro et al.[6] Output files for each BBB analyses can be found at https://doi.org/10.6084/m9.figshare.28271135.v1.

Dynamics of origination and extinction through time

We analyzed our fossil occurrence datasets using the Bayesian framework implemented in PyRate 3.[5,17] We analyzed our datasets under two different models: the birth-death model with constrained shifts (BDCS)[16] and the RJMCMC model (*-A 4* option).[17] This bimodal approach allows for confronting the results obtained under two different modes, *i.e.*, estimates of the parameters through time over specific time bins (BDCS) or in a more relaxed framework (RJMCMC), and follows the 'best practices' proposed to investigate diversity dynamics with PyRate.[77] Both models allow for a simultaneous estimate of (i) the parameters of the preservation process (q and α), (ii) the times of origination (*Ts*) and extinction (*Te*) of each taxon, (ii) the origination (λ) and extinction (μ) rates and their variation through time for each stage, and (iv) the number and magnitude of shifts in origination and extinction rates.

To test for parameter sensitivity, we analyzed our datasets with the BDCS model using two different time bins (*-fixShift* option). We constrained the λ and μ rates to be constant within time bins either corresponding to geological epochs (main analysis) or to 20-Myrs-bins (sensitivity analysis).

For all analyses, we tested for the best-fitting preservation process using the option (*-PPmodeltest*), which compares the homogeneous Poisson process (*-mHPP* option), the non-homogeneous Poisson process (default option), and the time-variable Poisson process (*-qShift* option). Under the time-variable Poisson process, preservation rates are assumed to be constant within a predefined time bin but may vary over time (here, set for bins corresponding to stages). This model is thus appropriate when rates are heterogeneous over time.[17]

We ran PyRate for either 100 or 200 million MCMC generations depending on the diversity of the clade, and the MCMC was sampled every 10,000 generations for the BDCS and RJMCMC models. We set the time bins to correspond either to geological stages or to 20-Myrs-bins (*-fixShift* option) for the BDCS model. All analyses were set with the best-fitting preservation process and a Gamma model (*-mG* option), through a gamma-distributed rate heterogeneity with four rate categories to account for variable preservation across taxa.[78] We set the minimal time interval (Myrs) between two shifts (*-min_dt* option) to 2 Myrs to avoid overparameterization in searching for too many shifts. Similarly, we tuned the model to directly estimate the rates of hyperpriors along with all other model parameters. As a result, the rate parameter is directly estimated from the data, using ''0'' in the (*-pP*), (*-pL*), and (*-pM*) options.

We monitored chain mixing and ESS by examining the MCMC log files in Tracer 1.7.2[53] after excluding the first 10% of the samples as a burn-in. The parameters are considered as converged when their ESSs are greater than 200. Finally, we computed the posterior mean and 95% credibility intervals of all parameters (λ_i, μ_i, and q_i rates, as well as all *Ts* and *Te* for the taxa included in each dataset). Output files for each PyRate analyses can be found at https://doi.org/10.6084/m9.figshare.28271135.v1.

QUANTIFICATION AND STATISTICAL ANALYSIS

BBB results were examined using Tracer 1.7.2[53] to assess the convergence of the runs, quantified through an effective sampling size (ESS) greater than 200 following the procedure in Silvestro et al.[6] For PyRate analyses, we monitored chain mixing and ESS by examining the MCMC log files in Tracer 1.7.2[53] after excluding the first 10% of the samples as a burn-in. The parameters are considered as converged when their ESSs are greater than 200. Finally, we computed the posterior mean and 95% credibility intervals of all parameters (λ_i, μ_i, and q_i rates, as well as all *Ts* and *Te* for the taxa included in each dataset) using the *-combLogRJ* function implemented in PyRate.