Supporting Information

Cardinal et al. 10.1073/pnas.1006299107

SI Materials and Methods

Partitioning of Data. The dataset was analyzed under various partitioning schemes. The model-based analyses were first done with the data partitioned by gene (seven partitions). However, examination of the parameter trace files after 50 million generations revealed that some parameters of the 18S partition had not converged on a stationary phase. Therefore, it seemed that the 18S partition did not contain enough data to adequately estimate all parameters in the model. To reduce the number of partitions, subsequent analyses combined the two ribosomal genes into one partition and kept the protein-coding genes separately partitioned or combined together into one partition. The dataset was also analyzed with the two ribosomal genes in one partition, and the protein-coding genes were combined together and partitioned by codon position. We did not partition each protein-coding gene separately by codon position, because this would have produced a 17-partition analyses. Such an analysis would have been computationally prohibitive and would have created some partitions with very little information for parameter estimation.

Model Selection. Selection of best-fit models of nucleotide substitution for each data partition used in a Bayesian or Maximum Likelihood (ML) analysis was based on the Akaike (AIC) and Bayesian Information Criteria (BIC) as implemented in JModelTest v.0.1.1 (1). Likelihood calculations were carried out for 88 models, which included 11 substitution schemes [of which the general time-reversible model (GTR) was the most complex], equal or unequal base frequencies, a proportion of invariable sites (I), and rate variation among sites with four rate categories (G) on an ML-optimized tree. Model-selection uncertainty was quantified using the AIC and BIC δ values (the difference in AIC or BIC score between the model with the smallest score and the model in question); values within 2 δ units of the best model were considered to have substantial support and therefore, considered. The AIC and BIC weights were used to approximate a 95% confidence set of models by summing the weights of the ranked models until the sum was 0.95 (as suggested in the manual). Any model that was part of that set of models was considered to have substantial support and was also considered. If a model selected under the above criteria could not be implemented in MrBayes, the least complex model that included all of the parameters of the selected model, but could be implemented in MrBayes, it was used instead.

Parsimony Analyses. We used Winclada v.10.00.08 (2) to remove parsimony uninformative sites and analyzed the concatenated dataset in Tree analysis using New Technology (TNT) (3). A traditional heuristic search was conducted first with tree bisection reconnection (TBR), 1,000 random taxon-addition replicates holding 10 trees per replicate, and sequence indels treated as missing data. Several new technology searches using Ratchet and Drift options were also done. The minimum number of steps found in these previous searches was used as the target for a new technology-driven search. This analysis was run until trees with this number of minimum steps were hit 1,550 times. The initial sequence additions were set to 10. Support for the branches in the strict consensus of all of the equally parsimonious trees found was assessed with 1,000 bootstrap pseudoreplicates. Each resampled matrix was searched 10 times using a combination of drifting, tree-fusing, and random sectorial searches, and the consensus tree of each iteration was saved. The 1,000 consensus trees from the resampled matrices were opened in Winclada for calculation of nodal support of the strict consensus of the most parsimonious trees.

ML Analysis. ML analyses were done in the program RaxML-VI-HPC (4) with the concatenated dataset partitioned into four partitions: (*i*) ribosomal genes, (*ii*) first codon positions, (*iii*) second codon positions, and (*iv*) third codon positions. We used RaxML's rapid bootstrap algorithm, which uses the GTR+CAT approximation to conduct 1,000 bootstrap replicates. Every fifth bootstrap tree was then used as a starting point to search for the highest-scoring ML tree under GTR+I+G for a total of 200 ML searches. The bootstrap support values were shown on the ML tree.

Bayesian Analysis. The Bayesian analyses were done in MrBayes v.3.1.2 (5). Each gene was individually analyzed under one or two different models according to the results of the model tests described above. The individual gene datasets were analyzed with two independent runs with four chains each for ∼20 million generations. The parameter trace files of each run were observed to verify that the runs had converged on the stationary distribution and to decide on the appropriate number of generations to discard as burn-in. Convergence was also assessed by observing the posterior probabilities of clades over generations in the program AWTY (6).

The concatenated dataset was analyzed under the different partitioning schemes described above. For the concatenated analyses, all partitions used the GTR+I+G model, with each parameter unlinked across partitions. Two individual runs of 20–50 million generations with four chains each were conducted for the variously partitioned concatenated dataset (|28S|18S|Ops|Win|Pol|Nak|EF-1α|, |28S,18S|Ops|Win|Pol|Nak|EF-1α|, and |28S|18S|pos1|pos2|pos3|). Our preferred partitioning scheme, where the two ribosomal genes are combined into one partition and the protein-coding genes are combined together and partitioned by codon position (|28S,18S|pos1|pos2|pos3|), was much more thoroughly analyzed. A total of 46 independent runs, 44 with four chains and 2 with eight chains, were conducted. The number of generations for each run varied from 4,096,000 to 23,353,000 generations. The tree files and parameter files with the burn-in removed from each run that had reached convergence were combined. Runs that had not converged were discarded. A maximum clade credibility tree was constructed from these 268,560,000 postburn-in generations in TreeAnnotator v1.4.8 (7). We chose to do numerous shorter runs instead of a few longer runs because of run-time restrictions on the computer clusters at Cornell University's Computational Biology Service Unit. Also, doing a large number of independent runs from different starting points allowed us to more fully explore tree space.

Divergence Time Estimates. We describe below the paleontological evidence on which our 10 calibration points are based.

Calibration 1. The fossil bee *Euglossa morenei*, which is from Miocene Dominican amber, was described and placed within *Euglossa* based on morphological characters (8). Several characters, however, differentiated this species from any of the described subgenera, suggesting that it represents a lineage of euglossine bees that is no longer extant (8). We, therefore, placed this fossil as part of the stem group of the genus *Euglossa*. Based on biostratigraphic and paleogeographic data, Dominican amber has been dated to be of late Early Miocene through early Middle Miocene (15–20 Mya) age (9). A lognormal distribution with a lognormal mean of 3 (20 Mya), SD of 1, and 0 offset of 11.6 Mya was applied as a prior for calibration point 1.

Note that *Eufriesea melissiflora* was first described under the name *Paleoeuglossa melissiflora* (10) but was later transferred to

Eufriesea (8). This fossil cannot be placed within the crown group of *Eufriesea* with confidence but only its stem group. It is also from Dominican amber of the same single sedimentary basin as *E. morenei.* Therefore, this fossil was not used to inform a calibration point, because *E. morenei* already informs a calibration node nearer the tips of the tree.

Calibration 2. *Apis lithohermaea* (11) is the oldest fossil record for crown-group *Apis*. The fossil is from the Chôjabaru Formation of Iki Island, Japan, which has been dated to be from the middle Miocene (11). Based on some key morphological characters, it was assigned to the *dorsata* species group and was described as being quite similar to modern *A. dorsata* (11). This fossil, therefore, provides a minimum age for the diversification of *A. dorsata* from *A. cerana*, a member of the mellifera species group. A lognormal distribution with a lognormal mean of 2.8 (17 Mya), SD of 1, and 0 offset of 11.6 Mya was applied as a prior for calibration point 2.

Calibration 3. There are several fossil honey-bee species that are thought to belong to the stem group of *Apis* (12). The oldest of these fossils is a compression fossil of *A. henshawi* from the Oligocene, Rott, Germany. The age of this famous European shale deposit is somewhat ambiguous, being either late Oligocene (Chattian) or early Miocene (Aquitanian) age (13). A lognormal distribution with a lognormal mean of 3.5 (33 Mya), SD of 1, and 0 offset of 20.4 Mya was applied as a prior for calibration point 3.

Note that the two fossil bee species of the armbrusteri species group, *A. armbrusteri* and *A. nearctica*, are thought to share the most recent common ancestor with extant *Apis* species (11, 14). The type specimen of *A. armbrusteri* was described from worker bees preserved on a slab of thermal limestone found near Böttingen in the Swabian Alb, Württemberg, Germany (11). The Randecker Maar is thought to be from the Early Miocene (13). *A. nearctica* was recently described from a worker preserved in paper shale from the Middle Miocene (Late Barstovian; 14–14.5 Mya) of the Stewart Valley Basin in west-central Nevada (14). Potassium-argon dating indicates an age of 16–10.5 Mya (13).

There have been a few reliable *Bombus* compression fossils described from the Miocene. These are *Bombus vetustus*, *B. luianus*, and *B. proavus.* However, these fossils do not reveal sufficient morphological synapomorphies of *Bombus* to be placed within the crown group of *Bombus* instead of the stem group (15). Grimaldi and Engel (13) mention the existence of various European Oligocene *Bombus* spp. but give no citation for these. We could not find any described *Bombus* fossils from the Oligocene that are still believed to belong to *Bombus*, and therefore, we assume that these fossils have not yet been described; we do not know if they would belong to the stem or crown group of *Bombus*. Therefore, the *Bombus* fossils could only be used as a minimum age on the node uniting Bombini with Meliponini. However, there have been much older fossils described that could be placed as a minimum age for this node, and therefore, we do not use any of the *Bombus* fossils in this analysis.

Calibration 4. *Kelneriapis eocenica* is a fossil bee from the Baltic amber thought to be most similar to, and perhaps sister to, *Hypotrigona* (16). Therefore, this fossil can be used to set a minimum age on the node of the most recent common ancestor of *Hypotrigona* and its sister lineage in the phylogeny. *Liotrigonopsis rozeni* is also a fossil bee from the Baltic amber and is thought to be morphologically similar to the extant genus *Liotrigona* (16), which is hypothesized to be the sister genus to *Hypotrigona* (17). *Liotrigona* was not included in this analysis, however, but this fossil also serves to inform a minimum age on the split of *Hypotrigona* from other extant stingless bees. K/Ar radiometric studies of the Baltic amber, which occurs in the blau Erde (blue Earth) throughout northern Europe, indicate it to be 44.1 ± 1.1 Mya (middle of the Lutetian stage of the Eocene) (18). A lognormal distribution with a lognormal mean of 3.85 (47 Mya), SD of 1, and 0 offset of 33.9 Mya was applied as a prior for calibration point 4.

Note that *Proplebeia dominicana*, *P. tantilla*, and *P. vetusta* are fossil bees from Dominican amber. A study of the extinct fauna of stingless bees in Dominican Amber corroborated the hypothesis that *Proplebeia* is a distinct group within the lineage of Neotropical *Plebeia* (s.s) (19). *Plebeia* was not included in this phylogeny but has been placed within a large Neotropical clade of stingless bees (17), which is represented in this analysis by the most recent common ancestor of *Trigona* and *Melipona*. Therefore, a minimum age of 15–20 Mya (*Calibration 1* discusses age of Dominican amber) could be placed on this node, but calibration point 4 already provides a much older age estimate for this node.

Calibration 5. *Cretotrigona prisca* is the oldest crown-group bee fossil recorded. It is from the Late Cretaceous amber of New Jersey (20). It was first placed within *Trigona* (20) but was later placed sister to the African genus *Dactylurina* (21). These two genera superficially resemble each other but are not phylogenetically close to one another (17). Therefore, this fossil can only be used as a minimum age on the node uniting all of the extant stingless bees.

The age of this fossil has been debated. It was first thought to be Campanian (ca. 80 Mya) in origin based on chemical comparison with other New Jersey ambers (20). However, based on the other insects found in this amber, it was suggested that it was of Tertiary age—probably from the Paleocene (22). It was once again argued that it was of Cretaceous age based on infrared spectroscopy and pyrolysis gas chromatography as well as the presence of other insects in the amber fragment dated to the Turonian (23). This fossil is usually now treated as being of Late Maastrichtian (ca. 65–70 Mya) age (21). A lognormal distribution with a lognormal mean of 4.25 (70 Mya), SD of 1, and 0 offset of 55 Mya was applied as a prior for calibration point 5.

Note that there are numerous fossils from the Baltic amber that have been placed within the corbiculate bees (16). Many of these do not fall within extant tribes but have been instead assigned to their own fossil tribes. Electrobombini, Electrapini, and Milikertini are corbiculate tribes containing only fossil bees. In a cladistic analysis with representatives of extant corbiculates and a few outgroups, the phylogenetic relationships among and within these tribes were analyzed (16). However, because of the difference in topology of the corbiculate tribes between most morphological and molecular studies described in the paper, these fossils can at best be used as a minimum age for the common ancestor of corbiculates. Pending a clearer resolution of extant and fossil corbiculate relationships, these fossil corbiculates provide limited information on the age of the extant groups. However, *C. prisca* is also placed within the corbiculates and is of much older age.

Calibration 6. There have been three fossil bee species described from the Baltic amber that have been placed in the fossil bee tribe Boreallodapini. Boreallodapini most resembles the extant Ceratinini and Allodapini (Xylocopinae). A phylogenetic analysis placed it as sister to the Allodapini, with Ceratinini being sister to Boreallodapini + Allodapini (16). These tribal relationships are in agreement with those found in this present phylogeny of extant taxa only, and therefore, these fossils were used to inform a minimum age for the node uniting Allodapini and Ceratinini (*Calibration 1* discusses the age of the Baltic amber). A lognormal distribution with a lognormal mean of 3.85 (47 Mya), SD of 1, and 0 offset of 33.9 Mya was applied as a prior for calibration point 6.

Calibration 7. *Paleohabropoda oudardi* from the lacustrine shales of Menat in France has been placed within the tribe Anthophorini based on a morphological cladistic analysis and wing morphometrics (24). Evidence from pollen and the mammals also found in Menat suggest it is of Paleocene age. K/Ar analysis proposes a date of ~56 Mya for the Menat fossils (25). This fossil is used to inform the calibration of the node uniting the Anthophorini. A lognormal distribution with a lognormal mean of 4.09 (60 Mya), SD of 1, and 0 offset of 23 Mya was applied as a prior for calibration point 7.

Calibration 8. *Xylocopa gabrielae* was described from a female bee preserved as a compression fossil from the Eocene–Oligocene boundary (34 Mya) of Florissant, Colorado (26). The combination of characters present in the fossil indicates that it is a species of *Xylocopa*; however, because of a general lack of preserved features, it is not possible to assign it to any recognizable subgenera (26). Therefore, there is the possibility that this fossil is part of the stem, rather than crown, group of *Xylocopa*. However, this age seems reasonable for the node of extant xylocopines, because there is a Miocene fossil from China closely resembling extant *Xylocopa (Nyctomelitta)* (27) and fossils from Switzerland of Middle Oligocene age (28) most likely related to *Xylocopa s.s.* and *Copoxyla* (29). A lognormal distribution with a lognormal mean of 3.53 (34 Mya), SD of 1, and 0 offset of 23 Mya was applied as a prior for calibration point 8.

Calibration 9. The oldest fossil of a melittid bee is *Paleomacropis eocenicus* from the early Eocene amber of Oise (France) (30). In a cladistic analysis, it formed the sister group to *Macropis* with another fossil bee *Eomacropis glaesaria*, falling as the sister to *Macropis* + *Paleomacropis* (30). *E. glaesaria* is from Baltic amber of the late Eocene (16). The Paris Basin amber of Oise, France, is thought to be of Ypresian age (13). This fossil is used to calibrate the node representing the common ancestor of *Macropis* and its sister group. A lognormal distribution with a lognormal mean of 4.01 (55 Mya), SD of 1, and 0 offset of 40.4 Mya was applied as a prior for calibration point 9.

Calibration 10. *Probombus hirsutus* was first described as a bumblebee. However, after doubts were placed on this (22), it was reexamined and transferred to the Megachilidae (31). The fossil is thought to be part of the Megachilinae, probably related to Osmiini or Megachilini (31), which together form a monophyletic group. We, therefore, use this fossil to calibrate the node uniting the Osmiini, Megachilini, and Anthidiini. This fossil is from a spongodiatomitic volcanic paleolake (maar) deposit in Menat, Puy-de-Dôme, France. K/Ar analysis proposes a date of ~56 Mya for the Menat fossils (25). A lognormal distribution with a lognormal mean of 4.09 (60 Mya), SD of 1, and 0 offset of 48.6 Mya was applied as a prior for calibration point 10.

Note that *Protolithurgus ditomeus* was described from Baltic amber (16). In a cladistic analysis, it was placed sister to all other lithurgines because of the plesiomorphic retention of some characters (16). This fossil can, therefore, only be used to calibrate the node uniting Lithurgini with the other Megachilinae.

Four fossil species described from the Baltic amber have been assigned to the fossil genus *Glyptapis* (16). The phylogenetic position of *Glyptapis* is uncertain and may possibly be more closely allied to Anthidiini than other Osmiini. Therefore, these fossils can tentatively be used as a minimum age on the node representing the most recent common ancestor of Anthidiini and extant Osmiini, which is the same node used for calibration point 10.

Four fossil species described from the Baltic amber and one from Rovno Amber (of roughly contemporaneous age) have been assigned to the fossil genus *Ctenoplectrella* of the subtribe Ctenoplectrellina (32, 33). *Glaesosmia genalis* has also been described from a fossil of the Baltic amber (16). A third genus was recently added to the tribe Ctenoplectrellini or subtribe Ctenoplectrellina. The type species, *Friccomelissa schopowi* (34), is described from a fossil recovered from a Messel pit near Darmstadt, Hesse, Germany. The Messel Formation is from the lower mid-Eocene, Geiseltalian, ca. 47 Mya (34). *Friccomelissa* might represent a stem group to Ctenoplectrellini (34). Ctenoplectrellina has been placed in the tribe Osmiini, which might be paraphyletic with respect to the Megachilini (16). Ctenoplectrellina is now its own tribe (Ctenoplectrellini) (34) and may be sister to extant Osmiini + Megachilini; therefore, it cannot provide a minimum age to the most recent common ancestor of Osmiini + Megachilini but to one node back from there, which is where calibration point 10 is placed.

1. Posada D (2008) jModelTest: Phylogenetic model averaging. *Mol Biol Evol* 25:1253–1256.
2. Nixon KC (1999) Winclada (BETA) ver. 0.9. 9. (K. C. Nixon, Ithaca, NY).
3. Goloboff PA, Farris JS, Nixon KC (2008) TNT, a free program for phylogenetic analysis. *Cladistics* 24:774–786.
4. Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
5. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
6. Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL (2008) AWTY (are we there yet?): A system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24:581–583.
7. Rambaut A, Drummond AJ (2008) TreeAnnotator v1.4.8. Available at http://tree.bio.ed.ac.uk/software/treeannotator/.
8. Engel MS (1999) The first fossil *Euglossa* and phylogeny of the orchid bees (Hymenoptera: Apidae; Euglossini). *Am Mus Novit* 3272:1–14.
9. Iturralde-Vinent MA, MacPhee RDE (1996) Age and paleogeographical origin of Dominican amber. *Science* 273:1850–1852.
10. Poinar G, Jr (1998) *Paleoeuglossa melissiflora* gen. n., sp. n.(Euglossinae: Apidae), fossil orchid bees in Dominican amber. *J Kans Entomol Soc* 71:29–34.
11. Engel MS (2006) A giant honey bee from the middle Miocene of Japan (Hymenoptera: Apidae). *Am Mus Novit* 3504:1–12.
12. Engel MS (1998) Fossil honey bees and evolution in the genus *Apis* (Hymenoptera: Apidae). *Apidologie* 29:265–281.
13. Grimaldi DA, Engel MS (2005) *Evolution of the Insects* (Cambridge Univ Press, New York).
14. Engel MS, Hinojosa-Díaz IA, Rasnitsyn AP (2009) A honey bee from the Miocene of Nevada and the biogeography of *Apis* (Hymenoptera: Apidae: Apini). *Proc Calif Acad Sci* 60:23–38.
15. Hines HM (2008) Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). *Syst Biol* 57:58–75.
16. Engel MS (2001) A monograph of the Baltic amber bees and evolution of the Apoidea (Hymenoptera). *B Am Mus Nat Hist* 259:1–192.
17. Rasmussen C, Cameron SA (2007) A molecular phylogeny of the old world stingless bees (Hymenoptera: Apidae: Meliponini) and the non-monophyly of the large genus *Trigona*. *Syst Entomol* 32:26–39.
18. Ritzkowski S (1997) K-ar-altersbestimmungen der bernsteinführenden sedimente des samlandes (paläogen, bezirk kaliningrad). *Metalla* 66:19–23.
19. Camargo JMF, Grimaldi DA, Pedro SRM (2000) The extinct fauna of stingless bees (Hymenoptera: Apidae: Meliponini) in Dominican amber: Two new species and redescription of the male of *Proplebeia dominicana* (Wille and Chandler). *Am Mus Novit* 3293:1–24.
20. Michener CD, Grimaldi DA (1988) A *Trigona* from late cretaceous amber of New Jersey (Hymenoptera: Apidae: Meliponinae). *Am Mus Novit* 2917:1–10.
21. Engel MS (2000) A new interpretation of the oldest fossil bee (Hymenoptera: Apidae). *Am Mus Novit* 3296:1–11.
22. Rasnitsyn A, Michener C (1991) Miocene fossil bumble bee from the Soviet Far East with comments on the chronology and distribution of fossil bees (Hymenoptera: Apidae). *Ann Entomol Soc Am* 84:583–589.
23. Grimaldi D (1999) The co-radiations of pollinating insects and angiosperms in the Cretaceous. *Ann Mo Bot Gard* 86:373–406.
24. Michez D, De Meulemeester T, Rasmont P, Nel A, Patiny S (2009) New fossil evidence of the early diversification of bees: *Paleohabropoda oudardi* from the French Paleocene (Hymenoptera, Apidae, Anthophorini). *Zool Scr* 38:171–181.
25. Russel DE (1982) The geology of the Menat basin, France. *Palaeontographica* 182: 87–150.
26. Engel MS (2001) The first large carpenter bee from the Tertiary of North America, with a consideration of the geological history of Xylocopinae (Hymenoptera: Apidae). *T Am Entomol Soc* 127:245–254.
27. Zhang JF (1990) New fossil species of Apoidea (Insecta: Hymenoptera). *Acta Zootaxonomica Sinica* 15:83–91.
28. Zeuner FE, Manning FJ (1976) A monograph on fossil bees (Hymenoptera: Apoidea). *Bull Br Mus (Nat Hist) Geol* 27:149–268.
29. Leys R, Cooper SJB, Schwarz MP (2000) Molecular phylogeny of the large carpenter bees, genus *Xylocopa* (Hymenoptera: apidae), based on mitochondrial DNA sequences. *Mol Phylogenet Evol* 17:407–418.
30. Michez D, Nel A, Menier JJ, Rasmont P (2007) The oldest fossil of a melittid bee (Hymenoptera: Apiformes) from the early Eocene of Oise (France). *Zool J Linn Soc* 150: 701–709.
31. Nel A, Petrulevicius JF (2003) New Palaeogene bees from Europe and Asia. *Alcheringa* 27:277–293.
32. Engel MS (2008) A new species of *Ctenoplectrella* in Baltic amber (Hymenoptera: Megachilidae). *Acta Zool Hung* 54:319–324.
33. Engel MS, Perkovsky EE (2006) An Eocene bee in Rovno amber, Ukraine (Hymenoptera: Megachilidae). *Am Mus Novit* 3506:1–11.
34. Wedmann S, Wappler T, Engel MS (2009) Direct and indirect fossil records of megachilid bees from the Paleogene of Central Europe (Hymenoptera: Megachilidae). *Naturwissenschaften* 96:703–712.

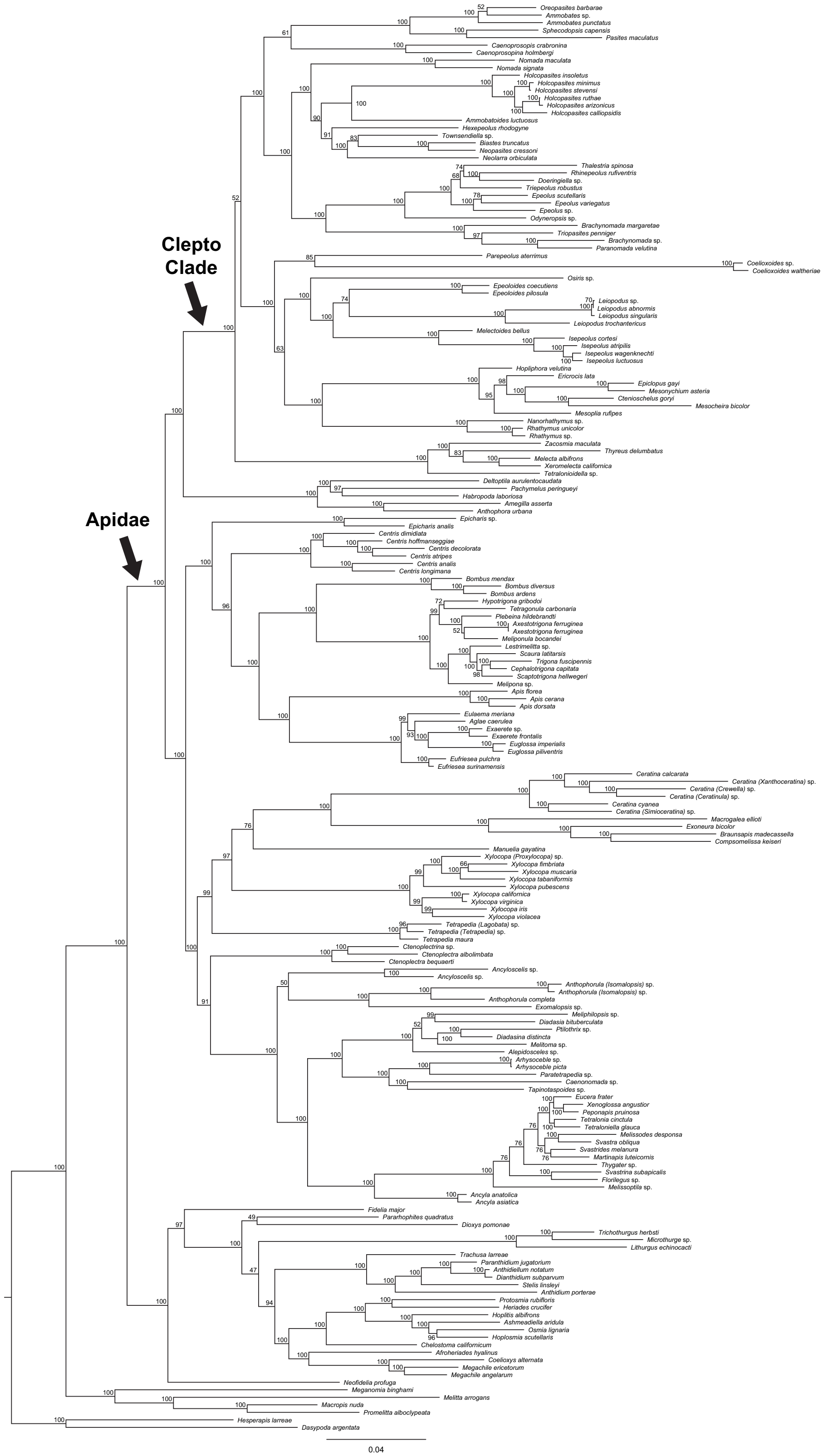

Fig. S1. Bayesian maximum clade credibility tree of the concatenated dataset showing relationships among apid bees and outgroups. Posterior probabilities are shown on the left of each node.

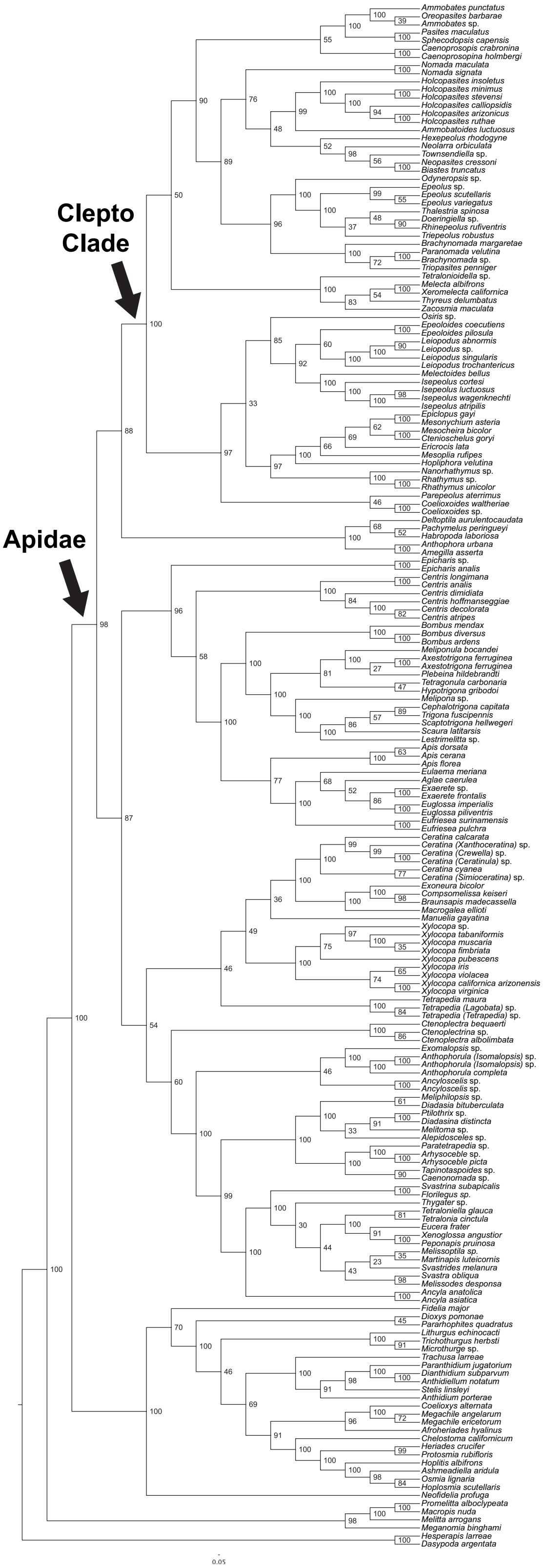

Fig. S2. ML clade credibility tree of the concatenated dataset showing relationships among apid bees and outgroups. Bootstrap support values are shown in percent values on the right of each node.

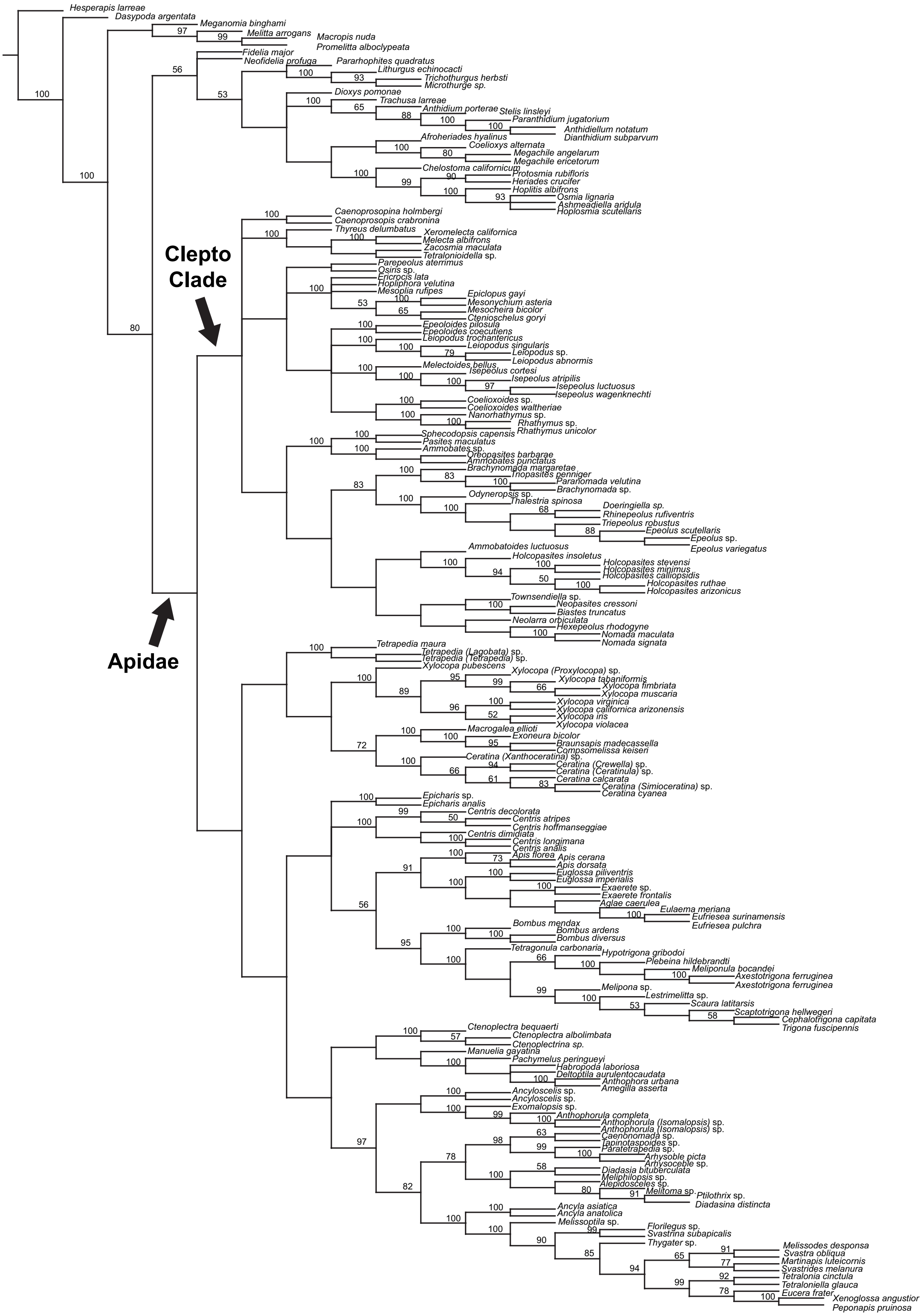

Fig. S3. Strict consensus of 32 equally parsimonious trees of the concatenated dataset showing relationships among apid bees and outgroups. Bootstrap support values are shown in percent values on the left of each node.

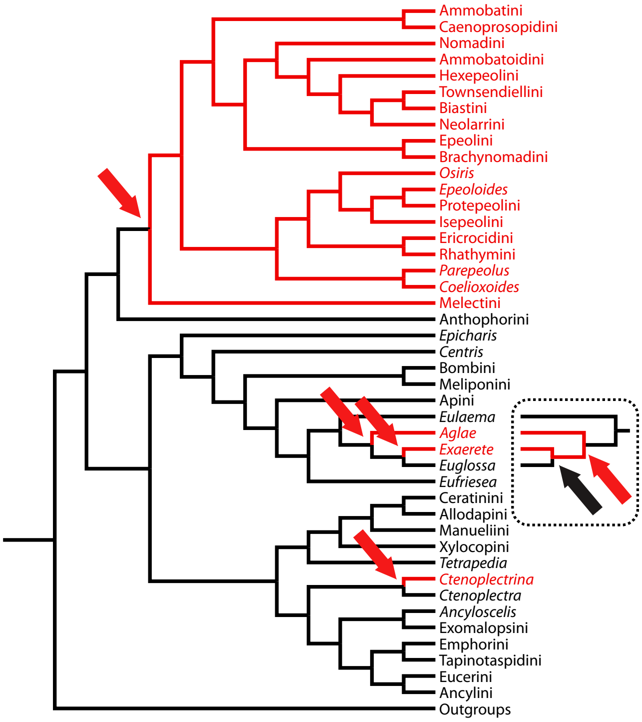

Fig. S4. Pruned version of the Bayesian maximum clade credibility tree of the concatenated dataset. Alternative equally parsimonious reconstructions of cleptoparasitism are mapped onto the tree. Cleptoparasitic taxa are shown in red, and nest-making taxa are in black. The main tree shows the hypothesis of four independent origins of cleptoparasitism, whereas *Inset* shows the reconstruction within orchid bees when three independent origins with one reversal back to nest making are hypothesized.

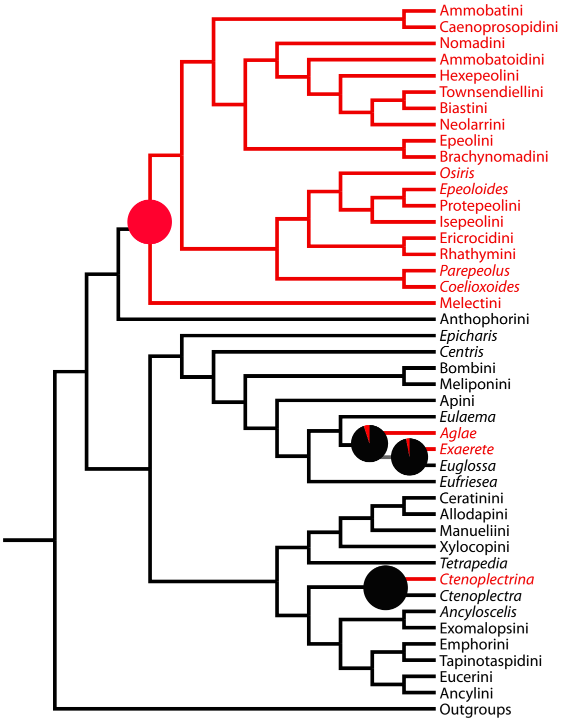

Fig. S5. Pruned version of the Bayesian maximum clade credibility tree of the concatenated dataset. Cleptoparasitic taxa are shown in red, and nest-making taxa are in black. Cleptoparasitism is mapped onto the tree according to the results of our model-based ancestral state reconstruction. Pie charts show the probability of the ancestral state being cleptoparasitic or nest making at four key nodes in the tree.

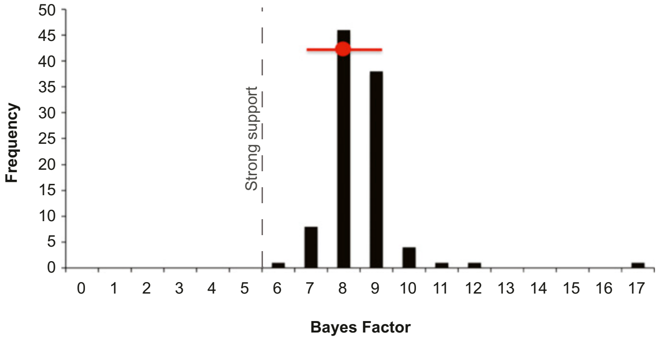

Fig. S6. Distribution, mean, and variance of the 100 Bayes Factor tests when the node uniting *Euglossa*, *Exaerete*, and *Aglae* was constrained to be nest making versus cleptoparasitic. Values above 0 indicate support for a nest-making ancestor, and those above 6 are considered to be strong support for this hypothesis.

Other Supporting Information Files

Table S1 (DOC)
Table S2 (DOC)
Table S3 (DOC)
Table S4 (DOC)