

UNIVERSIDADE FEDERAL DO PARANÁ

BRUNNO BUENO DA ROSA

FILOGENIA E TEMPO DE DIVERGÊNCIA DAS LINHAGENS DE
VESPAS APOIDEAS E ABELHAS COM BASE NO RELÓGIO
MORFOLÓGICO EM UMA ABORDAGEM BAYESIANA DOS GRUPOS
FÓSSEIS E RECENTES

CURITIBA

2019

FILOGENIA E TEMPO DE DIVERGÊNCIA DAS LINHAGENS DE VESPAS
APOIDEAS E ABELHAS COM BASE NO RELÓGIO MORFOLÓGICO EM UMA
ABORDAGEM BAYESIANA DOS GRUPOS FÓSSEIS E RECENTES

Tese apresentada à Coordenação do curso de Pós-Graduação em Ciências Biológicas, Área de Concentração em Entomologia, do Setor de Ciências Biológicas, Departamento de Zoologia da Universidade Federal do Paraná, como requisito parcial para obtenção de título de Doutor em Ciências Biológicas.

Orientador: Prof. Dr. Gabriel A. R. Melo
Co-orientador: Prof. Dr. Marcos S. Barbeitos

Universidade Federal do Paraná. Sistema de Bibliotecas.
Biblioteca de Ciências Biológicas.
(Giana Mara Seniski Silva – CRB/9 1406)

Rosa, Brunno Bueno da

Filogenia e tempo de divergência das linhagens de vespas apoideas e abelhas com base no relógio morfológico em uma abordagem bayesiana dos grupos fósseis e recentes. / Brunno Bueno da Rosa. – Curitiba, 2019. 111 + [15] p. : il.

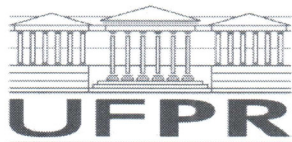
Orientador: Gabriel A. R. Melo

Coorientador: Marcos S. Barbeitos

Tese (doutorado) - Universidade Federal do Paraná, Setor de Ciências Biológicas. Programa de Pós-Graduação em Entomologia

1. Abelha. 2. Vespa. 3. Inseto – Filogenia. 4. Morfologia (Biologia) I. Título II. Melo, Gabriel Augusto Rodrigues de. III. Barbeitos, Marcos Soares IV. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Entomologia.

CDD (22. ed.) 595.79



MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS BIOLÓGICAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO CIÊNCIAS BIOLÓGICAS
(ENTOMOLOGIA) - 40001016005P5

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em CIÊNCIAS BIOLÓGICAS (ENTOMOLOGIA) da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **BRUNNO BUENO DA ROSA** intitulada: **Filogenia e tempo de divergência das linhagens de vespas apoideas e abelhas com base no relógio morfológico em uma abordagem bayesiana dos grupos fósseis e recentes**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 28 de Fevereiro de 2019.


GABRIEL AUGUSTO RODRIGUES DE MELO
Presidente da Banca Examinadora (UFPR)


LUIZ ROBERTO RIBEIRO FARIA JUNIOR
Avaliador Externo (UNILA)


EDUARDO ANDRADE BOTELHO DE ALMEIDA
Avaliador Externo (USP)


JOHN EDWIN LATTKE BRAVO
Avaliador Interno (UFPR)



Reconstrução artística de †*Burmasphex* sp. (Melo & Rosa em preparação).

“As abelhas são vespas [...] cujas fêmeas, em vez de capturarem outros artrópodes como alimento, coletam pólen e néctar diretamente nas flores para alimentarem suas larvas.”

G. A. R. Melo

Abelhas brasileiras (Silveira et al. 2002)

Este trabalho é dedicado a minha mãe Elisete Bueno e ao meu irmão
Matheus Bueno da Rosa.

AGRADECIMENTOS

A minha família pelo apoio conspícuo nas escolhas da minha vida. Essencialmente sou grato a minha mãe Elisete Bueno e meu irmão Matheus Bueno da Rosa, os quais amo acima de tudo. Agradeço também a Ana Gabriela Vantini Braga ao apoio e companheirismo durante esses anos.

Aos meus orientadores, Dr Gabriel A. R. Melo e Dr. Marcos S. Barbeitos pela amizade e sabedoria compartilhada durante estes anos.

Ao Dr. James Carpenter pela amizade, apoio durante o intercâmbio no American Museum of Natural History e empréstimo valioso de exemplares para a realização desse trabalho.

Ao Dr. Wojciech Pulawski pela amizade, apoio e empréstimo valioso de exemplares para a realização desse trabalho.

Ao Dr. David Grimaldi e ao Sr. Paul Nascimbene pelas preparações e suporte na coleção de fósseis do American Museum of Natural History.

A André Martins, Ândrio Zafalon Silva, David Muniz, David Luz e Thayrine Martins pela amizade e contribuições diretas a este trabalho ao longo dos anos de doutoramento.

Aos membros da banca, Dr. Claudio Carvalho, Dr. Eduardo Almeida, Dr. Luiz Roberto Faria Jr. e Dr. John Lattke.

Aos coletores, curadores e instituições que contribuíram com esse trabalho.

Ao PPG em Entomologia e professores que contribuíram com minha formação.

Ao CNPq pelo apoio financeiro durante todo o doutorado.

A CAPES pelo apoio financeiro durante o intercâmbio no American Museum of Natural History.

Grato!

RESUMO

Os Apoidea incluem, além das abelhas, uma grande diversidade de vespas predadoras, atualmente referidas informalmente como vespas apoideas. Os Apoidea formam um dos grupos mais diversos em Aculeata com mais de 27500 espécies descritas. Em suas relações filogenéticas, os maiores conflitos de relacionamentos estão no posicionamento das linhagens de Crabronidae *lato sensu*. O registro fóssil de Apoidea é muito rico contando com exemplares preservados tanto em compressões rochosas quanto em inclusões em âmbar. Além da grande diversidade de linhagens, esses fósseis estão distribuídos em uma ampla escala temporal contemplando um intervalo entre 10 a 130 Ma. As estimativas de idade para o grupo *crown* de Apoidea variam entre o Jurássico (185.8-153 Ma) e o Cretáceo inferior (~135 Ma). Mas as grandes linhagens de Apoidea já haviam se diferenciado até o final do Cretáceo inferior. O relógio morfológico é uma maneira de se calcular as estimativas do tempo de divergência entre linhagens. O método considera simultaneamente a informação temporal e a informação proveniente dos caracteres morfológicos para co-estimar a filogenia, as taxas de evolução e o tempo de divergência entre as linhagens. As análises de estimativas de diversificações de linhagens ao longo do tempo (LTTs, *Lineage Through Time plots*) são poderosas ferramentas macroevolutivas que utilizam filogenias datadas para fazer aproximações de padrões macroevolutivos ligados ao surgimento de espécies. Diversos autores argumentam que uma amostragem sem o registro fóssil é um panorama macroevolutivo pouco realista e apresenta um grande viés na estimativa de eventos de extinção. A utilização de um amplo registro fóssil em filogenias datadas como os de Apoidea permitem melhores aproximações para padrões macroevolutivos. As relações filogenéticas de Apoidea foram investigadas em uma abordagem Bayesiana aplicada a dados morfológicos. Foram incluídos 99 terminais viventes e 32 fósseis. As estimativas dos tempos de divergência das linhagens foram realizadas por relógio morfológico. Foram comparados 20 diferentes modelos de relógio morfológico através de valores de verossimilhança marginal. Os padrões macroevolutivos ligados às diversificações das grandes linhagens de Apoidea foram investigadas através de LTTs. No total, foram recuperadas pelo menos 17 linhagens independentes de Apoidea, sendo 11 linhagens viventes e seis extintas. O relógio morfológico foi eficiente para a estimativa do tempo de divergência dos Apoidea, entretanto, ainda existem desafios para entender e aprimorar suas aplicações. O grupo *stem* de Apoidea divergiu no final do Jurássico, 156.3 Ma (152.4-159.8 Ma). O Cretáceo inferior e a revolução terrestre do Cretáceo (*Cretaceous terrestrial revolution*, KTR) foram os momentos mais importante na diversificação das grandes linhagens de Apoidea. Os nossos dados indicaram que as altas taxas de diversificações e acúmulo gradativo das grandes linhagens de Apoidea corresponderam ao KTR. Conseguimos estimar, portanto, os padrões macroevolutivos das grandes linhagens de Apoidea utilizando somente dados morfológicos associados às linhagens fósseis em uma filogenia datada por relógio morfológico.

Palavras-chave: Apoidea, Crabronidae, Apidae, Cretáceo inferior, revolução terrestre do Cretáceo, KTR, LTTs, estatística gamma, Mr Bayes, particionamento morfológico.

ABSTRACT

Apoidea include, in addition to bees, a large diversity of predatory wasps, currently referred informally as apoid wasps. Apoidea form one of the most diverse groups in Aculeata with over 27500 species currently described. In the phylogenetic relationships, the greatest conflicts are in the positioning of the lineages of Crabronidae *sensu lato*. The fossil record of Apoidea is very rich with specimens preserved in both rock compression and amber inclusions. In addition to the great diversity of lineages, these fossils are distributed over a wide temporal scale contemplating a range between 10 and 130 Mya. The age estimates for the crown group of Apoidea vary between the Jurassic (185.8-153 Ma) and the Early Cretaceous (~135 Ma). But the major lines of Apoidea have already been differentiated by the end of the Early Cretaceous. The morphological clock is a way of calculating the estimates of the time of divergence between lineages. The method simultaneously considers the temporal information and the information coming from the morphological characters to co-estimate the phylogeny, the rates of evolution and the time of divergence between the lineages. The analyses of lineage diversification estimates over time (LTTs) are powerful macroevolutionary tools that use dated phylogenies to approximate macroevolutionary patterns linked to the emergence of species. Several authors argue that a sampling without the fossil record is an unrealistic macroevolutionary scenario and presents a great bias in the estimation of extinction events. The use of a broad fossil record in phylogeny dating in Apoidea allows better approximations to macroevolutionary patterns. The phylogenetic relationships of Apoidea were investigated in a Bayesian approach with morphological data. We included 99 living terminals and 32 fossils. The estimates of the divergence times of the lineages were carried out applying a morphological clock. Twenty different morphological clock models were tested and selected by marginal likelihood values. The macroevolutionary patterns linked to the diversification of the major Apoidea lineages were investigated through LTTs and gamma statistics. In total, at least 17 independent lineages of Apoidea, being 11 extant and six extinct clades, have been recovered. The morphological clock was efficient for the estimation of the time of divergence of the Apoidea, however, there are still challenges to understand and improve its applications. The stem group of Apoidea diverged at the end of the Jurassic, at about 156.3 Ma (152.4-159.8 Ma). The lower Cretaceous and the Cretaceous terrestrial revolution (KTR) were the most important moments in the diversification of the major lines of Apoidea. Our data indicated that the high rates of diversification and gradual accumulation of the main Apoidea lineages corresponded to the KTR. We therefore have been able to estimate the macroevolutionary patterns of these lineages using only morphological data associated with the fossil lineages in a dated phylogeny by morphological clock.

Key-words: Apoidea, Crabronidae, Apidae, Lower Cretaceous, Cretaceous terrestrial revolution, KTR, LTTs, gamma statistics, Mr Bayes, morphological partitioning.

SUMÁRIO

1. Introdução	3
1.1. Apoidea: diversidade, fósseis e relações filogenéticas	3
1.2. Estimativa do tempo de divergência com relógio morfológico	4
1.3. Estimativas de diversificação de linhagens.....	5
2. Materiais e método.....	7
2.1. Material examinado, terminologia e classificação	7
2.2.1 <i>Amostragem dos táxons viventes</i>	8
2.2.2 <i>Amostragem dos táxons fósseis</i>	11
2.2. Matriz de dados e delimitação dos caracteres	13
2.3. Análises filogenéticas.....	15
2.3.1. <i>Inferência Bayesiana</i>	15
2.3.2. <i>Particionamento</i>	15
2.3.3. <i>Parcimônia</i>	16
2.4. Testes de modelos	16
2.5. Estimativa dos tempos de divergência	18
2.5.1. <i>Calibração, prior da raiz e restrições topológicas</i>	18
2.5.2. <i>Modelos de relógio de ramos</i>	19
2.5.3. <i>Modelos de relógio de árvore</i>	19
2.5.4. <i>Comparações das estimativas de tempo de divergência dos diferentes modelos de relógio</i>	20
2.6. Análises de linhagens ao longo do tempo (LTTs) e estatística- γ	21
3. Resultados	23
3.1. Análises filogenéticas.....	23
3.1.1. <i>Análises em inferência Bayesiana</i>	23
3.1.2. <i>Particionamento</i>	23
3.1.3. <i>Análises sob parcimônia</i>	23
3.1.4. <i>Testes de hipóteses topológicas</i>	26
3.2 Estimativas do tempo de divergência.....	27
3.2.1. <i>Modelos testados</i>	27
3.2.2. <i>Estimativas do tempo de divergência em MCMC</i>	28
3.3. Análises de linhagens ao longo do tempo (LTTs) e estatística- γ	36
4. Discussão	39
4.1. Relações filogenéticas e tempo de divergência das linhagens de Apoidea.....	39
4.2. Estimativa do tempo de divergência com relógio morfológico	49
4.2.1. <i>Escolha de priors em relógio morfológico</i>	49
4.2.2. <i>Vantagens e desvantagens do relógio morfológico</i>	50
4.3. Taxas de diversificação e LTTs com a inclusão de fósseis.....	52
4.3.1. <i>Amostragem dos LTTs e interferência nas taxas de diversificação</i>	52
4.3.2. <i>Influência do registo fóssil nos LTTs</i>	54
4.4. Padrões de diversificação e extinção das grandes linhagens de Apoidea	55
5. Considerações finais.....	58
6. Referências.....	61
7. Apêndices	69
7.1 Apêndice 1. Figuras 1-6	69
7.2 Apêndice 2. Lista dos caracteres morfológicos e seus estados	75
7.3 Apêndice 3. Matriz de caracteres e seus estados.....	95
7.4 Apêndice 4. Árvores de parcimônia.....	102
7.5 Apêndice 5. Valores multiplicadores das partições morfológicas	107
7.6 Apêndice 6. Correlações entre as médias das idades estimadas para diferentes modelos de relógio	108
7.7 Apêndice 7. Cronograma das linhagens de Apoidea contraposta com o período KTR.....	109
7.8. Apêndice 8. Restrições do cronograma da Figura 5.	110
8. Anexo. Homoplasy-based partitioning outperforms alternatives in Bayesian analysis of discrete morphological data	111

1. Introdução

1.1 Apoidea: diversidade, fósseis e relações filogenéticas

Os himenópteros, popularmente conhecidos como abelhas, formigas e vespas, constituem um diverso e abundante grupo de insetos holometábolos. Atualmente existem mais de 150.000 espécies viventes descritas (Aguiar et al. 2013), embora estimativas apontem para um número muito maior, em torno de 375.000 a 480.000 espécies (Forbes et al. 2018). Entre as superfamílias de Hymenoptera, os Apoidea destacam-se por desempenhar importantes papéis ecológicos e econômicos, como o controle de populações de insetos de mais de 13 ordens através da predação (Bohart & Menke 1976) e a polinização das plantas com flores realizadas pelas abelhas (Roubik 1989; Martins et al. 2014).

Os Apoidea incluem, além das abelhas, uma grande diversidade de vespas predadoras, atualmente referidas informalmente como vespas apoideas. Durante o século XIX e grande parte do século XX, as vespas apoideas eram tratadas em uma superfamília à parte, Sphecoidea como em Ashmead (1899), ou em uma única família denominada Sphecidae (Fox 1894; Kohl 1897; Bohart & Menke 1976). No final do século XX, Alexander (1992a), Brothers (1999) e Melo (1999) demonstraram que Sphecidae *lato sensu* era na verdade parafilético em relação às abelhas. Com base nas hipóteses obtidas, Melo (1999) dividiu Apoidea em cinco famílias: Ampulicidae, Apidae *lato sensu*, Crabronidae, Heterogynaidae e Sphecidae.

Os Apoidea formam um dos grupos mais diversos em Aculeata com mais de 27500 espécies descritas atualmente (Michener 2007; Pulawski 2018). Apidae *lato sensu* é a família mais diversa com mais de 17500 espécies descritas (Michener 2007), seguida por Crabronidae (8988 espécies), Sphecidae (783 espécies), Ampulicidae (202 espécies) e Heterogynaidae, com nove espécies distribuídas em um único gênero (Pulawski 2018). Crabronidae *lato sensu* forma um grupo de vespas predadoras que apresentam não apenas uma grande riqueza, mas também uma variedade surpreendente em sua biologia e comportamento (Bohart & Menke 1976; Melo 2000; Melo et al. 2011). Essa linhagem consiste em vespas solitárias que capturam e paralisam uma grande diversidade de insetos, bem como de aranhas em alguns grupos, para alimentar a prole. Na maioria das vezes essas vespas exibem especializações ou preferências quanto aos táxons usados como presas (Bohart & Menke 1976; Melo 2000; Hanson & Menke 2006; Melo et al. 2011).

As relações filogenéticas de Apoidea já foram investigadas tanto a partir de dados morfológicos quanto moleculares (Alexander 1992a; Melo 1999; Devedec et al. 2012; Branstetter et al. 2017; Peters et al. 2017; Sann et al. 2018). Os maiores conflitos de relacionamentos estão no posicionamento das linhagens de Crabronidae *lato sensu*. Na hipótese de Melo (1999) essa linhagem de vespas forma um grupo monofilético estreitamente relacionado com as abelhas. Entretanto, em diferentes trabalhos esse grupo é parafilético em relação às abelhas (Alexander 1992a; Devedec et al. 2012; Branstetter et al. 2017; Peters et al. 2017; Sann et al. 2018). O relacionamento entre Sphecidae e os crabronídeos também é conflitante. Nas hipóteses propostas a partir de dados morfológicos Sphecidae é grupo irmão do clado Crabronidae *lato sensu* + Apidae *lato sensu* (Melo 1999). Por outro lado, nas hipóteses moleculares, Crabronidae *stricto*

sensu e por vezes *Mellinus* estão mais associados com Sphecidae que com as demais crabronídeos e abelhas (Devebec et al. 2012; Branstetter et al. 2017; Peters et al. 2017; Sann et al. 2018).

O registro fóssil de Apoidea é muito rico, contando com exemplares preservados tanto em compressões rochosas quanto em inclusões em âmbar (Darling & Sharkey 1990; Ohl 2004; Grimaldi & Engel 2005; Bennett & Engel 2006). Além da grande diversidade de linhagens, esses fósseis estão distribuídos em uma ampla escala temporal contemplando um intervalo entre 10 a 130 Ma. Diversos fósseis descritos do início do Cretáceo inferior são atribuídos a Apoidea e estão alocados em uma única família fóssil, os †Angarosphécidae (Rasnitsyn 1975; Rasnitsyn et al. 1998). A maioria desses fósseis foram descritos com base na morfologia da asa anterior devido à sua natureza de preservação. Entretanto, existem fortes evidências morfométricas que apontam para o posicionamento desses fósseis em outros grupos de Aculeata (Gouvêa, Rosa & Melo, em preparação). Os fósseis mais antigos que podem ser atribuídos a Apoidea são provenientes do âmbar libanês, com idade entre 130-125.5 Ma (Prentice 1993; Prentice 1994; Pinar & Milki 2001). Os dois fósseis conhecidos desse âmbar foram interpretados como ampulicídeos por Prentice (1993; 1994) mas como não foram descritos o seu posicionamento entre as linhagens de Apoidea permanece incerto.

As estimativas de idade para o grupo *crown* de Apoidea variam entre o Jurássico (185.8-153 Ma) (Peters et al. 2017; Sann et al. 2018) e o Cretáceo inferior (134.90 Ma) (Branstetter et al. 2017). De qualquer maneira, todas as grandes linhagens de Apoidea já haviam se diferenciado até o final do Cretáceo inferior (Branstetter et al. 2017; Peters et al. 2017; Sann et al. 2018). Em um contexto mais amplo da diversificação biológica muitos autores associam os recorrentes conflitos topológicos a uma rápida radiação em intervalos curtos do tempo geológico (Jackman et al. 1999; Lehtonen et al. 2012; Sterli et al. 2013; Suh et al. 2015; Suh et al. 2016; Longo et al. 2017; Rouard et al. 2018). Em vista das idades estimadas para a superfamília, uma rápida diversificação em um curto intervalo de tempo é uma boa hipótese para os conflitos nas relações das grandes linhagens de Apoidea.

1.2 Estimativa do tempo de divergência com relógio morfológico

A estimativa do tempo de divergência permite calcular o tempo transcorrido na divergência entre linhagens. Tradicionalmente essas estimativas são consideradas a partir de dados moleculares em que a taxa de substituição de nucleotídeos é a medida base da inferência, conhecido como método de relógio molecular (Kumar 2005; Renner 2005; Lemey & Posada 2009). Como a taxa de substituição de nucleotídeos apresenta-se bastante variável, dependendo de qual parte do genoma está sendo analisada, e variando significativamente também ao longo do tempo e entre as linhagens, a prática comum atualmente é acomodar essa variação em modelos de relógio relaxados (Ronquist et al. 2012a; Renner 2005; dos Reis et al. 2015).

Outro método de estimativa de tempos de divergência é o relógio morfológico, que como o nome sugere, é análogo ao relógio molecular, porém as taxas evolutivas são inferidas a partir de dados morfológicos (Ronquist et al. 2012a; Lee et al. 2014; Beck & Lee 2014). Nessa abordagem os fósseis são

incluídos como terminais na matriz e durante as análises são as peças chaves para a funcionalidade do método. O método considera simultaneamente a informação temporal e a informação proveniente dos caracteres morfológicos para co-estimar a filogenia, as taxas de evolução e o tempo de divergência entre as linhagens (Beck & Lee 2014). O método de calibração utilizado em uma abordagem de relógio morfológico é chamado de *tip-dating*. Esse método permite que as datas sejam informadas diretamente dos terminais evitando atribuições equivocadas com o emprego de *node-dating* (Lee & Palci 2015). Embora o relógio morfológico apresente excelentes resultados com estimativas equiparáveis ou ainda mais eficientes que o relógio molecular (Lee et al. 2013; Lee et al. 2014), para alguns grupos o método parece não funcionar corretamente (Beck & Lee 2014; Puttick et al. 2016).

Diversos trabalhos se propuseram a investigar o uso de diferentes *priors* em uma abordagem estritamente *tip-dating* com dados morfológicos (Lee et al. 2014; O'Reilly et al. 2015; Drummond & Stadler 2016; Puttick et al. 2016; Gavryushkina et al. 2017). Porém na maioria das vezes os trabalhos investigaram a eficiência entre *tip-dating* e *node-dating*, comparando diretamente os resultados com dados morfológicos com aqueles provenientes de dados moleculares e/ou por evidência total (O'Reilly et al. 2015 Drummond & Stadler 2016 Puttick et al. 2016). Outros trabalhos investigaram ainda a eficiência entre *priors* de relógio de ramos estritos e relaxados (Lee et al. 2014; Gavryushkina et al. 2017). Alguns poucos trabalhos exploraram um espaço amostral maior em relação às diferentes combinações de *priors* de relógios de árvores e ramos (Beck & Lee 2014; Matzke & Wright 2016). Mesmo assim, não existem trabalhos com delineamentos para escolha de modelos de *priors* de árvores e de ramos através de métodos comparativos em abordagens estritamente morfológicas.

1.3 Estimativas de diversificação de linhagens

As análises de estimativas de diversificação de linhagens ao longo do tempo (LTTs, *Lineage Through Time plots*) são poderosas ferramentas macroevolutivas que utilizam filogenias datadas para fazer aproximações de padrões macroevolutivos ligados ao surgimento de espécies (Nee et al. 1994; Ricklefs 2007). Diversos autores argumentam que uma amostragem sem o registro fóssil é um panorama macroevolutivo pouco realista e apresenta um grande viés na estimativa de eventos de extinção (Ricklefs 2007; Etienne et al. 2012). Adicionalmente, trabalhos demonstraram que filogenias moleculares sem a inclusão de fósseis são incapazes de estimar taxas e eventos de extinção (Quental & Marshall 2010; Rabosky 2010). A utilização de um amplo registro fóssil em filogenias datadas como os de Apoidea permitem estimativas mais acuradas de padrões macroevolutivos. Esse é, portanto, um cenário ideal para avaliar questões macroevolutivas ligadas a padrões de diversificações e extinções.

Como apontado anteriormente, os Apoidea tiveram uma clara diversificação de suas grandes linhagens durante o Cretáceo inferior. O Cretáceo inferior foi um período com altos níveis de CO₂ que contribuíram para o aumento das temperaturas globais causando pouca diferença entre as estações, o que é conhecido como ótimo climático (*Greenhouse*) do Cretáceo inferior (Price et al. 1998; Royer et al. 2004; Huber et al. 2018). As mudanças climáticas durante esse período interferiram diretamente na diversidade

biológica. Durante esse período houve uma explosão na diversificação dos grandes grupos de angiospermas (Boer et al. 2012; Montoya et al. 2018), diversos grupos de tetrápodes (Lloyd et al. 2008; Benton 2009; Benton 2010) e ainda numerosos grupos de insetos (Clapham et al. 2016; Nel et al. 2018). A parcela mais relevante desse período, em que as taxas de diversificação desses grupos foram mais altas, corresponde ao intervalo de 125-80 Ma e é conhecida como a revolução terrestre do Cretáceo (*Cretaceous terrestrial revolution*, KTR) (Lloyd et al. 2008).

Em resumo, os Apoidea possuem um amplo registro fóssil, diversas propostas filogenéticas, diferentes estimativas de tempo de divergência e uma clara diversificação das grandes linhagens durante o Cretáceo inferior. Portanto, os objetivos desse trabalho foram 1) estabelecer as relações filogenéticas das principais linhagens de Apoidea com ênfase nos grupos de crabronídeos; 2) estabelecer pela primeira vez as relações filogenéticas de fósseis de vespas apoideas; 3) propor estimativas de tempos de divergências usando o abundante registro fóssil de Apoidea através do relógio morfológico; 4) investigar os diferentes modelos de relógio de árvore e relógio de ramos estabelecendo modelos mais apropriados usando a verossimilhança marginal como parâmetro de escolha; e 5) avaliar pela primeira vez os processos macroevolutivos de Apoidea como um todo com ênfase nos processos de extinção e especiação ao longo das linhagens através do tempo, focando no Cretáceo inferior e nas relações dessas linhagens com o KTR.

2. Materiais e método

2.1. Material examinado, terminologia e classificação

Os táxons vivos examinados pertencem às seguintes instituições: Museu Americano de História Natural (AMNH, Nova Iorque, NY, EUA), Academia de Ciências da Califórnia (CAS, São Francisco, CA, EUA), Coleção de Entomologia Pe. Jesus Santiago Moure, Universidade Federal do Paraná (DZUP, Curitiba, Brasil) e Museu Nacional de História Natural do Paraguai (MNHNP, San Lorenzo, Paraguai).

Nos táxons vivos, a morfologia externa e interna foi estudada a partir da dissecação completa de um espécime adulto (preferencialmente uma fêmea) e/ou do exame de um exemplar preservado em via seca (Tabela 2), exceto por *Laphyragogus*, *Xenosphex*, *Handlirschia* e *Kohlia*. *Handlirschia* foi estudada a partir de um exemplar macho parcialmente dissecado por Dr. Byron Alexander (peças bucais e terminália), que se encontra depositado no AMNH. Os protocolos de dissecações seguiram Melo (1999) com modificações propostas por Porto et al. (2016). Também foram dissecadas as terminálias dos semafóritos opostos aos das dissecações completas (Tabela 1) e submetidas aos mesmos protocolos.

Os táxons fósseis (Tabela 3) examinados pertencem às seguintes instituições: Museu Americano de História Natural (AMNH, Nova Iorque, NY, EUA), Departamento de Zoologia da Universidade Federal do Paraná (DZUP, Curitiba, Brasil) e Instituto Paleontológico da Academia Russa de Ciências (PIN, Moscou, Rússia). As peças de âmbar que precisaram de preparações foram realizadas por Paul C. Nascimbene (AMNH) e pelo Dr. Gabriel Melo (DZUP), seguindo os protocolos correntes apresentados em Grimaldi (1993), Bisulca et al. (2012) e Lucena & Melo (2018).

Tanto os táxons vivos quanto fósseis foram estudados sob microscópio estereoscópio Leica M125 (DZUP) e Nikon SMZ 1500 (AMNH). As fotografias foram tiradas usando uma câmera acoplada no Leica M125 (DZUP) e em uma Canon Eos Rebel T6 (AMNH), usando um sistema de empilhamento a partir de uma sequência de fotos. Os empilhamentos das imagens foram realizados no programa Zerene Stacker®. As imagens das estruturas dos táxons vivos e dos fósseis encontram-se no Apêndice 1 Figuras 1-6.

Um dos fósseis usados, *Angarosphex magnus* Darling, foi também estudado com técnicas de tomografia computadorizada como em Barden et al. (2017). Os fósseis dessa espécie são provenientes da Chapada do Araripe, nordeste Brasileiro (Ceará) e preservados em carbonatos laminados (Grimaldi & Maisey 1990). Os espécimes AMNH-44017 (holótipo), AMNH SA 46322 e AMNH AS 46305 foram submetidos a micro tomografia computadorizada de raios-X (micro-CT) no Centro de Microscopia do Museu Americano de História Natural (*AMNH Microscopy*). As digitalizações foram feitas com o aparelho GE Phoenix v|tome|x s240 equipado com uma fonte de raios X de 240 ou 180 kV. As figuras encontram-se no Apêndice 1 Figuras 6g-h.

A terminologia morfológica adotada seguiu Bohart & Menke (1976), Michener (1944) e Snodgrass (1935) com modificações propostas em Melo (1999). A terminologia para os caracteres das asas foi baseada em Day (1998) e Goulet & Huber (1993). Outros termos foram citados diretamente na lista de caracteres.

Quanto à classificação adotada, consideramos as abelhas como pertencentes a uma única família (sensu Melo 1999 e Melo & Gonçalves 2005), sendo citada ao longo do texto como Apidae *lato sensu*. Na classificação de Melo (1999) os crabronídeos também são considerados como uma única família. Por outro lado, segundo os nossos resultados os crabronídeos são considerados como linhagens independentes. Seguimos parcialmente a classificação proposta por Sann et al. (2018) utilizando as seguintes famílias: Astatidae, Crabronidae *stricto sensu*, Bembicidae, Mellinidae, Pemphredonidae, Philanthidae e Psenidae. Em nossos resultados os Ammoplanidae *sensu* Sann et al. (2018) estão relacionados diretamente com Astatidae. Por essa razão suspendemos o uso dessa linhagem de maneira independente (veja a discussão em Astatidae). Por questões de esclarecimento usamos ao longo do texto os termos “Crabronidae *lato sensu*” ou “a linhagem dos crabronídeos” em referência à linhagem parafilética como um todo e “Crabronidae *stricto sensu*” para se referir a Crabroninae de Melo (1999).

2.2.1 Amostragem dos táxons viventes

Os terminais viventes foram escolhidos levando em consideração: 1) a disponibilidade de material para as dissecções; 2) informação sobre a biologia (para caracteres de larva e de comportamento); e 3) maior número de linhagens especialmente aquelas não amostradas ou pouco amostradas por Melo (1999). Foram utilizadas praticamente todas as tribos das grandes linhagens de crabronídeos, adicionando terminais em Astatidae l.s., Bembicidae, Crabronidae e Philanthidae dos quais foram menos representados por Melo (1999). Terminais complementares aos de Melo (1999) foram adicionados em Apidae l.s., Sphecidae, Pemphredoninae e no grupo externo. No total foram usados 99 terminais de táxons viventes distribuídos em 20 famílias de Aculeata, sendo nove delas representantes do grupo externo (Tabela 1).

Tabela 1. Táxons vivos utilizados como terminais para a reconstrução das relações filogenéticas em Apoidea. A coluna Dissecção indica o sexo do semaforonte submetido a dissecção completa; aqueles terminais para os quais não havia material para ser dissecado estão indicados como “—”. A coluna Referência remete aos exemplares usados como terminais em Melo (1999) ou dissecados pela primeira vez no presente estudo. Para todos os terminais, foram examinados tanto fêmeas quanto machos em via seca, exceto *Spilomena subterranea*, para o qual apenas fêmeas estavam disponíveis.

Taxa	Dissecção	Referência
Grupo externo		
Bethylidae		
<i>Bakeriella lindigi</i> (Kieffer)	Fêmea	Melo (1999)
Pompilidae		
<i>Notocyphus</i> sp. (da Costa Rica)	Fêmea	Melo (1999)
Rhopalosomatidae		
<i>Rhopalosoma nearticum</i> Brues	Fêmea	Melo (1999)
Sapygidae		
<i>Eusapyga proxima</i> (Cresson)	Fêmea	Melo (1999)
Scolecbythidae		
<i>Clystopsenella longiventris</i> Kieffer	Fêmea	Melo (1999)
Scoliidae		
<i>Trielis</i> sp. (do Brasil)	Macho	Este estudo
Sierolomorphidae		
<i>Sierolomorpha canadensis</i> Provancher	Fêmea	Melo (1999)
Tiphiidae		
<i>Anthosila</i> sp. (do Brasil)	Macho	Este estudo
Formicidae		
<i>Pachycondyla striata</i> Smith	Macho	Este estudo
Vespidae		
<i>Trimeria rubra</i> Hermes & Melo	Fêmea	Este estudo
Grupo interno		
Ampulicidae		
<i>Ampulex</i> sp. (da Costa Rica)	Fêmea	Melo (1999)
<i>Aphelotoma rufiventris</i> Turner	Macho	Melo (1999)
<i>Dolichurus</i> sp. (da Costa Rica)	Fêmea	Melo (1999)
Apidae lato sensu		
Andreninae		
<i>Arhysosage flava</i> Moure	Fêmea	Este estudo
Apinae		
<i>Anthophorula albata</i> (Timberlake)	Fêmea	Melo (1999)
<i>Bombus</i> sp. (da Grécia)	Fêmea	Este estudo
<i>Eulaema nigrita</i> Lepeletier	Fêmea	Este estudo
<i>Melipona asilvai</i> Moure	Fêmea	Este estudo
Colletinae		
<i>Caupolicana hirsuta</i> Spinola	Fêmea	Este estudo
<i>Lonchopria zonalis</i> (Reed)	Fêmea	Melo (1999)
Halictinae		
<i>Conanthalictus nigricans</i> Timberlake	Fêmea	Melo (1999)
<i>Dieunomia</i> sp. (dos EUA)	Fêmea	Este estudo
Megachilinae		
<i>Neofidelia</i> sp. (do Chile)	Fêmea	Este estudo
Melittinae		
<i>Hesperapis carinata</i> Stevens	Fêmea	Melo (1999)
<i>Melitta haemorrhoidalis</i> (Fabricius)	Fêmea	Este estudo
Stenotritinae		
<i>Ctenocolletes smaragdinus</i> (Smith)	Fêmea	Melo (1999)
“Crabronidae lato sensu”		
Astatidae		
Astatinae		
<i>Astata nevadica</i> Cresson	Fêmea	Melo (1999)
<i>Dryudella</i> sp. (dos EUA)	Fêmea	Este estudo
<i>Diplopectron</i> sp. (dos EUA)	Fêmea	Este estudo
Eremiasphecinae		
<i>Eremiasphecium sahelense</i> (Simon-Thomas)	Fêmea	Melo (1999)
Ammoplaninae		
<i>Ammoplanus</i> cfr. <i>apache</i> Pate	Fêmea	Melo (1999)
<i>Pulverro mescalero</i> Pate	Fêmea	Melo (1999)
<i>Mohavena yucaipa</i> Pate	Fêmea	Melo (1999)

Tabela 1. Continuação. (2/3).

Taxa	Dissecado	Referência
Bembicidae		
<i>Alysson melleus</i> (Say)	Fêmea	Este estudo
<i>Didineis texana</i> (Cresson)	Macho	Melo (1999)
<i>Zanysson</i> sp. (do Brasil)	Fêmea	Este estudo
<i>Nysson rusticus</i> (Cresson)	Fêmea	Melo (1999)
<i>Heliocausus larroides</i> (Spinola)	Fêmea	Melo (1999)
<i>Tiguiipa</i> aff. <i>fiebrigi</i>	Fêmea	Este estudo
<i>Handlirschia scoliaeformis</i> Arnold	—	Este estudo
<i>Bembecinus quinquespinosus</i> (Say)	Fêmea	Melo (1999)
<i>Stictia signata</i> (Linnaeus)	Fêmea	Este estudo
<i>Glenostictia scitula</i> (Fox)	Macho	Este estudo
<i>Ammatomus mesostenus</i> (Handlirsch)	Fêmea	Este estudo
<i>Argogorytes</i> sp (do Brasil)	Macho	Este estudo
<i>Austrogorytes</i> sp. (da Austrália)	Fêmea	Este estudo
<i>Exeirus lateritius</i> Shuckard	—	Este estudo
<i>Gorytes quinquefasciatus</i> (Panzer)	Fêmea	Este estudo
<i>Hoplisoides spilopterus</i> (Handlirsch)	Fêmea	Melo (1999)
<i>Kohlia coxalis</i> Morice	—	Este estudo
<i>Megistommum politum</i> (Smith)	Fêmea	Este estudo
<i>Ochleroptera bipunctata</i> (Say)	Fêmea	Melo (1999)
<i>Pterygorytes valens</i> (Fox)	Fêmea	Este estudo
<i>Sagenista scutellaris</i> (Spinola)	Fêmea	Este estudo
<i>Sphecius</i> sp. (da Grécia)	Macho	Este estudo
Crabronidae lato sensu		
<i>Dinetus pictus</i> (Fabricius)	Macho	Melo (1999)
<i>Xenosphex timberlakei</i> Williams	—	Melo (1999)
<i>Anacrabro ocellatus</i> Packard	Fêmea	Melo (1999)
<i>Lindenius columbianus</i> (Kohl)	Fêmea	Melo (1999)
<i>Oxybelus emarginatum</i> Say	Fêmea	Este estudo
<i>Willinkella argentina</i> (do Brasil)	Fêmea	Este estudo
<i>Scapheutes brasiliensis</i> Handlirsch	Macho	Este estudo
<i>Palarus latifrons</i> Kohl	Macho	Melo (1999)
<i>Paranysson melanopyrus</i> (Smith)	Fêmea	Este estudo
<i>Lyroda subita</i> (Say)	Fêmea	Melo (1999)
<i>Laphyragogus pictus</i> Kohl	—	Melo (1999)
<i>Gastrosericus waltlii</i>	Fêmea	Este estudo
<i>Tachytes</i> aff. <i>hades</i>	Fêmea	Este estudo
<i>Miscophus</i> sp. (da Alemanha)	Fêmea	Este estudo
<i>Nitela amazonica</i> Ducke	Fêmea	Melo (1999)
<i>Plenoculus davis</i> Fox	Macho	Melo (1999)
<i>Sericophorus relucens</i> Smith	Fêmea	Este estudo
<i>Solierella</i> aff. <i>minarum</i> Ducke	Fêmea	Este estudo
<i>Entomopison longicorne</i> Menke	Fêmea	Este estudo
<i>Pisonopsis</i> sp. (dos USA)		Este estudo
Mellinidae		
<i>Mellinus alpestris</i> Cameron	Macho	Melo (1999)
Psenidae		
<i>Odontosphex paradoxus</i> Menke	Macho	Melo (1999)
<i>Entomosericus concinnus</i> Dahlbom	Fêmea	Melo (1999)
<i>Mimesa cressonii</i> Packard	Fêmea	Melo (1999)
<i>Pluto minutus</i> (Malloch)	Fêmea	Melo (1999)
<i>Psenulus mayorum</i> Bohart & Grissell	Fêmea	Melo (1999)
Pemphredonidae		
<i>Arpactophilus steindachneri</i> Kohl	Fêmea	Melo (1999)
<i>Diodontus rugosus</i> Fox	Fêmea	Melo (1999)
<i>Parastigmus huecuvus</i> Finnamore	Fêmea	Melo (1999)
<i>Passaloecus areolatus</i> Vincent	Fêmea	Melo (1999)
<i>Pemphredon inornata</i> Say	Fêmea	Melo (1999)
<i>Spilomena catamarca</i> Antropov	Fêmea	Melo (1999)
<i>Spilomena subterranea</i> McCorquodale & Naumann	Fêmea	Este estudo

Tabela 1. Continuação. (3/3).

Taxa	Dissecado	Referência
<i>Stigmus temporalis</i> Kohl	Fêmea	Melo (1999)
Philanthidae		
<i>Aphilanthops frigidus</i> Smith	Macho	Melo (1999)
<i>Clypeadon evansi</i> Bohart	Fêmea	Este estudo
<i>Cerceris rufopicta</i> Smith	Fêmea	Este estudo
<i>Pseudoscolia simplicicornis</i>	Fêmea	Este estudo
<i>Philanthus gibbosus</i> (Fabricius)	Fêmea	Melo (1999)
Heterogynaidae		
<i>Heterogyna fantsilotra</i> Day	Macho	Melo (1999)
Sphecidae		
Ammophilinae		
<i>Podalonia communis</i> (Cresson)	Macho	Melo (1999)
Chloriontinae		
<i>Chlorion aerarium</i> Patton	Fêmea	Melo (1999)
Sceliphrinae		
<i>Penepodium</i> sp. (do Brasil)	Fêmea	Este estudo
Sphecinae		
<i>Palmodes rufiventris</i> (Cresson)	Macho	Melo (1999)
Stangeellinae		
<i>Stangeella cyaniventris</i> (Guerin-Meneville)	Fêmea	Melo (1999)

2.2.2 Amostragem dos táxons fósseis

Os terminais fósseis (Tabela 2) foram escolhidos levando em consideração: 1) disponibilidade de material nas coleções; 2) táxons já descritos; 3) a integridade dos exemplares nas peças; 4) representação do maior número de linhagens; e 4) maior abrangência temporal no registro fóssil. Em poucos casos os terminais foram codificados a partir de referências bibliográficas, a saber: os machos de *Trypoxylon* foram codificados a partir do trabalho Prentice & Poinar (1993); os machos de *Dolichurus* a partir de Nemkov (1988) e Ohl (2004); e o terminal †*Cretotrigona* foi codificado unicamente pela literatura com base nos trabalhos de Michener & Grimaldi (1988), Engel 2000 e Engel (2001). No total foram incluídos 33 terminais de táxons fósseis sendo 32 representantes de Apoidea, um terminal no grupo externo (Formicidae) e distribuídos em pelo menos 15 linhagens distintas (nível de família).

Tabela 2. Táxons fósseis utilizados como terminais para a estimativa das relações filogenéticas em Apoidea. Além dos depósitos são apresentados os dados de museu de cada peça/exemplar usado.

Terminais	Sexos usados	Depósito	Material usado para compor os terminais
Fomicidae			
† <i>Haidomyrmex</i>	Macho	Âmbar Burmês	† <i>Haidomyrmex</i> sp. (um macho DZUP Bur-337; um macho DZUP Bur-475)
Apoidea			
Fóssil A			
Fóssil A	Macho	Âmbar Burmês	Fóssil A (cinco machos DZUP Bur-534; um macho DZUP Bur-208; um macho DZUP Bur-002b)
Incertae Sedis			
Fóssil B	Macho	Âmbar Burmês	Fóssil B (um macho AMNH JZC-Bu56)
Fóssil C	Macho	Âmbar Burmês	Fóssil C (um macho DZUP Bur-622)

Tabela 2. Continuação. (2/3).

Terminais	Sexos usados	Depósito	Material usado para compor os terminais
“†Angarosphecidae”			
† <i>Angarosphex</i>	Fêmea/Macho	Formação Crato	† <i>Angarosphex magnus</i> (Darling) (Holótipo, macho? AMNH 44107; um macho (Holótipo); Parátipo, uma fêmea? AMNH 43267; um macho AMNH SA 46322; um macho AMNH SA 46305; um macho? AMNH SA 46307)
† <i>Burmasphex</i>	Macho	Âmbar Burmês	† <i>Burmasphex sulcatus</i> Melo & Rosa (Holótipo, macho DZUP Bur-503); <i>B. pilosus</i> Melo & Rosa (Holótipo, macho DZUP Bur-448); † <i>Burmasphex</i> sp. (um macho DZUP Bur-017a)
Fóssil D	Fêmea/Macho	Âmbar Burmês	Fóssil D (um macho DZUP Bur-502; um macho DZUP Bur-479; uma fêmea DZUP Bur-114)
Ampulicidae			
† <i>Cretampulex</i>	Fêmea/Macho	Âmbar Burmês	† <i>Cretampulex gracilis</i> Antropov (uma fêmea AMNH Bu-KL-4-2; um macho AMNH Bu-KL-8-19)
<i>Dolichurus</i>	Fêmea/Macho*	Âmbar Báltico	<i>Dolichurs</i> aff. <i>heevansi</i> Ohl (uma fêmea DZUP Bal-024) *Referências usadas para a codificação do macho: Nemkov (1988) e Ohl (2004).
Astatidae lato sensu			
Fóssil E	Fêmea/Macho	Âmbar Burmês	Fóssil E (uma fêmea DZUP- Bur-002; uma fêmea DZUP Bur-501; um macho DZUP Bur-010; uma fêmea AMNH JZC-Bu37)
Fóssil F	Macho	Âmbar Burmês	Fóssil F (uma fêmea DZUP Bur-236)
Heterogynaidae			
† <i>Apodolichurus</i>	Fêmea/Macho	Âmbar Burmês	† <i>Apodolichurus sphaerocephalus</i> Antropov (uma fêmea DZUP Bur-435); † <i>Apodolichurus</i> sp. (um macho DZUP Bur-611)
Fóssil G	Fêmea	Âmbar Burmês	Fóssil G (uma fêmea DZUP Bur-111)
Crabronidae			
Fóssil H	Fêmea/Macho	Âmbar Burmês	Fóssil H (uma fêmea DZUP Bur-540; um macho DZUP Bur-404; um macho DZUP-404)
† <i>Lindocerus</i>	Fêmea	Âmbar Dominicano	† <i>Lindocerus paleomystax</i> (Bennet & Engel) (Holótipo, fêmea AMNH DR-14-1091; Parátipo, fêmea AMNH DR-14-236; uma fêmea AMNH DR-15-229)
<i>Crossocerus</i>	Fêmea	Âmbar Báltico	<i>Crossocerus</i> sp. (uma fêmea DZUP Bal-002)
<i>Trypoxylon</i>	Fêmea/Macho*	Âmbar Dominicano	<i>Trypoxylon pallidiventre</i> Prentice & Poinar 1993 (uma fêmea DZUP Dom-021; uma fêmea DZUP Dom-022) *Referências usadas para a codificação do macho: Prentice & Poinar (1993)
Psenidae			
Fóssil I	Fêmea	Âmbar Burmês	Fóssil I (uma fêmea DZUP Bur-059)
Fóssil J	Fêmea/Macho	Âmbar Burmês	Fóssil J (uma fêmea DZUP Bur-099; uma fêmea DZUP Bur-289; um macho DZUP Bur-385; um macho DZUP Bur-535)
Fóssil K	Fêmea	Âmbar Burmês	Fóssil K (uma macho DZUP-064)
Pemphredonidae			
Fóssil L	Fêmea	Âmbar Burmês	Fóssil L (uma fêmea AMNH JZC-Bu36)
† <i>Cretospilomena</i>	Fêmea	Âmbar Burmês	† <i>Cretospilomena familiaris</i> Antropov (uma fêmea AMNH JZC-Bu46)
<i>Passaloecus</i>	Fêmea	Âmbar Báltico	<i>Passaloecus</i> sp. (uma fêmea DZUP Bal-004; uma fêmea DZUP Bal-005; uma fêmea DZUP Bal-012; uma fêmea DZUP Bal-019)
† <i>Psolimena</i>	Female	Âmbar Jersiano I	† <i>Psolimena electra</i> Antropov (Holótipo, fêmea AMNH NJ-268)
Bembicidae			
Fóssil M	Fêmea	Âmbar Burmês	Fóssil M (uma fêmea DZUP Bur-075)
Fóssil N	Macho	Âmbar Burmês	Fóssil N (um macho DZUP Bur-49)
Fóssil O	Fêmea	Âmbar Burmês	Fóssil O (uma fêmea DZUP Bur-338)
† <i>Palanga</i>	Fêmea	Âmbar Báltico	† <i>Palanga succinea</i> Budrys (Holótipo, fêmea PIN No 364/225)

Tabela 2. Continuação. (3/3).

Terminais	Sexos usados	Depósito	Material usado para compor os terminais
“†Cirrosphecidae”			
† <i>Cirrosphe</i>	Fêmea/Macho	Âmbar Burmês	† <i>Cirrosphe</i> <i>admirabilis</i> Antropov (uma fêmea DZUP Bur-380; uma fêmea DZUP Bur-474; um macho DZUP Bur-087; um macho AMNH Bu-KL-18-27)
Fóssil P	Fêmea	Âmbar Burmês	Fóssil P (uma fêmea DZUP Bur-438)
Fóssil Q	Fêmea	Âmbar Burmês	Fóssil Q (uma fêmea DZUP-318)
Apidae lato sensu			
† <i>Succinapis</i>	Fêmea	Âmbar Báltico	† <i>Succinapis</i> sp. (uma fêmea DZUP Bal-013)
† <i>Cretotrigona</i>	Fêmea	Âmbar Jersiano II	† <i>Cretotrigona</i> <i>prisca</i> (Michener & Grimaldi 1988)
Referências usadas para a codificação: Michener & Grimaldi (1988), Engel (2000) e Engel (2001)			

As estimativas de idade de cada depósito foram estabelecidas segundo a plataforma de dados paleontológicos *Fossilworks* (Alroy 2018). A Tabela 3 mostra as idades usadas no presente estudo bem como a idade média de cada um deles. A idade geológica de cada depósito foi estabelecida seguindo Walker et al. (2013).

Tabela 3. Idade dos depósitos conforme a base de dados *Fossilworks*. **Ma:** Milhões de anos.

Depósito	Idade estabelecida	Média	Idade geológica (Período, Época)
Âmbar Dominicano	20.4 – 13.7 Ma	17,05 Ma	Burdigaliano (Neógeno, Mioceno)
Âmbar Báltico	37.2 – 33.9 Ma	35,55 Ma	Priaboniano (Paleógeno, Eoceno)
Âmbar Jersiano II	70.6 – 66.0 Ma	68,3 Ma	Maastrichtiano (Cretáceo Superior)
Âmbar Jersiano I	94.3 – 89.3 Ma	91,8 Ma	Turoniano (Cretáceo Superior)
Âmbar Burmês	99.7 – 94.3 Ma	97 Ma	Cenomaniano (Cretáceo Superior)
Formação Crato	122.46 – 112.6 Ma	117,53 Ma	Apitiano (Cretáceo Inferior)

2.2. Matriz de dados e delimitação dos caracteres

Foi usada como base a matriz de Melo (1999) com as modificações feitas em Melo & Rosa (em preparação). Em poucos casos alguns caracteres foram modificados das matrizes originais, sejam por novas interpretações sejam por equívocos de codificações. A lista dos caracteres foi mantida na língua inglesa como a original de Melo (1999) e encontra-se no Apêndice 2. A matriz de caracteres encontra-se no Apêndice 3. Os seguintes caracteres foram modificados da matriz original (condição prévia em parênteses seguido da nova codificação): *Ampulex* (92:0) para (92: não aplicável); *Aphilanthops* (42:0) para (42:2); †*Burmasphex* (48:0) para (48:?); *Chlorion* (157:0) para (157:1); *Ctenocolletes* (2:2) para (2:0); *Didineis* (63:1) para (63: não aplicável); *Dinetus* (38:0) para (38:1); *Dolichurus* (92:0) para (92: não aplicável); *Heliocausus* (158:0) para (158:2); *Hesperapis* (132:1) e (134:0) para (132:0) e (134:1) respectivamente; *Laphyragogus* (59:2) para (59:0); *Nysson* (158:1) para (158:2); *Ochleroptera* (62:0) e (107:?) para (62:1) e (107:0) respectivamente; *Palmodes* (138:1) e (139:?) para (138:0) e (139:0) respectivamente; *Philanthus* (2:2) para (2:0); *Stangeela* (21:1) para (21:2).

Novos caracteres foram interpretados a partir dos trabalhos de Evans (1966), Brothers & Carpenter (1993) e Roig-Alsina & Michener (1993). Essas referências foram citadas ao longo da lista de caracteres. Adicionalmente foram construídos novos caracteres seguindo os pressupostos estabelecidos na delineação de caracteres morfológicos. Respeitamos a independência e hierarquia dos caracteres e as condições únicas de cada estado de caráter (Rieppel & Kearney 2002; Sereno 2007). Cada caráter foi tratado como uma

hipótese de agrupamento por similaridade ou correspondência topológica (homologias primárias *sensu* de Pinna (1991)), testadas por congruência (de Pinna 1991) e então reconhecidas como homologia filogenética (homologia secundária) (Nixon e Carpenter 2012). Em alguns casos, os caracteres de múltiplos estados foram substituídos por caracteres de codificação contingente usando dependência lógica (Forey & Kitching 2000).

Para os caracteres de larva foram seguidas as mesmas referências apresentadas em Melo (1999) com novas inclusões apresentadas na Tabela 4 para as espécies adicionadas nesse estudo. Desde a publicação de Melo (1999), os gêneros *Aphelotoma*, *Ctenocolletes*, *Eremiasphecium*, *Laphyragogus*, *Xenosphex*, *Timberlakena*, *Parastigmus* e *Heterogyna* ainda permanecem sem informações publicadas. Os caracteres de comportamento também seguem as referências listadas por Melo (1999) com adição de Evans (1966) e Evans and O'Neill (2007) para os Bembicinae estudados.

Tabela 4. Lista da bibliografia usada para obtenção de informação referente aos caracteres de larva, com indicação dos terminais e respectivos táxons usados como referência. Terminais empregados por Melo (1999) foram omitidos.

Terminal	Espécie estudada	Referência
<i>Trielis</i>	<i>Scolia ruficornis</i> Fabricius	Evans (1987)
<i>Anthosila</i>	<i>Anthosila arithropyga</i> Burmeister	Genise (1988)
<i>Trimeria</i>	<i>Trimeria howardi</i> Bertoni	Zucchi et al. (1976)
<i>Arhysosage</i>	<i>Arhysosage flava</i> Moure	Rozen (2013)
<i>Eulaema</i>	<i>Eulaema polychroma</i> (Mocsáry)	Rozen (2018)
<i>Melipona</i>	<i>Melipona scutellaris</i> Latreille	Carvalho et al. (2017)
<i>Caupolicana</i>	<i>Caupolicana</i>	Rozen & Rozen (1986)
<i>Dieunomia</i>	<i>Dieunomia triangulifera</i> (Vachal, 1897)	Cross & Bohart (1960)
<i>Neofidelia</i>	<i>Neofidelia profuga</i> Moure & Michener	Rozen (1973)
<i>Hesperapis</i>	<i>Hesperapis rhodocera</i> (Cockerell)	Rozen (2016)
<i>Dryudella</i>	<i>Dryudella immigrans</i> (Williams)	Williams (1946)
<i>Diplopectron</i>	<i>Diplopectron brunneipes</i> (Cresson)	Evans (1975)
<i>Alysson</i>	<i>Alysson mellus</i> Say	Evans & Lin (1956)
<i>Zanysson</i>	<i>Nysson daeckei</i> Viereck	Evans & Lin (1956)
<i>Tiguipa</i>	<i>Heliocausus larroides</i> (Spinola)	Evans (1972)
<i>Stictia</i>	<i>Stictia carolina</i> (Fabricius)	Evans (1956)
<i>Glenostictia</i>	<i>Glenostictia scitula</i> (Fox)	Evans (1964)
<i>Ammatomus</i>	<i>Ammatomus icaoides</i> (Turner)	Evans (1983)
<i>Argogorytes</i>	<i>Argogorytes carbonarius</i> (Smith)	Harris (1994)
<i>Austrogorytes</i>	<i>Austrogorytes bellicosus</i> (Smith)	Evans & Matthews (1971)
<i>Gorytes</i>	<i>Gorytes canaliculatus</i> Packard	Evans & Lin (1956)
<i>Megistommum</i>	<i>Psammaecius costalis</i> (Cresson)	Evans (1959)
<i>Sagenista</i>	<i>Hoplisoides latifrons</i> (Spinola)	Evans & Lin (1956)
<i>Sphecius</i>	<i>Sphecius speciosus</i> (Drury)	Evans & Lin (1956)
<i>Oxybelus</i>	<i>Oxybelus quadrimaculatus</i> Say	Evans (1957)
<i>Tachytes</i>	<i>Tachytes crassus</i> Patton	Evans (1964)
<i>Miscophus</i>	<i>Miscophus evansi</i> Krombein	Evans (1964)
<i>Plenoculus</i>	<i>Plenoculus davisii</i> Fox	Evans (1959)
<i>Sericophorus</i>	<i>Sericophorus viridis</i> Saussure	Evans (1971)
<i>Solierella</i>	<i>Solierella blaisdelli</i> (Bridwell)	Evans (1958)
<i>Entomopison</i>	<i>Pison argentatum</i> Shuckard	Evans (1957)
<i>Pisonopsis</i>	<i>Pisonopsis birkmanni</i> Rohwer	Evans (1959)
<i>Spilomena</i>	<i>Spilomena enslini</i> Blüthgen	Evans (1959)
<i>Clypeadon</i>	<i>Clypeadon evansi</i> Bohart	Evans (1964)
<i>Cerceris</i>	<i>Cerceris fumipennis</i> Say	Evans (1957)
<i>Pseudoscolia</i>	<i>Pseudoscolia martinezi</i> Suárez	Asís et al. (1991)
<i>Penepodium</i>	<i>Podium luctuosum</i> Smith	Evans (1964)

2.3 Análises filogenéticas

2.3.1 Inferência Bayesiana

As análises de inferência Bayesiana foram conduzidas no software MrBayes versão 3.2.6 (Ronquist et al. 2012b) através do portal CIPRES Science Gateway (Miller et al. 2010). Os caracteres com dados faltantes foram representados por “?” e aqueles considerados inaplicáveis, seja por codificação contingente ou por diferença morfológica dos semaforantes, foram codificados como “-”. As análises de MCMC foram realizadas com 5×10^6 gerações, quatro corridas independentes (*nrun*) e *burnin* de 25%. As temperaturas das cadeias foram ajustadas empiricamente a fim de garantir uma mistura adequada. Foi atribuído uma temperatura final de 0.025. As convergências das cadeias foram checadas no Tracer 1.6 (Rambaut et al. 2013), usando como parâmetros valores de ESS (*Estimated Sample Size*) maiores de 200 unidades e gráficos com distribuições de picos homogênea.

As árvores apresentadas foram produzidas usando-se um consenso com regra de maioria (*Contype=Halfcompat*). Esse procedimento de consenso é altamente recomendado para dados morfológicos (Brown et al. 2017; O'Reilly & Donoghue. 2018). Também se utilizou árvores com todos os grupos compatíveis adicionados (*Contype=allcompat*). Esse procedimento de consenso é comumente utilizado em dados morfológicos e principalmente naqueles utilizando relógio morfológico (Lee et al. 2014; Beck & Lee 2014; Bapst et al. 2016; Herrera & Dávalos. 2016; Matzke & Wright 2016; Puttick et al. 2016; Gavryushkina et al. 2017; Lee & Yates 2018). Todas as árvores são apresentadas com a probabilidade posterior como suporte de ramos.

2.3.2 Particionamento

O particionamento dos dados morfológicos foi realizado segundo o critério de homoplasias (Rosa et al. 2019, Anexo 1). Esse artigo representa uma importante parte desta tese e conduzido ao longo da mesma. No entanto como foi publicado antecipadamente optou-se por incluí-lo como um Anexo e não como um capítulo a parte. Os valores de homoplasias foram mensurados a partir de uma análise de parcimônia com pesagem implícita com valor de $k=3$ (*default*). Inicialmente cada valor de homoplasias foi considerado como uma partição independente. No total, o procedimento de particionamento resultou na sugestão de 25 partições. Para a árvore principal calculada por inferência Bayesiana, bem como para todos os testes de hipóteses de relacionamentos, foi empregado esse esquema de particionamento.

Para as análises de estimativa de tempo de divergência, devido à sua complexidade, optou-se por combinar algumas dessas partições otimizando consideravelmente o tempo total de análise. A partir de uma análise de inferência Bayesiana com as 25 partições foi calculado o valor multiplicador para cada partição (*rate multiplier per partition*) (Apêndice 5 Figura 1). Com a análise visual desses valores foram estipulados agrupamentos arbitrários sempre procurando atribuir reunir partições com valores multiplicadores próximos. Todos os modelos gerados foram submetidos a teste de modelos.

2.3.3 *Parcimônia*

Mesmo com as premissas sobre o desempenho incerto desse método em matriz com muitos dados faltantes, foram realizadas análises de parcimônia a fim de explorar comparativamente o desempenho desse método em nossos dados. As análises de parcimônia foram conduzidas no TNT versão 1.5 (Goloboff et al. 2008; Goloboff & Catalano 2016), com todos os caracteres tratados como não ordenados (Fitch 1971). Os caracteres para os quais não se tinha informação foram codificados como dados faltantes usando-se “?”. Os caracteres para os quais a codificação não se aplicava, seja por contingência ou por serem de semaforantes opostos, foram codificados como “*”.

Foram conduzidas análises com pesos iguais e com pesagem implícita. Para pesagem implícita (Goloboff 1993) foi utilizado o default do programa (K=3) e também um script proposto por Mirande (2009) para a busca do melhor valor de K para um determinado conjunto de dados. Para as três análises citadas acima foi utilizada uma busca heurística tradicional pelos cladogramas mais parcimoniosos, com os seguintes comandos: *Tradicional search*; *hold1000*; *mult*1000*; *hold/10*; e *multiple TBR+TBR*. Reconstrução dos estados ancestrais foi conduzida no software Winclada versão 1.0.8 (Nixon 2002), com otimizações não-ambíguas.

2.4 Testes de modelos

As análises desse trabalho foram conduzidas de maneira exploratória. Dessa forma foram realizados testes de hipóteses para as comparações de modelos de particionamento, testes de monofilia e também dos *priors* para a estimativa do tempo de divergência. Da mesma maneira que em Rosa et al. (2019; ver Anexo 1), os diferentes modelos foram comparados usando-se *Bayes factors*, interpretados de acordo com Kass & Raftery (1995), aqui denominado *Kass and Raftery statistics* (KRS). A verossimilhança marginal (MgL) foi calculada através do método de *stepping stone* – SS (Xie et al. 2011) no software MrBayes via CIPRES. Foram aplicados 50 passos (*steps*) na estimativa da MgL com 50×10^6 gerações em uma frequência amostral (*sampling frequency*) de 1000 unidades e com *burnin* de 25%. Segundo os critérios de Kass & Raftery, se os valores estiverem entre 0 e 2, as evidências contra H_0 “não valem mais do que uma simples menção”. Valores entre 2 e 6 indicam “evidência positiva” contra H_0 , de 6 a 10 devem ser interpretados como “evidência forte” e acima de 10, “evidência muito forte” contra o modelo com menor MgL.

Para os testes de hipóteses de relacionamentos filogenéticos foram seguidas as sugestões de Bergsten et al. (2013). A premissa dessa estratégia é que cada ponto de restrição (*constraint*) usado para simular uma hipótese funciona com um *prior* de viés muito forte que altera consideravelmente a MgL do modelo testado. Para evitar esse viés, em todos os testes foram restringidos os mesmos agrupamentos tanto para H_0 quanto para a hipótese testada, equiparando-se os *priors* de cada modelo. Por exemplo, dada uma relação (A + (B + (C + D))), em um teste de hipótese H_0 para C + D contra H_1 , com B + D, as restrições seriam H_0 (*constraint1* B + (*constraint2* C + D)) contra H_1 (*constraint1* C + (*constraint2* B + D)), dois *priors* de restrições atribuídos para ambas as hipóteses. Foram testadas cinco hipóteses divergentes para entender

como os dados respondiam a essas questões. Das hipóteses testadas uma está relacionada apenas com nossos dados e outras quatro relacionadas com trabalhos prévios.

1) **Posicionamento das linhagens basais.** O posicionamento do Fóssil A em relação aos demais Apoidea é importante para a compreensão morfológica e evolução do grupo como um todo. Esse fóssil possui muitas características plesiomórficas conjuntamente com outro fóssil, *†Burmasphex*, previamente estabelecido na base de Apoidea (Melo & Rosa, em preparação). Por outro lado, o Fóssil A também apresenta um relevante caráter compartilhado com os Apoidea *stricto sensu*. Testamos as hipóteses: H₀ Fóssil A grupo irmão de Apoidea *stricto sensu*; H₁ Fóssil A grupo irmão de *†Burmasphex*.

2) **Monofilia de Crabronidae lato sensu.** A monofilia dos crabronídeos foi discutida em Melo (1999) que defendeu a hipótese de que Crabronidae *lato sensu* seria uma linhagem monofilética e grupo irmão de Apidae *lato sensu*. Em contrapartida todos os trabalhos moleculares bem como o trabalho morfológico de Alexander (1992) apontaram para a parafilia de Crabronidae *lato sensu* em relação às abelhas. Por essa razão testamos: H₀ Crabronidae *lato sensu* parafilético em relação às abelhas; e H₁ Crabronidae *lato sensu* monofilético *sensu* Melo (1999).

3) **Grupo irmão das abelhas.** Testamos aqui a mais controversa hipótese de relacionamento de um grupo de vespas apoideas com as abelhas. A hipótese proposta por Sann et al. (2018) sugere que Ammoplaninae (Astatidae) seria grupo irmão das abelhas (Sann et al. 2018). Testamos os nossos resultados obtidos contra essa hipótese: H₀ nossos resultados, Ammoplaninae dentro de Astatidae e Psenini grupo irmão de Apidae *lato sensu*; contra H₁ Psenini + (Ammoplaninae + Apidae *lato sensu*) *sensu* Sann et al. (et al. 2018).

4) **Posicionamento de Mellinus.** Existe uma forte tendência para se considerar *Mellinus* como grupo irmão de Crabronidae *stricto sensu* (Evans 1964; Melo 1999). Porém em resultados moleculares apresentados por Peter et al. (2018) e Sann et al. (2018), esse táxon é tratado como uma linhagem independente. Testamos essas duas hipóteses: H₀ *Mellinus* como uma linhagem independente; H₁ *Mellinus* como grupo irmão de Crabronidae *stricto sensu*.

5) **Posicionamento de Sphecidae.** Os resultados conduzidos a partir de dados moleculares mostram que, além de parafilético em relação às abelhas, Crabronidae *lato sensu* também poderia ser polifilético. Nessas hipóteses, Crabronidae *stricto sensu* está mais relacionado com Sphecidae do que os demais crabronídeos e abelhas. Por outro lado, em trabalhos com dados morfológicos o clado Crabronidae *lato sensu* + Apidae *lato sensu* (mesmo que Crabronidae seja parafilético em relação às abelhas) é muito bem suportado (Alexander 1992a; Melo 1999). Portanto nós testamos H₀ a monofilia do clado Crabronidae *lato sensu* + Apidae *lato sensu* (nesse cenário, as restrições foram ajustadas para a parafilia de Crabronidae em relação às abelhas) contra H₁ a polifilia de Crabronidae *lato sensu* com o posicionamento Crabronidae *stricto sensu* + Sphecidae (Devebec et al. 2012; Branstatter et al. 2017) e ainda H₂ *Mellinus* + Crabronidae *stricto sensu* + Sphecidae (Sann et al. 2018).

2.5 Estimativa dos tempos de divergência

Para estimar os tempos de divergência das linhagens de Apoidea foi utilizada a abordagem de estimativas de tempos de divergência por *tip-dating*. Nessa abordagem assume-se que os pontos de calibração fazem parte da estimativa da árvore datada, ou seja, os fósseis estão inseridos na matriz morfológica em questão. Todas as análises foram conduzidas no MrBayes 3.2.6. No total foram estabelecidos 20 modelos diferentes, testados utilizando-se a abordagem de ‘teste de modelos’ descrita anteriormente.

O modelo com o melhor desempenho, aquele com o melhor KRS, foi submetido a uma análise de MCMC e os resultados escolhidos como hipótese central e usados na discussão principal. Adicionalmente, a fim de explorar o desempenho na estimativa de tempos de divergência de diferentes modelos de relógio (idade média, intervalos de confiança, estrutura filogenética, etc.) foram escolhidos os melhores modelos de cada perfil de relógio e submetidos ao mesmo procedimento em MCMC.

As análises de inferência Bayesiana estimando o tempo de divergência das linhagens foram conduzidas no software MrBayes versão 3.2.6 (Ronquist et al. 2012b) através do portal CIPRES Science Gateway (Miller et al. 2010). As análises de MCMC foram conduzidas com 20×10^6 gerações (para os relógios de ramos com IGR) ou 50×10^6 gerações (para os relógios de ramos com TK02), quatro corridas independentes e *burnin* de 25%. As temperaturas das cadeias foram ajustadas empiricamente a fim de garantir uma mistura adequada. Foi atribuído uma temperatura final de 0.025. As convergências das cadeias foram checadas no Tracer 1.6 (Rambaut et al. 2013) usando como parâmetros valores de ESS (*estimated sample size*) superior a 200 unidades e os gráficos com distribuições de picos homogênea.

Todas as árvores apresentadas foram produzidas com todos os grupos compatíveis adicionados (*Contype=allcompat*) e com a estimativa da probabilidade posterior como suporte de ramos. Adicionalmente foi realizada uma representação gráfica da densidade de sobreposições de árvores utilizando métodos de agrupamento hierárquico Bayesiano no software DensiTree versão 2.2.5 (Remco et al. 2014). Foi usada a árvore datada resultante do melhor modelo para estimativa do tempo de divergência.

2.5.1 Calibração, prior da raiz e restrições topológicas

Os 33 fósseis amostrados foram calibrados com um prior uniforme a partir das idades previstas na Tabela 3. Para os modelos que precisaram de um *prior* de raiz (restrição datada da raiz) foi atribuído um intervalo uniforme (*prset treeagepr = uniform (155.7, 161.2)*). Esse *prior* foi definido com base no fóssil mais antigo de Bethylonymoidea Rasnitsyn, que segundo o mesmo autor se refere a um grupo vespas Apocrita mais próxima de Aculeata (provavelmente grupo irmão de Aculeata) (Rasnitsyn 1975). Os fósseis de impressões de siltito mais antigos dessa linhagem são provenientes da formação Karabastau (Jurássico 161.2 – 155.7), da região de Shymkent, Cazaquistão (*Fossilworks*; Alroy 2018). Para os relógios de árvore do tipo uniforme o *prior* de raiz não é obrigatório como no caso de relógios que utilizam modelos *birth-death* (Health et al. 2014). Nesse caso foram testados modelos de relógios de árvore uniformes sem o *prior* de raiz.

Para todas as análises de estimativa de tempo de divergência foram atribuídas restrições topológicas dentro da subfamília Apinae (Apidae *lato sensu*). Essas restrições se referem à monofilia do clado das abelhas corbiculadas e do clado composto por Meliponini + Bombini. Essas restrições estão de acordo com as hipóteses moleculares previamente estabelecidas (Romiguier et al. 2015).

2.5.2 Prior de comprimento de ramos (*Priors on the branch rates*)

Os *priors* de comprimento de ramos são comumente divididos em *priors* estritos, aqueles que possuem uma única taxa de estimativa de tempos de divergência ao longo de toda a árvore, ou relaxados que possuem diferentes taxas ao longo dos ramos de uma árvore. Assim como em dados moleculares, os caracteres morfológicos possuem mudanças variáveis ao longo de uma linhagem (Close et al. 2015; King et al. 2017), bem como em diferentes grupos de caracteres (Rosa et al. 2019, Anexo 1). Essas variações refletem nos relógios dos ramos de uma árvore levando à conclusão de que dados morfológicos são acomodados de maneira muito mais eficiente em relógios relaxados do que estritos (Lee et al. 2014; King et al. 2017; Herrera & Dávalos 2016; Gavryushkina et al. 2017).

Partindo dessa premissa foram testados dois *priors* relaxados. Para o *prior* autocorrelacionado foi usado o modelo de Thorne & Kishino (2002) (TK02) que utiliza modelo de movimento Browniano, com o comando *prset clockvarpr = tk02*, atribuindo-se um *prior* $\exp(1)$. Para o *prior* não correlacionado foram usadas distribuições gamas independentes (Independent Gamma-rates – IGR), com o comando *prset clockvarpr = igr* (Ronquist et al. 2012a) e um *prior* de $\exp(10)$.

Para o *prior* de taxas de evolução (*Clockratepr*) foi utilizado o *default*, valor fixo de 1.0 (*fixed=1.0*), ou seja, o número de mudança de cada estado é correspondente a uma unidade de tempo (Ronquist et al. 2012a). Segundo Pyron (2017), em uma matriz de dados morfológicas onde todos os caracteres se comportam de maneira variável, pode-se assumir *a priori* que cada um deles mudem pelo menos uma vez na escala do tempo da árvore. No particionamento conduzido no presente estudo nenhum dos caracteres se comportou de maneira não informativa. Por essa razão usou-se $1/158.45$ (média da idade do *prior* da raiz da árvore) = 0.006311 mudanças por milhão de ano. Em uma distribuição lognormal esse valor corresponde à média de -1.217235 com um desvio padrão de expoente médio de $\exp(1.006331)$. Portanto, para comparações da eficiência do *prior default* foi usado o *prior* *clockratepr = lognorm (-1.217235, 1.006331)*.

2.5.3 Prior de modelo de árvore (*Priors on the tree model*)

Ao se estimar os tempos de divergência das linhagens em uma abordagem *tip-dating* é necessário a aplicação de um *prior* de modelo de árvore uma vez que as estimativas topológicas são um dos pontos importantes do procedimento. Resumidamente existem dois grandes grupos de *priors* de modelos de árvore, os chamados mecanicistas e os não-mecanicistas (Health et al. 2014; Matzke & Wright 2016). Nos *priors* não-mecanicistas não existem premissas “biológicas” do processo do qual a árvore é gerada, em outras palavras por exemplo, nesses modelos não existem parâmetros para acomodar as taxas de especiações e extinções durante o processo (Matzke & Wright 2016). Para essa abordagem é comumente

utilizado um *prior* de modelo de árvores do tipo uniforme, que considera todas as topologias possíveis como igualmente prováveis (Huelsenbeck et al. 2012). Para a simulação desse *prior* foi atribuído o código *prset brlenspr=clock:uniform* sem mais proposições, conforme Lee et al. (2014). Essa abordagem foi a mais utilizada em *tip-dating* com dados morfológicos até meados de 2015 (Matzke & Wright 2016).

Os *priors* de modelo de árvore mecanicistas por sua vez descrevem eventos biológicos como as taxas de especiação (e/ou de extinção) distribuídos ao longo do tempo (Health et al. 2014; Zhang 2016; Bromham et al. 2018). A premissa mais simplista é o modelo de Yule que assume apenas a diversificação das linhagens na estimativa de tempos de divergência (Aldous 2001; Steel & McKenzie 2001). Já o modelo *birth-death* assume não só as taxas de diversificação como também as de extinções ao longo de suas estimativas (Bromham et al. 2018). O *prior fossilized birth-death*, uma derivação do modelo *birth-death*, unifica terminais extintos (fósseis) e vivos em um único modelo macroevolutivo, eliminando a necessidade de *priors* de calibração ad hoc (Health et al. 2014). Por essa razão esse é o modelo mais indicado para análises de evidência total ou *tip-dating* apenas com dados morfológicos (Health et al. 2014; Matzke & Wright 2016; Zhang 2016; Pyron 2017; Bromham et al. 2018). Esse foi o *prior* de modelo de árvore escolhido para contrapor ao modelo uniforme anteriormente descrito.

Para tanto, foi atribuído *prset brlenspr=clock:fossilized* no MrBayes sob amostragem de diversificação. Conforme Zhang (2015) e Pyron (2017), foram atribuídos os *priors* para especiação (*prset speciationpr = exp(10)*), extinção (*extinctionpr = beta(1,1)*) e fossilização (*prset fossilizationpr = beta(1,1)*) do *default* do modelo. Para proporção amostral de grupos vivos foi calculado o número de terminais vivos pela proporção real de diversidade do grupo interno (Apoidea), ou seja, 99 terminais vivos por aproximadamente 27.500 espécies. Portanto, *prset sampleprob = 0.004* (0.0036). Por fim, o *prior prset samplestrat* possui três estratégias distintas e todas foram abordadas.

O *sampleStrat* é o parâmetro que define a estratégia sob a qual as espécies serão amostradas durante a análise. O parâmetro apresenta três estratégias: 1) “*Fossiltip*” assume que os táxons vivos são amostrados aleatoriamente e os táxons extintos são amostrados com base nos fósseis da matriz (no caso de um *tip-dating* estrito). Nesse *prior* os táxons extintos são os fósseis amostrados (*tips*) que após a amostragem são considerados como linhagens mortas e que não produziram mais descendentes; 2) “*random*”, que é o *default* do modelo, assume que os táxons vivos são amostrados aleatoriamente, enquanto que os fósseis são amostrados com taxas constantes e em série ao longo da árvore (*birth-death tree*) para que então possam ser considerados ‘*tips*’ ou ‘*ancestors*’; e por fim, 3) “*diversity*” que assume que os táxons vivos são amostrados para maximizar a diversidade enquanto os fósseis são amostrados aleatoriamente.

2.5.4 Comparações das estimativas de tempo de divergência dos diferentes modelos de relógio

Para entender as variações na estimativa do tempo de divergência dos diferentes modelos de relógio para os Apoidea, duas métricas foram usadas: a idade média de cada clado e o intervalo entre o mínimo e

máximo de 95%HPD (*Highest Posterior Density*, Maior densidade posterior). Essas métricas foram comparadas entre diferentes clados e diferentes modelos de relógios.

Relógios. Foram selecionados seis diferentes modelos com base em três *priors*. Essa escolha foi feita com base no padrão encontrado no melhor modelo. Foi sempre usada a taxa de relógio dos ramos em lognormal (*Clockratepr* lognormal) por ser melhor que a *default* na maioria dos modelos (veja resultados). Para os modelos com *prior* de relógio de árvores FBD foram escolhidos aqueles com *Samplestrat* (*fossiltip*). Para os modelos com *priors* de relógio de árvores uniformes foram escolhidos aqueles em que combinavam a designação do *prior* de raiz e aqueles modelos uniformes sem essa designação. Portanto, as combinações de *priors* escolhidos para as comparações seguiram os seguintes perfis: Modelo 6 *Brlenspr* (FBD), *Clockvarpr* (IGR), *Treeagepr* (lognormal); Modelo 8 *Brlenspr* (uniforme), *Clockvarpr* (IGR), *Treeagepr* (lognormal); Modelo 10 *Brlenspr* (uniforme), *Clockvarpr* (IGR), *Treeagepr* (sem *prior* raiz); Modelo 16 *Brlenspr* (FBD), *Clockvarpr* (TK02), *Treeagepr* (lognormal); Modelo 18 *Brlenspr* (uniforme), *Clockvarpr* (TK02), *Treeagepr* (lognormal); e Modelo 20 *Brlenspr* (uniforme), *Clockvarpr* (TK02), *Treeagepr* (sem *prior* raiz).

Clados. Foram escolhidos os principais clados referentes às principais linhagens de Apoidea, sempre levando em consideração a presença do mesmo clado na maior parte dos conjuntos de dados comparados. Os 26 clados (13 *stem* e 13 *crown*) comparados com os modelos de relógio morfológicos foram: Apoidea (*stem* e *crown*), Angarosphecidae I (*stem* e *crown*), Ampulicidae (*stem* e *crown*), Heterogynaidae (*stem* e *crown*), Sphecidae (*stem* e *crown*), Angarosphecidae II (*stem* e *crown*), Crabronidae *stricto sensu* (*stem* e *crown*), Bembicidae (*stem* e *crown*), Pemphredonidae (*stem* e *crown*), Astatidae (*stem* e *crown*), Philantidae (*stem* e *crown*), Psenidae (*stem* e *crown*), Apidae *lato sensu* (*stem* e *crown*).

Os relógios morfológicos também foram comparados com as hipóteses a partir de dados moleculares de Branstetter et al. (2017), Peters et al. (2017) e Sann et al. (2018). Para essas comparações foram escolhidos os 24 clados vivos (13 *stem* e 13 *crown*): Apoidea (*stem* e *crown*), Ampulicidae (*stem* e *crown*), Heterogynaidae (*stem* e *crown*), Sphecidae (*stem* e *crown*), Crabronidae *stricto sensu* (*stem* e *crown*), Bembicidae (*stem* e *crown*), Pemphredonidae (*stem* e *crown*), Astatidae (*stem* e *crown*), Philanthidae (*stem* e *crown*), Psenidae (*stem* e *crown*), Apidae *lato sensu* (*stem* e *crown*).

Além disso as médias das idades estimadas entre os modelos morfológicos e as hipóteses moleculares foram correlacionadas. Para essas correlações foi usado o pacote *Hmisc* do R (Harrell 2018). Os 16 clados correlacionados foram: Apoidea (*stem* e *crown*), Ampulicidae (*stem*), Sphecidae (*stem* e *crown*), Crabronidae *stricto sensu* (*stem* e *crown*), Bembicidae (*stem* e *crown*), Pemphredonidae (*stem* e *crown*), Philantidae (*stem* e *crown*), Psenidae (*stem*) e Apidae *lato sensu* (*stem* e *crown*).

2.6 Análises de linhagens ao longo do tempo (LTTs) e estatística- γ

Para entender os processos de surgimento e extinção das grandes linhagens de Apoidea durante o Cretáceo e principalmente durante o KTR foram realizados alguns procedimentos empíricos em

macroevolução. Para entender esses padrões através do tempo, através de uma árvore datada, as linhagens foram representadas graficamente usando gráficos LTTs (*Lineage Through Time plots*) e analisadas em estatística- γ (Pybus & Harvey 2000). As duas abordagens foram conduzidas utilizando os pacotes Ape (Paradis 2014) e Phytools (Revell 2012) implementados em R.

A nossa amostragem contém aproximadamente 0.36% (99 espécies) do total das espécies viventes de Apoidea (aproximadamente 27.500 espécies somando os dados de Michener 2007 e Pulwaski 2018). Outros trabalhos com números expressivos de terminais chegam a 0.49% (136 espécies) em Sann et al. (2018) e 0.50% (138 espécies) em Devebec et al. (2012) do total de espécies da superfamília. Em razão de que amostragens taxonômicas incompletas interferem negativamente nessas abordagens (Pybus & Harvey 2000), nós optamos por utilizar apenas as grandes linhagens de Apoidea a nível de subfamília/ família. Foi mantido um único terminal em cada linhagem e sempre mantendo um terminal vivente para aquelas linhagens que ainda estão presentes nos dias de hoje. A amostragem usada representa 100% das grandes linhagens viventes a nível de subfamília/ família. Como os nossos objetivos focaram na diversificação inicial das grandes linhagens (mais antigas) essa amostragem foi considerada suficiente.

Depois dessa poda inicial (*terminal pruning delimitation*) retiramos o grupo externo e realizamos todas as análises estabelecendo o seguinte delineamento amostral: 1) para estabelecer os efeitos das linhagens extintas foram conduzidas todas as análises alternando uma árvore incluindo e outra excluindo os terminais fósseis. 2) os LTTs gerados foram comparados com 500 simulações realizadas através do modelo *Pure-Birth* (PB) e *Birth-Death* (BD) com as premissas: a) taxa de especiação constante e sem extinção (*Pure-Birth*); b) taxa de extinção constante com o dobro de especiação (*Birth-Death*); e c) taxa de extinção constante e com o triplo de especiação (*Birth-Death*). 3) realizamos estatística- γ (Pybus & Harvey 2000) comparando com as mesmas 500 simulações realizadas no procedimento anterior seguindo os critérios de interpretação: a) $\gamma = 0$, indica que os nós têm uma distribuição balanceada, em relação às expectativas simuladas em PB; b) $\gamma < 0$, sugere que a taxa de diversificação foi inicialmente alta, mas depois diminuiu com o tempo; e c) $\gamma > 0$ sugere que a taxa de diversificação foi inicialmente lenta e aumentou ao longo do tempo.

3. Resultados

3.1 Análises filogenéticas

3.1.1 Análises em inferência Bayesiana

As análises de inferência Bayesiana realizadas com dados morfológicos não trouxeram dificuldades na convergência das corridas. Todas as corridas tiveram valores de ESS superiores a 200 e foram bem-sucedidas em explorar o espaço amostral. A Figura 1 mostra as relações encontradas para os Apoidea em uma análise com 25 partições e com todos os grupos compatíveis adicionados (*Contype=allcompat*). Os clados indicados são discutidos na seção 4.2 *Discussão dos principais clados e os aspectos morfológicos*. A Figura 2 apresenta o mesmo perfil analítico, mas com um consenso em regra de maioria (*Contype=Halfcompat*).

3.1.2 Particionamento

O particionamento dos dados dos dados morfológicos resultou em sete modelos diferentes (Tabela 5) com base no valor multiplicador de cada partição (Apêndice 5 Figura 1). O esquema de particionamento com o melhor valor de MgL foi o Modelo 4 com nove partições. O Modelo sem particionamento foi fortemente rejeitado em relação ao melhor modelo (KRS=1348.22).

Tabela 5: Modelos de particionamento com os respectivos valores de verossimilhança marginal (MgL) e *Kass and Rafilery statistics* (KRS). Em negrito o melhor modelo (de maior verossimilhança marginal). N: número de partições incluídas para cada modelo.

Modelos	Partições	N partições	MgL	KRS
Modelo 0	Sem particionamento	0	-7882.34	1348.22
Modelo 1	Todas as categorias do Apêndice 5 Figura 1	25	-7238.74	61.02
Modelo 2	1; 2; 3; 4-7; 8-12; 13-15; 16-25	7	-7262.3	108.14
Modelo 3	1; 2; 3; 4-7; 8-12; 13-15; 16-19; 20-25	8	-7250.22	83.98
Modelo 4	1; 2; 3; 4-7; 8-12; 13-15; 16-19; 20-24; 25	9	-7208.23	0
Modelo 5	1; 2; 3; 4-5; 6-7; 8; 9-10; 11; 12; 13-15; 16-25	11	-7265.9	115.34
Modelo 6	1; 2; 3; 4-5; 6-7; 8; 9-10; 11; 12; 13-15; 16-20; 21-25	12	-7254.68	92.9
Modelo 7	1; 2; 3; 4-5; 6-7; 8; 9-10; 11; 12; 13-15; 16-20; 21-24; 25	13	-7215.37	14.28

3.1.3 Análises sob parcimônia

A análise de parcimônia com pesos iguais resultou em um comprimento de 1541 passos, 380 árvores igualmente parcimoniosas, 2214 passos, índice de consistência de 13 e um índice de retenção igual a 66. O consenso estrito das árvores resultantes é mostrado no Apêndice 4 Figura 1. A árvore com valor de K=3 (*default*) resultou em uma única árvore (Apêndice 4 Figura 2) com um *best score* de 94.49, 2089 passos, índice de consistência de 14 e um índice de retenção igual a 68. E por fim, a análise com o valor melhor valor de K sugerido para esse conjunto de dados, K= 27.226563, resultou em uma única árvore mais parcimoniosa (Apêndice 4 Figura 3) com *best score* igual a 31.60, 1995 passos, índice de consistência de 14 e um índice de retenção igual a 70.

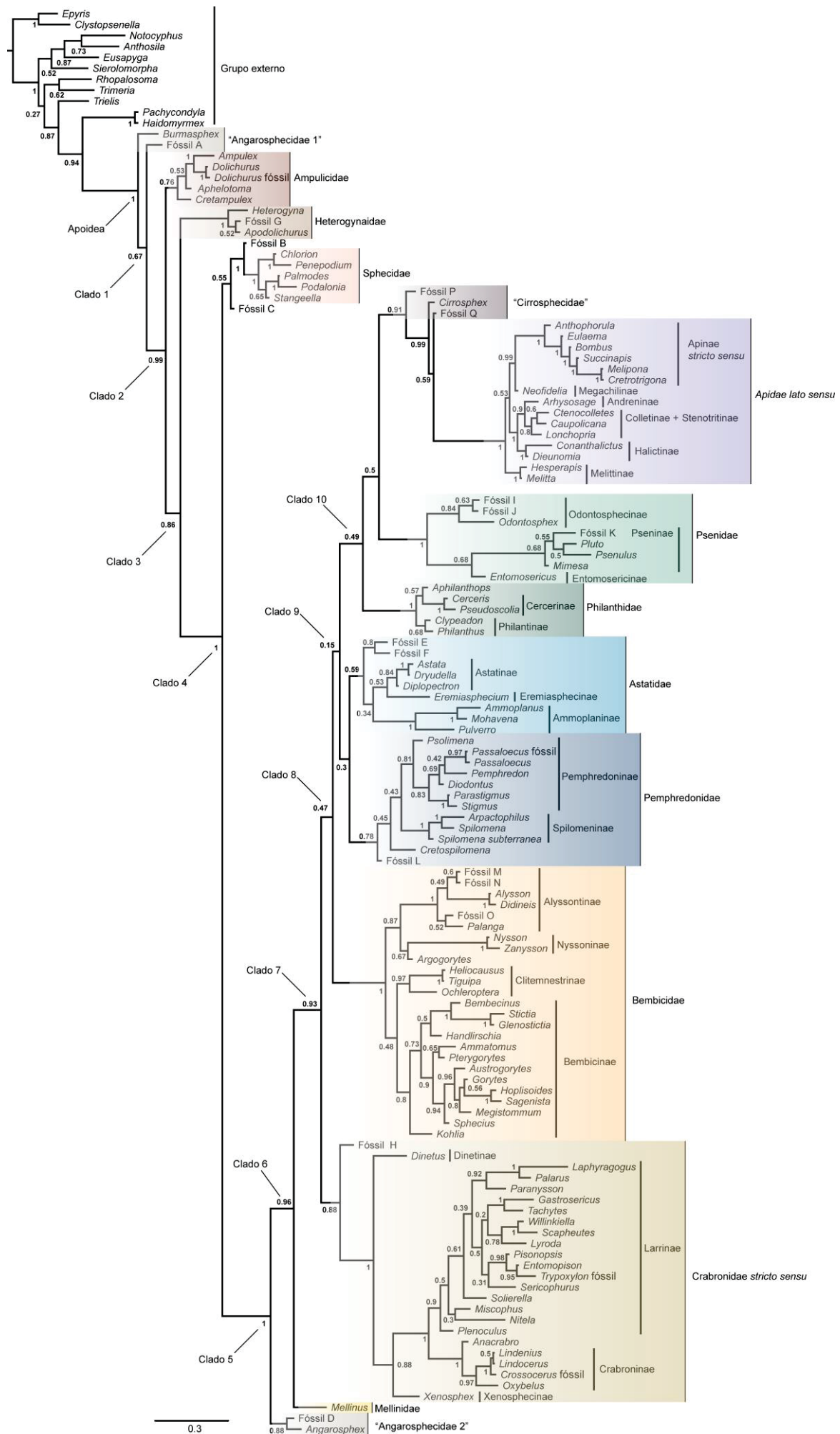


Figura 1. Relações filogenéticas de Apoidea com base em dados morfológicos em análise de inferência Bayesiana com 25 partições. Todos os grupos compatíveis adicionados (*Contype=allcompat*). Probabilidade posterior como suporte dos ramos.

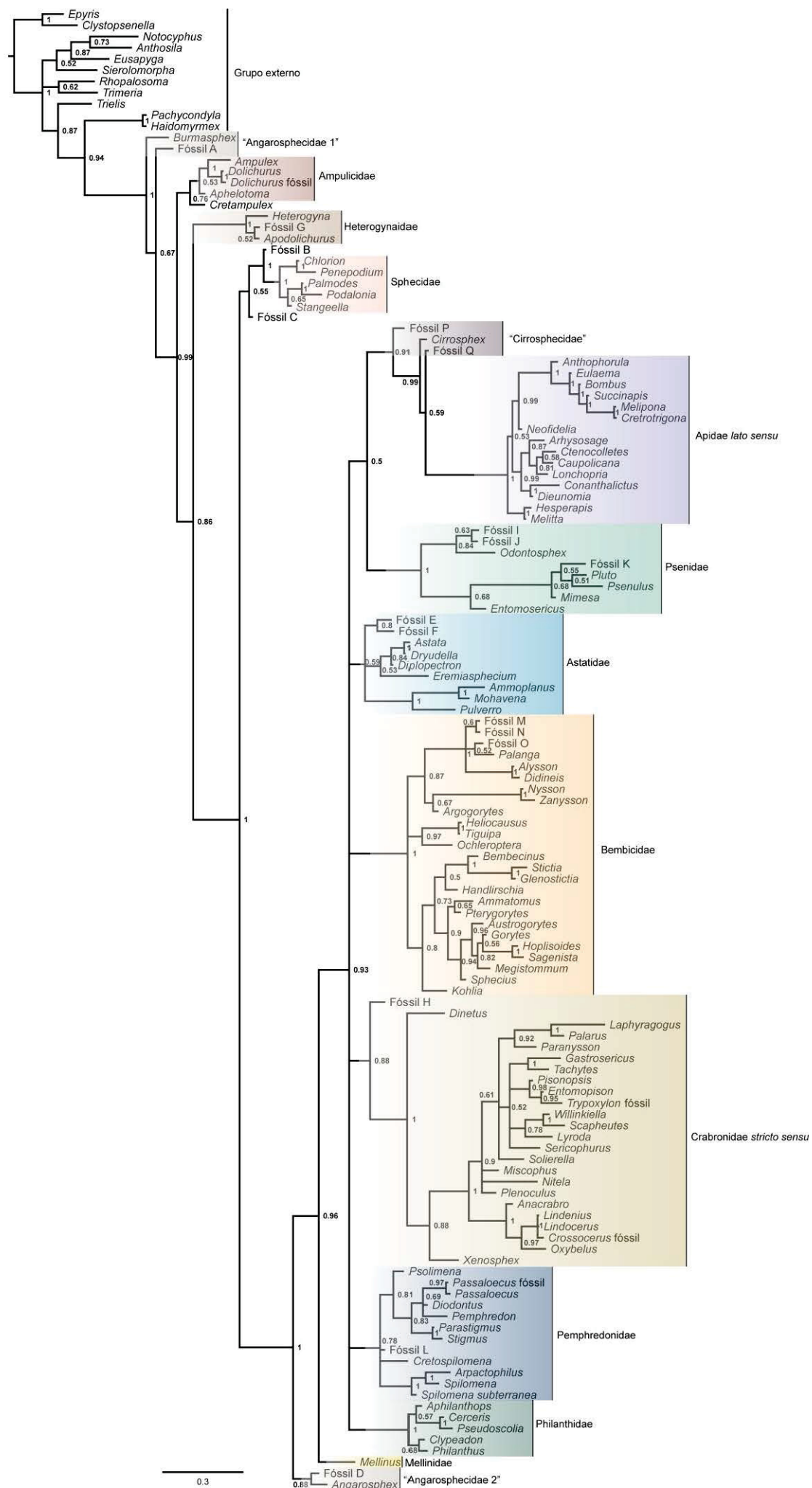


Figura 2. Relações filogenéticas de Apoidea com base em dados morfológicos em análise de inferência Bayesiana com 25 partições. Consenso em regra de maioria (*Contype=Halfcompat*). Probabilidade posterior como suporte dos ramos.

3.1.4 Testes de hipóteses topológicas

Todos os testes topológicos descritos na seção 2.4 *Testes de modelos* estão na Tabela 6 com seus respectivos valores de MgL e KRS. Nas relações dos grupos basais de Apoidea foi testado o posicionamento do Fóssil A em relação a *†Burmasphex* e a aos Apoidea *stricto sensu*. Em nossos resultados (Figura 1 e 2), o Fóssil A está relacionado com os Apoidea *sensu stricto* (H₀). A H₁ foi fortemente rejeitada (KRS=34.66).

Os resultados apresentados nas Figuras 1 e 2 mostraram a parafilia de Crabronidae *lato sensu* em relação aos Apidae *lato sensu*. A hipótese proposta em de Melo (1999) (H₀) onde os Crabronidae *lato sensu* formam um grupo monofilético foi fortemente rejeitada (KRS=371.64). As relações encontradas nas Figuras 3 e 4 sugerem que o grupo vivo de vespas apoideas mais próximo das abelhas são os Psenidae (H₁). A hipótese proposta em Sann et al. (2008) em que o grupo vivo mais próximo das abelhas são os Ammoplaninae (Astatidae) (H₀) foi fortemente rejeitada (KRS=10.74).

A proposta de que *Mellinus* está diretamente associado com os Crabronidae *stricto sensu* (Evans 1964; Melo 1999) foi fortemente rejeitada (KRS= 13.94). Em nossos resultados o gênero *Mellinus* (Mellinidae) representa uma linhagem independente das demais linhagens de vespas apoideas (Figura 1 e 2). Por fim, as hipóteses que sugerem que os Sphecidae estão relacionados com os Crabronidae *stricto sensu* foram fortemente rejeitadas (KRS=384.14 e KRS=377.04). Em nossos resultados os Crabronidae *stricto sensu* formam um clado com os demais crabronídeos e as abelhas.

Tabela 6. Testes de hipóteses topológicas realizados nas relações filogenéticas de Apoidea com seus respectivos valores de verossimilhança marginal (MgL) e *Kass and Raftery statistics* (KRS).

Hipóteses	MgL	KRS
Posicionamento do Fóssil A		
H ₀ Fóssil A + Apoidea <i>stricto sensu</i>	-7200.78	0
H ₁ Fóssil A + <i>†Burmasphex</i>	-7235.44	34.66
Monofilia de Crabronidae		
H ₀ Crabronidae <i>parafilético</i>	-7003.84	0
H ₁ Crabronidae <i>monofilético</i>	-7162.66	317.64
Grupo irmão das abelhas (grupos vivos)		
H ₀ Psenini + Apidae <i>lato sensu</i>	-7176.53	0
H ₁ Psenini + (Ammoplanini + Apidae <i>lato sensu</i>)	-7181.9	10.74
Posicionamento de <i>Mellinus</i>		
H ₀ <i>Mellinus</i> como linhagem independente	-7003.84	0
H ₁ <i>Mellinus</i> como grupo irmão de Crabronidae ss.	-7010.81	13.94
Posicionamento de Crabroninae e Sphecidae		
H ₀ Crabronidae <i>sensu stricto</i> irmão dos demais “Crabronidae” <i>lato sensu</i>	-7003.84	0
H ₁ Crabronidae ss. + Sphecidae	-7195.91	384.14
H ₂ Crabroninae ss. + Sphecidae + <i>Mellinus</i>	-7192.36	377.04

3.2 Estimativas do tempo de divergência

3.2.1 Modelos testados

A combinação de diferentes *priors* para a estimativa do tempo de divergência das linhagens de Apoidea resultaram em 20 modelos diferentes (Tabela 7). A melhor combinação de *priors* para os nossos dados foi encontrada no modelo 6, com *Brlenspr* (FBD), *Samplestrat* (fossiltip), *Clockvarpr* (IGR), *Clockratepr* (lognormal) e *Treeagepr* (lognormal). Com exceção de dois modelos (12 e 15) todos as combinações que utilizaram as taxas de relógio de ramos (*Clockratepr*) com distribuição lognormal foram significativamente superiores às taxas *default*. Exceto pelas taxas de relógio de ramos (*Clockratepr*) os outros *priors* não demonstraram tendências no aumento da MgL isoladamente.

Por outro lado, os modelos FBD + *diversity* com uma taxa de relógio de ramos fixa (*default*) foram sistematicamente interrompidas durante o primeiro *step* no processo de *stepping-stone*. Esse erro pode ter ocorrido devido a uma falha na implementação dessa combinação de *priors* no MrBayes via CIPRES (Mark Miller, comunicação pessoal) ou ainda uma falha na implementação do próprio MrBayes (testes conduzidos em diferentes máquinas por longas horas resultaram no mesmo erro). De qualquer maneira não foi possível estimar a eficiência dos modelos 1 e 11.

Tabela 7. Modelos testados para a estimativa do tempo de divergência das linhagens de Apoidea com seus respectivos valores de verossimilhança marginal (MgL) e *Kass and Raftery statistics* (KRS). N/A: não aplicável. FBD: *Fossilized birth death*. IGR: Independent Gamma-rates. TK02: Thorne & Kishino (2002). Em negrito o melhor modelo. *modelos sem designações de *prior* de raiz (*Treeagepr*).

Modelos	Prior relógio (árvores) <i>Brlenspr</i>	<i>Samplestrat</i>	Prior relógio (ramos) <i>Clockvarpr</i>	Taxa relógio (ramos) <i>Clockratepr</i>	MgL	KRS geral
Modelo 1	FBD	diversity	IGR	default	Erro	Erro
Modelo 2	FBD	diversity	IGR	lognormal	-7354.95	96.24
Modelo 3	FBD	random	IGR	default	-7594.75	575.84
Modelo 4	FBD	random	IGR	lognormal	-7324.94	36.22
Modelo 5	FBD	fossiltip	IGR	default	-7488.04	362.42
Modelo 6	FBD	fossiltip	IGR	lognormal	-7306.83	0.00
Modelo 7	Uniforme	N/A	IGR	default	-7474.85	336.04
Modelo 8	Uniforme	N/A	IGR	lognormal	-7319.33	25.00
Modelo 9*	Uniforme	N/A	IGR	default	-7503.44	393.22
Modelo 10*	Uniforme	N/A	IGR	lognormal	-7340.97	68.28
Modelo 11	FBD	diversity	TK02	default	Erro	Erro
Modelo 12	FBD	diversity	TK02	lognormal	-7464.87	316.08
Modelo 13	FBD	random	TK02	default	-7439.98	266.30
Modelo 14	FBD	random	TK02	lognormal	-7467.4	321.14
Modelo 15	FBD	fossiltip	TK02	default	-7464.22	314.78
Modelo 16	FBD	fossiltip	TK02	lognormal	-7347.54	81.42
Modelo 17	Uniforme	N/A	TK02	default	-7426.84	240.02
Modelo 18	Uniforme	N/A	TK02	lognormal	-7349.24	84.82
Modelo 19*	Uniforme	N/A	IGR	default	-7446.33	279.00
Modelo 20*	Uniforme	N/A	IGR	lognormal	-7375.50	137.34

3.2.2 Estimativas do tempo de divergência em MCMC

Nas análises de estimativa de tempo de divergência das linhagens de Apoidea não se detectou dificuldades na convergência das corridas. As análises com os modelos de relógio de ramos usando TK02 necessitaram corridas mais longas que aquelas usando IGR (50x10⁶ gerações e 20x10⁶ gerações respectivamente). Porém todas as corridas tiveram valores de ESS superiores a 200 e foram bem sucedidas em explorar o espaço amostral.

O resultado do melhor modelo de relógio (modelo 6) está representado na Figura 3 com as idades médias de cada clado, bem como a indicação em barras dos valores máximos e mínimos com 95% HPD (*Highest Posterior Density*, Maior densidade posterior). A Figura 4 apresenta o mesmo cronograma, porém com os valores de probabilidade posterior para cada clado. De maneira geral as análises conduzidas estimando o tempo de divergência entre as linhagens resultaram em topologias ligeiramente diferentes daquela apresentada na Figura 1. Apenas para manter a coerência durante a discussão dos relacionamentos filogenéticos entre as linhagens de Apoidea através do tempo foram realizadas restrições topológicas no Modelo 6 tornando a topologia mais próxima daquela apresentada na Figura 1 (veja Apêndice 7.8). O cronograma da Figura 5 é resultado de 20 restrições topológicas com base nos resultados de inferência Bayesiana.

A Tabela 8 resume as médias de idades encontradas para as grandes linhagens de Apoidea bem como os valores máximos e mínimos em 95% HPD. A diferença mais significativa encontrada nos diferentes modelos foi o efeito do *prior* de raiz (*Treeagepr*). Os modelos sem atribuição do *prior* de raiz (Modelo 10 e 20) resultaram em estimativas mais recentes que aqueles com o *prior* designado, principalmente nas divergências na base da árvore. Os modelos com *prior* de relógio de árvores TK02 resultaram em estimativas com menor amplitude de intervalo (95%HPD) quando comparados com os modelos IGR. Por outro lado, os modelos sem atribuição de *prior* de raiz não apresentaram diferenças no intervalo quando combinados com TK2 ou IGR.

Os nossos resultados comparados com as hipóteses moleculares estão apresentados na Tabela 9. Comparativamente, os intervalos de 95% HPD foram muito maiores para as hipóteses moleculares (Branstatter et al. 2017; Peters et al. 2017; Sann et al. 2018) que aqueles encontrados para os modelos morfológicos. Por fim, as médias das idades estimadas para os 16 clados de Apoidea apresentaram forte correlação ($R^2 > 0.61$) entre as 10 hipóteses avaliadas (os sete modelos morfológicos e as três hipóteses moleculares) (Apêndice 6 Figura 1). Essa correlação foi significativa para a maior parte dos resultados ($p > 0.05$). Foram encontradas correlações não significativas nos resultados de Branstetter et al. (2017) em relação à maioria dos modelos morfológicos (todos exceto Modelo 16).

Tabela 8. Tabela com os diferentes modelos de tempo de divergência para as principais linhagens de Apoidea com as médias de idades estimadas bem como os valores máximos e mínimos em 95% HPD. Inter: intervalo total de 95%HPD.

Clado	Modelo 6 (Melhor modelo) (FBD+IGR+com <i>prior</i> de raiz)					Modelo 8 (UNI+IGR+ com <i>prior</i> de raiz)					Modelo 10 (UNI+IGR+ sem <i>prior</i> raiz)				
	media	min	max	inter		media	min	max	inter		media	min	max	inter	
Apoidea <i>stem</i>	153.34	152.15	159.81	7.66		157.72	154.59	160.35	5.76		148.07	136.82	158.76	21.22	
Apoidea <i>crown</i>	154.51	149.58	158.87	9.29		156.28	152.73	159.39	6.66		145.92	143.90	156.37	20.35	
Angarosphecidae I <i>stem</i>	154.51	149.58	158.87	9.29		138.94	123.96	153.70	29.74		132.37	117.97	147.05	24.41	
Angarosphecidae I <i>crown</i>	120.04	100.62	141.45	40.83		121.08	123.96	153.70	29.74		117.77	100.82	135.92	24.29	
Ampulicidae <i>stem</i>	140.75	124.64	153.91	29.27		138.94	123.96	153.70	29.74		132.37	117.97	147.05	24.41	
Ampulicidae <i>crown</i>	115.03	98.45	134.11	35.66		120.03	101.29	140.67	39.38		116.50	100.01	134.13	26.64	
Heterogynaidae <i>stem</i>	140.75	124.64	153.91	29.27		150.67	141.83	157.88	16.05		141.51	129.95	153.89	20.41	
Heterogynaidae <i>crown</i>	112.66	99.77	127.99	28.22		116.71	101.25	135.53	34.28		114.64	100.67	130.54	23.82	
Sphecidae <i>stem</i>	103.67	95.28	114.63	19.35		106.52	95.52	121.48	25.96		105.40	95.43	118.40	21.02	
Sphecidae <i>crown</i>	59.18	30.12	86.68	56.56		67.22	37.77	97.69	59.92		67.12	37.76	97.02	65.13	
Angarosphecidae II <i>stem</i>	148.33	141.93	154.5	12.57		152.31	147.75	156.50	8.75		142.52	132.27	153.24	17.30	
Angarosphecidae II <i>crown</i>	129.12	116.4	142.88	26.48		130.13	116.25	145.12	28.87		127.16	115.58	139.36	11.81	
Crabronidae ss. <i>stem</i>	139.82	130.03	148.45	18.42		146.80	140.00	153.40	13.40		137.19	126.64	148.20	17.30	
Crabronidae ss. <i>crown</i>	129.42	114.57	143.48	28.91		139.87	128.08	150.56	22.48		131.41	117.84	144.84	22.01	
Bembicidae <i>stem</i>	139.82	130.03	148.45	18.42		146.80	140.00	153.40	13.40		137.19	126.64	148.20	17.96	
Bembicidae <i>crown</i>	129.8	116.22	142.59	26.37		139.54	128.00	150.37	22.37		131.81	119.03	144.29	19.69	
Pemphredonidae <i>stem</i>	138.18	128.22	147.96	19.74		144.80	136.29	152.28	15.99		135.87	124.35	147.79	17.30	
Pemphredonidae <i>crown</i>	124.99	110.18	139.36	29.18		137.28	124.46	149.30	24.84		129.70	116.59	143.00	20.12	
Astatidae <i>stem</i>	138.18	128.22	147.96	19.74		144.80	136.29	152.28	15.99		135.87	124.35	147.79	17.30	
Astatidae <i>crown</i>	126.99	112.01	141.47	29.46		131.44	114.85	147.39	32.54		125.70	110.79	140.75	20.76	
Phlanthidae <i>stem</i>	131.95	121.24	142.49	21.25		138.97	128.92	148.34	19.42		131.48	119.79	142.68	17.76	
Phlanthidae <i>crown</i>	54.42	27.71	84.15	56.44		64.70	32.85	100.09	67.24		63.39	31.43	96.77	45.59	
Psenidae <i>stem</i>	131.95	121.24	142.49	21.25		138.97	128.92	148.34	19.42		131.48	119.79	142.68	17.76	
Psenidae <i>crown</i>	125.92	114.23	138.04	23.81		132.64	120.48	144.66	24.18		126.48	114.46	138.13	16.31	
Apidae ls. <i>stem</i>	107.37	97.71	118.41	20.7		112.66	99.26	126.48	27.22		118.38	105.40	131.68	18.08	
Apidae ls. <i>crown</i>	98.25	85.02	111.65	26.63		104.00	87.61	121.12	33.51		105.14	90.61	120.24	19.48	
	Média do intervalo				25.09	Média do intervalo				24.51	Média do intervalo				20.24

Tabela 8. Continuação. (2/2). Tabela com os diferentes modelos de tempo de divergência para as principais linhagens de Apoidea com as médias de idades estimadas bem como os valores máximos e mínimos em 95% HPD. Inter: intervalo total de 95%HPD.

Clado	Modelo 16				Modelo 18				Modelo 20			
	(FBD+TK02+com <i>prior</i> raiz)				(UNI+TK02+com <i>prior</i> raiz)				(UNI+TK02+sem <i>prior</i> raiz)			
	media	min	max	inter	media	min	max	inter	media	min	max	inter
Apoidea <i>stem</i>	159.34	156.86	161.14	4.28	157.56	154.31	160.31	6.00	143.34	133.35	154.57	21.22
Apoidea <i>crown</i>	152.62	147.67	157.42	9.75	153.86	149.61	157.69	8.08	140.42	130.51	150.86	20.35
Angarosphecidae I <i>stem</i>	132.06	117.54	145.55	28.01	138.19	124.94	150.28	25.34	128.42	116.11	140.52	24.41
Angarosphecidae I <i>crown</i>	111.54	99.59	125.92	26.33	115.84	100.88	131.96	31.08	111.31	99.82	124.11	24.29
Ampulicidae <i>stem</i>	132.06	117.54	145.55	28.01	138.19	124.94	150.28	25.34	128.42	116.11	140.52	24.41
Ampulicidae <i>crown</i>	117.50	102.37	132.86	30.49	126.46	110.13	141.27	31.14	119.08	105.54	132.18	26.64
Heterogynaidae <i>stem</i>	147.48	139.49	154.89	15.40	150.06	143.61	156.01	12.40	137.43	127.61	148.02	20.41
Heterogynaidae <i>crown</i>	110.22	98.52	124.25	25.73	115.55	100.13	131.15	31.02	110.37	98.84	122.66	23.82
Sphecidae <i>stem</i>	102.52	94.80	113.30	18.50	107.13	95.19	122.91	27.72	104.24	94.96	115.98	21.02
Sphecidae <i>crown</i>	70.81	40.81	96.96	56.15	82.74	45.82	121.92	76.10	79.97	47.64	112.77	65.13
Angarosphecidae II <i>stem</i>	132.15	125.13	139.05	13.92	139.06	133.15	144.94	11.79	132.00	123.62	140.92	17.30
Angarosphecidae II <i>crown</i>	120.12	113.81	127.30	13.49	122.35	114.42	130.65	16.23	118.81	113.37	125.18	11.81
Crabronidae ss. <i>stem</i>	135.27	127.89	142.05	14.16	140.80	134.82	146.72	11.90	132.00	123.62	140.92	17.30
Crabronidae ss. <i>crown</i>	123.39	111.24	134.90	23.66	133.24	123.12	142.42	19.30	123.02	112.08	134.09	22.01
Bembicidae <i>stem</i>	141.04	134.84	147.09	12.25	145.20	140.08	150.22	10.14	133.57	124.94	142.90	17.96
Bembicidae <i>crown</i>	121.31	110.76	131.59	20.83	130.36	120.59	139.83	19.24	121.57	111.88	131.57	19.69
Pemphredonidae <i>stem</i>	133.47	125.14	141.19	16.05	139.41	132.34	146.15	13.81	132.00	123.62	140.92	17.30
Pemphredonidae <i>crown</i>	116.94	106.75	127.78	21.03	125.61	113.50	137.04	23.54	116.03	106.18	126.30	20.12
Astatidae <i>stem</i>	133.47	125.14	141.19	16.05	139.41	132.34	146.15	13.81	132.00	123.62	140.92	17.30
Astatidae <i>crown</i>	121.50	110.09	141.19	31.10	131.18	119.97	141.39	21.42	121.96	111.42	132.18	20.76
Philanthidae <i>stem</i>	127.59	119.89	135.25	15.36	128.46	118.34	137.89	19.55	127.44	118.75	136.51	17.76
Philanthidae <i>crown</i>	74.47	41.50	104.02	62.52	94.70	66.15	119.81	53.66	91.91	67.53	113.12	45.59
Psenidae <i>stem</i>	127.59	119.89	135.25	15.36	133.51	126.72	140.36	13.64	127.44	118.75	136.51	17.76
Psenidae <i>crown</i>	122.42	114.49	130.14	15.65	126.57	118.66	134.40	15.74	119.61	111.59	127.90	16.31
Apidae ls. <i>stem</i>	123.48	114.40	132.46	18.06	137.54	131.15	143.82	12.67	121.24	112.28	130.36	18.08
Apidae ls. <i>crown</i>	102.67	93.06	112.67	19.61	112.38	102.09	122.40	20.31	105.88	96.26	115.74	19.48
	Média do intervalo 18.28				Média do intervalo 19.27				Média do intervalo 20.23			

Tabela 9. Tabela comparativa entre o Modelo 6 (com restrições) em comparação com as estimativas de tempo de divergência para as linhagens de Apoidea encontradas nos trabalhos atuais. Comparativamente é usado as médias de idades estimadas bem como os valores máximos e mínimos em 95% HPD. Inter: intervalo total de 95%HPD. N/A: dados não aplicáveis.

Clado	Modelo 6 (Melhor modelo) (Com restrições)				Peters et al. (2017)				Sann et al. (2018)				Branstatter et al. (2017)	
	média	min	max	inter	média	min	max	inter	média	min	max	inter	média	média
Apoidea stem	156.31	152.39	159.77	7.38	162	136	193	56.66	190.3	175.08	229.86	54.78	144.59	144.59
Apoidea crown	154.02	149.38	158.16	8.78	153	128	182	53.94	185.8	171.63	229.09	57.47	134.90	134.90
Angarosphecidae I stem	152.02	149.38	158.16	8.78	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Angarosphecidae I crown	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ampulicidae stem	150.24	144.82	155.27	10.45	153	128	182	53.94	185.8	171.63	229.09	57.47	134.90	134.90
Ampulicidae crown	117.83	98.62	138.61	39.99	130	100	160	59.43	152.9	125.27	202.66	77.39	0.00	0.00
Heterogynaidae stem	148.57	143.02	154	10.98	N/A	N/A	N/A	0.00	N/A	N/A	N/A	0.00	116.89	116.89
Heterogynaidae crown	113.39	99.9	129.11	29.21	N/A	N/A	N/A	0.00	N/A	N/A	N/A	0.00	N/A	N/A
Sphecidae stem	103.52	95.45	114.43	18.98	129	107	154	46.57	126.7	93.86	155.54	61.68	112.19	112.19
Sphecidae crown	59.21	32.33	88.74	56.41	83	65	104	38.99	99.9	79.30	123.74	44.44	81.16	81.16
Angarosphecidae II stem	144.06	137.68	150.12	12.44	N/A	N/A	N/A	0.00	N/A	N/A	N/A	N/A	N/A	N/A
Angarosphecidae II crown	129.1	116.81	141.3	24.49	N/A	N/A	N/A	0.00	N/A	N/A	N/A	N/A	N/A	N/A
Crabronidae ss. stem	140.02	133.28	146.55	13.27	129	107	154	46.57	141.4	107.27	169.71	62.45	112.19	112.19
Crabronidae ss. crown	127.12	110.77	140.98	30.21	107	87	129	42.07	124.8	92.71	149.79	57.08	89.19	89.19
Bembicidae stem	138.15	131.35	145.03	13.68	133	111	159	47.62	152.8	127.96	180.82	52.87	117.17	117.17
Bembicidae crown	127.69	116.19	139.02	22.83	119	96	144	48.89	135.7	103.44	168.18	64.74	64.26	64.26
Pemphredonidae stem	131.87	123.36	140.58	17.22	121	100	145	44.57	128.6	113.78	153.62	39.84	100.27	100.27
Pemphredonidae crown	120.65	108.66	132.69	24.03	109	88	133	45.05	110.4	98.46	135.62	37.16	99.17	99.17
Astatidae stem	131.87	123.36	140.58	17.22	119	96	144	48.89	172.3	154.01	209.56	55.55	N/A	N/A
Astatidae crown	121.07	108.89	133.06	24.17	52	30	81	51.80	71.2	52.48	161.29	108.80	N/A	N/A
Philantidae stem	131.85	123.75	139.76	16.01	121	100	145	44.57	128.6	113.78	153.62	39.84	100.27	100.27
Philantidae crown	53.83	27.13	82.45	55.32	92	65	121	56.60	92.7	62.83	119.14	56.32	82.27	82.27
Psenidae stem	128.69	120.6	137.11	16.51	124	104	147	43.72	136.7	125.66	161.67	36.01	100.27	100.27
Psenidae crown	121.89	112.4	131.6	19.2	N/A	N/A	N/A	N/A	111.6	88.50	137.92	49.42	99.17	99.17
Apidae ls. stem	105.95	97.23	115.99	18.76	124	104	147	43.72	128	107.27	149.41	42.14	108.58	108.58
Apidae ls. crown	97.54	85.97	110.16	24.19	111	93	132	39.56	113	95.39	135.62	40.23	98.89	98.89
		Média do intervalo			17.99	Média do intervalo			44.57	Média do intervalo			46.93	N/A

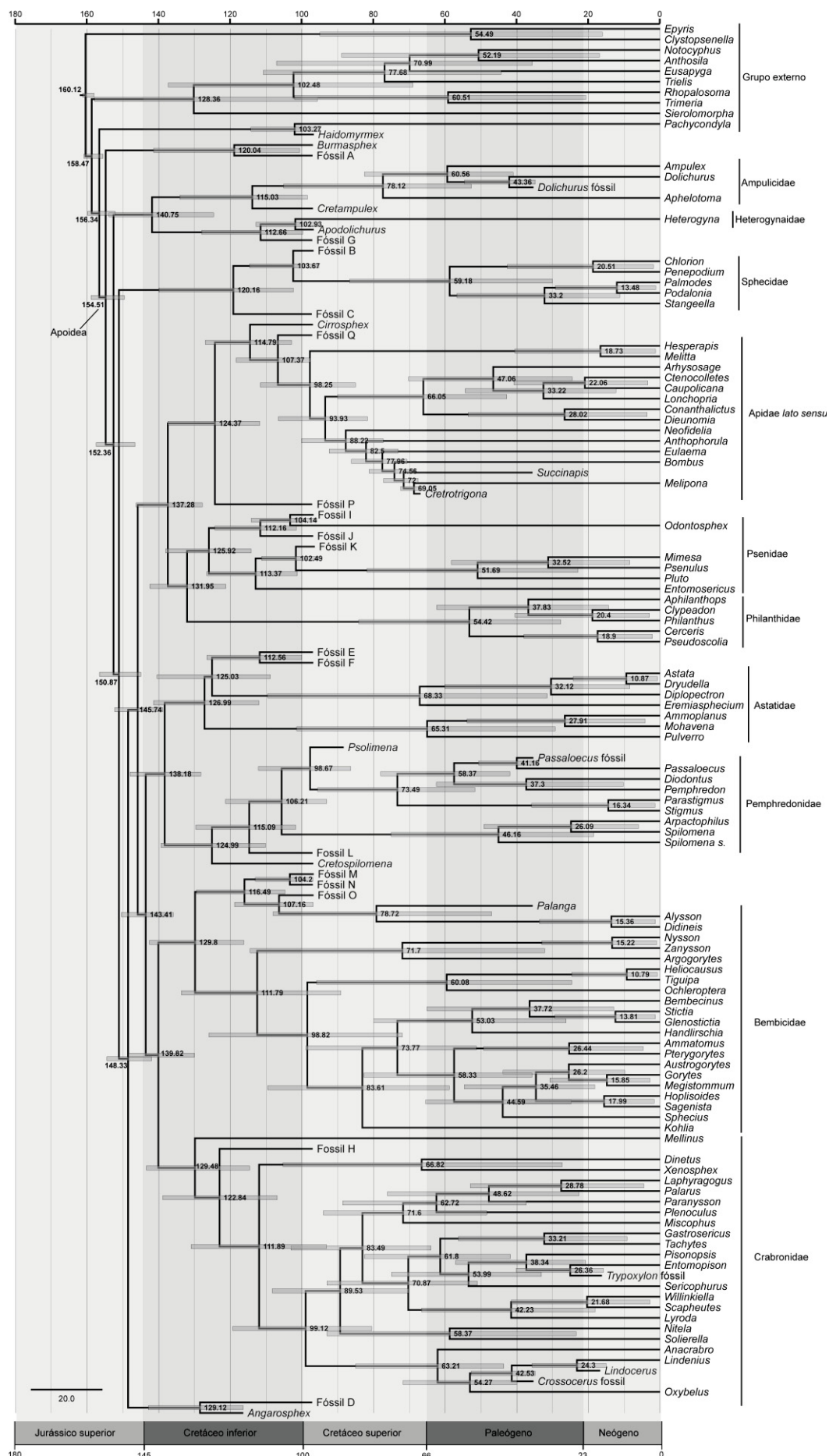


Figura 3. Cronograma com base em dados morfológicos em abordagem Bayesiana. Relógio morfológico do Modelo 6. Nove partições. Todos os grupos compatíveis adicionados (*Contype=allcompat*). Escala em milhões de anos.

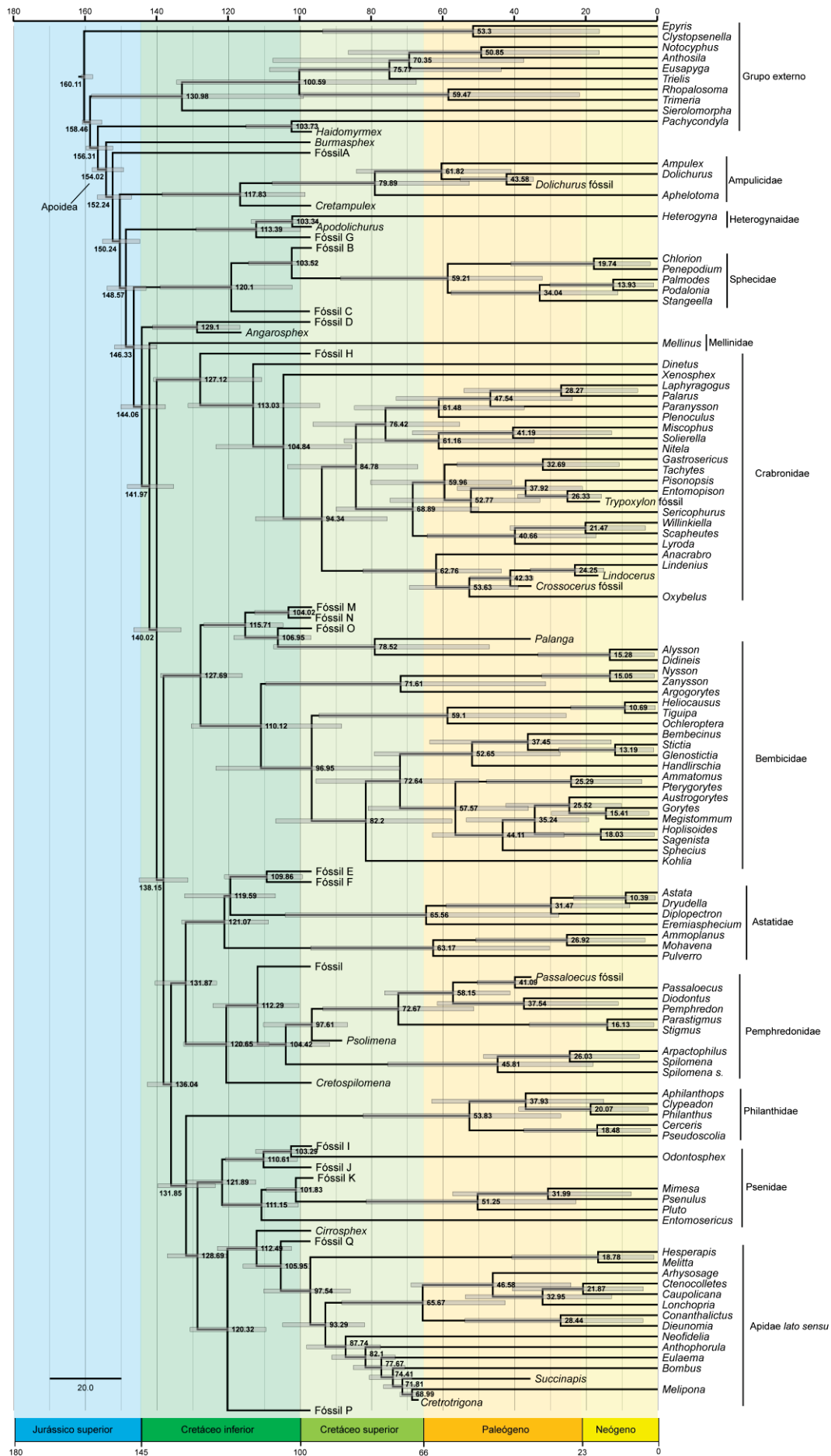


Figura 5. Cronograma com base em dados morfológicos em abordagem Bayesiana com 20 restrições. Relógio morfológico do Modelo 6. Nove partições. Todos os grupos compatíveis adicionados (*Contype=allcompat*). Probabilidade posterior como suporte dos ramos. Escala em milhões de anos.

A Figura 6 apresenta uma representação gráfica da densidade de sobreposições de árvores (*densitree*) composta por 2×10^6 árvores geradas a partir de uma análise de tempos de divergência com o modelo 6 e 20 restrições (o mesmo que a Figura 5). As áreas mais escuras representam maiores sobreposições de árvores, portanto maior suporte para uma determinada topologia. Por outro lado, áreas com menor sobreposições de árvores representam o oposto, ou seja, menor suporte para uma hipótese topológica.

O intervalo de tempo entre aproximadamente 150 e 120 milhões de anos (barras vermelhas) representa o momento da diversificação das grandes linhagens de Apoidea. De maneira geral esse intervalo contém uma baixa sobreposição de árvores quando comparadas com os momentos mais recentes da história evolutiva de Apoidea. A sobreposição desse intervalo também não é uniforme. Existe uma sobreposição muito maior nas primeiras divergências das linhagens de Apoidea que no restante. Além disso, o momento de divergência das linhagens de Crabronidae *lato sensu* apresentam a menor sobreposição de árvores na história evolutiva como um todo. Dentro do clado Crabronidae *lato sensu* + Apidae *lato sensu* a maior sobreposição ocorre nas divergências entre as linhagens de Astatidae + Pemphredonidae e também entre Philanthidae + (Psenidae + Apidae + *lato sensu*). Essa sobreposição é maior por exemplo que a vista dentro de Bembicidae, mais especificamente entre Alyssontinae e os demais Bembicidae. Por fim, a partir dos 90 milhões de anos, as relações internas nos clados de vespas apoideas e abelhas parecem mais estáveis com um alto número de sobreposições de árvores.

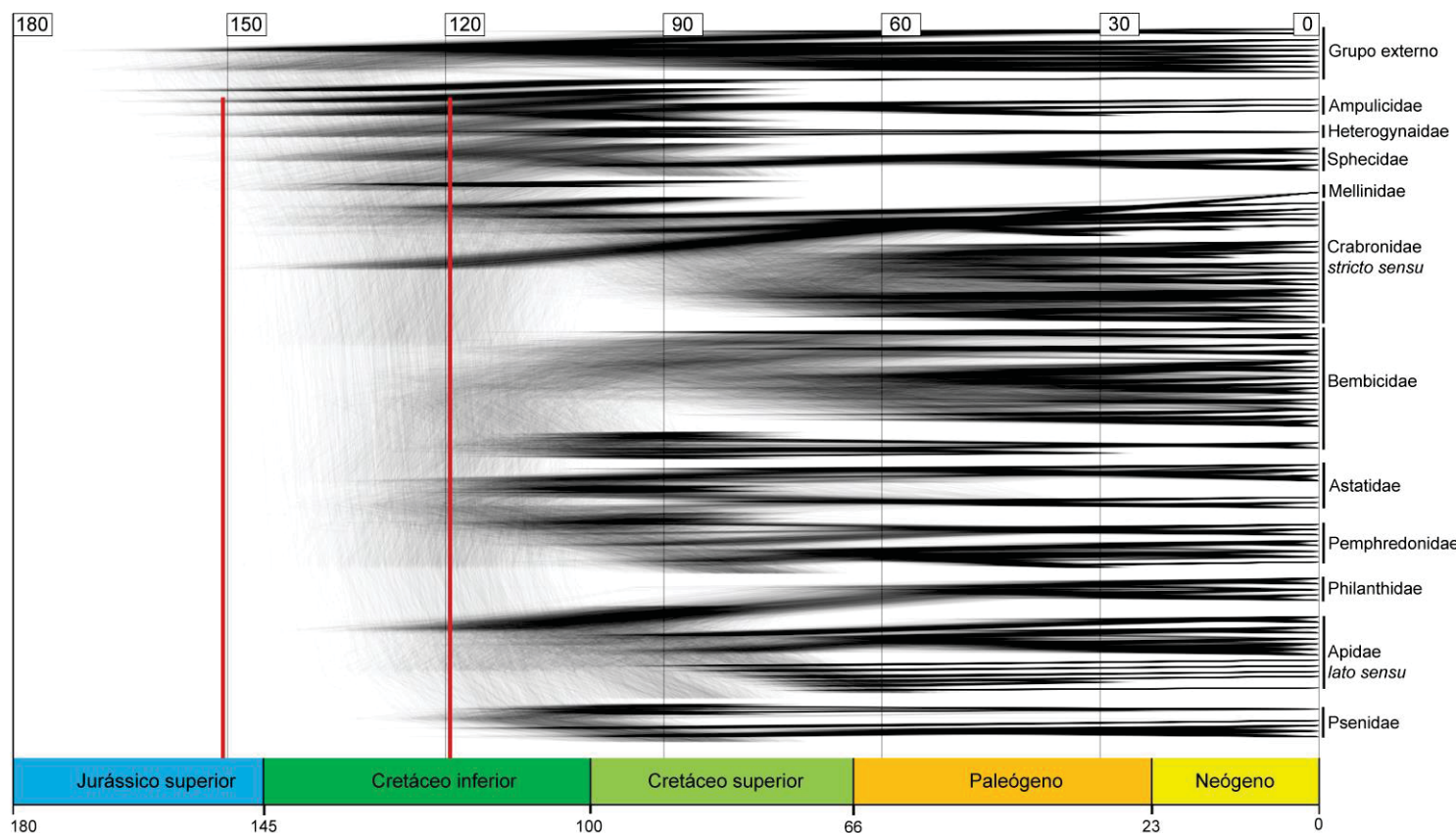


Figura 6. Representação gráfica da densidade de sobreposições de árvores (*densitree*) composta por 2×10^6 árvores geradas a partir de uma análise de tempos de divergência do modelo 6 e 20 restrições. Escala em milhões de anos.

3.3 Análises de linhagens ao longo do tempo (LTTs) e estatística- γ

As análises de diversificações ao longo do tempo foram conduzidas com base no cronograma (árvore datada) apresentado na Figura 5. A Figura 7 apresenta as análises de diversificações das grandes linhagens de Apoidea incluindo os fósseis. Cada um dos ramos do cronograma da Figura 7a representa uma das subfamílias de Apoidea e ao lado das linhagens estão indicadas as 11 famílias viventes. A Figura 7b apresenta as linhagens de Apoidea ao longo do tempo (LTT). As Figuras 7c, 7e e 7g apresentam as contraposições dos LTTs com as 500 simulações em modelos de *Pure-Birth* (7c) e *Birth-Death* (7e e 7g). As Figuras 7d, 7f e 7h apresentam as distribuições gamma das mesmas simulações para os mesmos modelos.

As simulações do modelo com *Pure-Birth* (7c) indicou uma média positiva do valor de estatística- γ ($\gamma = 0.7753663$), sugerindo uma taxa de especiação inicialmente lenta e aumentando ao longo do tempo até o presente. Já as simulações com extinções usando um modelo *Birth-Death* (7e e 7g) indicaram valores negativos de estatística- γ ($\gamma = -0.734756$ e -1.422717 respectivamente), portanto sugerindo o contrário do modelo *Pure-Birth*, com uma taxa de diversificação inicialmente alta e queda ao longo do tempo até o presente. A estatística- γ dos LTTs para as linhagens de Apoidea incluindo os fósseis obtiveram $\gamma = -2.5584$. O valor de $\gamma < 0$ sugere que, assim como nas simulações com *Birth-Death*, houve uma taxa de diversificação inicialmente alta e com uma queda ao longo do tempo até o presente.

A Figura 8 apresenta as análises de diversificações das grandes linhagens de Apoidea excluindo os fósseis. Cada um dos ramos da Figura 8a representam as subfamílias de Apoidea viventes indicadas sob as 11 famílias viventes. A Figura 8b apresenta as linhagens de Apoidea ao longo do tempo (LTT). As Figuras 8c, 8e e 8g apresentam as contraposições dos LTTs com as 500 simulações em modelos de *Pure-Birth* (8c) e *Birth-Death* (8e e 8g). As Figuras 8d, 8f e 8h apresentam as distribuições gamma das mesmas simulações para os mesmos modelos. Os resultados das simulações nesse caso indicaram os mesmos padrões encontrados da árvore com os fósseis, o modelo *Pure-Birth* (8c) com $\gamma > 0$ ($\gamma = 0.9136612$) e os modelos *Birth-Death* (8e e 8g) com $\gamma < 0$ ($\gamma = -1.01703$ e -0.5426654 respectivamente). Da mesma maneira, os dados avaliados também resultaram em $\gamma < 0$ ($\gamma = -4.9559$) indicando, portanto, uma taxa de diversificação inicialmente alta com uma queda ao longo do tempo até o presente.

Por outro lado, os LTTs apresentaram resultados discrepantes entre as análises dos dados com a inclusão e com a exclusão dos fósseis. Os gráficos LTTs com os fósseis apresentaram um evento de extinção em massa para as grandes linhagens de Apoidea logo após o Cenomaniano (Cretáceo superior) aproximadamente entre 90 a 84 milhões de anos (durante o Turoniano e Coniaciano, Cretáceo superior). Em contraste, os resultados sem a inclusão dos fósseis não apresentaram eventos significativos de extinções ao longo da história evolutiva do grupo.

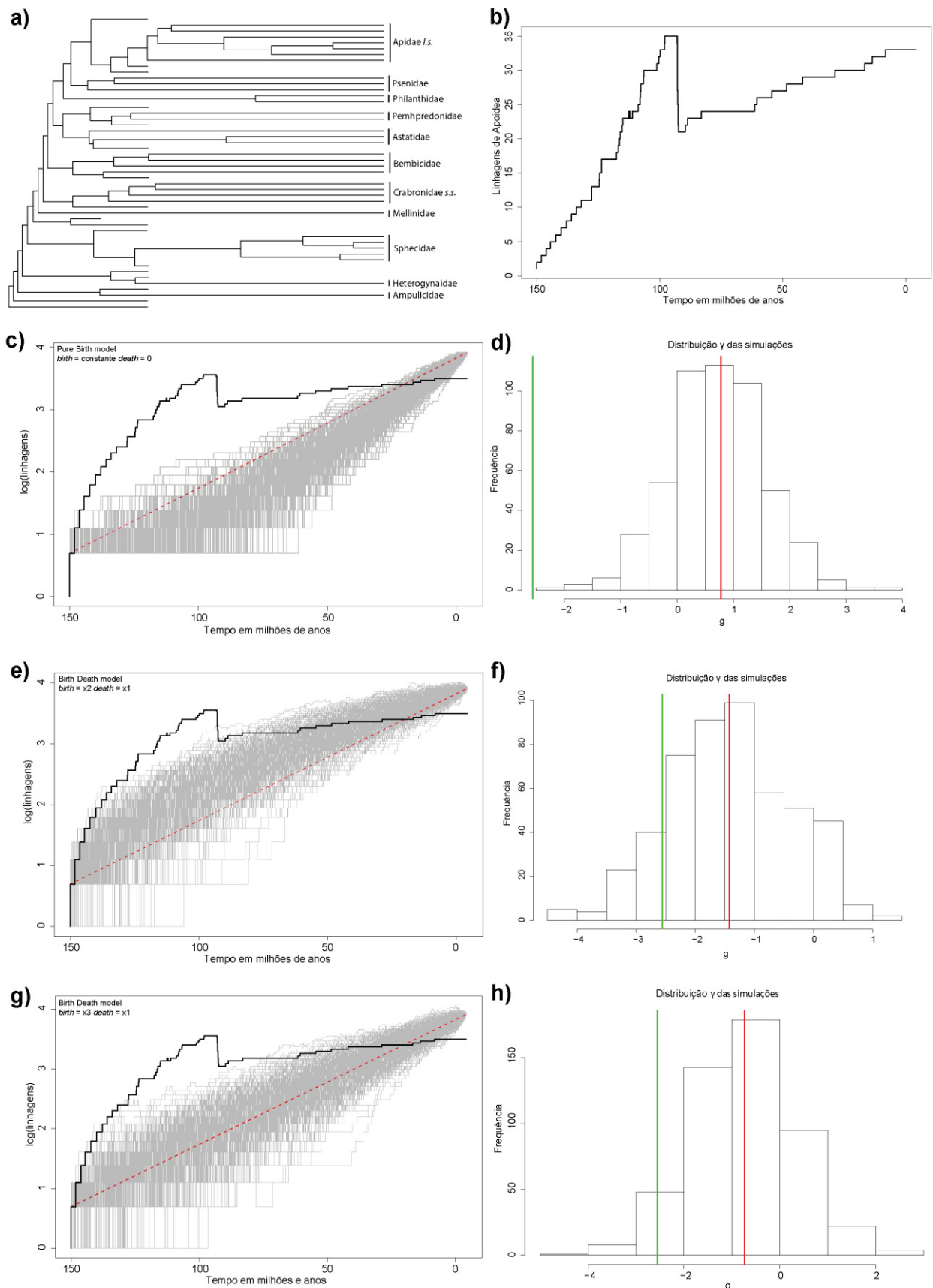


Figura 7. Análises de diversificações das grandes linhagens de Apoidea incluindo os fósseis. **a)** cronograma das grandes linhagens de Apoidea resultante de uma análise de IB com dados morfológicos. **b)** gráfico das linhagens de Apoidea ao longo do tempo (LTT). **c)** LTT contraposta com simulações com taxas de especiação constante e sem extinção (*Pure-Birth*). **d)** histograma com a distribuição das frequências das 500 simulações em *Pure-Birth* em uma distribuição gamma. **e)** LTT contraposta com 500 simulações com taxa de extinção constante com o dobro de especiação (*Birth-Death*). **f)** histograma com a distribuição das frequências das 500 simulações em *Birth-Death* em uma distribuição gamma. **g)** LTT contraposta com 500 simulações com taxa de extinção constante e com o triplo de especiação (*birth-death*). **h)** histograma com a distribuição das frequências das 500 simulações em *Birth-Death* em uma distribuição gamma. Linhas verdes representam o valor γ para os resultados analisados. Linhas vermelhas representam o valor γ médio das simulações.

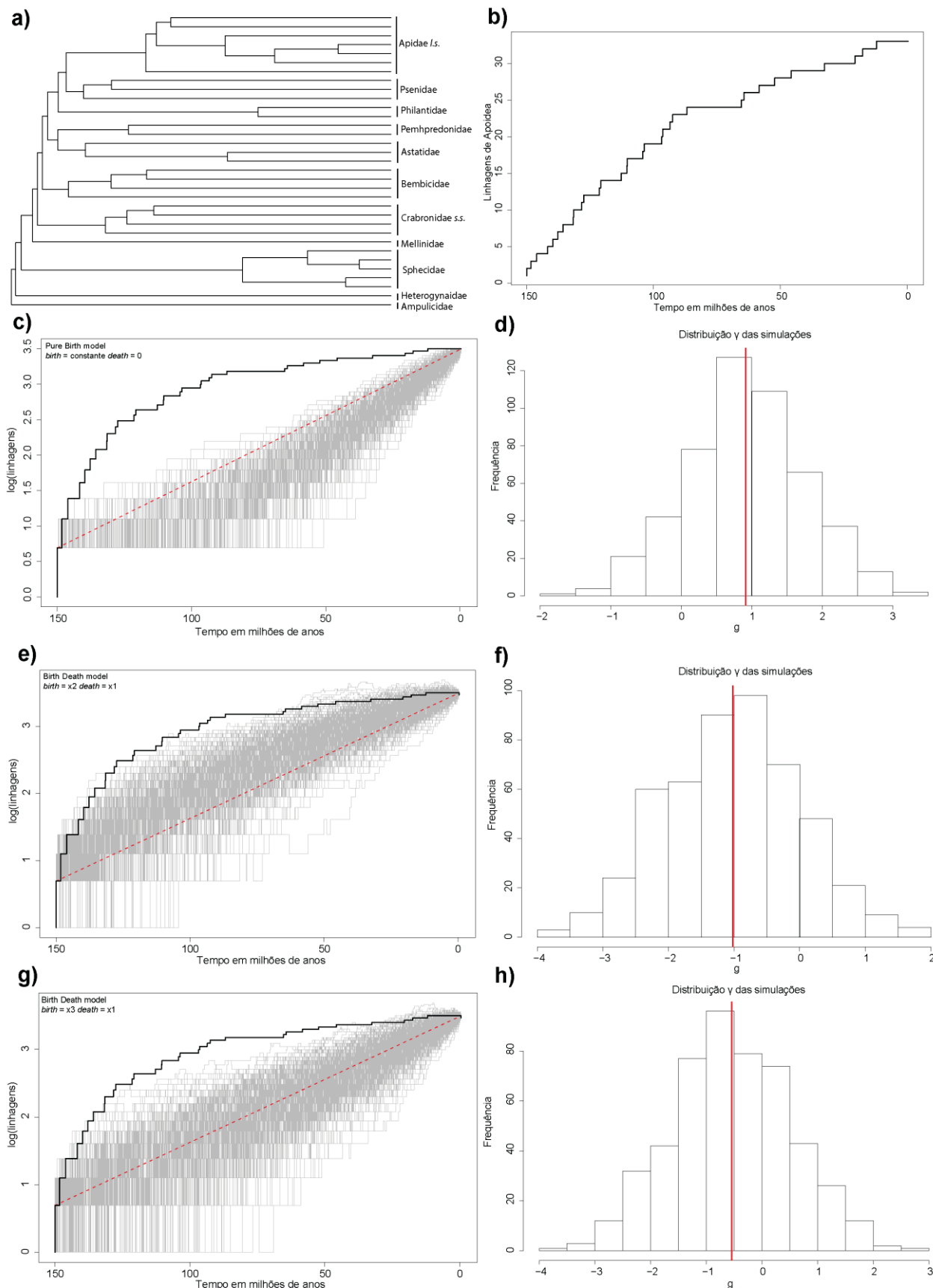


Figura 8. Análises de diversificações das grandes linhagens de Apoidea incluindo os fósseis. **a)** cronograma das grandes linhagens de Apoidea resultante de uma análise de IB com dados morfológicos. **b)** gráfico das linhagens de Apoidea ao longo do tempo (LTT). **c)** LTT contraposta com simulações com taxas de especiação constante e sem extinção (*Pure-Birth*). **d)** histograma com a distribuição das frequências das 500 simulações em *Pure-Birth* em uma distribuição gamma. **e)** LTT contraposta com 500 simulações com taxa de extinção constante com o dobro de especiação (*Birth-Death*). **f)** histograma com a distribuição das frequências das 500 simulações em *Birth-Death* em uma distribuição gamma. **g)** LTT contraposta com 500 simulações com taxa de extinção constante e com o triplo de especiação (*Birth-Death*). **h)** histograma com a distribuição das frequências das 500 simulações em *Birth-Death* em uma distribuição gamma. Linhas vermelhas representam o valor γ médio das simulações. Linhas verdes não estão representadas pois o valor γ para esses resultados foi muito menor que a representação gráfica da distribuição gamma ($\gamma = -4.9559$).

4. Discussão

4.1 Relações filogenéticas e tempo de divergência das linhagens de Apoidea

As análises filogenéticas conduzidas em nosso estudo combinando táxons vivos com táxons fósseis foram as mais inclusivas já realizadas para os Apoidea (99 vivos e 32 fósseis). Nossos resultados apontam para pelo menos 17 linhagens independentes de Apoidea sendo 11 linhagens vivos e seis extintas. De maneira geral os nossos resultados reconhecem as grandes linhagens de Apoidea e são congruentes com trabalhos prévios (Alexander 1992a; Melo 1999; Devebec et al. 2012; Branstetter et al. 2017; Peters et al. 2017; Sann et al. 2018). Entretanto os relacionamentos dessas grandes linhagens são discrepantes entre as diversas fontes de dados e métodos filogenéticos aplicados (Figura 9). Essas incongruências foram o foco do nosso trabalho e serão discutidas pontualmente.

Os nossos resultados mostram uma forte tendência à parafilia de Vespoidea *lato sensu* com Scoliidae e Formicidae mais próximos a Apoidea que os demais Aculeata. Esses resultados são congruentes com as hipóteses filogenéticas realizadas a partir de dados moleculares (Ohl & Bleidorn 2006; Pilgrim et al. 2008; Devebec et al. 2012; Johnson et al. 2013; Faircloth et al. 2014; Branstetter et al. 2017; Peters et al. 2017). Em nossos resultados, o clado formado por Scoliidae grupo irmão do clado Formicidae + Apoidea apresentou um alto suporte de ramos (PP: 0.87). O prepecto fusionado ao mesepisterno com a sutura entre eles completamente obliterada (83:2) e a fusão ventro-lateral do mesepisterno (90:1) são sinapomorfias desse clado sem nenhuma reversão nos grupos analisados.

Em nossos resultados, os Formicidae formam um clado com os Apoidea com um alto suporte de ramos (PP = 0.94). O reconhecimento desse clado é congruente com as hipóteses moleculares de Faircloth et al. (2014), Branstetter et al. (2017), Peters et al. (2017). Existem hipóteses alternativas para o relacionamento dessas linhagens, como: (Scoliidae + Formicidae) + Apoidea ou mesmo Formicidae + (Scoliidae + Apoidea) (Pilgrim et al. 2008; Johnson et al. 2013). Porém a hipótese Scoliidae + (Formicidae + Apoidea) parece mais estável (Branstetter et al. 2017; Camacho et al. 2018). A realocação do alimento da larva realizada pela fêmea (197:1), certamente faz parte do plano básico de Formicidae + Apoidea. O surgimento e fixação deste comportamento em diferentes linhagens de Aculeata indubitavelmente abriu importantes caminhos que levaram a comportamentos complexos associados à construção de ninhos e à socialidade (Melo et al. 2011). Por outro lado, a biologia de Scoliidae não é totalmente compreendida e existem boas evidências que em algumas linhagens dessa família também ocorre a realocação de presas (Melo et al. 2011). Se esse cenário for esclarecido, essa característica pode ter surgido na base do clado Scoliidae + (Formicidae + Apoidea).

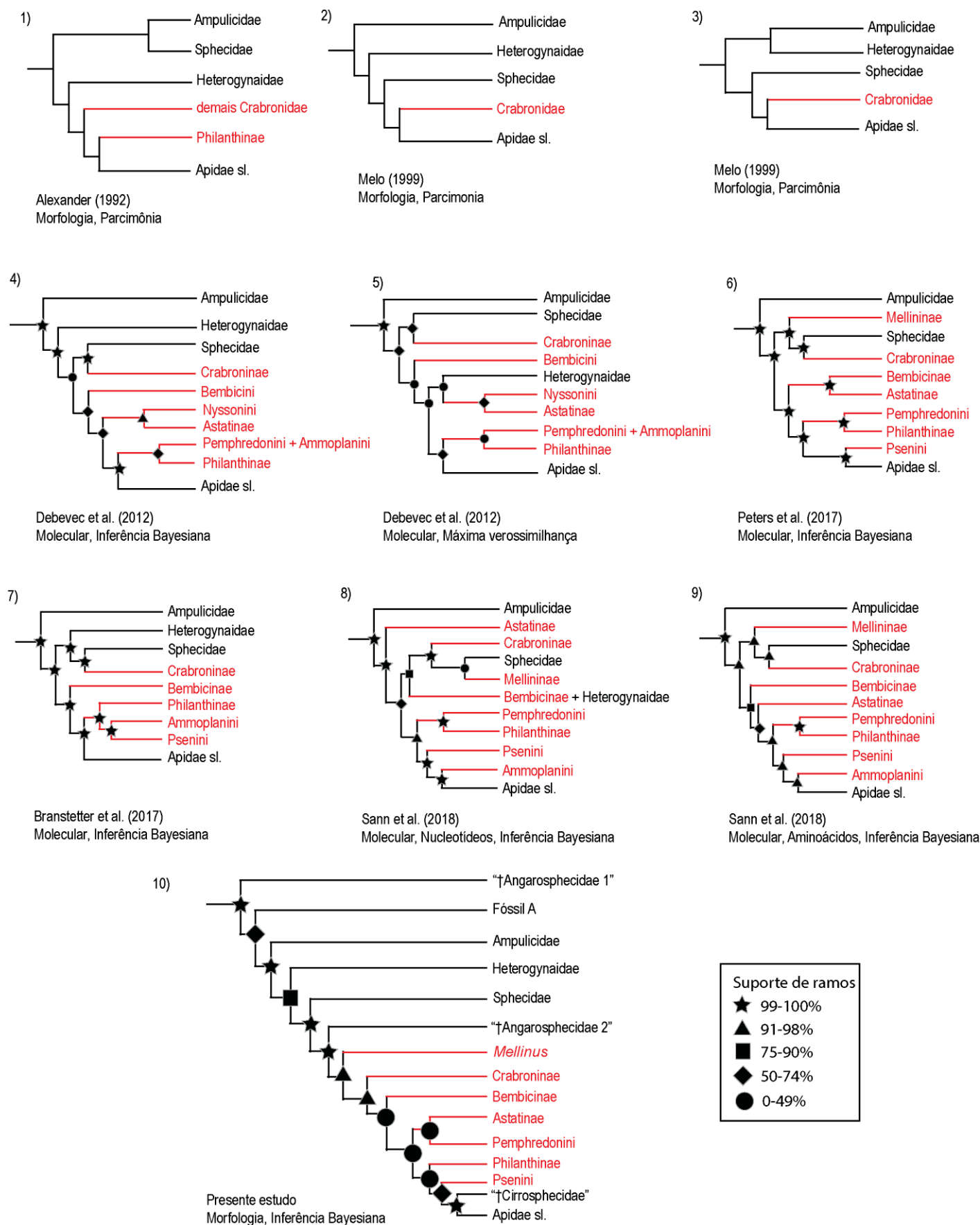


Figura 9. Hipóteses filogenéticas de Apoidea segundo a classificação de Melo (1999) com suas respectivas referências. Suporte de ramos (probabilidade posterior). Em vermelho as linhagens de crabronídeos.

A monofilia de Apoidea foi recuperada em diversos trabalhos tanto por dados morfológicos quanto moleculares (Brothers & Carpenter 1993; Melo 1999; Pilgrim et al. 2008; Devedec et al. 2012; Johnson et al. 2013; Faircloth et al. 2014; Branstetter et al. 2017; Peters et al. 2017). Em nossos resultados essa hipótese também foi altamente suportada (PP=1). Segundo os nossos resultados o grupo *stem* de Apoidea divergiu no final do Jurássico, 156.31 Ma (152.39-159.77Ma). Essas datas são um pouco mais antigas que aquela encontrada por Branstetter et al. (2017) (144.59 Ma), porém mais recentes que as estimadas por Peters et al. (2017) (162, 136-193 Ma). Por outro lado, Sann et al. (2018) estimou a divergência do grupo *stem* de Apoidea no Jurássico inferior (190.3, 175.08-229.86 Ma). Comparando com as demais estimativas essa última hipótese é pouco realista. As nossas estimativas de idade do grupo *crown* de Apoidea foram 154.02 Ma (149.38-158.16 Ma). Essa hipótese é parcialmente compatível com as estimativas de Peters et al. (2017) (153, 128-182 Ma) e mais antigas que as de Branstetter et al. (2017) (134.90 Ma). E por fim, as estimativas de Sann et al. (2018) também foram artificialmente antigas (185.8, 171.63-229.09 Ma).

As primeiras linhagens do grupo *crown* de Apoidea a surgirem foram “Angarosphecidae 1” e Fóssil A. Essas duas linhagens divergiram no final do Jurássico a 156.02 Ma (149.38-158.16 Ma) e 154.02 Ma (147.16-156.74 Ma) respectivamente. “†Angarosphecidae 1” foi aqui representada por †*Burmasphex*, um grupo de vespas apoideas recentemente descritas por Melo & Rosa (2018). As espécies desse gênero são conhecidas do âmbar burmês e possivelmente é o grupo fóssil mais importante para a compreensão da evolução dos Apoidea. Nesse táxon é possível observar uma grande variedade de plesiomorfias que foram completamente perdidas no decorrer da história evolutiva dos Apoidea. Um Apoidea ancestralmente hipotético seria uma vespa pequena, esguia e de aparência formicóide, com coloração metálica (189:1), o pronoto com uma constrição distinta (52:1) bem como com ‘lobos pronotais’ (50:1), o metapostnoto expandido, formando o ‘triângulo propodeal’ (150:1), e muito possivelmente com fêmeas capazes de realocar as provisões de suas larvas (197:1).

Além de todas essas condições estarem presentes em †*Burmasphex* (talvez com exceção é claro do caráter 197, para o qual não temos evidências diretas) e que o posicionam inequivocamente em Apoidea, as espécies desse gênero possuem uma condição incomum que os separa das outras linhagens da superfamília. Em um Apoidea “típico” a margem posterior da porção dorsal do pronoto está sempre inclinada para baixo e em um plano distinto comparado à porção adjacente do pronoto (51:1). Em †*Burmasphex* a margem posterior da porção dorsal do pronoto está no mesmo nível de sua porção anterior adjacente (51:0). Essa condição do pronoto remete diretamente aos aculeados não-apoideas como em Pompilidae, Sierolomorphidae, Sapygidae e muitos dos Chrysidoidea (Melo & Rosa em preparação).

O terminal nomeado como Fóssil A representa um grupo de apoideas basais que estão presentes no âmbar burmês com pelo menos duas espécies e dois gêneros distintos. Além disso é possível que estejam presentes também na formação Crato (Rosa & Melo em preparação) provavelmente formando uma família distinta e de distribuição paleotropical durante o Cretáceo médio e inferior. Da mesma maneira que †*Burmasphex*, Fóssil A possui a maior parte das condições plesiomorfas de um Apoidea ancestral

hipotético, como o corpo esguio e de aparência formicóide, coloração metálica (189:1), as linhas parapsidiais em contato com a sutura transcutal (78:0), e os notaulos divergentes em sua porção anterior ao mesoscutum (75:0). Adicionalmente, outra condição plesiomórfica de destaque presente nesses dois grupos é o ângulo ventral do pronoto que mal excede a base das coxas anteriores (53:0).

Apesar de todas essas condições, o Fóssil A possui a sinapomorfia 51:1 compartilhada com todos os demais Apoidea e formando o Clado 1. No entanto na topologia a Figura 3 apresenta uma estreita relação do Fóssil A com *†Burmarphex*. Se esse cenário alternativo fosse levado em consideração (*†Burmarphex* + Fóssil A), ou a condição 51:1 teria surgido duas vezes, uma nos Apoidea *stricto sensu* e outra paralelamente no grupo Fóssil A; ou a condição 51:0, presente em todos os demais Aculeata, teria surgido independentemente em *†Burmarphex* e a condição 51:1 faria parte do plano básico de Apoidea. No entanto, essas hipóteses são altamente rejeitadas (KRS = 34.66). Portanto, o clado 1 incluindo Fóssil A parece mesmo ser uma linhagem distinta em relação a *†Burmarphex* e com um suporte de ramos razoável (PP = 67). É possível que o estreito relacionamento entre *†Burmarphex* e Fóssil A seja resultado da atração de ramos decorrente à proximidade temporal mútua no registro fóssil.

O clado 2 é formado pelos demais Apoidea e é altamente suportado em nossos resultados (PP=99). Esse clado divergiu ainda no final do Jurássico com estimativas de 150.24 Ma (144.82-155.27Ma). A primeira divergência nesse clado envolve os ampulicídeos e os grupos mais derivados de Apoidea. Os Ampulicidae são definidos aqui como contendo os seis gêneros vivos tradicionalmente reconhecidos mais *†Cretampulex*. Em nossos resultados o clado Ampulicidae é bem suportado (PP=0.76) e é corroborado tanto a partir de dados morfológicos (Melo 1999; Ohl & Spahn 2010) quanto moleculares (Peters et al. 2017; Sann et al. 2018). *†Cretampulex* é incluído nesse clado por apresentar a condição em que a parte superior da borda externa do esclerito antenal está projetada distintamente para frente em relação à porção inferior deixando o alvéolo antenal direcionado para baixo (33:1). Essa é uma condição comum a todos os Ampulicidae, porém surgiu várias vezes dentro de Aculeata e apenas mais uma outra vez em Apoidea (em Nyssonini, Bembicidae).

Em nossos resultados, o clado 3 é formado por Heterogynaidae, Sphecidae, “*†Angarosphécidae* 2”, Crabronidae *lato sensu* e Apidae *lato sensu*. Esse clado apresentou um bom suporte de ramos (PP = 0.86) e é representando tanto em filogenias morfológicas quanto moleculares (Melo 1999; Devebec et al. 2012; Branstetter et al. 2017; Sann et al. 2018). O posicionamento de Heterogynaidae é bastante instável dentro do clado 3 (Figura 10). O posicionamento desse táxon pode ser resumido em três hipóteses maiores: 1) Heterogynaidae dentro de Crabronidae *lato sensu* (Ohl & Bleidorn 2006; Devebec et al. 2012 em máxima verossimilhança; Sann et al. 2018): Essa hipótese não possui base morfológica e as evidências moleculares propostas são fracas. Por exemplo, os resultados de Ohl & Bleidorn (2006) foram baseados em um único gene nuclear e no trabalho de Sann et al. (2018) o terminal usado para representar Heterogynaidae foi considerado um táxon instável (*rogue taxon*) em 36 das 36 simulações conduzidas para estimar a estabilidade de um terminal (veja a tabela suplementar desse trabalho em “Additional table S3”). 2) Heterogynaidae grupo

irmão de Ampulicidae (Melo 1999): Essa hipótese é proposta somente por dados morfológicos e é detalhadamente discutida em Melo (1999). 3) Heterogynaidae + (Sphecidae + (“Crabronidae” s.l. + Apidae): Os nossos resultados suportam essa hipótese com alto valor de probabilidade posterior (PP=0.86) e é corroborada tanto por dados morfológicos quanto moleculares (Melo 1999; Devebec et al. 2012). Outras hipóteses variantemente próximas são equiparáveis, como a partir de dados genômicos Heterogynaidae + (Sphecidae + Crabronidae s.s.) (Branstetter et al. 2017) ou ainda dados morfológicos, Heterogynaidae + (“Crabronidae” s.l. + Apidae) (Alexander 1992a).

Heterogynaidae é a menor família dentre os Apoidea viventes com nove espécies descritas em um único gênero, *Heterogyna*. A distribuição dessas espécies está restrita a poucos locais de coleta ao longo da região Etiópica (Botswana e Madagascar) e região Paleártica (Mediterrâneo e Oriente Médio), apresentando um padrão relictual no Velho Mundo (Ohl & Bleidorn 2006; Pulwaski 2018). Em nossos resultados *Heterogyna* está altamente relacionada com dois fósseis provenientes do âmbar burmês formando um grupo monofilético com alto suporte de ramos (PP=1). Essa família é aqui definida incluindo *Heterogyna*, e outras dois táxons fósseis, †*Apodolichurus* e Fóssil B. †*Apodolichurus* foi descrito em Antropov (2000) em Ampulicidae e o Fóssil B representa uma linhagem nova ainda não descrita.

Esses táxons foram agrupados por caracteres que descrevem uma drástica redução na venação das asas anteriores e que não é compartilhada com mais nenhum outro grupo de Apoidea. Nas espécies viventes de Heterogynaidae, as fêmeas são braquípteras e os machos possuem asas funcionais com venação reduzida. Curiosamente todas as fêmeas fósseis dessa linhagem possuem asas funcionais. O interessante aqui é que as asas das fêmeas fósseis são muito reduzidas, ainda mais reduzidas que as asas dos machos de *Heterogyna*. Além disso, as asas do macho conhecido de †*Apodolichurus* possui venação mais “completa” que os dos machos atuais da linhagem. A nossa interpretação é que a redução da venação é comum a todos os Heterogynaidae sendo que uma redução ainda maior e concomitante surgimento da condição braquíptera ocorreu ao longo da linhagem de *Heterogyna*.

No clado 4 estão inclusos os Sphecidae, “†Angarosphecidae 2”, todos os Crabronidae *lato sensu* e as abelhas. Em nossos resultados, o clado 4 apresentou alto suporte de ramos (PP = 1) e é corroborado na maioria das hipóteses propostas previamente (Figura 9). Entretanto, as relações internas desse clado são incongruentes entre os dados morfológicos e moleculares. Nas hipóteses com dados morfológicos, os Sphecidae são grupo irmão das demais linhagens (Melo 1999). Em compensação, nas hipóteses moleculares Crabronidae *strictu sensu* e por vezes Mellinidae estão mais associados com Sphecidae que com as demais linhagens de crabronídeos e abelhas (Devebec et al. 2012; Peters et al. 2017; Branstetter et al. 2017; Sann et al. 2018). Os nossos resultados suportam o posicionamento de Sphecidae como grupo irmão do restante das linhagens que formam o clado 5 (KRS = 384.14 e 377.04, veja Tabela 6). Em nossos resultados, o clado 5 apresentou alto suporte de ramos (PP = 1) e inclui “†Angarosphecidae 2”, todos os Crabronidae *lato sensu* e as abelhas.

A diversificação inicial de Apoidea foi rápida em um curto espaço de tempo e é possível que esse evento tenha interferido nas hipóteses filogenéticas de Apoidea e gerando hipóteses frágeis ou mesmo incongruências entre diferentes dados filogenéticos. Entretanto, as relações internas dos clados 4 e 7 (veja adiante) parecem ter sofrido maior interferência. O tempo de divergência da linhagem de Sphecidae (incluindo os fósseis associados), “†Angarosphecidae 2”, Mellinidae e os demais Crabronidae *lato sensu* (incluindo as abelhas) correspondem a 146.33, 144.06 e 140.02 Ma, respectivamente. Consequentemente as relações internas do clado 4 foram estabelecidas em menos de 10 Ma durante o Berriasiano (Cretáceo). Além de uma rápida diversificação das grandes linhagens o evento ocorreu nos primeiros milhões de anos do Cretáceo inferior. Além disso, a linhagem *stem* de Sphecidae possui um ramo muito longo em relação aos demais Apoidea. O grupo *crown* de Sphecidae (táxons vivos) surgiu mais de 60 Ma depois desse evento. Portanto, é muito possível que as incongruências nas relações internas do clado 4 estejam relacionadas a uma rápida diversificação em um curto intervalo de tempo somados a um acontecimento antigo da história evolutiva do grupo e ainda uma interferência de uma linhagem com ramo muito longo.

O clado Sphecidae apresentou alto suporte de ramos (PP = 1) e é corroborado em todos os trabalhos conduzidos com Apoidea (Alexander 1992a; Melo 1999; Devebec et al. 2012; Peters et al. 2017; Branstetter et al. 2017; Sann et al. 2018). Como citado anteriormente, o grupo *crown* de Sphecidae apresenta uma idade relativamente recente quando comparada com os demais Apoidea. Em nossos resultados, essa linhagem surgiu no Paleógeno (59.21, 32.33-88.74 Ma). Nas estimativas apresentadas em Peters et al. (2017) e Branstetter et al. (2017) o grupo *crown* de Sphecidae divergiu na metade do Cretáceo superior, pouco mais de 80 Ma. Também como discutido anteriormente, as estimativas de Sann et al. (2018) apresentam idades artificialmente mais antigas 99.9 (79.30-123.74) Ma. Além disso, nossos resultados mostraram que os fósseis Fóssil B e Fóssil C estão associados à linhagem dos Sphecidae.

O Fóssil C não possui caracteres relevantes que possam posicioná-lo com segurança nessa linhagem e o seu posicionamento pode ter sido um artefato. Entretanto, esse fóssil certamente pertence ao clado 4 e muito possivelmente representa uma linhagem independente. Por outro lado, o Fóssil B apresentou um alto suporte de ramos na posição de grupo irmão de Sphecidae (PP = 1). Esse táxon apresenta a fossa tentorial distintamente acima da sutura epistomal (30:1) e mesmo que esse caráter tenha surgido outras vezes em Apoidea (Crabronidae *strictu sensu* e Apidae *lato sensu*), esse caráter é uma importante transformação compartilhada com os Sphecidae. A razão pela qual não incluímos esses dois táxons em Sphecidae foi a ausência de um pecíolo metassomal (155:1) e a ausência da principal sinapomorfia de Sphecidae, o alveolo da coxa posterior fechado principalmente por um aumento da porção medial do metacatepisterno (103:2). Ademais, esses dois fósseis são provenientes do âmbar burmês (99.7–94.3 Ma), muito antes do surgimento do grupo *crown* de Sphecidae.

Os clados 5, 6 e 7 apresentaram alto suporte de ramos com probabilidade posterior igual a 1, 0.96 e 0.93 respectivamente. A primeira divergência do clado 5 foi a família fóssil “Angarosphecidae 2”. Em nossos

resultados essa família é composta pelo Fóssil D e †*Angarosphex magnus*. A espécie de †*Angarosphex* que usamos é proveniente do Crato (Brasil) e o Fóssil D é proveniente do âmbar burmês, com pelo menos três espécies, todas ainda não descritas. Essa família é muito interessante do ponto de vista biogeográfico pois demonstra que as linhagens derivadas de vespas apoideas apresentavam distribuição paleopantropical durante o Cretáceo inferior.

Mellinidae representa uma linhagem vivente que divergiu dos demais crabronídeos ainda no Cretáceo inferior. Essa família é representada por um único gênero com 16 espécies com distribuição principalmente Holártica. Entretanto, algumas espécies também são encontradas na América Central e uma única na Venezuela (*Mellinus andinus* Menke). Em análises dados moleculares essa família representa uma linhagem independente (Peters et al. 2017; Sann et al. 2018), mas foi incluída em Crabronidae *strictu sensu* nos dados morfológicos com parcimônia (Melo 1999). Nos resultados de Melo (1999), *Mellinus* é grupo irmão dos demais Crabronidae *strictu sensu* sendo que a principal sinapomorfia atribuída a esse clado é a posição ventral do ânus da larva (192:1). Essa condição não é encontrada em mais nenhum outro Aculeata avaliado. Entretanto, os nossos resultados apresentam *Mellinus* como uma linhagem independente com alto suporte de ramos (PP = 0.96, clado 6) e KRS = 13.94. Uma interpretação para esse caráter seria que essa transformação (192:1) faça parte do plano básico da linhagem Crabronidae *lato sensu* + Apidae *lato sensu*. Essa condição seria mantida em Crabronidae *strictu sensu* e perdida uma única vez nos demais crabronídeos e abelhas (clado 6).

Como citado anteriormente, a diversificação inicial de Apoidea foi rápida em um curto espaço de tempo e é possível que esse evento tenha comprometido a resolução filogenética. As relações internas do clado 7 são poucas suportadas. Os cladogramas 8, 9 e 10 apresentaram baixo suporte de ramos com probabilidade posterior igual a 0.47, 0.15 e 0.49 respectivamente. As relações desse clado podem facilmente ser interpretadas como uma grande politomia (Figura 2). Nesse cenário bastante conflitante, Melo (1999) propõe a monofilia de Crabronidae *lato sensu* como base em fortes evidências morfológicas e cenários biogeográficos hipotéticos. Entretanto, os nossos resultados apontam para a parafilia de Crabronidae (KRS = 317.64) com algumas linhagens de crabronídeos estando mais relacionadas com as abelhas (veja abaixo). Melo (1999) argumenta que durante a divergência do clado Apidae *lato sensu* + Crabronidae *lato sensu* essas linhagens estariam ocupando desertos e zonas temperadas. Essa hipótese foi proposta em vista da alta diversidade atual de diversos táxons de diferentes linhagens de crabronídeos nesses ambientes. No entanto, considerando o registro fóssil associado às nossas hipóteses filogenéticas, existe uma clara associação dessas linhagens com ambientes florestados úmidos. O registro fóssil do depósito de Myanmar é rico e diverso o suficiente para propor uma diversificação das linhagens de crabronídeos em ambientes florestados. Todavia, essa nova proposta não exclui o cenário de diversificação das abelhas em ambientes desérticos proposta por Michener (1979).

Crabronidae *strictu sensu* é a maior linhagem de vespas apoideas com mais de 4300 espécies atualmente (Pulawski 2018). Esse clado apresentou alto valor de suporte de ramos com probabilidade posterior igual a 0.88 e o grupo *crown* da família foi estimado em 127.12 (110.77-140.98) Ma. O Fóssil H foi

a primeira linhagem a divergir dentro desse clado, seguida de Dinetinae (113.03, 94.47-131.35), Xenosphecinae (104.84, 85.57-123.57 Ma) e posteriormente o clado Crabroninae + Larrinae (94.34, 75.68-112.47 Ma). Fóssil H é grupo irmão de todos os Crabronidae *strictu sensu* atuais. Esse fóssil proveniente do âmbar burmês é o registro mais antigo dessa família. Fóssil H pode ser considerado como representante do grupo *crown* pois possui a principal sinapomorfia da família, onde a carena mesepisternal é completa e em contato com a borda anterior do mesepisterno longe da linha média do corpo (84:0).

Bembicidae é a segunda maior família de vespas apoideas em número de espécies, mais de 1600 espécies atualmente (Pulawski 2018). Em nossas estimativas a idade do grupo *crown* foi estimada em 127.69 Ma (116.19-139.02 Ma). Os únicos fósseis atribuídos a essa família foram descritos do Cenozóico (Pulawski 2018). Os nossos resultados mostraram que pelo menos três fósseis do Cretáceo podem ser atribuídos a essa família. Adicionalmente, nossos resultados apontam que †*Palanga* também está associado a essa linhagem. †*Palanga* é proveniente do âmbar báltico e foi originalmente descrita em Pemphredonidae (Budrys 1993). Assim como o Fóssil M e Fóssil O, †*Palanga* possui os arólios das pernas anteriores muito maiores que as demais pernas (68:1). Essa transformação é única dos Bembicidae e aparentemente surgiu diversas vezes dentro da família. As relações entre esses táxons são instáveis em nossos resultados e merecem investigações futuras. Entretanto, a ausência da carena escutal oblíqua (82:1) nesses fósseis remete a um posicionamento basal na linhagem como um todo.

Astatidae é definido aqui incluindo *Eremiasphecium* (Eremiasphecinae), Ammoplaninae, Fóssil E e Fóssil F. Esse clado foi apresentado pela primeira vez em Melo (1999) e é corroborado em nossos dados. As nossas estimativas de idade para essa família foi de 121.07 Ma (108.89-133.06 Ma). Essas estimativas são muito mais antigas que aquelas apresentadas em Peters et al. (2017) e Sann et al. (2018), 52 Ma (30-81 Ma) e 71.2 Ma (52.48-161.29 Ma), respectivamente. Entretanto, essas idades são mais congruentes com as nossas estimativas para o clado Eremiasphecinae + Astatidae *strictu sensu*, 65.56 Ma (27.68-104.19 Ma). Definimos Astatidae aqui para todos aqueles táxons próximos a essa linhagem e que possuem a fossa tentorial das fêmeas acima da linha das bordas inferiores dos alvéolos antenais (26:1) e a célula marginal mais curta que o pterostigma (130:1). A primeira transformação também surge convergentemente em Spilomeninae (Pemphredonidae) e diversas vezes em Crabronidae *strictu sensu*. Por outro lado, a célula marginal mais curta que o pterostigma (130:1) surge apenas em Astatidae dentre as linhagens do clado 8. Os resultados de Sann et al. (2018) sugerem que os Ammoplaninae estão relacionados com as abelhas. Essa hipótese foi altamente rejeitada em nossos resultados (KRS = 10.74). Adicionalmente, não existem atribuições morfológicas para essa relação, que também foi rejeitada em outros trabalhos moleculares (Devebec et al. 2012; Branstetter et al. 2017).

Os resultados de Melo (1999) sugerem a monofilia de Pemphredonidae + Psenidae. Em nossos resultados essas duas linhagens são apresentadas independentemente. Pemphredonidae é definido aqui incluindo o Fóssil L mais todas aquelas subtribos incluídas em Pemphredonini por Melo (1999).

†*Cretospilomena* foi descrita do âmbar burmês e †*Psolimena* do âmbar jersiano, ambas por Antropov (2000a; 2000b). Esses dois gêneros foram originalmente descritos em Pemphredonidae, uma posição corroborada por nossos resultados. O grupo *crown* de Pemphredonidae foi estimado para o Cretáceo inferior 120.65 (108.66-132.69) Ma. Essas estimativas foram relativamente mais antigas que aquelas estimadas por Branstetter et al. (2017), Peter et al. (2017) e Sann et al. (2018). Os fósseis de Pemphredonidae são muito comuns ao longo do Cretáceo superior e Cenozóico, com mais de 30 registros segundo o Fossilworks (Alroy 2018), sendo o grupo com o mais abundante e diverso registro fóssil dentre as vespas apoideas.

Philanthidae é definido aqui contendo os oito gêneros tradicionalmente atribuídos a essa família (Pulawski 2018). Nós utilizamos cinco dos gêneros e que representaram as três linhagens de Alexander (1992b). Nos resultados de Alexander (1992b), *Cerceris*, *Eucerceris* e *Pseudoscolia* formam a tribo Cercerini, *Aphilanthops* e *Clypeadon* formam Aphilanthopsini e *Philantus*, *Philanthinus* e *Trachypus* formam Philanthini. Essa hipótese foi corroborada por Sann et al. (2018). Os nossos resultados não indicaram essas relações (Figura 1, 2, 3 e 5). Entretanto os nossos dados não parecem substanciais para elucidar essas relações. Philanthidae é a família de grupo *crown* mais jovem entre os Apoidea. Nossas estimativas apontam para um surgimento no Cenozoico, 53.83 Ma (27.13-82.45 Ma). Contudo, as estimativas moleculares sugerem um surgimento no final do Cretáceo superior (Branstatter et al. 2017; Peters et al. 2017; Sann et al. 2018). Os nossos dados provavelmente sofreram interferência pela não inclusão de fósseis dessa linhagem e também de caracteres adicionais que suportariam as relações internas dessa linhagem (veja discussão no tópico 4.2.2). Diversos fósseis de rocha do Mioceno são atribuídos a Philanthidae, mas não tivemos a oportunidade de incluí-los em nossas análises.

Psenidae é definido aqui incluindo *Odontosphex*, *Entomosericus*, Pseninae, Fóssil I, Fóssil J e Fóssil K. Em nossos resultados esse clado apresentou alto valor de suporte de ramos (PP = 1). A inclusão de *Odontosphex* em Psenidae também é sugerida por Sann et al. (2018) em dados moleculares. No trabalho de Sann et al. (2018) os autores argumentam que é possível que *Entomosericus* esteja mais relacionado com as abelhas que Ammoplaninae. Os autores comentam ainda que como não houve a inclusão desse táxon, o seu posicionamento permanece incerto e que seria melhor manter essa “linhagem enigmática” em *incertae sedis*. A hipótese de inclusão de *Odontosphex* e *Entomosericus* em Psenidae foi proposta pela primeira vez por Melo (1999). *Entomosericus* não apresenta nenhuma característica “enigmática” frente aos demais Psenidae e considerar o posicionamento desse gênero em *incertae sedis* não faria sentido. Todos os táxons representados em Psenidae em nossas análises possuem inequivocamente as seguintes sinapomorfias: o côndilo externo do fêmur posterior das fêmeas é amplo e truncado (112:1); os pelos espiniformes da perna posterior em formato espatulado e distribuídos em duas fileiras bem definidas (117:1); a porção basal do grádulo do esterno II de forma laminar e direcionada para dentro, formando uma superfície articular especializada com a porção posterior diferenciada da borda lateral do esterno I (161:1) (Fig. 77 em Melo 1999); e o integumento da larva com espículas pequenas e densas (190:1). Nos fósseis analisados a codificação do caráter 190 não foi possível

porque as larvas desses táxons não são conhecidas. O Caráter 161 também só foi possível determinar para o Fóssil K. Portanto, existem fortes evidências para considerar esse clado com a inclusão desses táxons.

De todos as linhagens viventes de vespas apoideas os Psenidae estão mais próximos das abelhas (PP = 0.5). Esse posicionamento também é corroborado por alguns dados moleculares (Peters et al. 2017). O grau de fósseis aqui determinado como “†Cirrosphecidae” apresenta uma relação ainda mais estreita com as abelhas. Assim como as abelhas os “†Cirrosphecidae” possuem as garras com um dente subbasal (71:1). Podemos destacar ainda as principais transformações relacionadas aos Psenidae, “†Cirrosphecidae” e as abelhas. No trabalho de Michener (2007) o autor apresenta as fimbrias pigidiaias e pré-pigidiaias das abelhas (Figura 10.13 em Michener 2007 e Apêndice 1 Figura 2). Além das abelhas, as fimbrias presentes no T5 e que cobrem o T6 (pré-pigidiaias, 165:1) também são compartilhadas com *Odontosphex*, Fóssil I, Fóssil J e Fóssil Q. Entretanto as fimbrias que cobrem lateralmente a placa pigidiaial no T6 (pigidiaias, 166:1) podem ser consideradas uma sinapomorfia para o clado Psenidae + “†Cirrosphecidae” + Apidae *lato sensu*. A transformação 166:1 é perdida apenas na subfamília Pseninae, aqui representados como Fóssil K, *Mimesa*, *Psenulus* e *Pluto*.

O clado Apidae *lato sensu* apresentou alto suporte de ramos (PP = 1) e é corroborado em todos os trabalhos conduzidos com Apoidea (Alexander 1992a; Melo 1999; Devebec et al. 2012; Peters et al. 2017; Branstetter et al. 2017; Sann et al. 2018). Em nossos resultados as abelhas de língua longa e as de língua curta formaram um grupo monofilético irmão de Melittinae. Esse resultado também é coerente com trabalhos prévios (Devebec et al. 2012; Peters et al. 2017; Branstetter et al. 2017; Sann et al. 2018). Entretanto, como o relacionamento interno do clado das abelhas não foi o foco de nosso estudo, os dados não parecem substanciais para elucidar essas relações, principalmente no relacionamento das abelhas de língua curta. Em nossos resultados o grupo *crown* de abelhas foi estimado em 97.54 Ma (85.97-110.16 Ma). As nossas estimativas são muito mais jovens que aquelas propostas por Cardinal & Danforth (2013) (123, 113–132 Ma). Entretanto as nossas estimativas são mais coerentes com as apresentadas em Peters et al. (2017) (111, 93-132 Ma), Branstetter et al. (2017) (98.89 Ma) e Sann et al. (2018) (113, 95.39-135.62 Ma). Finalmente, as abelhas são classificadas tradicionalmente como famílias distintas (Michener 2007). Entretanto, Melo (1999) propõe o uso de Apidae *lato sensu* e incluí todas as famílias de abelhas como subfamílias. Considerando as idades dos grupos *crown* das linhagens de crabronídeos e as abelhas, algo em torno de 100 Ma e 130Ma, parece uma proposta coerente. Nossos resultados mostram que as linhagens de Apoidea evoluíram em momentos compatíveis e para uma classificação ‘igualitária’ entre as demais famílias a utilização de Apidae *lato sensu* é adequada.

4.2 Estimativa do tempo de divergência com relógio morfológico

4.2.1 Escolha de priors em relógio morfológico

Como indicado na introdução foram poucos os trabalhos que se propuseram a investigar o uso de diferentes combinações de *priors* de relógios de árvores e de relógios de ramos em uma abordagem estritamente *tip-dating* com dados morfológicos (Lee et al. 2014; Beck & Lee 2014; Matzke & Wright. 2016; Puttick et al. 2016). O nosso trabalho é o mais extenso na investigação comparativa desses dois tipos de *priors* com 20 modelos avaliados. Segundo nossos resultados, algumas combinações de *prior* resultaram em aumento considerável na MgL.

Os resultados de MgL sugeriram como melhor modelo (Modelo 6) a combinação de um *prior* de modelo de árvores FBD com uma amostragem *fossiltip*, *prior* de comprimento de ramos IGR com taxas *lognormal* e *prior lognormal* de raiz. O melhor modelo apresentou evidência muito forte ($KRS = 25.00$) contra o segundo melhor modelo (Modelo 8). Esse sugere uma combinação muito diferente da encontrada no Modelo 6: um *prior* de modelo de árvores *uniforme*, *prior* de comprimento de ramos IGR com uma taxa *lognormal* e um *prior* de raiz *lognormal*. Surpreendentemente, o pior modelo de todos (Modelo 3, $KRS = 575.84$) apresentou uma parametrização mais próxima do Modelo 6 que do Modelo 8: *Brlenspr* (FBD), *Samplestrat* (*random*), *Clockvarpr* (IGR), *Clockratepr* (*default*) e *Treeagepr* (*lognormal*). O principal padrão encontrado é que nossos resultados refutam fortemente o uso de taxas de relógio de ramos *default* (taxas fixas) em detrimento de taxas com distribuição *lognormal*. Nossos resultados também apresentaram preferência por modelos com *prior* de comprimento de ramos FBD e *prior* de comprimento de ramos IGR.

No trabalho de Matzke & Wright (2016), os autores encontraram que os modelos com *prior* de árvore *uniforme* tendem a produzir estimativas muito antigas e pouco realistas. De maneira similar, em Puttick et al. (2016) tanto os modelos FBD quanto uniformes geraram idades muito antigas e não realistas na base da árvore de mamíferos placentários (do ponto de vista dos dados moleculares). Os nossos resultados apontaram para valores maiores de verossimilhança marginal nos modelos FBD (modelo 6), porém ao contrário do apresentado em Matzke & Wright (2016) os nossos modelos uniformes obtiveram alta correlação com os modelos FBD ($R^2 > 0.80$). Além disso, ao contrário do proposto por Puttick et al. (2016), os nossos resultados apontaram uma forte correlação para com as hipóteses moleculares utilizando ambos os *priors* ($R^2 > 0.60$). Por fim, os modelos com *prior* de árvores *uniforme* sem a designação do *prior* de raiz obtiveram as estimativas mais jovens, principalmente nas linhagens na base da árvore. Isso mostra que a combinação do *prior* de árvores com outros *priors* podem trazer diferentes perspectivas na estimativa de tempos de divergência em relógio morfológico.

Assim como nos trabalhos de Lee et al. (2014) e Beck & Lee. 2015, os nossos resultados indicaram maior valor de MgL para um modelo com relógio de ramos IGR. Beck & Lee (2015) afirmam que a despeito

dos modelos IGR e TK02 possuírem premissas distintas no ajuste das taxas, os dois modelos apresentam resultados concordantes. Assim como no trabalho citado anteriormente, os nossos resultados demonstraram alta correlação entre os dois modelos ($R^2 > 0.80$). Os modelos com *prior* de comprimento de ramos TK02 resultaram em estimativas com menor amplitude de intervalo (95%HPD) quando comparados com os modelos IGR. O viés encontrado na designação do *prior* de raiz não surtiu efeito entre o relógio de ramos. Ambas designações não apresentaram diferenças no intervalo quando combinados com TK2 ou IGR.

Como argumentado anteriormente as idades médias da estimativa do tempo de divergência para os 16 clados de Apoidea apresentaram forte correlação entre as 10 hipóteses avaliadas, tanto por relógio morfológico quanto molecular ($R^2 > 0.61$). A despeito desses resultados indicarem forte correlação entre os modelos de relógios avaliados, os valores de MgL foram altamente discrepantes. O uso comparativo de MgL via SS proporciona a escolha de modelos eficientes evitando modelos superparametrizados, ou seja, evitando-se modelos com menor correspondência entre o *prior* e a distribuição posterior (Rosa et al. 2019). Por essa razão, as análises de sensibilidade (testes de modelos) são altamente recomendadas por diversos autores para diferentes abordagens Bayesianas pois o reconhecimento de parâmetros eficientes em um extenso espaço amostral é essencial para a inferência filogenética final (Nylander et al. 2004; Clarke & Middleton 2008; Fan et al. 2011; Xie et al. 2011; Harrison & Larsson 2015; Rosa et al. 2019). Portanto, mesmo que a correlação entre os modelos tenha sido alta, a escolha dos melhores modelos para estimativa de tempo de divergência para dados morfológicos deve ser realizada em um amplo teste amostral e com seleção por MgL.

4.2.2 Vantagens e desvantagens do relógio morfológico

O relógio morfológico apresentou vantagens e desvantagens na estimativa do tempo de divergência de linhagens. Algumas desvantagens são óbvias como a necessidade de um registro fóssil muito rico tanto em quantidade quanto em escala de tempo geológico. Ou vantagens inerentes ao procedimento de relógio como a estimativa simultânea dos tempos de divergência e das relações filogenéticas. Nesse seguimento discutiremos alguns pontos importantes sobre esse método de análise.

Alguns autores destacam a ineficiência do relógio morfológico usando *tip-dating* frente ao relógio molecular com *node-dating*. O'Reilly et al. (2015) afirmam por exemplo que o *node-dating* gera maior precisão na estimativa pontual dos clados usando como base o intervalo 95%HPD (O'Reilly et al. 2015). Os nossos resultados mostram o oposto. Os modelos morfológicos utilizando *tip-dating* apresentaram a metade dos valores de intervalos de 95% HPD quando comparados com as hipóteses moleculares (Branstetter et al. 2017; Peters et al. 2017; Sann et al. 2018).

Frequentemente o relógio morfológico gera uma grande diferença nas topologias em comparação com os relógios moleculares (Beck & Lee. 2015; Puttick et al. 2016; O'Reilly et al. 2015). De fato, em nossos resultados as análises conduzidas com estimativa de tempo de divergência entre as linhagens de Apoidea resultaram em topologias diferentes daquelas sugeridas por dados moleculares (Branstetter et al. 2017; Peters

et al. 2017; Sann et al. 2018). Porém isso é inerente à incongruência resultante do tipo de informação filogenética. Até mesmo os nossos resultados em inferência Bayesiana sem a aplicação de relógio diferem consideravelmente das hipóteses moleculares.

Na análise dos Apoidea foram encontradas hipóteses discrepantes como o posicionamento de Sphecidae + Crabroninae *stricto sensu* sugeridas por dados moleculares, o clado Crabronidae *lato sensu* + Apidae sugerido por dados morfológicos ou mesmo os frágeis suportes de relacionamento das famílias da grande linhagem dos crabronídeos sugeridos por ambas as fontes de dados. Entretanto essa questão não interferiu nas estimativas de tempo de divergência das grandes linhagens de Apoidea. Visto que todas as grandes linhagens de Apoidea possuem estimativas médias de tempo de divergência equivalentes, independente da hipótese topológica, não houve interferências consideráveis na estimativa de tempo com diferentes topologias. Esse fenômeno pode ser facilmente atribuído ao padrão de diversificação encontrado para as grandes linhagens da superfamília. A linhagem de Apoidea como um todo apresentou rápida diversificação em um período relativamente curto para uma linhagem de mais de 150 Ma. Esse rápido processo encurtou os relacionamentos da base da árvore interferindo fortemente na instabilidade da topologia, mas não interferindo na estimativa do tempo de divergência das linhagens.

De maneira geral, o relógio morfológico pode apresentar estimativas de idades muito antigas e não realistas quando comparadas com aquelas geradas a partir de dados moleculares (Beck & Lee. 2015; O'Reilly et al. 2015; Matzke & Wright 2016; Puttick et al. 2016). Essa diferença não foi encontrada em nossos resultados. As estimativas atribuídas para as grandes linhagens de Apoidea estão fortemente correlacionados com os dados moleculares. Além de que, os resultados moleculares de Peters et al. (2017) e Sann et al. (2018) apresentaram idades ligeiramente mais antigas.

Beck & Lee (2014) atribuem essas grandes discrepâncias entre as estimativas baseadas em *tip-dating* com dados morfológicos e *node-dating* com dados moleculares a mudanças relativamente pequenas nas taxas evolutivas ao longo do tempo. Além disso, as datas do relógio morfológico são notavelmente sensíveis a inclusão ou exclusão de táxons ao longo da amostragem (Lee et al. 2014). De fato, em certos pontos da árvore de Apoidea algumas estimativas parecem sofrer maior distorção (i.e., estimativas muito jovens) decorrente a falta de informação filogenética atribuída a uma determinada linhagem (número de caracteres). Ao mesmo passo em que a sensibilidade na inclusão e exclusão dos táxons foi um grande desafio para a escolha dos terminais fósseis ao longo do nosso trabalho. Aqueles clados que não apresentavam fósseis, como por exemplo o grupo *crown* de Sphecidae, apresentaram estimativas muito jovens e com intervalo 95%HPD muito maiores. Portanto, a sensibilidade do relógio morfológico ao número de caracteres mutáveis ao longo de uma linhagem bem como a inclusão e exclusão de táxons são as desvantagens mais perceptíveis do método. Da mesma maneira em que são atribuições pouco compreendidas e merecem investigações futuras.

Dentre as vantagens, a reconstrução das relações filogenéticas ao mesmo passo em que o tempo de divergência das linhagens é estimado parece a mais atrativa. Esse ponto é muito relevante pois uma estimativa

simultânea gera um viés menor no posicionamento do ponto de calibração e evitando atribuições errôneas de *node-dating* (Lee & Palci 2015). Em nossos resultados existe um contraste nas atribuições de calibração quando comparamos o posicionamento dos fósseis de Apoidea em nossos resultados filogenéticos com aqueles realizados por *node-dating* em trabalhos moleculares recentes. As atribuições de †*Apodolichurus* em Peters et al. (2017) e de †*Cretampulex* em Sann et al. (2018) divergem dos nossos resultados. Entretanto, os nossos resultados mostraram que as análises conduzidas com estimativa de tempo de divergência entre as linhagens sugeriram topologias consideravelmente diferentes daquela sugerida sem a aplicação do relógio. As relações filogenéticas geradas com o relógio morfológico sugerem um relacionamento simplificado em relação à topologia atribuída por inferência Bayesiana, principalmente nos grupos basais de Apoidea. Aparentemente esses clados são artificialmente atraídos uns pelos outros devido à sua proximidade no tempo geológico bem como por um baixo número de caracteres que os diferencia. Por fim, através do relógio morfológico também conseguimos estimar o tempo de divergência de grupos inteiramente fósseis como as linhagens extintas de Heterogynaidae ou ainda dos grupos basais de Apoidea que não possuem contraparte na fauna moderna. Associado a esse fator conseguimos estimar padrões macroevolutivos adicionais com a estimativa de diversidade de grandes linhagens de Apoidea desconhecidas atualmente bem como períodos de extinção em massa para a linhagem como um todo.

4.3 Taxas de diversificação e LTTs com a inclusão de fósseis

4.3.1 Amostragem dos LTTs e interferência nas taxas de diversificação

Como citado anteriormente as análises de estimativas de diversificação de linhagens ao longo do tempo (LTTs) utilizam filogenias datadas para fazer aproximações de padrões macroevolutivos ligados ao surgimento de espécies. Por essa razão a principal premissa desse método é a disponibilidade de uma filogenia que além de datada comporte uma grande amostragem de espécies das linhagens a serem estimadas. Essa premissa é importante pois os LTTs são muito suscetíveis a amostragens incompletas gerando estimativas artificiais (Pybus & Harvey 2000; Ricklefs 2007; Pie 2009; Cusimano et al. 2010). Por outro lado, obter uma filogenia completa de linhagens muito diversas torna-se um grande desafio.

Para estimar esses processos macroevolutivos com filogenias incompletas existem maneiras de minimizar esse viés. É possível incluir a diversidade taxonômica não representada em uma determinada árvore utilizando um *prior* com o número de espécies por gêneros ou linhagens maiores (Paradis 2003; Pie 2009; Almeida et al. 2011; Almeida et al. 2018) ou ainda fazer um “enxerto” com árvores mais completas provenientes de trabalhos secundários (Arcila & Tyler 2017). Essas atribuições são importantes para estimar taxas de diversificações mais realistas para as linhagens. Para o primeiro caso é necessária uma árvore com boa parte ou todos os gêneros da linhagem para que o viés de aproximação seja zero (Couvreur et al. 2011a;

Couvreur et al. 2011b). No segundo caso a árvore “enxerto” deve ter uma amostragem praticamente completa de táxons. Como mencionado anteriormente a nossa amostragem representa aproximadamente 0.36% (99 espécies) do total das espécies viventes de Apoidea (mais de 27.500 espécies somando as estimativas de Michener 2007 e Pulwaski 2018). E outros trabalhos com números expressivos de terminais chegam a 0.49% (136 espécies) e 0.50% (138 espécies) do total de espécies nos trabalhos de Sann et al. (2018) e Devebec et al. (2012) respectivamente. Infelizmente nenhuma das duas abordagens puderam ser implementadas para a estimativa da diversificação de Apoidea como um todo.

Para reduzir o viés atribuído à amostragem incompleta o nosso objetivo aqui foi investigar a diversificação das grandes linhagens de Apoidea ao longo do Cretáceo, em outras palavras, concentramos nossas análises na diversificação inicial dos grandes grupos da superfamília. Portanto a amostragem usada representou 100% das grandes linhagens viventes a nível de subfamília/ família. Isso significou que a amostragem das linhagens viventes foi completa na porção mais antiga da árvore e tornando-se progressivamente mais incompleta em relação ao presente (Ricklefs 2007; Cusimano et al. 2010; Couvreur et al. 2011a; Couvreur et al. 2011b). Essa abordagem foi baseada em trabalhos que visaram a estimativa de padrões macroevolutivos de linhagens megadiversas como as monocotiledôneas (mais de 60 mil espécies) (Hertweck et al. 2015) ou ainda todas as plantas terrestres (mais de 280 mil espécies) (Fiz-Palacios et al. 2011). Nesses dois trabalhos foram usadas abordagens similares a nível de família ou mesmo de ordem. Os nossos resultados demonstraram-se factíveis ao estimar os eventos de diversificação e extinção das grandes linhagens de Apoidea durante o Cretáceo tendo em vista a amplitude de amostragem, o registro fóssil avaliado e os padrões dos LTTs frente às simulações ao acaso.

O implemento da diversidade taxonômica atual para calcular as taxas de diversificações é uma estimativa previsível para as linhagens viventes de maneira mais ou menos realistas. Por outro lado, uma estimativa minimamente razoável da real diversidade dos grupos fósseis é bastante difícil. As simulações com os modelos PB e BD sugeriram que as linhagens de Apoidea continuam crescendo ao longo do tempo até o presente. O nosso gráfico LTT incluindo os fósseis (ou acúmulo de linhagens) mostrou uma clara diferença na diversidade de linhagens no Cretáceo médio em relação ao presente. No Cretáceo médio houve o maior acúmulo de linhagens a nível de subfamília da história evolutiva de Apoidea. Esse resultado seria falseado se a diversidade atual total tivesse sido implementada.

O mesmo vale para as estimativas das taxas de diversificações. Nas taxas de diversificação ao longo do tempo os nossos resultados mostraram um aumento significativo no início da diversificação das grandes linhagens com uma diminuição da mesma e final estabilidade até o presente. No panorama atual tem-se mais de 8.9000 espécies de vespas apoideas distribuídas em 10 famílias. No Cretáceo médio existem pelo menos 18 famílias sendo oito delas extintas. Os nossos resultados mostram que a diversidade de vespas apoideas durante o Cretáceo pode ser sido muito maior que a diversidade atual. Nesse contexto, quaisquer estimativas da diversidade real das linhagens fósseis seriam não realistas. Essa questão é muito importante pois se as taxas de

diversificação tivessem sido estimadas com base na diversidade atual provavelmente essa estabilidade das grandes linhagens não aconteceria. Como destacado para o acúmulo de linhagens em nossos resultados, a não inclusão de uma estimativa plausível da diversidade do registro fóssil provavelmente levaria a estimativas de taxas de diversificações contrárias as dos resultados apresentados. A inclusão da diversidade atual e não inclusão da diversidade fóssil provavelmente levaria a menores taxas de diversificação na base da história evolutiva de Apoidea. Para perguntas a níveis taxonômicos mais altos (família/ subfamília) utilizando filogenias com fósseis, sugerimos a não inclusão da diversidade atual.

Outro ponto importante na estimativa de diversificação de Apoidea é que a grande porção da diversidade atual é relativamente nova, ou ao menos pós Cretáceo médio. A Tabela 10 apresenta as idades estimadas do grupo *crown* dos táxons mais diversos em Apoidea conforme as nossas estimativas de tempo de divergência. Mesmo que não consideremos a diversidade total da superfamília, todas as generalizações realizadas para as grandes linhagens de Apoidea ao longo do Cretáceo inferior apresentaram o mínimo de viés. Com exceção de Apidae *lato sensu*, esses táxons não representam as grandes linhagens avaliadas. As sete linhagens atuais de abelhas (subfamílias) estão presentes em nossos LTTs e aparentemente não causaram ruído nas análises. Trabalhos futuros com amostragem completa desses táxons podem avaliar essas estimativas mínimas apresentadas aqui de maneira empírica.

Tabela 10. Táxons mega diversos em Apoidea com suas idades de grupo *crown* estimadas empiricamente. N: número atual de espécies com base em Michener 2007 e Pulawski (2018).

Táxon	N	Épocas estimadas de diversificação (mínima)	Clados de referências
Sphecidae	~783	Paleoceno	Após ~59.21 ma. Grupo <i>crown</i> de Sphecidae.
<i>Ammophila</i>	~240	Mioceno	Após ~13 ma. Idade entre <i>Palmodes</i> e <i>Podalonia</i> .
<i>Bembix</i>	~330	Mioceno	Após ~13 ma. Idade entre <i>Stictia</i> e <i>Glenostictia</i> .
Crabronini (48 gêneros)	~1500	Eoceno	Após ~53 ma. Idade entre Crabronini e Oxybelini.
<i>Oxybelus</i>	~260	Eoceno	Após ~53 ma. Idade entre Crabronini e Oxybelini.
Larrini (15 gêneros)	~1320	Paleoceno	Após 59.96 ma. Idade entre Larrini e Trypoxylini.
<i>Tachysphex</i>	~450	Oligoceno	Após ~32.69 ma. Idade entre <i>Gastrosericus</i> e <i>Tachytes</i> .
<i>Tachytes</i>	~300	Oligoceno	Após ~32.69 ma. Idade entre <i>Gastrosericus</i> e <i>Tachytes</i> .
<i>Liris</i>	~310	Oligoceno	Após ~32.69 ma. Idade entre <i>Gastrosericus</i> e <i>Tachytes</i> .
Trypoxylini (7 gêneros)	~950	Eoceno	Após ~52.77 ma. Idade entre Trypoxylini e <i>Sericophorus</i> .
<i>Pison</i>	~280	Paleoceno	Após ~52.77 ma. Idade entre Trypoxylini e <i>Sericophorus</i> .
<i>Trypoxylon</i>	~630	Paleoceno	Após ~52.77 ma. Idade entre Trypoxylini e <i>Sericophorus</i> .
<i>Cerceris</i>	~870	Mioceno	Após ~18.48 ma. Idade entre <i>Cerceris</i> e <i>Pseudoscolia</i> .
Apidae <i>lato sensu</i>	+17500	Cretáceo superior	Após ~97.54. Grupo <i>crown</i> de Apidae <i>lato sensu</i> .

4.3.2 Influência do registro fóssil nos LTTs

Se nas análises de diversificações os vieses inerentes ao método comprometem as avaliações dos padrões, as estimativas de taxas ou eventos de extinções são ainda mais difíceis de prever. Uma primeira aproximação da estimativa de taxas e eventos de extinção foi realizada utilizando filogenias moleculares sem a inclusão de fósseis (Nee et al. 1994; Harvey et al. 1994; Nee 2006). Esses autores utilizaram modelos *birth-death* para estimar as taxas de extinções simuladas ao longo das linhagens onde o resultado dessa LTT apresenta o padrão

chamado “*pull-of-the-present*”. Esse padrão descreve uma aceleração no acúmulo de linhagens próximo ao presente revelando linhagens que passaram por um grande evento de extinção (Nee et al. 1994; Nee 2006). Os nossos resultados sem a inclusão de fósseis (Figura 8) apresentou uma acumulação contínua de linhagens ao longo do tempo sem exibir o “*pull-of-the-present*”. Em outras palavras, os nossos resultados sem os fósseis não foram capazes de determinar qualquer evento de extinção ao longo da história evolutiva de Apoidea.

Ainda nesse contexto, uma amostragem sem o registro fóssil é um panorama macroevolutivo pouco realista e apresenta um grande viés na estimativa de eventos de extinção (Ricklefs 2007; Etienne et al. 2012). Diversos autores demonstraram que filogenias moleculares sem a inclusão de fósseis são incapazes de estimar taxas e eventos de extinção (Quental & Marshall 2010; Rabosky 2010). Para os Apoidea, nossos resultados com a inclusão de fósseis corroboram essa premissa, pois os LTTs da Figura 7 exibem um claro evento de extinção durante o período Cretáceo. Esse evento foi imperceptível nas análises sem os fósseis corroborando com trabalhos que realizaram um experimento inclusivo e exclusivo semelhante para um mesmo conjunto de dados (Morlon et al. 2011; Xing et al. 2014; Arcila & Tyler 2017). Contudo é de se destacar que se as estimativas de taxas e eventos de extinção são altamente suscetíveis ao registro fóssil, tais padrões podem ser falseados se o registro fóssil pós evento não for avaliado. Em outras palavras uma descontinuidade amostral no registro fóssil pode gerar falsos eventos de extinção em uma determinada linhagem.

4.4 Padrões de diversificação e extinção das grandes linhagens de Apoidea

Segundo os nossos resultados, os Apoidea surgiram a aproximadamente 150 Ma no final do período Jurássico. As taxas de diversificação das grandes linhagens foram inicialmente altas ao longo Cretáceo inferior com uma diminuição durante o Cretáceo superior e permanecendo mais ou menos constantes até o presente ($\gamma < 0$). Durante o Cretáceo inferior também houve o maior acúmulo de linhagens a nível de subfamílias em toda a história evolutiva de Apoidea. Esse acúmulo é maior que a diversidade atual de grandes linhagens de Apoidea. Ainda durante o início do Cretáceo inferior surgiram todas as linhagens que deram origem às famílias atuais de Apoidea. No início do Cretáceo inferior provavelmente também surgiram os grupos *crown* de “†Angarosphecidae 1” e Fóssil A. Durante o Barremiano, aproximadamente entre 129-135 Ma (Cretáceo inferior), surgiram as linhagens *crown* de “†Angarosphecidae 2”, Crabronidae *strictu sensu* e Bembicidae. Posteriormente, com exceção de Sphecidae e Philanthidae, todos os demais grupos *crown* das linhagens atuais de Apoidea surgiram entre o Aptiano e o Albiano, aproximadamente entre 125-100 Ma (Cretáceo inferior).

O Cretáceo inferior e o KTR foram os momentos mais importante na diversificação das grandes linhagens de Apoidea (Figura 10 e Apêndice 7 Figura 1). Os nossos dados indicam que as altas taxas de diversificações e acúmulo gradativo das grandes linhagens de Apoidea corresponde ao KTR (Figura 10c). Se a história evolutiva de Apoidea for dividida em cinco intervalos de 30 Ma, como na Figura 6, a diversificação das grandes linhagens aconteceu no primeiro intervalo, ou seja, no quinto inicial da história evolutiva total.

Portanto, os nossos dados indicam que o surgimento das grandes linhagens de Apoidea foi uma rápida diversificação em um espaço de tempo relativamente curto tendo em vista a história total da superfamília (~150Ma). Muitos autores indicam que grupos que apresentam rápida radiação em intervalos relativamente curtos de tempo podem apresentar conflitos topológicos decorridas desse padrão de diversificação (Jackman et al. 1999; Lehtonen et al. 2012; Sterli et al. 2013; Suh et al. 2015; Suh et al. 2016; Longo et al. 2017; Rouard et al. 2018). Nesse contexto, é possível que os conflitos topológicos entre os grupos basais de Apoidea e principalmente entre as famílias da linhagem de crabronídeos possam ser explicados com um cenário de rápida diversificação em curto intervalo de tempo.

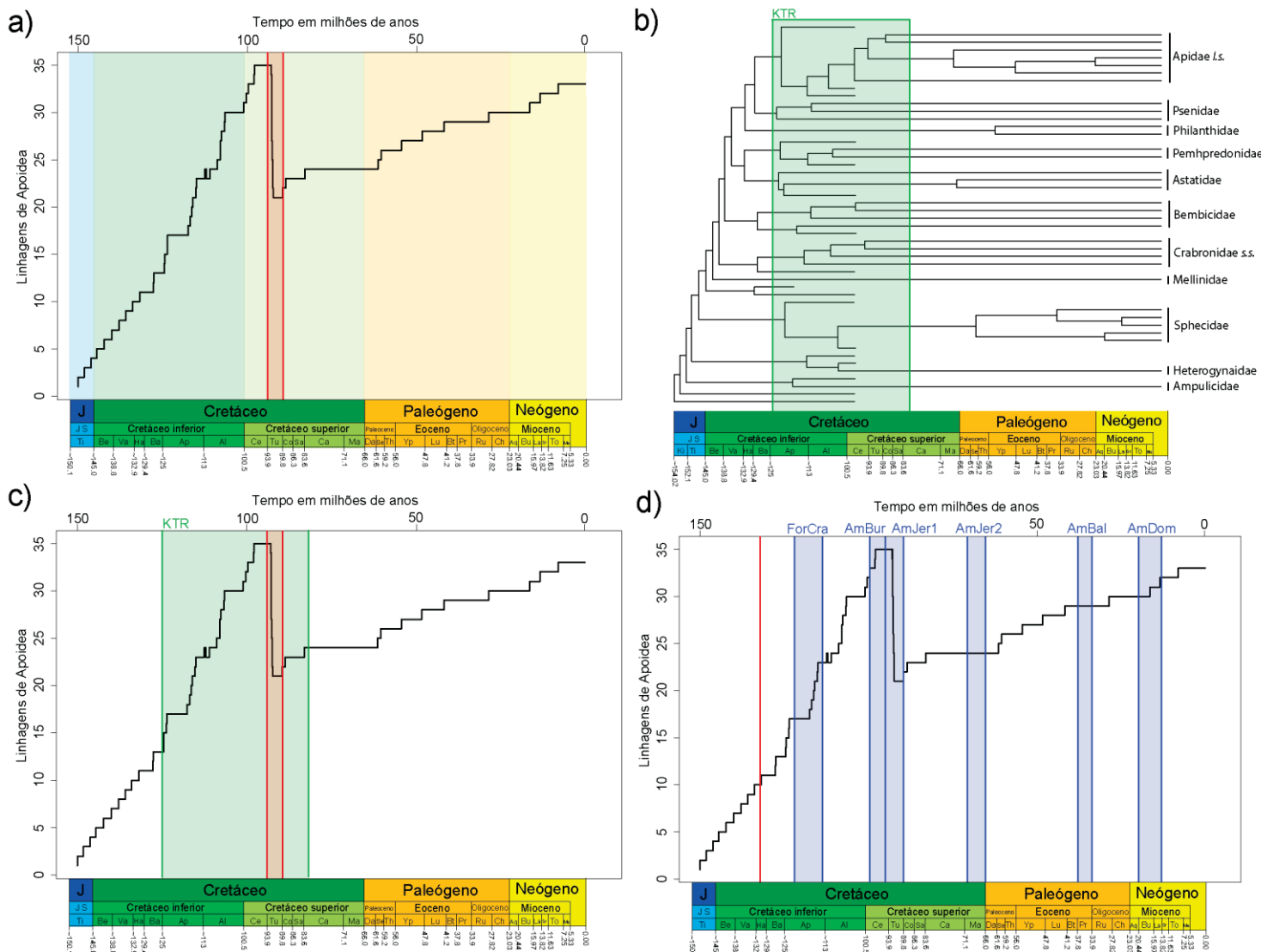


Figura 10. Análises de diversificações das grandes linhagens de Apoidea incluindo os fósseis. **a)** LTT contraposta com a escala de tempo geológico. Em vermelho o episódio de extinção das linhagens de Apoidea. **b)** cronograma por subfamília resultante de uma análise de IB com dados morfológicos. Em verde o intervalo correspondente ao KTR (125–80 Ma). **c)** LTT contraposta com o KTR (em verde) e com o episódio de extinção (em vermelho) das linhagens de Apoidea. **d)** LTT contraposta com os depósitos fósseis usados em nosso estudo (em roxo) e o fóssil mais antigo de Apoidea (linha vermelha). ForCra: Formação Crato. AmBur: âmbar Burmês. AmJer1: âmbar Jersiano 1. AmJer2: âmbar Jersiano 2. AmBal: âmbar Báltico. AmDom: âmbar Dominicano. Abreviações da escala de tempo geológico: J: Jurássico. JS: Jurássico superior. Ti: Tithoniano. Be: Berriasiano. Va: Valanginiano. Ha: Hauteriviano. Ba: Barremiano. Ap: Aptiano. Al: Albiano. Ce: Cenomaniano. Tu: Turoniano. Co: Coniaciano. Sa: Santoniano. Ca: Campaniano. Ma: Maastrichtiano. Da: Daniano. Se: Selandiano. Th: Thanetiano. Yp: Ypresiano. Lu: Lutetiano. Bt: Bartoniano. Pr: Priaboniano. Ru: Rupeliano. Ch: Chattiano. Aq: Aquitaniano. Bu: Burdigaliano. La: Langhiano. Sr: Serravalliano. To: Tortonian. Me: Messiniano.

Outro interessante padrão na história evolutiva dos Apoidea foi um grande evento de extinção na metade do Cretáceo (Figura 10a e 10b). As nossas LTTs mostraram uma queda abrupta na acumulação de linhagens durante o Turoniano, aproximadamente entre 89-93 Ma (Cretáceo superior). Boa parte das linhagens basais que surgiram no início do Cretáceo ou no final do Jurássico e que se diversificaram e permaneceram durante o Cretáceo inferior provavelmente foram extintas na parcela final do KTR. Adicionalmente, algumas linhagens derivadas de Apoidea também se extinguíram nesse momento. No total, pouco mais de 40% das linhagens avaliadas (subfamílias) foram extintas. Mesmo que esse evento seja importante no contexto geral da compreensão da história evolutiva dos Apoidea, as LTTs são muito suscetíveis ao registro fóssil (veja a discussão, em influência do registro fóssil em LTTs). Como previamente destacado, uma descontinuidade amostral no registro fóssil pode gerar falsos eventos de extinção. Por essa razão, previsões de eventos desse tipo devem ser tomadas com cautela.

O evento de extinção exibido em nossos resultados está diretamente ligado à nossa amostragem do depósito do âmbar burmês, pontualmente anterior ao evento (Figura 10d). Entretanto, nós amostramos quatro depósitos após o evento de extinção. Nesse contexto, existem três hipóteses plausíveis para esse fenômeno: 1) essas linhagens foram de fato extintas durante o Turoniano; 2) as linhagens não se extinguíram no final do KTR e foram extintas em algum momento desconhecido do Cretáceo superior; ou ainda 3) essas linhagens foram extintas ao longo do Cenozóico. A terceira hipótese é fortemente incongruente com o registro do Cenozóico. O número de espécimes estudados tanto em nosso estudo quanto em trabalhos prévios é equivalente àquele estudado para o âmbar de Myanmar, validando, portanto, que nenhuma dessas linhagens estão presentes ao longo do Cenozóico. Os nossos resultados apontam fortemente para o primeiro cenário visto que as linhagens que provavelmente foram extintas nesse evento não estão presentes em nenhum outro depósito ao longo do Cretáceo superior (Figura 10d). Entretanto a segunda hipótese não pode ser descartada antecipadamente pois os depósitos desse período possuem um baixo número de espécimes de vespas apoideas quando comparados com Myanmar ou com o âmbar báltico. A formação Klondike Mountain do Eoceno da América do Norte (48.6-37.2 Ma) apresenta muitos grupos de himenópteros conhecidos apenas do Mesozóico (Archibald et al. 2018). Trabalhos futuros em formações semelhantes podem elucidar se os eventos observados ao longo do Cretáceo foram de fato padrões de extinção em massa para grupos de Apoidea.

De qualquer maneira, eventos de extinção ou mesmo de mudanças repetidas de grupos ao longo das linhagens são bem conhecidos para o final do KTR. Todos os grandes grupos de plantas e animais terrestres que tiveram ascensão durante o ótimo climático do Cretáceo inferior tiveram um considerável decaimento na diversidade no final do KTR (Lloyd et al. 2008; Benton 2010; Boer et al. 2012; Grossnickle & Polly. 2013; MacLeod. 2015; Condamine et al. 2016; Clapham et al. 2016; Nel et al. 2018). O objetivo desse trabalho não foi investigar as causas desse evento de extinção para as linhagens de Apoidea. Porém algumas considerações podem ser pontuadas para trabalhos futuros.

O registro fóssil do depósito de Myanmar associado às nossas hipóteses filogenéticas sugerem que a diversificação das grandes linhagens de Apoidea (exceto abelhas) ocorreu em ambientes florestados. No final do KTR houve grandes mudanças estruturais nas florestas decorrente da substituição das angiospermas nos ambientes outrora dominados por gimnospermas (Lidgard et al. 1990; Berendse et al. 2009; Boer et al. 2012). É possível que as mudanças na estrutura das florestas arcaicas do Cretáceo inferior tenham contribuído para o desaparecimento de diversas linhagens basais de Apoidea. Além da estrutura geral, as estruturas foliares das florestas também mudaram durante essa intercessão de período (Feild et al. 2011; Boer et al. 2012). Essa mudança de flora é atribuída como causa para as mudanças de fauna e/ou extinções de diferentes grupos de herbívoros, bem como alguns de grupos de mamíferos (Grossnickle & Polly 2013) ou mesmo de insetos (Peris et al. 2017; Nel et al. 2018). Também é possível que as linhagens de vespas apoideas, por serem vespas predadoras, possam ter participado de eventos de cascata com a mudança de flora-herbívoro-predador.

Após o evento de extinção do Turoniano, as taxas de diversificações e o acúmulo de linhagens de Apoidea diminuíram durante o Cretáceo superior e permaneceram mais ou menos constante durante o Cenozoico. No Cretáceo superior surgiram os grupos mais diversos de Apoidea. Nessa época as grandes linhagens de abelhas se diversificaram dando origem posteriormente ao longo do Cenozoico a mais de 17500 espécies. Os grupos mais diversos de vespas apoideas também surgiram nessa época, os *Crabroninae strictu sensu* e os *Bembicinae strictu sensu*. Essas duas subfamílias possuem hoje aproximadamente 1380 e 4835 espécies respectivamente e correspondem a um pouco mais de 60% da diversidade atual de vespas apoideas. Somente esses três grupos citados correspondem a um pouco mais de 80% da diversidade atual de Apoidea.

No início do Cenozoico, durante o Paleógeno, surgiram os grupos *crown* de Sphecidae e Philanthidae, as famílias mais recentes de vespas apoideas. Também é durante o Paleógeno que ocorreu a maior diversificação de espécies de Apidae *lato sensu* (Cardinal 2018). O âmbar báltico apresenta uma notável diversidade de abelhas com muitos grupos extintos após esse período (Engel 2001). Para as vespas apoideas o Cenozoico foi um acúmulo de espécies e gêneros. Ademais, apenas alguns poucos gêneros são conhecidos apenas do âmbar báltico (Budrys 1993).

5. Considerações finais

As análises filogenéticas conduzidas em nosso estudo foram as mais inclusivas já realizadas para os Apoidea (99 viventes e 32 fósseis), podendo ser considerado pioneiro por incluir linhagens fósseis de Apoidea como terminais. No total, foram estabelecidas pelo menos 17 linhagens independentes de Apoidea, sendo 11 linhagens viventes e seis extintas. De maneira geral os nossos resultados reconhecem as grandes linhagens de Apoidea e são congruentes com os trabalhos prévios (Alexander 1992a; Melo 1999; Devebec et al. 2012; Branstetter et al. 2017; Peters et al. 2017; Sann et al. 2018). Entretanto os relacionamentos dessas grandes linhagens foram discrepantes entre as diversas fontes de dados e métodos filogenéticos aplicados. As principais considerações topológicas foram: 1) Formicidae + Apoidea; 2) Ampulicidae + demais Apoidea viventes; 3) Heterogynaidae + (Sphecidae + (Crabronidae *lato sensu* + Apidae); 4) Sphecidae + (Crabronidae *lato sensu* + Apidae); 5) Crabronidae *lato sensu* + Apidae; 6) parafilia dos Crabronidae *lato sensu* com Psenidae mais próximo das abelhas (dentre as linhagens viventes); 7) “†Angarosphecidae” como definido por Rasnitsyn et al. (1998) apresentou posicionamento polifilético com linhagens na base de Apoidea e outras próximas aos grupos derivados. 8) o grupo “†Cirrosphecidae” forma a linhagem mais próxima das abelhas.

O relógio morfológico foi eficiente para a estimativa do tempo de divergência dos Apoidea. As médias das idades estimadas para 16 clados de Apoidea apresentaram forte correlação entre as 10 hipóteses avaliadas (sete modelos morfológicos e as três hipóteses moleculares). Dentre os modelos de relógio morfológicos testados os resultados também demonstraram alta correlação. Entretanto, mesmo que as médias das idades indiquem forte correlação entre os modelos de relógios avaliados, os valores de MgL foram altamente discrepantes. Os resultados de MgL sugeriram como melhor modelo a combinação de um relógio de árvores FBD com uma amostragem *fossiltip*, relógio de ramos IGR com taxas *lognormal* e *prior lognormal* de raiz. De maneira geral, os modelos de relógios de árvore e relógios de ramos não apresentaram padrões claros para escolhas *a priori*. Na escolha de modelos sugerimos que as taxas de relógio de ramos (*Clockratepr*) sejam escolhidas *a priori* em trabalhos futuros. Nesse contexto, as combinações dos *priors* e posterior teste empírico são fundamentais para a escolha de um modelo eficiente na estimativa do tempo de divergência.

Entretanto, ainda existem desafios para entender e aprimorar as aplicações do relógio morfológico. A aplicação de modelos de particionamento em buscas exaustivas dos melhores *priors* foram o primeiro passo. A sensibilidade na inclusão e exclusão dos táxons foi um grande desafio para a escolha dos terminais fósseis ao longo do nosso trabalho. Aqueles clados que não apresentavam fósseis consequentemente apresentaram estimativas muito jovens e com intervalo 95%HPD muito maiores. Os nossos resultados mostraram também que as análises conduzidas com estimativa de tempo de divergência entre as linhagens sugeriram topologias parcialmente diferentes daquela sugerida sem a aplicação do relógio. As relações filogenéticas geradas com o relógio morfológico sugerem um relacionamento simplificado em relação à topologia atribuída por inferência

Bayesiana, principalmente nos grupos basais de Apoidea. Todos os testes de modelos conduzidos mostram valores altos de KRS rejeitando as relações encontradas no relógio do Modelo 6 (veja a seção 3.2.1 Modelos testados). Aparentemente esses clados são artificialmente atraídos uns pelos outros devido à sua proximidade no tempo geológico, bem como por um baixo número de caracteres diferenciais. Entretanto, apesar dos desafios, conseguimos através do relógio morfológico estimar o tempo de divergência de grupos inteiramente fósseis ou grupos associados a linhagens viventes que não possuem contraparte na fauna moderna.

O grupo *stem* de Apoidea divergiu no final do Jurássico, 156.31 Ma (152.39-159.77Ma). Essas datas foram um pouco mais antigas que aquela encontrada por Branstetter et al. (2017) (144.59 Ma), porém mais recentes que as estimadas por Peters et al. (2017) (162, 136-193 Ma). Até o final do Cretáceo inferior todas as grandes linhagens de Apoidea já haviam se diferenciado. As taxas de diversificação das grandes linhagens foram inicialmente altas ao longo do Cretáceo inferior com uma diminuição durante o Cretáceo superior e permanecendo mais ou menos constante até o presente. Durante o Cretáceo inferior também houve o maior acúmulo de linhagens a nível de subfamílias em toda a história evolutiva de Apoidea. Esse acúmulo é maior que a diversidade atual de grandes linhagens de Apoidea. Portanto, a diversificação inicial de Apoidea foi rápida em um curto espaço de tempo e é possível que esse evento tenha interferido nas hipóteses filogenéticas de Apoidea e gerando hipóteses frágeis ou mesmo incongruências entre diferentes dados filogenéticos. Também concluímos que o Cretáceo inferior e o KTR foram os momentos mais importante na diversificação das grandes linhagens de Apoidea. Os nossos dados indicaram que as altas taxas de diversificações e acúmulo gradativo das grandes linhagens de Apoidea corresponderam ao KTR. Outro interessante padrão na história evolutiva dos Apoidea foi um grande evento de extinção na metade do Cretáceo. Boa parte das linhagens basais que surgiram no início do Cretáceo ou no final do Jurássico, e que se diversificaram e permaneceram durante o Cretáceo inferior, provavelmente foram extintas na parcela final do KTR. Adicionalmente, algumas linhagens derivadas de Apoidea também se extinguíram nesse momento. No total, pouco mais de 40% das linhagens avaliadas (subfamílias) foram extintas. Portanto, conseguimos estimar os padrões macroevolutivos das grandes linhagens de Apoidea utilizando somente dados morfológicos associados às linhagens fósseis em uma filogenia datada por relógio morfológico.

5. Referências

- 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., Mikó I., Ohl M., Rasmussen C., Taeger A., Yu D.S.K.Y. 2013. Order Hymenoptera, p. 051–062. In Zhang, Z.-Q. (Ed.) *Animal Biodiversity: An Outline of Higher-level Classification and Survey of Taxonomic Richness* (Addenda 2013). *Zootaxa*, 3703, 1–82.
- Aldous D.J. 2001. Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Stat. Science*. 16(1):23–34.
- Alexander, B.A. 1992a. An exploratory analysis of cladistic relationships within the superfamily Apoidea, with special reference to sphecoid wasps (Hymenoptera). *J. Hym. Res.* 1: 25–61.
- Alexander, B.A. 1992b. A cladistic analysis of the subfamily Philantinae (Hymenoptera: Sphecidae). *Syst. Entomol.* 17: 91–108.
- Almeida E.A.B., Packer L., Melo G.A.R., Danforth B.N., Cardinal S.C., Quinteiro F.B., Pie M.R. 2018. The diversification of neopasiphaeinae bees during the Cenozoic (Hymenoptera: Colletidae). *Zool. Scr.* 1–17.
- Almeida E.A.B., Pie M.R., Brady S.G., Danforth B.N. 2011. Biogeography and diversification of colletid bees (Hymenoptera: Colletidae): Emerging patterns from the southern end of the world. *J. Biogeogr.* 39:526–544.
- Alroy, J. (2018). Fossilworks: gateway to the paleobiology database. Disponível em: <http://www.fossilworks.org>.
- Antropov A.V. 2000. Digger wasps (Hymenoptera, Sphecidae) in burmese amber. *Bull. Nat. Hist. Museum. London*. 56:59–77.
- Arcila D., Tyler J.C. 2017. Mass extinction in tetraodontiform fishes linked to the Palaeocene-Eocene thermal maximum. *Proc. R. Soc. B Biol. Sci.* 284.
- Ashmead, W.H. 1899. Classification of the entomophilous wasps, or the superfamily Sphegoidea. *The Canad. Entomol.* 31: 145–155.
- Asís J.D., Gayubo S.F., Tormos J. 1991. Description of mature larva and nesting behavior of *Pseudoscolia martinezii* Suárez (Hymenoptera: Sphecidae). *J. New York Entomol. Soc.* 99(2):261–266.
- Bapst D.W., Wright A.M., Matzke N.J., Lloyd G.T. 2016. Topology, divergence dates, and macroevolutionary inferences vary between different tip-dating approaches applied to fossil theropods (Dinosauria). *Biol. Lett.* 12:10–13.
- Barba-Montoya J., dos Reis M., Schneider H., Donoghue P.C.J., Yang Z. 2018. Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial Revolution. *New Phytol.* 218:819–834.
- Barden P., Herhold H.W., Grimaldi D.A. 2017. A new genus of hell ants from the Cretaceous (Hymenoptera: Formicidae: Haidomyrmecini) with a novel head structure. *Syst. Entomol.* 42:837–846.
- Beck, R.M.D. & M.S.Y. Lee. 2014. Ancient dates or accelerated rates? Morphological clocks and the antiquity of placental mammals. *P. Royal Soc.* 281:1–9.
- Bennett, D.J., & M.S. Engel. 2006. A new moustache wasp in Dominican amber, with an account of apoidea wasp evolution emphasizing Crabroninae (Hymenoptera: Crabronidae). *Am. Museum Novit.* 3529: 1–10.
- Benton M.J. 2009. The Red Queen and the Court Jester: and Abiotic Factors Through Time. *Nature*. 323:728–732.
- Benton M.J. 2010. The origins of modern biodiversity on land. *Philos. Trans. R. Soc. B Biol. Sci.* 365:3667–3679.
- Berendse F., Scheffer M. 2009. The angiosperm radiation revisited, an ecological explanation for Darwin’s “abominable mystery”. *Ecol. Lett.* 12:865–872.
- Bergsten J., Nilsson A.N., Ronquist F. 2013. Bayesian tests of topology hypotheses with an example from diving beetles. *Syst. Biol.* 62:660–673.
- Bisulca C., Nascimbene P.C., Elkin L., Grimaldi D.A. 2012. Variation in the Deterioration of Fossil Resins and Implications for the Conservation of Fossils in Amber. *Am. Museum Novit.* 3734:19.

- Boer H.J., Eppinga M.B., Wassen M.J., Dekker S.C. 2012. A critical transition in leaf evolution facilitated the Cretaceous angiosperm revolution. *Nat. Commun.* 3:1–11.
- Bohart R.R., Menke A.S. 1976. *Sphecid wasps of the world: A generic revision*. University of California Press, Berkeley.
- Branstetter M.G., Danforth B.N., Pitts J.P., Faircloth B.C., Ward P.S., Buffington M.L., Gates M.W., Kula R.R., Brady S.G. 2017. Phylogenomic Insights into the Evolution of Stinging Wasps and the Origins of Ants and Bees. *Curr. Biol.* 27:1019–1025.
- Bromham L., Duchêne S., Hua X., Ritchie A.M., Duchêne D.A., Ho S.Y.W. 2018. Bayesian molecular dating: opening up the black box. *Biol. Rev.* 93:1165–1191.
- Brothers, D.J. 1999. Phylogeny and evolution of wasps, ants and bees (Hymenoptera, Chrysidoidea, Vespoidea and Apoidea). *Zoologica Scripta* 28: 233–249.
- Brothers, D.J., Carpenter J.M. 1993. Phylogeny of Aculeata: Chrysidoidea and Vespoidea (Hymenoptera). *J. Hymenopt. Res.* 2:227–304.
- Brown J.W., Parins-Fukuchi C., Stull G.W., Vargas O.M., Smith S.A. 2017. Bayesian and likelihood phylogenetic reconstructions of morphological traits are not discordant when taking uncertainty into consideration: a comment on Puttick et al. *Proc. R. Soc. B* 284: 20170986.
- Budrys, E.R. 1993. Digger wasps of the subfamily Pemphredoninae (Hymenoptera, Sphecidae) from the Baltic and Taimyr amber. *Acta Entomol. Lituanica* 2:34–56.
- Camacho G.P., Pie M.R., Feitosa R.M., Barbeitos M.S. 2018. Exploring gene tree incongruence at the origin of ants and bees (Hymenoptera). *Zool. Scr.*:1–11.
- Cardinal S. 2018. Bee (Hymenoptera: Apoidea: Anthophila) Diversity Through Time Behavioral Diversity. *Insect Biodivers. Sci.* 2:851–867.
- Cardinal S., Danforth B.N. 2013. Bees diversified in the age of eudicots. *Proc. R. Soc. B Biol. Sci.* 280:851–867.
- Carvalho W.J., Fujimura P.T., Bonetti A.M., Goulart L.R., Cloonan K., Da Silva N.M., Araújo E.C.B., Ueira-Vieira C., Leal W.S. 2017. Characterization of antennal sensilla, larvae morphology and olfactory genes of *Melipona scutellaris* stingless bee. *PLoS One*. 12:1–22.
- Clapham M.E., Karr J.A., Nicholson D.B., Ross A.J., Mayhew P.J. 2016. Ancient origin of high taxonomic richness among insects. *Proc. R. Soc. B Biol. Sci.* 283:1–6.
- Clarke J.A., Middleton K.M. 2008. Mosaicism, modules, and the evolution of birds: results from a Bayesian approach to the study of morphological evolution using discrete character data. *Syst. Biol.* 57:185–201.
- Close R.A., Friedman M., Lloyd G.T., Benson R.B.J. 2015. Evidence for a mid-Jurassic adaptive radiation in mammals. *Curr. Biol.* 25:2137–2142.
- Condamine F.L., Clapham M.E., Kergoat G.J. 2016. Global patterns of insect diversification: Towards a reconciliation of fossil and molecular evidence? *Sci. Rep.* 6:1–13.
- Couvreur T.L.P., Forest F., Baker W.J. 2011a. Origin and global diversification patterns of tropical rain forests: Inferences from a complete genus-level phylogeny of palms. *BMC Biol.* 9:1–12.
- Couvreur T.L.P., Pirie M.D., Chatrou L.W., Saunders R.M.K., Su Y.C.F., Richardson J.E., Erkens R.H.J. 2011b. Early evolutionary history of the flowering plant family Annonaceae: Steady diversification and boreotropical geodispersal. *J. Biogeogr.* 38:664–680.
- Cross E.A., Bohart G.E. 1960. The biology of *Nomia triangulifera* with comparative notes on other species of *Nomia*. *Univ. Kansas Sci. Bull.* 41:761–792.
- Cusimano N., Renner S.S. 2010. Slowdowns in diversification rates from real phylogenies may not be real. *Syst. Biol.* 59:458–464.
- Darling D.C., Sharkey M.J. 1990. Order Hymenoptera. In: Grimaldi D.A. *Insects from the Santa Formation, Lower Cretaceous, of Brazil*. *B. Am. Mus. Nat. Hist.* 195:123–153.
- de Pinna M.G.G. 1991. Concepts and Tests of Homology in the Cladistic Paradigm. *Cladistics.* 7:367–394.
- Debevec, A.H., S. Cardinal, B.N. Danforth. 2012. Identifying the sister group to the bees: a molecular phylogeny of Aculeata with an emphasis on the superfamily Apoidea. *Zoologica Scripta.* 527–535.

- dos Reis M., Donoghue P.C.J., Yang Z. 2015. Bayesian molecular clock dating of species divergences in the genomics era. *Nat. Rev. Genet.* 17:1–10.
- Drummond A.J., Stadler T. 2016. Bayesian phylogenetic estimation of fossil ages. *Philos. Trans. R. Soc. B Biol. Sci.* 371.
- Engel M.S. 2000. A New Interpretation of the Oldest Fossil Bee (Hymenoptera: Apidae). *Am. Museum Novit.* 3296:1–11.
- Engel M.S. 2001. a Monograph of the Baltic Amber Bees and Evolution of the Apoidea (Hymenoptera). *Bull. Am. Museum Nat. Hist.* 259:1–192.
- Etienne R.S., Haegeman B., Stadler T., Aze T., Pearson P.N., Purvis A., Phillimore A.B. 2012. Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc. R. Soc. B Biol. Sci.* 279:1300–1309.
- Evans H.E. 1956. Studies on the Larvae of Digger Wasps (Hymenoptera, Sphecidae) Part III: Philanthinae, Trypoxyloninae, and Crabroninae. *Trans. Am. Entomological Soc.* 83:79–117.
- Evans H.E. 1958. Studies on the Larvae of Digger Wasps (Hymenoptera, Sphecidae). Part IV: Astatinae, Larrinae, and Pemphredoninae. *Trans. Am. Entomol. Soc.* 84:109–139.
- Evans H.E. 1959. Studies on the Larvae of Digger Wasps (Hymenoptera, Sphecidae) Part V: Conclusion. *Trans. Am. Entomol. Soc.* 81:131–153.
- Evans H.E. 1964. The Classification and Evolution of digger wasps as suggested by larval characters (Hymenoptera: Sphecoidea). *Entomol. News.* 9:1–592.
- Evans H.E. 1971. The larva of *Heliocausus larroides*. *Psyche.* 78:166–168.
- Evans H.E. 1975. The nest and larva of *Diploplectron brunneipes* (Cresson) (Hymenoptera: Sphecidae). *Gt. Basin Nat.* 35:123–125.
- Evans H.E., Lin C.S. 1956. Studies on the Larvae of Digger Wasps (Hymenoptera, Sphecidae). Part II: Nyssoninae. *Trans. Am. Entomol. Soc.* 82:35–66.
- Evans H.E., O'Neill K.M. 2007. The sand wasps: natural history and behaviour. Harvard University Press. Cambridge. 340p.
- Evans, H. E. 1966. The comparative ethology and evolution of the sand wasps. Harvard Univ. Press, Cambridge.
- Faircloth, B.C., M.G. Branstetter, N.D. White, S.G. Brady. 2014. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. *Mol. Eco. Resor.* 15:489–501.
- Fan Y., Wu R., Chen M.H., Kuo L., Lewis P.O. 2011. Choosing among partition models in Bayesian phylogenetics. *Mol. Biol. Evol.* 28:523–532.
- Fitch, W. N. 1971. Toward defining the course of evolution: minimum change for a specified tree topology. *Syst. Zool.* 20:406–416.
- Fiz-Palacios O., Schneider H., Heinrichs J., Savolainen V. 2011. Diversification of land plants: Insights from a family-level phylogenetic analysis. *BMC Evol. Biol.* 11:1–12.
- Forbes A.A., Bagley R.K., Beer M.A., Hippee A.C., Widmayer H.A. 2018. Quantifying the unquantifiable: Why Hymenoptera, not Coleoptera, is the most speciose animal order. *BMC Ecol.* 18:1–11.
- Forey, P.L., Kitching, I.J., 2000. Experiments in coding multistate characters. In: Scotland, R.W., Pennington, T. (Eds.), *Homology and Systematics: Coding Characters for Phylogenetic Analysis*. Taylor & Francis, London.
- Fox, W.J. 1894. A proposed classification of the fossorial Hymenoptera of North America. *P. of Acad. Nat. Sci. Philadelphia.* 1894: 292–307.
- Gavryushkina A., Heath T.A., Ksepka D.T., Stadler T., Welch D., Drummond A.J. 2017. Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* 66:57–73.
- Goloboff & Catalano. 2016. TNT version 1.5, including a full implemetantion of phylogenetic morphometrics. *Cladistics* 32:221–238.

- Goloboff P., Farris S., Nixon K. 2008. TNT, a free program for phylogenetic analysis. *Cladistics*. 24:774–786.
- Goloboff, P. A. 1993. Estimating character weights during tree search. *Cladistics*. 9: 83–91
- Grimaldi D.A., Engel M.S. 2005. *Evolution of the insects*. Cambridge University Press: New York.
- Grimaldi D.A., Maisey J.G. 1990. Introduction. In: Grimaldi D.A. 1990. *Insects from the Santa Formation, Lower Cretaceous, of Brazil*. B. Am. Mus. Nat. Hist. 195:5–14.
- Grimaldi, D. 1993. The care and study of fossiliferous amber. *Curator*. 36(1):31–40
- Grossnickle D.M., David Polly P. 2013. Mammal disparity decreases during the Cretaceous angiosperm radiation. *Proc. R. Soc. B Biol. Sci.* 280: 1–8.
- Hanson P., Menke A.S. 2006. Capítulo 17: Las avispas apoideas: Ampulicidae, Sphecidae, Crabronidae, p. 694–733 In Hanson P.E., Gauld I.D. (Eds.). *Hymenoptera de la Región Neotropical*. Memoirs of the American Entomological Institute, Florida.
- Harrell F.E. 2018. Package ‘Hmisc’: Harrell Miscellaneous. Disponível em: <https://github.com/harrelfe/Hmisc>.
- Harrison L.B., Larsson H.C.E. 2015. Among-character rate variation distributions in phylogenetic analysis of discrete morphological characters. *Syst. Biol.* 64:307–324.
- Harvey H.H., May R.M., Nee S. 1994. Phylogenies without fossils. *Evolution*. 48(3):523–529.
- Heath T.A., Huelsenbeck J.P., Stadler T. 2014. The Fossilized Birth-Death Process: A Coherent Model of Fossil Calibration for Divergence Time Estimation. E2957–E2966.
- Herrera J.P., Dávalos L.M. 2016. Phylogeny and Divergence Times of Lemurs Inferred with Recent and Ancient Fossils in the Tree. *Syst. Biol.* 65:772–791.
- Hertweck K.L., Kinney M.S., Stuart S.A., Maurin O., Mathews S., Chase M.W., Gandolfo M.A., Pires J.C. 2015. Phylogenetics, divergence times and diversification from three genomic partitions in monocots. *Bot. J. Linn. Soc.* 178:375–393.
- Huber B.T., MacLeod K.G., Watkins D.K., Coffin M.F. 2018. The rise and fall of the Cretaceous Hot Greenhouse climate. *Glob. Planet. Change*. 167:1–23.
- Jackman T.R., Larson A., de Queiroz, K., Losos J.B. 1999. Phylogenetic Relationships and Tempo of Early Diversification in *Anolis* lizards. *Syst. Biol.* 48(2):254–285.
- Johnson B.R., Borowiec M.L., Chiu J.C., Lee E.K., Atallah J., Ward P.S. 2013. Phylogenomics resolves evolutionary relationships among ants, bees, and wasps. *Curr. Biol.* 23:2058–2062.
- Kass R.E., Raftery A.E. 1995. Bayes factors. *J. Am. Stat. Assoc.* 18:773–795.
- King B., Qiao T., Lee M.S.Y., Zhu M., Long J.A. 2017. Bayesian Morphological Clock Methods Resurrect Placoderm Monophyly and Reveal Rapid Early Evolution in Jawed Vertebrates. *Syst. Biol.* 66:499–516.
- Kohl, F.F. 1897. Die Gattungen der Sphegiden. *Annalen des k.k. Naturhistorischen Hofmuseums* 11:233–516.
- Kumar S. 2005. Molecular clocks: four decades of evolution. *Nat. Rev. Genet.* 6:654–662.
- Lee M.S.Y., Palci A. 2015. Morphological phylogenetics in the genomic age. *Curr. Biol.* 25:R922–R929.
- Lee M.S.Y., Soubrier J., Edgecombe G.D. 2013. Rates of phenotypic and genomic evolution during the Cambrian explosion. *Curr. Biol.* 23:1889–1895.
- Lee, S.Y., A. Cau, D. Naish, G. Dyke. 2014. Morphological clocks in paleontology, and a Mid-Cretaceous origins of crown Aves. *Syst. Bio.* 63 (1):1–8.
- Lee M.S.Y., Yates A.M. 2018. Tip-dating and homoplasy: Reconciling the shallow molecular divergences of modern gharials with their long fossil record. *Proc. R. Soc. B Biol. Sci.* 285.
- Lehtonen S., Wahlberg N., Christenhusz M.J.M. 2012. Diversification of lindsaeoid ferns and phylogenetic uncertainty of early polypod relationships. *Bot. J. Linn. Soc.* 170:489–503.
- Lemey. P., D. Posada. 2009. Molecular clock analysis, p. 362–379. In Lemey P., Salemi M., Vandamme. A. *The phylogenetic handbook*, Cambridge University Press, Cambridge.

- Lidgard S., Crane P.R. 1990. Angiosperm diversification and Cretaceous floristic trends: A comparison of palynofloras and leaf macrofloras. *Paleobiology*. 16:77–93.
- Lloyd G.T., Davis K.E., Pisani D., Tarver J.E., Ruta M., Sakamoto M., Hone D.W.E., Jennings R., Benton M.J. 2008. Dinosaurs and the Cretaceous terrestrial revolution. *Proc. R. Soc. B Biol. Sci.* 275:2483–2490.
- Longo S.J., Faircloth B.C., Meyer A., Westneat M.W., Alfaro M.E., Wainwright P.C. 2017. Phylogenomic analysis of a rapid radiation of misfit fishes (Syngnathiformes) using ultraconserved elements. *Mol. Phylogenet. Evol.* 113:33–48.
- Lucena D.A.A., Melo G.A.R. 2018. Chrysidid wasps (Hymenoptera: Chrysididae) from Cretaceous Burmese amber: Phylogenetic affinities and classification. *Cretac. Res.* 89:279–291.
- MacLeod N. 2015. Overview of End Cretaceous Extinctions. Module in Earth Systems and Environmental Sciences. Disponível em: <http://dx.doi.org/10.1016/B978-0-12-409548-9.09371-4>.
- Martins A.C., Melo G.A.R., Renner S.S. 2014. The corbiculate bees arose from New World oil-collecting bees: Implications for the origin of pollen baskets. *Mol. Phylogenet. Evol.* 80:88–94.
- Matthews R.W., Evans H.E. 1971. Biological notes on two species of *Sericophorus* from Australia. *Psyche*. 413–429.
- Matzke N.J., Wright A. 2016. Inferring node dates from tip dates in fossil Canidae: The importance of tree priors. *Biol. Lett.* 12:2016–2019.
- Melo G.A.R., Gonçalves R.B. 2005. Higher-level bee classifications (Hymenoptera, Apoidea, Apidae sensu lato). *Rev. Bras. Zool.* 22:153–159.
- Melo G.A.R., Hermes M.G., Garcete-Barett B.R. 2011. Origin and occurrence of predation among Hymenoptera: A phylogenetic perspective. In Polidori C. *Predation in the Hymenoptera: An Evolutionary Perspective*. *An. Evol. Perspect.* 661:1–22.
- Melo, G.A.R. 1999. Phylogenetic relationships and classification of the major lineages of Apoidea (Hymenoptera), with emphasis on crabronid wasps. *Scientific Papers Nat. Hist. Mus. Univ. Kansas*. 14:1–55.
- Melo, G.A.R. 2000. Comportamento social em vespas da família Sphecidae (Hymenoptera, Apoidea), p. 85–130 In: Martins, R. P.; T. M. Lewinsohn & M. S. Barbeitos (eds). *Ecologia e comportamento de insetos*. *Oec. Bras.* 8:85–130.
- Melo G.A.R., Rosa B.B. 2018. New genus of fossil apoid wasps (Hymenoptera, Apoidea) from the Cretaceous amber of Myanmar. *Rev. Bras. Entomol.* 62:319–323.
- Michener C.D. 1944. Comparative external morphology, phylogeny, and a classification of the bees (Hymenoptera). *B. Am. Mus. Nat. Hist.* 82: 151–326.
- Michener C.D., Grimaldi D.A. 1988. A *Trigona* from Late Cretaceous Ambre of New Jersey (Hymenoptera: Apidae: Meliponinae). *Am. Museum Novit.* 1–10.
- Michener, C.D. 2007. *The bees of the world*. Johns Hopkins Univ. Press. Baltimore.
- Miller M.A., Pfeiffer W., Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE)*, 14 Nov. 2010, New Orleans, LA.
- Mirande J. 2009. Weighted parsimony phylogeny of the family Characidae (Teleostei: Characiformes). *Cladistics*. 25:574–613.
- Montoya J., dos Reis M., Schneider H., Donoghue P.C.J., Yang Z. 2018. Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial Revolution. *New Phytol.* 218:819–834.
- Morlon H., Parsons T.L., Plotkin J.B. 2011. Reconciling molecular phylogenies with the fossil record. *Proc. Natl. Acad. Sci.* 108:16327–16332.
- Nee S. 2006. Birth-Death Models in Macroevolution. *Annu. Rev. Ecol. Evol. Syst.* 37:1–17.
- Nee S., Holmes E.C., May R.M., Harvey P.H. 1994. Extinction rates can be estimated from molecular phylogenies. *Philos. Trans. R. Soc. London, B.* 344:77–82.

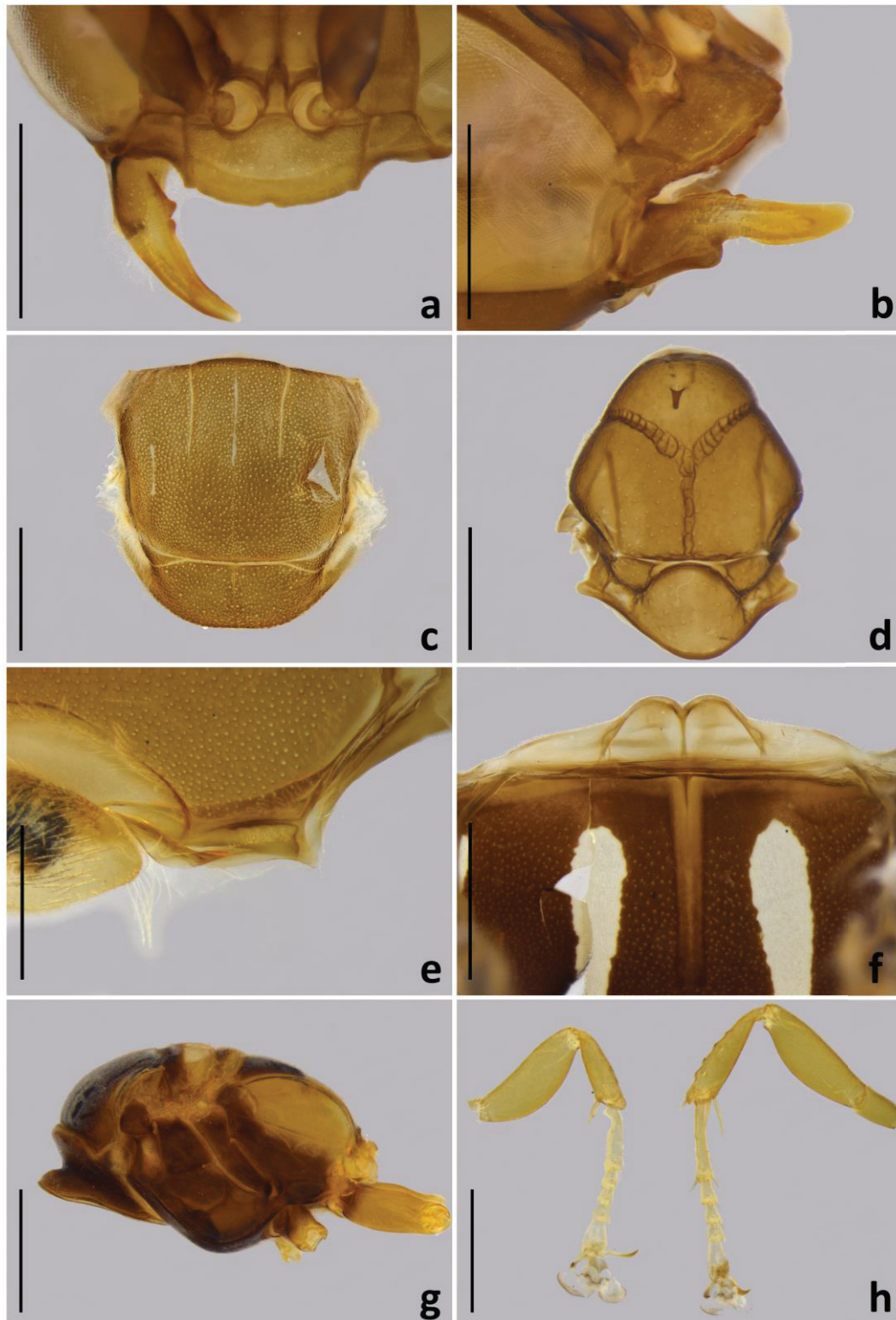
- Nel P., Bertrand S., Nel A. 2018. Diversification of insects since the Devonian: A new approach based on morphological disparity of mouthparts. *Sci. Rep.* 8:1–10.
- Nemkov P.G. 1988. Новый род роющих ос из Балтийского янтаря [= A new genus of digger wasps from the Baltic amber]. *Paleontological Journal.* 22:119–121.
- Nixon K.C., Carpenter J.M. 2012. On homology. *Cladistics.* 28: 160–169.
- Nylander J.A.A., Ronquist F., Huelsenbeck J., Nieves-Aldrey J.L. 2004. Bayesian Phylogenetic Analysis of Combined Data. *Syst. Biol.* 53:47–67.
- O'Reilly J.E., dos Reis M., Donoghue P.C.J. 2015. Dating Tips for Divergence-Time Estimation. *Trends Genet.* 31:637–650.
- O'Reilly J.E., Donoghue P.C.J. 2018. The Efficacy of Consensus Tree Methods for Summarizing Phylogenetic Relationships from a Posterior Sample of Trees Estimated from Morphological Data. *Syst. Biol.* 67:354–362.
- Ohl M. 2004. The First Fossil Representative of the Wasp Genus *Dolichurus*, with a Review of Fossil Ampulicidae (Hymenoptera: Apoidea) *J. Kans. Entomol. Soc.* 77(4): 332–342.
- Ohl M., Spahn P. 2010. A cladistic analysis of the cockroach wasps based on morphological data (Hymenoptera: Ampulicidae). *Cladistics.* 26: 49–61.
- Ohl, M. & C. Bleidorn. (2006). The phylogenetic position of the enigmatic wasp family Heterogynaidae based on molecular data, with description of a new, nocturnal species (Hymenoptera: Apoidea). *Systematic Entomology* 31, 321–337.
- Paradis E. 2003. Analysis of diversification: Combining phylogenetic and taxonomic data. *Proc. R. Soc. B Biol. Sci.* 270:2499–2505.
- Paradis E., Claude J., Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics.* 20:289–290.
- Peris D., Pérez-de la Fuente R., Peñalver E., Delclòs X., Barrón E., Labandeira C.C. 2017. False Blister Beetles and the Expansion of Gymnosperm-Insect Pollination Modes before Angiosperm Dominance. *Curr. Biol.* 27:897–904.
- Pie M.R., Tschá M.K. 2009. The macroevolutionary dynamics of ant diversification. *Evolution (N. Y).* 63:3023–3030.
- Pilgrim E.M., Von Dohlen C.D., 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.
- Poinar G.O., Jr, Milki R. 2001. Lebanese Amber: The oldest insect ecosystem in fossilized resin. Oregon State Univeristy Press. 96p.
- Porto D., Melo G.A.R., Almeida E.A.B. 2016. Clearing and dissecting insects for internal skeletal morphological research with particular reference to bees. *Rev. Bras. Entomol.* 60:109–113.
- Prentice M. 1993. Early Cretaceous Aculeata from Lebanese amber. *Sphecos.* 26: 8.
- Prentice M. 1994. Some further notes on Lebanese Aculeata. *Sphecos.* 27: 12.
- Prentice M.A., Poinar G.O. 1993. Three species of *Trypoxylon* Latreille dominican amber (Hymenoptera, Sphecidae). *J. Kansas Entomol. Soc.* 66(3): 280–291.
- Price G.D., Valdes P.J., Sellwood B.W. 1998. A comparison of GCM simulated Cretaceous “greenhouse” and “icehouse” climates: Implications for the sedimentary record. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 142:123–138.
- Pulawski W.J. 2018. Catalog of Sphecidae sensu lato. Disponível em: http://research.calacademy.org/ent/catalog_sphecidae.
- Puttick M.N., Thomas G.H., Benton M.J. 2016. Dating placentalia: Morphological clocks fail to close the molecular fossil gap. *Evolution (N. Y).* 70:873–886.
- Puttick M.N., Thomas G.H., Benton M.J. 2016. Dating placentalia: Morphological clocks fail to close the molecular fossil gap. *Evolution.* 70:873–886.

- Pybus O.G., Harvey P.H. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. B Biol. Sci.* 267:2267–2272.
- Pyron R.A. 2017. Novel approaches for phylogenetic inference from morphological data and total-evidence dating in squamate reptiles (Lizards, Snakes, and Amphisbaenians). *Syst. Biol.* 66:38–56.
- Quental T.B., Marshall C.R. 2010. Diversity dynamics: Molecular phylogenies need the fossil record. *Trends Ecol. Evol.* 25:435–441.
- Rabosky D.L. 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution.* 64:1816–1824.
- Rasnitsyn A.P. 1975. Vysshie pereponchatokrylye mezozoya (= Hymenoptera Apocrita of Mesozoic). *T. Paleont. Inst.* 147:1–133.
- Rasnitsyn A.P., Jarzembowski E.A., Ross A.J. 1998. Wasps (Insecta: Vespida = Hymenoptera) from the Purbeck and Wealdon (Lower Cretaceous) of southern England and their biostratigraphical and palaeoenvironmental significance. *Cret. Res.* 19:329–391.
- Renner S.S. 2005. Relaxed molecular clocks for dating historical plant dispersal events. *Trends Plant Sci.* 10:550–558.
- Revell L.J. 2012. Phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3:217–223.
- Ricklefs R.E. 2007. Estimating diversification rates from phylogenetic information. *Trends Ecol. Evol.* 22:601–610.
- Rieppel O., Kearney M. 2002. Similarity. *Biol. J. Linn. Soc.* 75:59–82.
- Roig-Alsina A., Michener C.D. 1993. Studies of the phylogeny and classification of long-tongued bees (Hymenoptera: Apoidea). *U. Kansas Scie. Bull.* 55:123–160.
- Romiguier J., Cameron S.A., Woodard S.H., Fischman B.J., Keller L., Praz C.J. 2015. Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate bees. *Mol. Biol. Evol.* 33:670–678.
- Ronquist F., Klopstein S., Vilhelmsen L., Schulmeister S., Murray D.L., Rasnitsyn A.P. 2012a. A total-evidence approach to dating with fossils, applied to the early radiation of the hymenoptera. *Syst. Biol.* 61:973–999.
- Ronquist F., M. Teslenko, P. van der Mark, D.L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M.A. Suchard, J.P. Huelsenbeck. 2012b. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539–542.
- Rosa B.B., Melo G.A.R., Barbeitos M.S. 2019. Parsimony-based partitioning outperforms alternatives in Bayesian analysis of discrete morphological data. *Syst. Biol.* No prelo.
- Rouard M., Droc G., Martin G., Sardos J., Hueber Y., Guignon V., Cenci A., Geigle B., Hibbins M.S., Yahiaoui N., Baurens F.-C., Berry V., Hahn M.W., D'Hont A., Roux N. 2018. Three new genome assemblies support a rapid radiation in *Musa acuminata* (wild banana). *Genome Biol. Evol.* 10:3129–3140.
- Roubik D. W. 1989. Ecology and natural history of tropical bees. New York: Cambridge University Press.
- Royer D.L., Berner R.A., Montañez I.P. 2004. CO₂ as a primary driver of Phanerozoic climate. *Geol. Soc. Am. Today.* 14(3):4–10.
- Rozen J.G. 2013. Mature Larvae of Calliopsine Bees: *Spinoliella*, *Callonychium*, and *Arhysosage* Including Biological Notes, and a Larval Key to Calliopsine Genera (Hymenoptera: Apoidea: Andrenidae: Panurginae). *Am. Museum Novit.* 3782:1–28.
- Rozen J.G. 2016. *Hesperapis rhodocera*: Behavioral Biology, Egg, and Larval Instars, Including Behavioral and Larval Comparisons with *H. larrea* (Hymenoptera: Melittidae: Dasypodinae). *Am. Museum Novit.* 3856:1–20.
- Rozen J.G. 2018. On Egg Eclosion and Larval Development in Euglossine Bees. *Am. Museum Novit.*:1–16.
- Rozen Jr. J.G. 1973. Life history and immature stages of the bee *Neofidelia* (Hymenoptera, Fidiidae). *Am. Museum Novit.*:1–14.

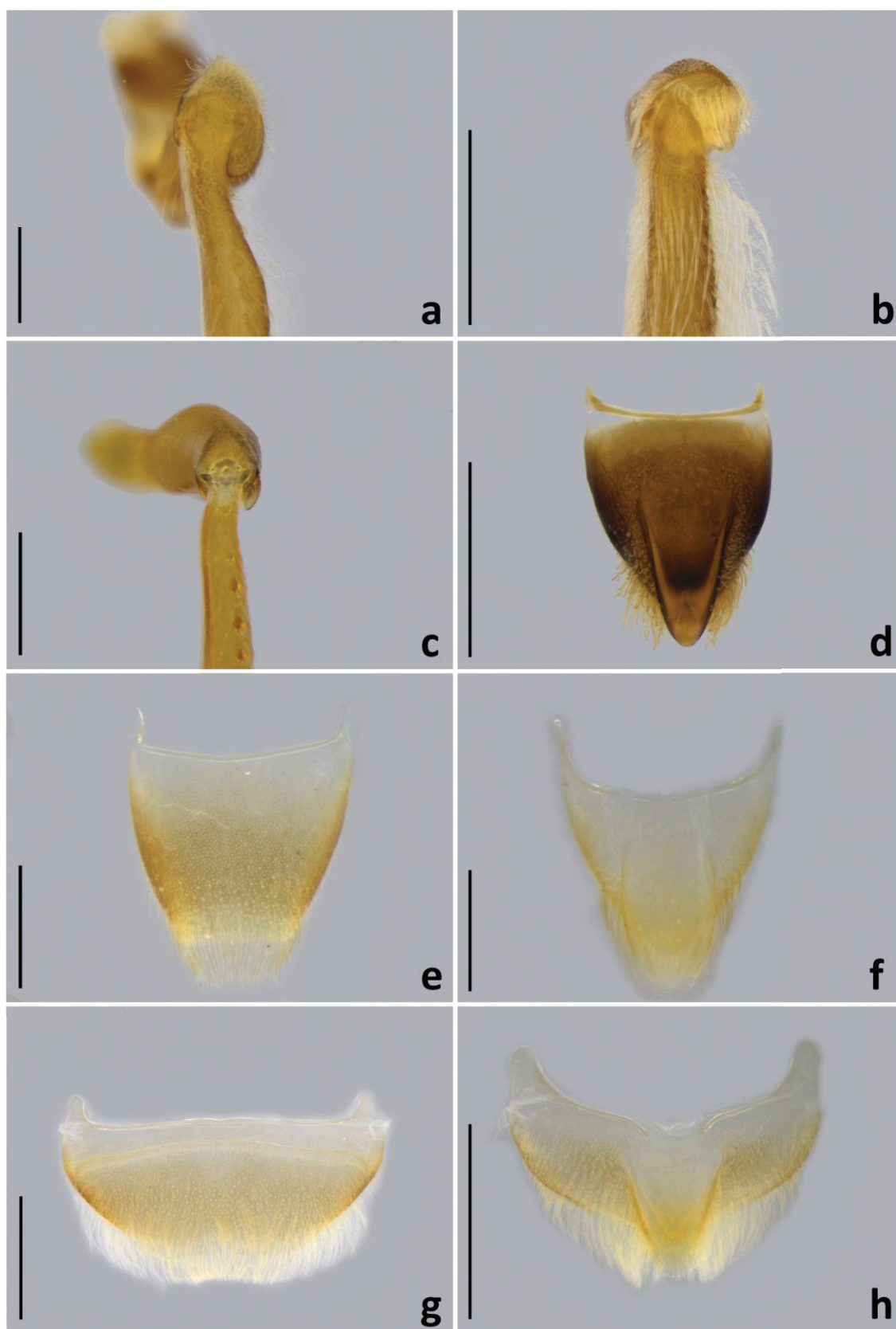
- Sann M., Niehuis O., Peters R.S., Mayer C., Kozlov A., Podsiadlowski L., Bank S., Meusemann K., Misof B., Bleidorn C., Ohl M. 2018. Phylogenomic analysis of Apoidea sheds new light on the sister group of bees. *BMC Evol. Biol.* 18:1–15.
- Sereno P.C. 2007. Logical basis for morphological characters in phylogenetics. *Cladistics*. 23: 565–587.
- Snodgrass R.E. 1935. *Principles of Insect Morphology*. McGraw-Hill. Book Co., New York, New York.
- Steel M., McKenzie A. 2001. Properties of phylogenetic trees generated by yule-type speciation models. *Math. Biosci.* 170:91–112.
- Sterli J., Pol D., Laurin M. 2013. Cladistics palaeontological dating and the timing of turtle diversification. *Cladistics*. 29:233–246.
- Suh A. 2016. The phylogenomic forest of bird trees contains a hard polytomy at the root of Neoaves. *Zool. Scr.* 45:50–62.
- Suh A., Smeds L., Ellegren H. 2015. The dynamics of incomplete lineage sorting across the ancient adaptive radiation of neoavian birds. *PLoS Biol.* 13:1–18.
- Thorne J.L., Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* 51:689–702.
- Walker J.D., Geissman J.W., Bowring S.A., Babcock L.E. 2013. The geological society of America geologic time scale. *Bull. Geol. Soc. Am.* 125:259–272.
- Williams F.X. 1946. Two new species of Astatinae, with notes on the habits of the group. *Proc. Haw. Ent. Soc.* 12:641–650.
- Xie W., Lewis P.O., Fan Y., Kuo L., Chen M.H. 2011. Improving marginal likelihood estimation for bayesian phylogenetic model selection. *Syst. Biol.* 60:150–160.
- Xing Y., Onstein R.E., Carter R.J., Stadler T., Linder H.P. 2014. Fossils and a large molecular phylogeny show that the evolution of species richness, generic diversity, and turnover rates are disconnected. *Evolution*. 68:2821–2832.
- Zhang C., Stadler T., Klopstein S., Heath T.A., Ronquist F. 2016. Total-Evidence Dating under the Fossilized Birth-Death Process. *Syst. Biol.* 65:228–249.
- Zucchi R., Yamane S., Sakagami F. 1976. Preliminary note on the habits of *Trimeria howardi*, a Neotropical communal masarid wasp. *Insec. Matsumurana*. 8:47–57.

7. Apêndices

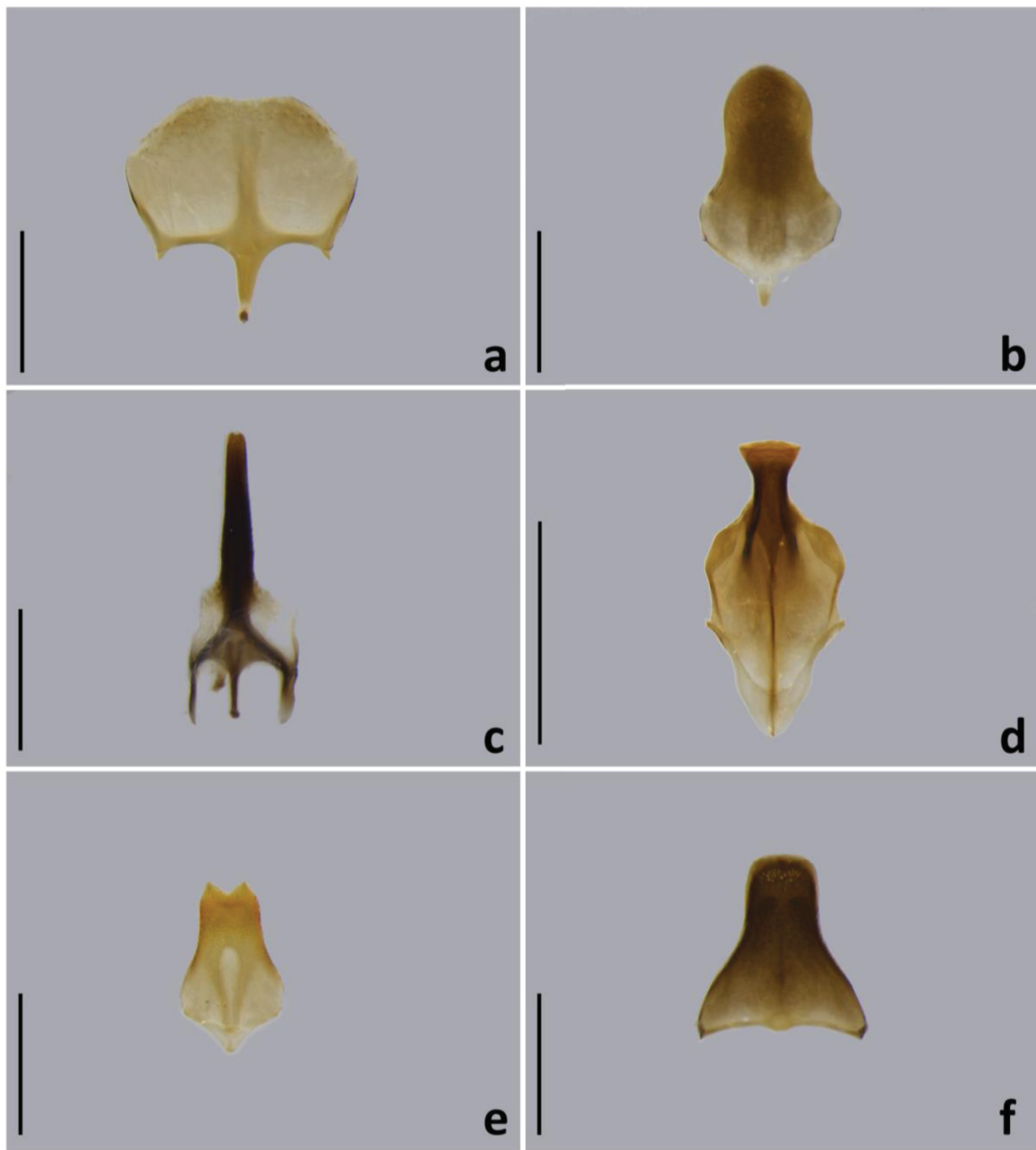
7.1 Apêndice 1. Figuras.



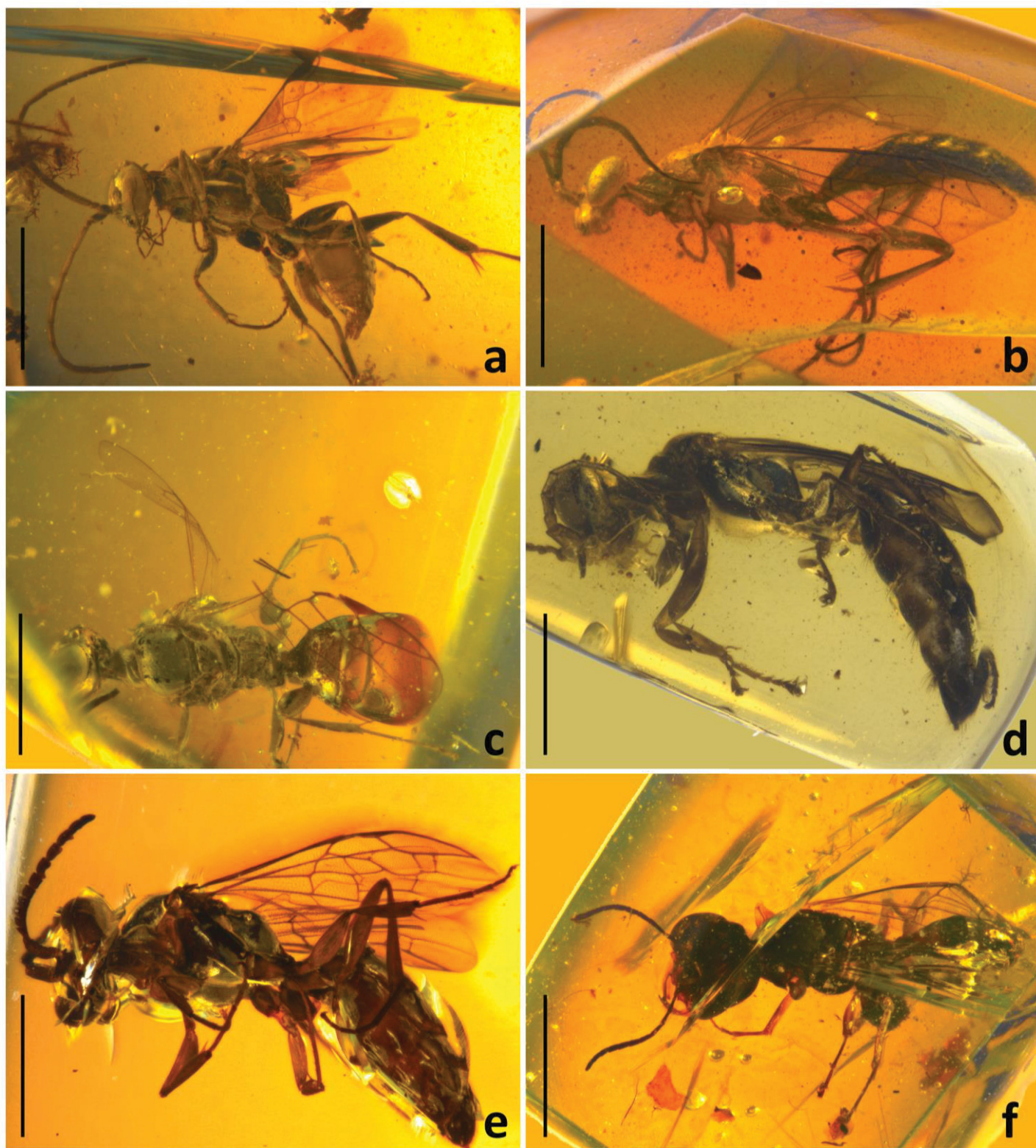
Apêndice 1 Figura 1. Morfologia interna e externa dos grupos vivos. **a)** Cabeça em vista frontal de *Liris* sp. **b)** Cabeça em vista lateral de *Liris* sp. **c)** Mesoscuto e escutelo em vista dorsal de *Dieunomia* sp. **d)** Mesoscuto e escutelo em vista dorsal de *Fulakora* sp. **e)** Mesoscuto em vista latero-dorsal de *Hesperapis carinata* Stevens. **f)** Mesoscuto em vista interna e ventral de *Stictia signata* (Linnaeus). **g)** Mesosoma em vista lateral de *Stenogorytes foxii* (Handlirsch). **h)** Perna anterior (esquerda) e perna média (direita) de uma fêmea de *Stenogorytes foxii* (Handlirsch). Escala a-c, f-h: 0.5mm. Escala d-e: 0.2mm.



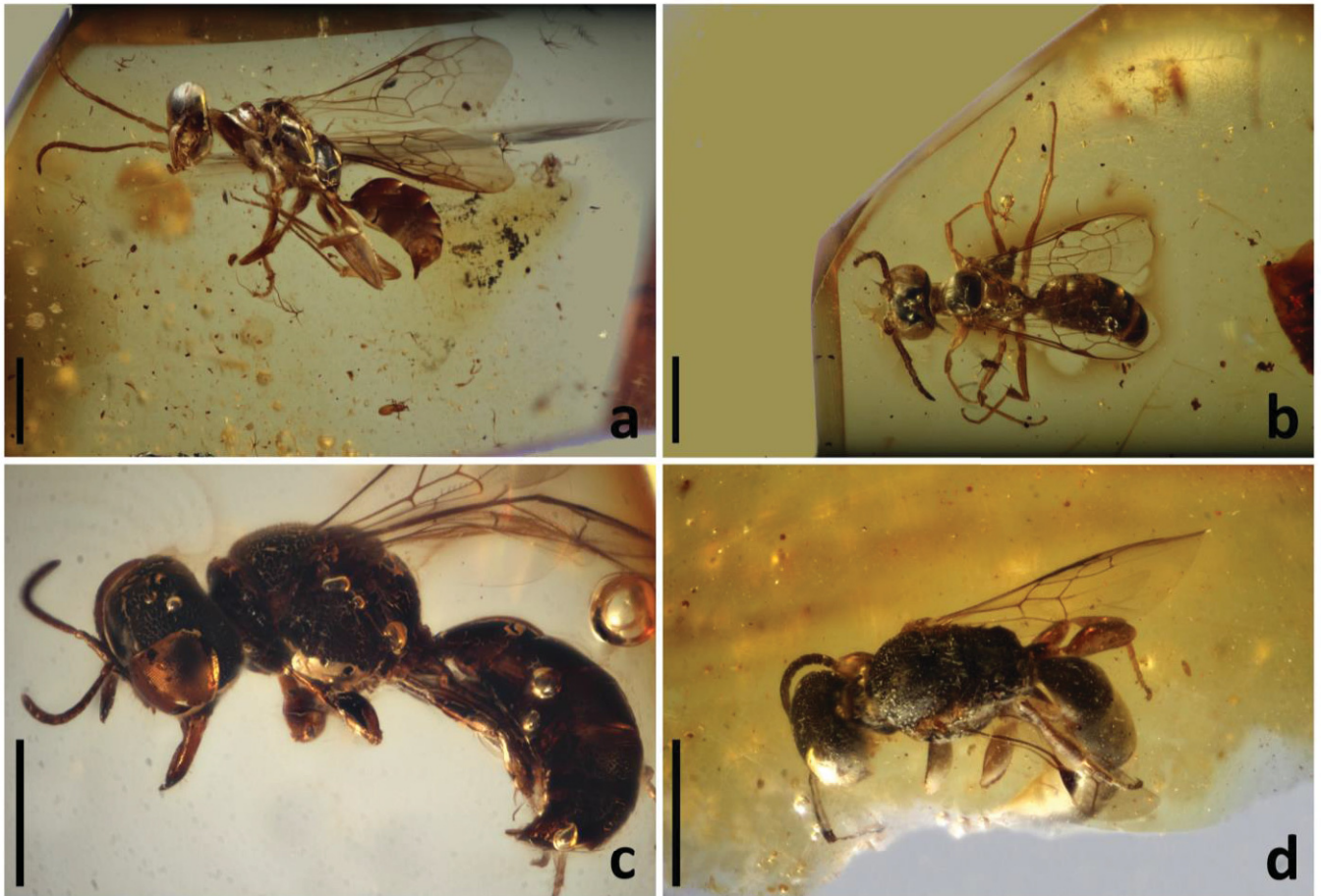
Apêndice 1 Figura 2. Morfologia externa dos grupos vivos (fêmeas). **a)** Ápice do fêmur posterior e base da tíbia de *Entomosericus concinnus* Dahlbom. **b)** Ápice do fêmur posterior e base da tíbia de *Hesperapis carinata* Stevens. **c)** Ápice do fêmur posterior e base da tíbia de *Zanysson* sp. **d)** Tergo VI de *Entomosericus concinnus* Dahlbom. **e)** Tergo V de *Odontosphex paradoxus* Menke. **f)** Tergo VI de *Odontosphex paradoxus* Menke. **g)** Tergo V de *Hesperapis carinata* Stevens. **h)** Tergo VI de *Hesperapis carinata* Stevens. Escala a, d, g-h: 0.5mm. Escala b-c, e-f: 0.2mm.



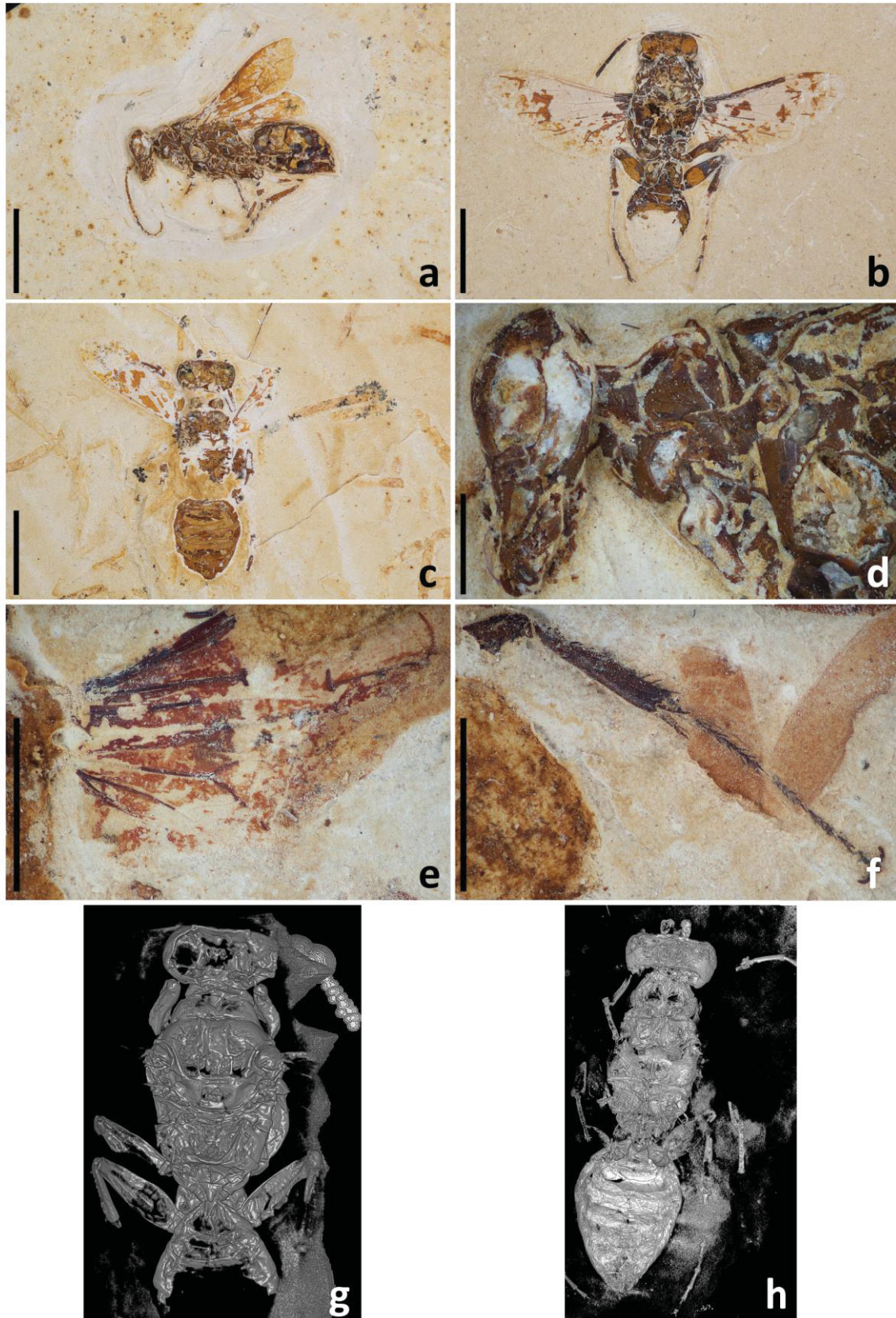
Apêndice 1 Figura 3. Esterno V dos grupos viventes(machos). **a)** *Ampulex* sp. **b)** *Clypeadon* sp. **c)** *Argogorytes* sp. **d)** *Hesperapis carinata* Stevens. **e)** *Odontosphex paradoxus* Menke. **f)** *Willinkiella argentina* (Schrottky). Escala: 0.2mm.



Apêndice 1 Figura 4. Morfologia externa dos grupos fósseis depositados no DZUP. **a)** Vista lateral de *Apodolichurus* sp. (macho). **b)** Vista lateral de Fóssil D (macho). **c)** Vista lateral de Fóssil N (macho). **d)** Vista lateral de Fóssil I (fêmea). **e)** Vista lateral de *Cirrosphex* sp. (macho). **f)** Vista lateral de Fóssil P (fêmea). Escala: 0.5mm.



Apêndice 1 Figura 5. Morfologia externa dos grupos fósseis depositados no AMNH. **a)** Vista lateral de *Cretampulex gracilis* Antropov (fêmea). **b)** Vista lateral de *Cirrosphex* sp. (macho). **c)** Vista lateral de *Lindocerus paleomystax* (Bennet & Engel). **d)** Vista lateral de *Psolimena electra* Antropov. Escala a-b: 0.5mm. Escala c-d: 0.2mm.



Apêndice 1 Figura 6. Morfologia externa de *Angarosphex magnus*. **a)** Vista lateral, AMNH SA 46322. **b)** Vista ventral, AMNH 44107 (Holótipo). **c)** Vista dorsal, AMNH SA 46305. **d)** Cabeça e pronoto em vista lateral. **e)** Asa posterior, AMNH 43267 (Parátipo). **f)** Perna posterior em vista lateral. **g)** CT-Scan em vista dorsal, AMNH 44107 (Holótipo). **h)** CT-Scan em vista ventral de AMNH SA 46305.

7.2 Apêndice 2. Lista dos caracteres morfológicos e seus estados.

1. Labrum:

- (0) at least one and a half times wider than long.
- (1) less than one and a half times wider than long, very flat (Figs. 34 and 40 in Melo 1999).

Melo (1999): Labral width measured across its base. In taxa with state (0), the labral apex usually has numerous bristles, whereas in taxa with state (1) the labrum has few or no bristles. *Podalonia* has a flat, long labrum, but with several apical bristles; *Ampulex* also has an elongate and somewhat flat labrum; both were assigned state (1). In *Epyris* and *Eusapyga*, the labrum is vestigial and these taxa were assigned (?) for characters 1-3.

2. Labral apex:

- (0) entire and broadly rounded, or at most slightly emarginated in the middle.
- (1) entire and pointed.
- (2) notched in the middle (Figs. 34 and 40 in Melo 1999).

Melo (1999): In *Arpactophilus steindachneri* the apex has several small teeth, but the labrum is considered notched in the middle. The common condition in the genus seems to be a shallow notch in the middle and small lateral teeth.

3. Labral insertion:

- (0) more or less in the middle of the labrum.
- (1) entire in the labrum base (usually short labrum).

4. A pair of rounded or oval spots on base of labrum:

- (0) absent.
- (1) present.

Melo (1999): It is not known what those spots represent, but they do not seem to be some type of sensillum; in the cleared specimens, they look like areas where the dorsal and ventral surfaces of the labrum are fused, being similar to the margins of the labrum.

5. Mandibular sexual dimorphism:

- (0) absent.
- (1) present.

6. Female mandibular apex:

- (0) apical plus one dorsal subapical tooth.
- (1) apical tooth only, i.e., simple.
- (2) apical plus two or more dorsal subapical teeth.

Melo (1999): Some groups, like *Psenulus* and some species of *Parastigmus*, have a small, more basal tooth along the inner margin that it is not taken into consideration here.

7. Male mandibular apex:

- (0) apical plus one dorsal subapical tooth.
- (1) apical tooth only.
- (2) apical plus two or more dorsal subapical teeth.

Melo (1999): The number of teeth varies among species of *Heterogyna*. Males of *H. fantsilotra* have two subapical teeth, whereas in *H. protea* has only one tooth. *Heterogyna* is assigned both states, i.e., (0) and (2).

8. Subbasal cleft on inner edge of mandible (female):

- (0) absent.
- (1) present (Apêndice 1 Figura 1a).

Melo (1999): This cleft or incision is found in several Crabroninae. For an illustration of this character, see Fig. 2 in Pulawski (1995).

9. Outerventral margin of mandible:

- (0) simple.
- (1) notched (Apêndice 1 Figura 1b).

Melo (1999): This is the same as the externoventral notch of Bohart & Menke (1976).

10. Labial palps:

- (0) with more or less similar size to one another.
- (1) with first two segments very elongate (Fig. 33-1a in Michener 2007).

11. Glossa:

- (0) less than twice as long as wide.
- (1) at least twice as long as wide.

12. Glossal apex:

- (0) glossa rounded, truncate, or bilobed at apex.
- (1) glossa pointed at apex.

Michener (2005) proposes that pointed glossal apex it is not homologous between long-tongue bees and

Pseudoscolia. Additionally the bilobaded glossal apex it is not homologous between the Colletinae and the apoid wasps. The primary homology used here did not make such distinction, but the secondary homology found in the analyses (see results) corroborates Michener's (2005) proposal.

13. Prementum:

- (0) less than three times as long as apical width.
- (1) at least three times as long as apical width.

14. Ventral surface of prementum:

- (0) continuous with lateral surfaces or separate by rounded angles.
- (1) separated from lateral surfaces by sharp angles or carinae.

Melo (1999): The ventral surface of the prementum in *Podalonia* and *Eusapyga* is excavated longitudinally, forming a sulcus margined by two ridges; *Lindenius* has a similar condition (although the middle part is slightly elevated). In *Diodontus* and *Passaloecus*, the separation between ventral and lateral surfaces is quite abrupt, but there are no carinae. These are all coded as (0).

15. Basal margins of lateral arms of prementum:

- (0) approximately perpendicular to ventral surface. (angle over 60°) (Figs. 10 and 11 in Melo 1999).
- (1) slanting, forming an acute angle (45° or smaller with ventral surface (Fig. 12 in Melo 1999).
- (2) lateral arms absent.

Melo (1999): The lateral arms of the prementum are considered absent in the bees. In this group, the basal part (articulated with the prementum) of the anterior conjunctival thickenings of Michener (1944) may be homologous to the lateral arms of the prementum that became detached from the rest of the prementum. Another possibility is that the basal portion of these thickenings could have originated from a detached part of the stipes and that the lateral arms disappeared.

16. Paramandibular process:

- (0) absent or very short, well separated from back of clypeus.
- (1) closing at least 3/4 of mandibular socket, but not reaching back of clypeus.
- (2) reaching back of clypeus, but not fused to it.
- (3) fused to clypeus.

Melo (1999): In some taxa, for example *Philanthus*, *Aphilanthops*, *Palmodes* and *Podalonia*, males and females differ in the degree of development of the paramandibular process. I chose to use the state found in females, because for most taxa only female specimens were available for complete dissection. In *Ampulex* females, the mandibular socket is closed by an extension of the gena, and not of the hypostoma. In the species dissected, the hypostoma gets close to but does not reach the clypeus. In the males, however, the mandibular socket is closed by a hypostomal process. *Ampulex* is assigned state (3).

17. Wall of the proboscoidal fossa fused to tentorium:

- (0) not at all or only near posterior wall of head (Figs. 113 and 115 in Michener 1944; Fig. 1 in Alexander and Michener 1995).
- (1) forward almost to clypeus (Fig. 114 in Michener 1994; Fig. 3 in Alexander and Michener 1995).

This character was modified of the character 19 of Alexander and Michener (1995).

18. Posterior wall of pharynx (between arms of oral plate):

- (0) not expanded.
- (1) forming two bulging sacs (walls usually covered with numerous acanthae) (Figs. 29-33 in Melo 1999).

Melo (1999): These structures are sometimes hard to see in specimens treated in KOH, especially when the pharyngeal walls are relatively thin. For this reason, specimens preserved in fixative were dissected and examined; due to lack of suitable specimens, most of the material dissected belongs to species (or even genera) not included in the phylogenetic analyses. Fig. 31 and 33 show cross-sections of the pharyngeal sacs prepared from material of *Ammoplanus* preserved in fixative. The cleared specimen of *Bembecinus* has what seems to be small expansions on the pharyngeal wall, but no fixed material was available for dissection. It was scored as (?). The male of *Heterogyna fantsilotra* has a distinct expansion of the pharynx. The posterior wall of the pharynx is dilated, but does not form a pair of sacs, and is continuous with two short expansions on the upper part of the pharynx. It is assigned (1) despite these differences.

19. Upper part of pharynx (at the tip of the oral plate arms):

- (0) simple, not expanded.
- (1) forming a pair of elongate, sometimes very large and branched, diverticula (Fig. 33 in Melo 1999).

20. Apical inflection of clypeus:

- (0) joining epistomal ridge lateral to tentorial pit (Fig. 35 in Melo 1999).
- (1) joining at tentorial pit.
- (2) joining considerably mesal to tentorial pit (Figs. 13, 14 and 34 in Melo 1999).

Melo (1999): The term "apical inflection" was taken from Roig-Alsina and Michener (1993); see their paper for additional illustrations. The segment of the inflection being considered here seems to be, in most taxa, internal to the membrane that connects the base of the mandible to the inflection. In *Eremiasphecium sahelense*, the joining of the inflection at the tentorial pits seems to result from an apparent outward displacement of the tentorial pits, and not from an expansion of the inflection; it is assigned state (0). In *Philanthus*, the inflection joins a lower branch of the tentorial arm which is broadly fused to a laminar epistomal ridge. In *Entomosericus* and *Mimesa*, the apical inflection joins the epistomal ridge slightly lateral to the tentorial pits, but these taxa are coded as having state (1). The male of *Heterogyna fantsilotra* has an internal longitudinal ridge connecting each internal rim of the antennal sclerite to the apical inflection of the clypeus. This peculiar condition is not treated as equivalent to state (2) above. This taxon is assigned state (0).

21. Eye-clypeus contact:

- (0) none.
- (1) extending for the diameter of one antennal socket or less.
- (2) extending for more than the diameter of one antennal socket.

22. Subantennal sutures:

- (0) absent.
- (1) present.

23. Subantennal sutures II:

- (0) not connected to tentorial arms.
- (1) connected to tentorial arms.

This character applies only to taxa assigned states (1) for character 22.

Melo (1999): *Laphyragogus* and *Xenosphex* have subantennal sutures, but they are assigned (?) because no internal observations were made.

24. Distance between antennal socket and clypeus (female):

- (0) more than one half of socket diameter.
- (1) one half of socket diameter or less, but not nil.
- (2) nil.

25. Epistomal suture, between antennal sockets, in taxa wherein the antennal sockets are in contact with clypeus (female):

- (0) not above transverse median line across antennal sockets (Fig. 34 in Melo 1999).
- (1) above transverse median line across antennal sockets, but not reaching tangent to upper rims (Fig. 13 in Melo 1999).
- (2) extending above tangent to upper rims of antennal sockets (Fig. 14 in Melo 1999).

This character applies only to taxa with state (2) in the preceding character.

26. Tentorial pit (female):

- (0) below or level with tangent to lower rims of antennal sockets (Fig. 13 in Melo 1999).
- (1) above tangent to lower rims of antennal sockets (Fig. 14 in Melo 1999).

27. Distance between antennal socket and clypeus (male):

- (0) more than one half of socket diameter.
- (1) one half of socket diameter or less, but not nil.
- (2) nil.

28. Epistomal suture, between antennal sockets, in taxa wherein the antennal sockets are in contact with clypeus (male):

- (0) not above transverse median line across antennal sockets.
- (1) above transverse median line across antennal sockets, but not reaching tangent to upper rims.
- (2) extending above tangent to upper rims of antennal sockets.

This character applies only to taxa with state (2) in the preceding character.

29. Tentorial pit (male):

- (0) below or level with tangent to lower rims of antennal sockets.
- (1) above tangent to lower rims of antennal sockets.

30. Tentorial pit II:

- (0) situated on epistomal ridge.
- (1) distinctly placed above epistomal ridge.

31. Distance between tentorial pit and eyes:

- (0) more than one half of socket diameter.
- (1) one half of socket diameter or less.

32. Internal rim of antennal sclerite:

- (0) level with internal surface of head or projecting only slightly (Fig. 34 in Melo 1999).
- (1) expanded toward the center and covering most of the socket (portion containing antennifer not expanded) (Fig. 35 in Melo 1999).
- (2) expanded and forming a short, but distinct cylinder (portion containing antennifer expanded together with rest of rim) (Figs. 36 and 37 in Melo 1999).

Melo (1999): The term antennifer is taken from Michener (1944). In state (2), the antennifer remains at the edge of the rim.

33. Upper portion of outer rim of antennal sclerite:

- (0) level with lower portion, antennal alveolus directed forward;
- (1) distinctly projected forward in relation to lower portion, antennal alveolus directed downward.

Melo & Rosa (in prep.): This character is being used to characterize the condition found in extant Ampulicidae, in which the antennal alveolus is directed downward, due to the expansion of its upper rim. In frontal view, less than half of the alveolus is visible. There is variation within Ampulicidae in the degree of projection and in the formation of an interantennal plate due to medial fusion of the adjacent rims. Ohl & Spahn (2010) attributed to *Aphelotoma* (as well as to *Riekefella*) the same condition of the outgroup, but in this taxon the upper portion of the antennal rim is clearly projected and the alveolus is directed downward. In *Burmasphex*, the lower portion of the frons has a relatively strong curvature, but the antennal rim is not projected dorsally. In *Notocyphus* and *Sierolomorpha*, the interantennal region is raised in

relation to the remainder of the frons, a condition that causes a slight tilt in the orientation of the alveoli; despite this, they were coded as (0). In *Pachycondyla*, there is strong sex dimorphism, with the male exhibiting the alveoli slightly directed downward, while the queen exhibits a condition similar to that in Ampulicidae, in which the antennae moves up and down laterally. It was coded (0) based on the male condition.

34. Antennal socket cover:

- (0) absent.
- (1) with a platform.

This character was modified from character 4 of Ohl and Spahn (2010).

35. Pedicel attachment:

- (0) basically centric.
- (1) eccentric (Figs. 38 and 39 in Melo 1999).

36. Socket on apex of scape with an eccentric pedicel:

- (0) entirely membranous.
- (1) sclerotized in the center (Fig. 38 in Melo 1999).

This character applies only to taxa with state (1) for the preceding character.

37. Sexual dimorphism in number of antennomeres:

- (0) none.
- (1) male with 13 and female with 12 antennomeres.

38. Female 2nd flagellomere:

- (0) at least 3X longer than pedicel.
- (1) less than 2X longer than pedicel.

Melo (1999): This character is used to characterize the relative length of the antenna. The female of *Heterogyna protea* has an unusually long pedicel; it is assigned state (0) despite the fact that its 2nd flagellomere is not 3X longer than the pedicel.

39. Facets of compound eyes (female):

- (0) approximately uniform in size.
- (1) frontal facets much larger than remaining ones.

40. Facets of compound eyes (male):

- (0) more or less uniform in size.
- (1) frontal facets much larger than remaining ones (Fig. 39 in Melo 1999).

41. Ocelli:

- (0) lens unmodified, lens globous.
- (1) lens modified in scars.

This character corresponds to the ocelli modifications found in some groups of Crabronidae *lato sensu*.

42. Inner orbits of compound eyes (female):

- (0) straight or slightly concave, more or less parallel.
- (1) concave, diverging below.
- (2) more or less straight, diverging below.
- (3) convex, diverging below.
- (4) sinuate (upper portion concave, lower portion convex), strongly converging below.
- (5) straight, converging below.
- (6) concave, converging below.

43. Inner orbits of compound eyes II:

- (0) not emarginated.
- (1) emarginated.

44. Compound eyes in males:

- (0) distant each other in the vertex.
- (1) touching or really near in the vertex.

45. Integument of paraocular area of female:

- (0) not differentiated from more median part of frons.
- (1) with a specialized area, sometimes very distinct and forming a fovea (Figs. 40-43 in Melo 1999).

Melo (1999): This specialized area represents the surface for release of the secretions of an underlying epidermal gland [see Schubert and Schonitzer (1993)]. It differs considerably in development and position among the various taxa; I consider all different forms as homologous.

46. Preoccipital carina:

- (0) complete (Fig. 44 in Melo 1999).
- (1) interrupted ventrally.
- (2) interrupted ventrally but reaching hypostomal carina.
- (3) interrupted dorsally.
- (4) completely absent.

Melo (1999): This is the same as occipital carina of Bohart and Menke (1976). The species of *Dolichurus* dissected does not have a preoccipital

carina, but it is assigned state (1) based on a female of *D. corniculatus*.

47. Periforaminal depression:

- (0) absent.
- (1) present (Fig. 44 in Melo 1999).

Melo (1999): This character concerns a distinct depression on the occipital region. It is usually more developed dorsally, as well as laterally, and in some taxa, it is marked dorsally and laterally by a marginal sulcus and/ or carina. The anterior dorsal portion of the pronotum seems to fit in this depression; in the taxa with a well-developed marginal sulcus, one could imagine that head and pronotum are locked together when the anterior dorsal margin of the pronotum is inside the marginal sulcus. In *Entomosericus*, the depression is weakly indicated, but this taxon is considered as having state (1).

48. Cervical sclerite:

- (0) absent.
- (1) present.

49. Fore coxa in males:

- (0) unmodified.
- (1) with a deeply, longitudinally channeled mesad and bearing an apical process, trochanter concave basally (Fig. 115E in Bohart & Menke 1976).

This character corresponds to the fore coxa modifications found in the Bothynotesthini (Crabronidae *stricto sensu*).

50. Posterolateral angle of pronotum:

- (0) evenly rounded (or modified differently from state 1).
- (1) reduced dorsally above and anterior to differentiated spiracular operculum.

Melo (1999): This character corresponds to character 35 of Brothers and Carpenter (1993).

51. Posterior margin of the dorsal portion of the pronotum:

- (0) level with adjacent anterior portion.
- (1) slanted down and in a distinct plane compared to the adjacent anterior portion of the pronotum.

Melo & Rosa (in prep.): This character is being used to differentiate the Apoidea, exclusive of †*Burmasphex*, for their specialized articulation in the region between the dorsal portion of the pronotum and the mesoscutum. Their pronotum has its posterior margin slanted down toward the mesoscutum and at the same time the anterior border of the mesoscutum has a step along its entire extension in which the margin of the pronotum fits in. Therefore, the contact zone between the posterior margin of the pronotum and the anterior border of the mesoscutum is rigidly articulated. In †*Burmasphex* the posterior margin is not slanted down and extends over the anterior margin of the mesoscutum, a condition reminiscent of that found in Pompilidae, Sierolomorphidae, Sapygidae and most Chrysidoidea. In Formicidae the posterior margin is also not slanted down, but similarly to Apoidea exclusive of †*Burmasphex* it fits in a sulcus along the anterior margin of the mesoscutum. As indicated by Brothers (1975; character 19), the condition of Formicidae is also present in Scoliidae, Vespidae and Bradynobaenidae.

52. Shape of lateral portion of pronotum:

- (0) uniformly convex, without constriction.
- (1) exhibiting a distinct constriction between the anterior and posterior portions of the pronotum.

Melo & Rosa (in prep.): We introduced this character for a specialized condition exhibited by the Apoidea, inclusive of *Burmasphex*, in which the lateral portions of the pronotum are compressed and form an evident constriction or strangulation, in dorsal view, right anteriorly to the pronotal lobes.

53. Ventral angle of pronotum:

- (0) scarcely exceeding base of procoxa.
- (1) greatly produced mesad and closely approaching its counterpart midventrally.

This character is being used to indicate the distinct condition found in all Apoidea, irrespective of the small variation found among them in how closely the two ventral halves of the pronotum approach each other. Brothers and Carpenter's (1993) assignment of a distinct state to bees is unjustified, since a similar condition to that found in bees is present in several other Apoidea.

54. Pronotal collar:

- (0) anterior edge rounded, or a collar not differentiated from anterior portion of pronotum.
- (1) delimited anteriorly by a transverse carina (sometimes interrupted in the middle) (Fig. 45 in Melo 1999).
- (2) delimited by a carina only laterally.

55. Shape of dorsal portion of pronotum:

- (0) pronotal collar large and delimited.
- (1) pronotal collar virtually absent.

This character applies only to taxa with state (1) for the character 51.

56. Pronotum (internally):

- (0) without lateral ridges.
- (1) with a pair of lateral, oblique ridges (converging anteriorly).

Melo (1999): *Ammoplanus* and *Hesperapis* have only a pair of weak carinae (very short in *Hesperapis*); both are coded as (1). In some taxa, e.g., *Astata*, *Ctenocolletes* and *Stangeella*, the ridges are continuous in the middle. The male of *Aphelotoma rufiventris* has no ridges, but the female of *A. nigricula* has a distinct external sulcus in the place where the ridge is situated; this taxon is assigned state (1). *Laphyragogus* and *Xenosphex* are also assigned (1) based only on external examination (presence of a sulcus in *Xenosphex* and a line in *Laphyragogus*).

57. Outer ventral posterior corner of prothoracic episternum:

- (0) not differentiated from rest of episternum.
- (1) more or less protuberant in lateral view, lateral carina of episternum not differentiated.
- (2) as (1), but lateral carina of episternum, above the protuberance, forming a distinct lamella.

58. Short sulcus dorsal to outer ventral posterior corner of prothoracic episternum:

- (0) absent.
- (1) present (Fig. 47 in Melo 1999).

Melo (1999): A correspondingly short segment of the anterior margin of the pronotum fits inside this sulcus, although it does not show any particular modification in relation to the remainder of the

anterior margin. This sulcus is present only in *Heliocausus* and *Ochleroptera*.

59. Prothoracic basisternum:

(0) lateral segments of posterior edge oblique, converging in the middle.

(1) posterior edge basically straight, except for small medial projection, lateral corners pointed.

(2) as (1), but lateral corners rounded (basisternum very small).

Melo (1999): This structure shows considerable variation among the taxa examined, making difficult the delimitation of discrete characters and states. This character is used to recognize what seems to be a distinct morphology found in the Bembicinae examined.

60. Medial portion of prothoracic basisternum:

(0) at same level as rest of basisternum, strongly pointed posteriorly (Fig. 46 in Melo 1999).

(1) declivous in relation to anterior portion (sometimes only slightly), rounded or weakly pointed posteriorly.

Melo (1999): The comment for character [59] also applies here; however, the present character is being used to represent the distinct posterior reduction of the medial portion of the basisternum in Apidae s.l. and Crabronidae in relation to the other Apoidea. In *Clystopsenella* and *Eusapyga*, the medial portion is not strongly pointed, but they are assigned state (0).

61. Apophyseal arms of prothoracic endosternum:

(0) separate (Figs. 48 and 49 in Melo 1999).

(1) fused (forming a bridge) (Fig. 50 in Melo 1999)

62. Bases of apophyseal arms of prothoracic endosternum (internally):

(0) not connected by divergent plates (Fig. 48 in Melo 1999).

(1) connected by two continuous, divergent plates originating at base of furcasternum (broadening dorsally) (Fig. 49 in Melo 1999).

Melo (1999): In *Palarus*, these plates are very broad and close half of the coxal cavity. In *Didineis* and *Ochleroptera*, the plates are apparently absent, but I assume that the plates fused completely to the furcasternum (the medial line in the furcasternum is absent, contrary to what occurs in groups where the plates are originally absent).

63. Internal divergent plates of prothoracic endosternum:

(0) separate from furcasternum by a medial ridge.

(1) partially (dorsally) or completely fused to furcasternum, medial ridge absent at least dorsally (Fig. 50 in Melo 1999)

This character applies only to taxa with state (1) for the preceding character.

64. Fore basitarsus:

(0) apex not modified, or only with a short apical lobe.

(1) with a distinct apical lobe, extending at least to half the length of second tarsomere.

Melo (1999): The condition described in state (1) is found only in females of *Laphyragogus* and in most species of *Eremiasphecium* [see Figs. 118-123 in Marshakov (1976)].

65. Foretarsal rake (females):

(0) absent.

(1) present, bristles longer than diameter of basitarsus.

(2) present, bristles as long as or shorter than diameter of basi tarsus.

66. Socket of foreleg spur:

(0) broadly connected to basitarsal socket.

(1) narrowly connected to basitarsal socket and away from tibial apex.

(2) as (1) or even farther from tibial apex and spur socket almost or completely closed (Fig. 51 in Melo 1999).

Melo (1999): In *Lyroda* and *Laphyragogus*, the spur socket is somewhat closed, but it is situated near the tibial apex. Both taxa are coded (0).

67. Leg form of female:

(0) all similar, slender and generalized (or modified differently from state 1).

(1) all femora inflated and fusiform although midfemur often less so; tibiae and tarsi fairly slender.

68. Arolia II (Female):

(0) more or less with the same size in all legs.

(1) foretarsal arolia much larger than reminds legs (Apêndice 1 Figura 1h).

This character reference to an unusual condition present in some groups of Bembicidae. Usually, in the taxa that has the enlargement of the arolia, the foretarsal and foreclaws also are large than reminds.

69. Fore coxa:

- (0) without setae.
- (1) with specialized setae.

70. Mid coxa:

- (0) without setae
- (1) with specialized setae.

71. Claws:

- (0) with a subapical or at least one subbasal tooth.
- (1) simple, without subapical or subbasal teeth.

72. Claws II:

- (0) more or less angulated.
- (1) long with the apex almost truncate.

Long claws with the apex almost truncate at the apex are found in Alyssontini and others basal fossil groups in Bembicidae.

73. Notauli:

- (0) indicated externally by a sulcus (Apêndice 1 Figura 1d).
- (1) indicated externally by a line (Apêndice 1 Figura 1c).
- (2) vestigial (slightly indicated in the anterior margin of mesoscutum).
- (3) no indication (absent).

Laphyragogus and *Xenosphex* are assigned (1) based on external examination only.

74. Notauli in sulcus:

- (0) indicated externally by a sulcus and internally by a ridge.
- (1) indicated only externally by a sulcus.

This character applies only to taxa assigned state (0) for character 73.

75. Notauli in the anterior portion of mesoscutum:

- (0) divergents (Apêndice 1 Figura 1c).
- (1) parallel (Apêndice 1 Figura 1d).

This character refers to the delimitation of the notauli. For the taxa in which this condition is

absent it was coded "?". In *Anthosila* the notauli are subparallel, but in the apical portion the distance between them remains greater than their basal portion. For this taxon was assigned (0).

76. Notauli in the posterior portion of the mesoscutum:

- (0) not in contact with the transcutal suture.
- (1) in contact with transcutal suture (complete notauli).
- (2) fused to the base (Apêndice 1 Figura 1d).

This character applies only to taxa assigned state (0) and (1) for character 73.

77. Admedial lines:

- (0) parallel and spaced from each other.
- (1) fused lines but not completely obliterated. Always forming an internal ridge (Apêndice 1 Figura 1c).
- (2) fused lines and complete obliterated. Sometimes forming an internal ridge (Apêndice 1 Figura 1f).

78. Parapsidial lines:

- (0) present and in contact to the transcutal suture (Apêndice 1 Figura 1d).
- (1) present but not in contact to the the transcutal suture (Apêndice 1 Figura 1c).
- (2) absent.

79. Supra-alar carina:

- (0) absent or if present, not meeting tegular ridge (preaxilla open anteriorly; Fig. 25 in Melo 1999).
- (1) curving down anteriorly and fused to the anterior segment of the tegular ridge (preaxilla closed off anteriorly; Fig. 26 in Melo 1999).

Melo (1999): The terms supra-alar carina and tegular ridge were taken from Michener (1944). This carina has also been named scutal flange by Menke (1988).

80. Supra-alar carina II:

- (0) present.
- (1) absent.

81. Setal patch on anterior segment of tegular ridge:

- (0) present.
- (1) absent (Apêndice 1 Figura 1e).

Melo (1999): This setal patch seems to be a proprioceptor field and is present in all aculeate outgroup taxa I studied. It is absent from a braconid, evaniid, trigonalid and a xiphydriid examined. Judging from Ronquist and Nordlander (1989), it is also apparently absent from Ibalidae. This might be a synapomorphy for Aculeata.

82. Oblique scutal carina:

(0) absent.

(1) present.

83. Prepectus:

(0) not immovably fused to the mesepisternum.

(1) immovably fused to mesepisternum, suture between them not obliterated.

(2) as (1), but suture completely obliterated.

Melo (1999): I recognized only three states for the taxa included here, taking into consideration only the degree of fusion to the mesepisternum. Brothers (1975) assumed that the prepectus in Apoidea extended completely across the anterior margin of the mesepisternum, being fused in the midline, as well as fused to and forming the depressed anterior margin of the mesepisternum. Considering the condition in basal Chrysidoidea and Vespoidea, an alternative interpretation for the Apoidea would be a prepectus not contiguous medially and fused to the mesepisternum only along its dorsal half; the depressed anterior margin of the mesepisternum would be a modification of the mesepisternum itself in response to a modified pronotum ventral angles greatly produced mesad). The prepectus in *Rhopalosoma* is very narrow, but it has a distinct fovea also present in *Eusapyga* and *Sierolomorpha*.

84. Mesepisternal ridge:

(0) complete, reaching anterior edge of mesepisternum away from body's midline.

(1) complete, reaching body's midline ventrally.

(2) reaching ventral portion of mesepisternum, but absent from middle of mesepisternum.

(3) restricted to ventral portion of mesepisternum (absent laterally).

(4) restricted to lateral portion of mesepisternum (absent ventrally).

(5) absent.

Melo (1999): This is an internal ridge present laterally and/ or ventrally on the mesepisternum (Fig. 52). In most cases it is marked externally by a sulcus (see next character). In *Dinetus* and *Lyroda*, the ridge is interrupted ventrally and only a small segment is present along the anterior edge of the mesepisternum; both taxa are coded as (0). In *Bembecinus*, the mesepisternal ridge is vestigial since only a very short segment is present below the subalar fossa; however, it is still coded (4).

85. Mesepisternal sulcus:

(0) complete, reaching anterior edge of mesepisternum away from body's midline.

(1) complete, reaching body's midline ventrally (Fig. 53 in Melo 1999).

(2) restricted to ventral portion of mesepisternum (absent laterally).

(3) restricted to lateral portion of mesepisternum (absent ventrally).

(4) absent.

Melo (1999): Sometimes the sulcus is only weakly indicated ventrally. In *Ctenocolletes*, the sulcus is only weakly indicated on the upper part of the mesepisternum; it is assigned (3).

86. Omaular sulcus:

(0) absent.

(1) present (Fig. 54 in Melo 1999).

87. Omaular carina:

(0) absent.

(1) present.

Melo (1999): This carina corresponds to the "omaulus" of Bohart and Menke (1976). Omaulus here is used to designate the area of the mesepisternum where its anterior and lateral surfaces meet. When both a carina and a sulcus are present, the carina is always anterior to the sulcus (Fig. 54 in Melo 1999).

88. Omaular carina II:

(0) omaular carina restrict to mesepisternum;

(1) omaular carina going to metaepisternum (Apêndice 1 Figura 1g).

This character applies only to taxa assigned state (1) for character 87. The state (1) correspond to the "sternaulus" of Bohart & Menke (1976). In the Fossil M the omaular carina seems to runs to the

metaspisternum. However this area in the integument of the exemplar seems like to an artifact of preservation, for this reason it was attributed (?).

89. Scrobal sulcus:

- (0) absent.
- (1) present but incomplete.
- (2) present and in contact to the ventral mesepisternum.
- (3) present and in contact to the mesepisternum sulcus and/or carina.
- (4) present and in contact to the omaulus.

The condition (1) attributed to the Apinae and Mellitinae does not appear homologous. In the Apinae the scrobal sulcus runs to upward toward to the tegula. In Mellitinae it runs to back toward the metaepisternum. In both cases the condition (1) was attributed because in neither of the taxa does the scrobal sulcus contact with the mesepisternal sulcus/carina or with the omaular sulcus/carina occurs. In *Odontosphex* the condition is the same as that seen in Mellitinae, for this taxon was also attributed (1).

90. Interfurcal muscle:

- (0) present.
- (1) absent.

Melo (1999): Absence of this muscle is considered a synapomorphy for Apoidea by Heraty et al. (1994). Only a male of *Sierolomorpha canadensis* and of *Dolichurus* sp. were examined using the technique described by these authors. The remaining taxa were assumed to have the ground plan condition postulated by Heraty et al. (1994). The loss of this muscle is probably correlated with the ventral fusion of the meso- and meta thoraces in Apoidea (condition not used here as an independent character).

Melo & Rosa (in prep.): Differently from Melo (1999), this character is being used here to represent the ventrolateral fusion of the mesepisternum with the metepisternum, and not the absence of the interfurcal muscle. He argued that the fusion of these two segments ventrolaterally probably led to the loss of the interfurcal muscle, a synapomorphy of Apoidea documented by Heraty et al. (1994). Here we also did not use Melo's character 66 due to difficulties in determining the degree of fusion along the pleural suture. In

†Burmasphex, the fusion between the sclerites cannot be directly examined. One could assume that the sclerites are not fused in *†Burmasphex* considering the small notch present right above the mid coxae at the limit between the mese- and the metepisternum laterally. A similar notch can also be observed in *Pachycondyla*, *Aphelotoma* and *Dolichurus*, but despite it the sclerites are fused in these terminals. Therefore, we attributed state (1) to *†Burmasphex*.

91. Arms of meso- and meta thoracic furca:

- (0) fused.
- (1) not fused or only weakly fused (separate in KOH cleared specimens; Figs. 15 and 16 in Melo 1999).

Melo (1999): Together with loss of the interfurcal muscle, Heraty et al. (1994) considered fusion of the arms of the meso- and metathoracic furca a synapomorphy for the Apoidea. In two apoid groups, however, the arms are not immovably fused and become separated after KOH treatment. This condition was found in Ampulicidae and in the crabronid genera *Astata*, *Eremiasphexium* and *Pulverro* (as well as *Ammoplanops*). In the Ampulicidae, this probably represents a plesiomorphic condition, while in those crabronid genera, it is clearly a secondary derived condition. In the cleared specimen of *Astata nevadica*, the metafurcal arm has a cup-like expansion and the mesofurcal arm has a callus-like structure (finely fibrous) in the region where they are supposed to be fused (Fig. 15). The lateral arm of the metafurca has also an additional cup-like expansion on its tip and it is weakly attached to the larger cup-like expansion projecting from the endophragmal (= upper meta pleural) apophysis. Dissection of a specimen of *Astata* sp. (Costa Rica) preserved in alcohol showed that the two furcal arms are firmly attached to each other and that the cup-like expansions of the metafurcal arms are covered with tendon-like material (finely fibrous and whitish). The KOH treatment probably breaks down this material, causing the separation of the furcal arms. In the cleared specimens of *Pulverro* and *Eremiasphexium sahelense*, the lateral arms of the metafurca are broad (laminar) and also separate from the mesofurcal arms (Fig. 16). Apparently the two cup-like expansions observed in *Astata* fused together and became one large expansion entirely

covering each lateral arm of the metafurca in these taxa.

92. Upper margin of discriminallamella (segment posterior to furcal arms):

- (0) narrow, as broad as remainder of lamella.
- (1) expanded, forming a horizontal lamella perpendicular to vertical portion.

This character applies only to taxa assigned state (1) for character 94.

93. Pseudophragma of second phragma:

- (0) absent (Fig. 17 in Melo 1999).
- (1) present (Fig. 18 in Melo 1999).

94. Medial portion of mesometepisternal suture (between midcoxae):

- (0) clearly visible (Fig. 59 in Melo 1999).
- (1) mostly obliterated (Figs. 23, 24, 60-62 in Melo 1999).

Melo (1999): In all Apoidea, the mesepisternum and metepisternum are fused ventrally. The morphology of this area is very complex and variable, making character delimitation difficult. Brothers and Carpenter (1993) consider loss of any sulcus between meso- and metepisterna, ventrally, as part of the Apoidea ground plan. However, in Ampulicidae and *Heterogyna*, the suture is clearly visible (the two lateral halves converge forward) and the midcoxal sockets are small and widely separated; also the mesal articulation is closer to the lateral articulation than to the body's midline. The coxal sockets are large and the suture is mostly obliterated only in the remaining apooids. The expansion of the coxal sockets was apparently accompanied by a posterior expansion of the mesokatepisternum, forming a broad flap covering the sockets medially (Fig. 23); concomitantly, the suture is directed posteriorly (in lateral view; Fig. 24). Brothers and Carpenter (1993) also inappropriately consider this condition as part of the Apoidea groundplan (see their Fig. 11 and state 57-2 in their Appendix IX).

Eremiasphecium and the Ammoplanini have a somewhat distinct condition. Their metepisternum is not projected in the middle as a strong keel continuous with the mesepisternum (see Fig. 62); there is a transverse line that looks like the mesometepisternal suture; this line does not seem to be homologous to the suture and it is probably a

structure unique to these taxa. Despite these modifications, they are coded as having state (1).

95. Relative position of the mesal condyle of the mid coxa:

- (0) condyle away from the body's midline and closer to the lateral articulation of the coxa.
- (1) situated in a small projection of the posterior border of the mesepisternum.
- (2) situated in a broad flap of the mesepisternum which covers the coxal sockets medially.

Melo & Rosa (in prep.): In the terminals attributed state (0), the mesepisternum, right in front of the mid coxae, forms a vertical surface separate from its anterior portion by a carina.

96. Medial flap of mesokatepisternum and condyle of mesal midcoxal articulation:

- (0) flap well developed and broadly continuous in the middle; portion containing condyle not or only slightly projecting (Fig. 23 and 60 in Melo 1999).
- (1) flap well developed, interrupted in the middle by a deep cleft; condyle close to midline of body, but not situated at apex of flap projection.
- (2) flap well developed, interrupted in the middle by a deep cleft; condyle situated at apex of flap projection and close to midline of body (Fig. 61 and 62 in Melo 1999).
- (3) flap narrow, interrupted in the middle by a deep cleft; condyle situated at apex of flap projection and well separated from midline of body (Fig. 53 in Melo 1999).

This character applies only to taxa assigned state (1) for the preceding character.

Melo (1999): This region probably represents the fusion of the true katepisternum with the trochantin [see discussion in Gibson (1993)], since it seems more parsimonious to assume that the mesal articulation in the Hymenoptera is homologous to the trochantinal articulation of other insects, and not a new articulation.

97. Medial portion of metepisternum:

- (0) narrow, forming a strong keel and narrowly fused to medial portion of mesokatepisternum (anterior portion of keel extending through vertical medial portion of mesokatepisternum) (Fig. 59-61 in Melo 1999).

(1) narrow, forming a carina, perpendicular to vertical medial portion of mesokatepisternum (anterior portion of carina not extending through vertical medial portion of mesokatepisternum) (Fig. 62 in Melo 1999).

(2) narrow, but not forming a strong keel or carina, narrowly fused to mesokatepisternum.

(3) wide and flat, broadly fused to mesokatepisternum (mesometepisternal suture transverse or V-shaped in ventral view) (Fig. 23 in Melo 1999).

(4) wide and flat, broadly fused to mesokatepisternum (mesometepisternal suture indistinct) (Fig. 53 in Melo 1999).

This character applies only to taxa assigned state (1) for Character 90 (ventral fusion of meso- and metathoraces).

Melo (1999): *Xenosphex* is assigned (3) despite the fact that the medial portion of its metepisternum is mostly vertical.

98. Mesocoxal carina 1:

(0) absent.

(1) present (Fig. 55-57 in Melo 1999).

Melo (1999): This carina runs from the lateral articulation obliquely across the coxa to the ventral articulation with the trochanter posteriorly (Michener 1981). In some taxa, for example *Aphilanthops* and *Philanthus*, it is present only basally, on the rest of the coxa being indicated by a rounded ridge; contrary to Alexander (1992a), I assigned state (1) to these taxa (but see next character). Michener (1981) assumed, without further argumentation, that the basal groove on the coxa of the Apoidea represents the separation of the basicoxite from the rest of the coxa (Michener's disticoxite) and that the mesocoxal carina is present only in those taxa whose basal groove has been displaced distally, i.e. taxa in which an enlargement of the "basicoxite" and a correlated reduction of the "disticoxite" occurred. Michener's terminology and homology assessments were also adopted by Johnson (1988) for the rest of Hymenoptera. The Xyelidae examined here has a distinct basal suture, certainly homologous to the basicostal suture of other insects as defined by Snodgrass (1993), delimiting a very short basicoxite. The other Hymenoptera examined have no such basicoxite

and the basicostal suture seems to correspond to the region of attachment of the membrane connecting the base of the coxa to its thoracic socket. It is assumed here that a basicoxite is absent from the Apoidea and that the mesocoxal carina, therefore, does not mark the limit between the basicoxite and a disticoxite.

99. Mesocoxal carina 2:

(0) weak, sometimes restricted to upper half or indicated only by a ridge (Fig. 55 in Melo 1999).

(1) well defined, more or less uniform throughout (Fig. 56 in Melo 1999).

(2) conspicuously enlarged, becoming a lamella toward lower half (Fig. 57 in Melo 1999).

This character applies only to taxa assigned state (1) in the preceding character.

100. Basal part of mesocoxa:

(0) more or less continuous with rest of coxa (Fig. 55-57 in Melo 1999).

(1) forming a narrow pedicel (coxa pedunculate) (Fig. 58 in Melo 1999).

101. Mesocoxa:

(0) upper portion of the midcoxa exposed.

(1) upper portion of the midcoxa hidden beneath the mesepisternum (hemicryptic midcoxa).

The term hemicryptic midcoxa is taken from Michener (1981). This character refers to the unique condition found in the short-tough bees.

102. Number of mid tibial spurs:

(0) two.

(1) one.

103. Hind coxal socket:

(0) closed by membrane only.

(1) closed by a narrow sclerotized bridge connecting the propodeum to the metakatepisternum.

(2) closed mostly by an enlargement of the medial portion of the metakatepisternum (Fig. 23 in Melo 1999).

104. Socket of mesal articulation on hindcoxa:

(0) away from ventral surface of hindcoxa (distance to ventral surface at least one half of distance between socket of mesal articulation and condyle of lateral articulation).

(1) on ventral surface of hindcoxa or very close to it (Fig. 27 in Melo 1999).

105. Hind coxal carina:

(0) absent.

(1) present, well developed but not forming a lamella.

(2) present, very strong, forming a lamella (Fig. 63 in Melo 1999).

Melo (1999): This hind coxal lamella is present in the Ammoplanini. In several taxa in this tribe, the lamella is stronger anteriorly and sometimes absent posteriorly, forming a spinelike projection.

106. Paired lobes on inner side of hind coxal apex:

(0) subequal in size and separated by a narrow cleft.

(1) dorsal lobe large, usually forming a spatulate process, ventral lobe small or absent (Fig. 64 in Melo 1999).

Melo (1999): *Eremiasphecium* and the Ammoplanini are assigned state (1) despite the fact that their coxal apices are practically straight, without any lobes (see Fig. 63). In *Ochleroptera*, the lobes are subequal in size, but the dorsal one is spatulate; it is assigned state (1). In *Heterogyna*, *Clystospenella* and *Sierolomorpha*, the lobes are very small and separated by a shallow notch in the female and practically nil in the male; these taxa are assigned state (0).

107. Dorsal apical process on inner surface of hindcoxa:

(0) relatively small, trochanter without any conspicuous depression.

(1) well developed, articulating with a basal depression on inner surface of trochanter, apical edge of depression not delimited by a weak crest (Fig. 64 in Melo 1999).

(2) as (1) or even more developed, apical edge of trochanterical depression delimited by a weak crest.

(3) nil (Fig. 63 in Melo 1999).

This character applies only for taxa assigned state (1) in the preceding character.

Melo (1999): These modifications of the coxa and trochanter are probably related to the truncation of the femoral apex, since they are more developed in taxa with such modified femora (see next character), like *Entomosericus* and *Odontosphex* (also *Bothynostethus*, *Cerceris*, *Oxybelus* and a few other crabronid taxa). These structures probably allow the femur to be locked in one position.

108. Apex of hind femur I (females):

(0) apex unmodified.

(1) apex with an expansion forming a plate.

The condition described in state (1) is found only in Philantinae (Cercerninae).

109. Apex of hind femur II (females):

(0) without pilosity.

(1) with pilosity cover the apex of the femur and the hindtibia condile (Apêndice 1 Figura 2a e 2b).

110. Apex of hind fêmur III:

(0) unmodified.

(1) with a large lateral expansion (Fig. 115F in Bohart & Menke 1976).

The condition described in state (1) is found only in Bothynostethini (Crabronidae *stricto sensu*).

111. External condyle of the hind femur (females):

(0) external condyle subigual to the inner condyle.

(1) external condyle expanded in relation to the inner condyle (Apêndice 1 Figura 2a, 2b e 2c).

112. External condyle of the hind femur (females) II:

(0) unmodified.

(1) broadened, truncate (Apêndice 1 Figura 2a).

(2) with an apical spatulate process (Apêndice 1 Figura 2a, 2b e 2c).

This character applies only for taxa assigned state (1) in the preceding character.

113. External condyle of the femur (males):

(0) external condyle subigual to the inner condyle.

(1) external condyle expanded in relation to the inner condyle. Such as the females.

114. Base of the hind tibia (females):

(0) without any carina or plate.

(1) with a longitudinal carina (Apêndice 1 Figura 2a, 2b e 2c).

115. Basitibial plate (females):

(0) absent.

(1) presente (Apêndice 1 Figura 2b e 2c).

This character applies only for taxa assigned state (1) in the character 114.

116. Hind tibial bristles:

(0) present (at least one).

(1) absent.

Melo (1999): These bristles correspond to relatively large and spiniform setae, in contrast to the fine and usually shorter setae also present on the hind tibia. Despite the large and plumose setae on the hind tibia of bees, bristles are considered absent from this taxa; their plumose setae probably correspond only to enlarged fine setae.

117. Hind tibial bristles II:

(0) unmodified or absent.

(1) bristles modified in a foliaceous-like bristles in two parallel rows.

118. Hind tarsal bristles:

(0) absent or thin, never wider or larger than the apical spurs of the tarsi.

(1) present and wider or larger than the apical spurs of the tarsi.

The condition described in state (1) is found only in Astatinae (Astatidae), Eremiasphecinae (Astatidae) and *Angarospheex magnus*.

119. Pollen collection apparatus:

(0) absent.

(1) scopa or corbícula present.

The condition described in state (1) is found only in Apidae *lato sensu*.

120. Hind tibia:

(0) unmodified.

(1) modified in a corbícula.

The condition described in state (1) is found only in Apidae *lato sensu*.

121. External hind tibial spur:

(0) present.

(1) absent.

122. Inner hind tibial spur:

(0) present.

(1) absent.

123. Hind tibia with a pollen rake:

(0) absent.

(1) present.

The condition described in state (1) is found only in Apidae *lato sensu*.

124. Hind basitarsus:

(0) longer than wide.

(1) wider than long.

The condition described in state (1) is found only in Apidae *lato sensu*.

125. Base of the hind basitarsus:

(0) unmodified.

(1) modified in a pollen press.

The condition described in state (1) is found only in Apidae *lato sensu*.

126. Gland on posterior surface of hind tibia:

(0) absent.

(1) present (Figs. 65-67 in Melo 1999).

Melo (1999): The presence of this gland was initially inferred from the distinct micropore field [term taken from Finnamore (1995)] on the posterior surface of the tibia (Figs. 65 and 66). Its presence was confirmed with histological sectioning only for *Ammoplanus* (Fig. 67). This gland possibly is absent from *Timberlakena*, because there is no micropore field on its hind tibia, except for a transverse darker area in the region where the field is situated in the other *Ammoplanini*. It is assigned (?), because a histological study might reveal the presence of the gland.

127. Pterostigma:

(0) flat, not conspicuously thickened.

(1) thickened dorso-ventrally (Figs. 72-74 in Melo 1999), in dry specimens postero-medial portion conspicuously convex ventrally.

Melo (1999): A thickened pterostigma is present in members of the subtribes Pemphredonina and Stigmina of the Pemphredonini. In *Stigmus*, the pterostigma is greatly enlarged and both dorsal and ventral surfaces possess a distinct micropore field (Figs. 68-71 in Melo 1999); under these fields, there is a thick glandular epidermis (Figs. 72 and 73 in Melo 1999). This gland is probably present in all Stigmina, because most of them also have a distinct micropore field on the pterostigma. Also a somewhat diffuse micropore field is present in *Diodontus* (Fig. 75 in Melo 1999). The pterostigma of *Passaloecus* is similarly swollen, but no glandular tissue was detected (Fig. 74 in Melo 1999).

128. Width of forewing costal cell:

- (0) at least as wide as width of vein C.
- (1) linear, narrower than vein C.

Maximum width measured perpendicular to costal margin of wing.

Melo (1999): *Dinetus* was assigned state (0), but its condition is somewhat intermediate between the two states. *Laphyragogus* is assigned (?) because its vein C is unusually slender.

129. Pterostigma width:

- (0) less than or subequal to length of prestigma (Sc + R distal to Rs).
- (1) at least one and a third the length of prestigma.

Maximum width measured perpendicular to costal wing margin.

130. Marginal cell:

- (0) longer than pterostigma (measured along vein C) (Fig. 19 in Melo 1999).
- (1) shorter than pterostigma (Figs. 20-22 in Melo 1999).

Melo (1999): Some species of *Astata* have a marginal cell longer than the pterostigma, but the remaining Astatini have state (1), including *A. nevadica*. Considering that *Astata* does not seem to be the most basal lineage within the tribe, the relatively long marginal cell in some of its species might represent a derived condition. A positive correlation between body size and relative length of the wing cells has been documented for the Hymenoptera (Danforth 1989). Since *Astata*

contains the largest forms in the tribe, it is expected that they have longer cells as well.

131. Segment of Rs separating 1st and 2nd submarginal cells:

- (0) present.
- (1) absent.

Melo (1999): The characters of forewing venation for *Eremiasphecium* were taken from *E. longiceps* and not from *E. sahelense*, because this latter species has some reductions that are probably not part of the ground plan for the genus [see also Fig. 184G in Bohart and Menke (1976) for the wing venation of *E. schmiedeknechtii* Kohl]. For *Heterogyna*, wing characters taken from male of *H. protea*.

132. Forewing 2rs-m:

- (0) present.
- (1) absent.

Melo (1999): Present only as nebulous vein in *Heterogyna*.

133. Forewing M (distal to 2rs-m) and 3rs-m:

- (0) present.
- (1) absent.

Melo (1999): Present only as spectral veins in *Sierolomorpha*.

134. Forewing CuA1 and 2m-cu:

- (0) present.
- (1) absent (discal cell2 absent).

Melo (1999): Present only as spectral veins in *Sierolomorpha*.

135. Recurrent vein 1m-cu:

- (0) touching the first submarginal cell.
- (1) touching at Rs.
- (2) touching the second submarginal cell.

136. Recurrent vein 2m-cu:

- (0) touching the second submarginal cell.
- (1) touching at 2rs-m.
- (2) touching the third submarginal cell.

137. Forewing Rs (anterior to 2r-rs) and 2rs-m:

(0) separated by Rs (segment distal to 2r-rs) anteriorly (Fig. 19 and 21 in Melo 1999).

(1) touching anteriorly (i.e., 2nd submarginal cell pointed anteriorly).

(2) fused anteriorly (i.e., 2nd submarginal cell petiolate; Figs. 20 and 22 in Melo 1999).

Melo (1999): *Nitela*, *Lindenius* and *Anacrabro* are assigned (?) for this character because some of the veins involved are absent, but there is good evidence that their reduced venation pattern is derived from lineages with a petiolate 2nd submarginal cell.

138. Forewing M and CuA:

(0) diverging distal to cu-a.

(1) diverging at cu-a.

(2) diverging basal to cu-a.

139. Forewing M + CuA (distal to cu-a):

(0) subequal to or shorter than cu-a (Fig. 21 in Melo 1999).

(1) longer than cu-a (Figs. 19 and 20 em Melo 1999).

This character applies only to taxa assigned states
(0) in the preceding character.

140. Forewing M (basal to Rs):

(0) gently curved or straight (Figs. 19,20 and 22 in Melo 1999).

(1) strongly bent (Fig. 21).

141. Forewing vein CuA2:

(0) present, reaching vein 1A.

(1) much reduced or absent, not reaching vein 1A.

142. Hindwing C:

(0) present.

(1) absent.

Melo (1999): In *Heterogyna*, *Epyris* and *Clystopsenella*, it is not possible to determine if the vein along the costal margin of the hindwing represents only Sc+R or a fusion of C with Sc+R. They are assigned(?).

143. Hind wing M:

(0) diverging from CuA before or at cu-a (Fig. 19 in Melo 1999).

(1) diverging from CuA after cu-a.

Melo (1999): *Nitela* and *Timberlakena* are assigned (?) because their hind wing venation is very reduced and parts of the veins involved are lacking.

144. Hind wing vein 2A:

(0) indicated at least as a short spur on basal portion of 1A.

(1) absent.

See the figura 5 in Bohart and Menke (1976). This vein is absent in Sphecidae. Essa veia está ausente em Sphecidae. Unlike Melo (1999) it was assigned (1) to all terminals of the Sphecidae.

145. Hind wing vein 3A:

(0) absent.

(1) present.

See the figura 5 in Bohart and Menke (1976).

146. Hind wing clavus(= plical lobe):

(0) indicated posterodistally by moderate incision.

(1) indicated by short incision or only a shallow notch.

(2) not indicated posterodistally on wing margin.

147. Jugal lobe:

(0) absent.

(1) small to moderately long and indicated by distinct incision.

(2) large and not indicated by an incision on wing margin (fused to clavus).

148. Metanotum:

(0) complete.

(1) incomplete on the sides.

This character is interpreted from Evans (1966).

149. Metanotum II:

(0) flate or elevated but not forming a contrition.

(1) elevated forming a contrition between scutellum-metanotum and metanotum-metapostnotum.

150. Metapostnotum I:

(0) transverse, depressed and distinct mesally between metanotum and propodeum (or shortened).

(1) strongly expanded posteromesally to form "propodeal triangle."

151. Metapostnotum II (propodeal enclosure):

(0) restricted to dorsal surface of propodeum, apex rounded.

(1) extending as a narrow triangle into posterior surface of propodeum, but for less than half the length of the posterior surface.

(2) extending as a narrow triangle for more than half the length of the posterior surface, but not reaching posterior apex of propodeum.

(3) extending as a narrow triangle to posterior apex of propodeum.

Melo (1999): This character applies only to taxa assigned state (1) in the preceding character. It is being used to represent an apparent progressive shortening of the propodeum and a simultaneous relative elongation and posterior narrowing of the metapostnotum within Apoidea.

152. Third mesosomal phragma:

(0) forming a transverse, vertical flange, continuous or narrowly interrupted in the middle.

(1) forming only a narrow medial, transverse flap, situated at the apex of expanded metapostnotum.

(2) as (1), but flap longitudinal.

(3) as (1), but forming a spine-like projection.

This character applies only to taxa assigned state (1) for character 150.

153. Triangular posterior extension of metapostnotum:

(0) flat or forming a broad, shallow sulcus.

(1) forming a narrow, deep sulcus (Fig. 76 in Melo 1999).

This character applies only to taxa assigned states (1) to (3) for character 150.

154. Lateral carina on the propodeum:

(0) absent.

(1) present.

This condition (1) refers to a carina that limits the sides of the propodeum and crosses the propodeal spiracle. In many taxa there may be projections adjacent to this belief in the posterior region of the propodeum. In *Mellinus* this carina is attached to the metapostnotum line and not above the propodeal spiracle, for this taxon was assigned (0). In *Ochleroptera* and *Argogorytes* the carena is restricted to the basis of the propodeum, for this

taxon was assigned (0). In *Anacrabro* does not have the carena towards the spiracle, for this taxon was assigned (0).

155. Metasomal petiole:

(0) absent, tergum I and sternum I not immovably fused, suture between them clearly visible.

(1) present, tergum I and sternum I fused anteriorly, portion of tergum I forming petiole very reduced, suture between tergum I and sternum I along petiole mostly obliterated.

(2) present, tergum I and sternum I not fused, sclerites subequal in size, suture between them clearly visible.

Melo (1999): *Dolichurus* has what looks like a very short petiole and is coded as state (1).

156. Stout bristles in the terga apex:

(0) absent.

(1) presente.

The condition described in state (1) is found only in Nyssonini and *Argogorytes* (Bembicidae).

157. Lateral line on tergum I:

(0) present, sometimes marked as a weak carina.

(1) absent.

158. Lateral carina on base of tergum I (dorsal to lateral line):

(0) absent.

(1) present, basal portion simple.

(2) present, basal portion protuberant.

159. Medial longitudinal ridge on base of sternum I:

(0) absent.

(1) present, more developed basally.

160. Second sternum posteromedially:

(0) slightly convex to flat, basal portion more or less at same level as remainder of sclerite.

(1) strongly convex, basal portion distinctly in a different level in relation to remainder of sclerite, surface separating these two portions almost vertical (Fig. 28 in Melo 1999).

Melo (1999): This character is used to recognize the distinct condition found in the Ampulicidae. *Palarus* and males of *Heliocausus* have sternum II modified (with transverse and thick keels), but

differently from the condition found in the Ampulicidae. These two taxa are assigned state (0) to avoid creating autapomorphic states for them.

161. Basal portion of lateral gradulus of sternum II:

- (0) simple, not modified, or gradulus absent.
- (1) laminar and directed inward, forming a specialized articulating surface with the differentiated posterior portion of the lateral edge of sternum I (Fig. 77 in Melo 1999).

162. Anterior lateral epidermal gland on sternum II:

- (0) situated mesal to lateral gradulus.
- (1) situated lateral to lateral gradulus (Fig. 78 in Melo 1999).

Melo (1999): Taxa in which the lateral gradulus is absent, as well as those in which the gland could not be detected, were coded as (?).

163. Adult female silk glands:

- (0) absent.
- (1) present, associated with tergum VI.
- (2) present, associated with sterna IV and V.

Melo (1999): See Melo (1997) for more details on the structure, function and taxonomic distribution of these glands.

164. Female pygidial plate (tergum VI):

- (0) absent.
- (1) presente (Apêndice 1 Figura 2d, 2f e 2h).

165. Frimbria in the tergum V:

- (0) absent.
- (1) present and cover part of the tergum VI (prepygidial frimbria) (Fig. 10.13 in Michener 2007) (Apêndice 1 Figura 2e e 2g).

The term prepygidial frimbria is taken from Michener (2007).

166. Frimbria in the tergum VI:

- (0) absent.
- (1) present forming a lateral cover in the pygidial plate (pygidial frimbria) (Fig. 10.13 in Michener 2007) (Apêndice 1 Figura 2d, 2f e 2h)

The term pygidial frimbria is taken from Michener (2007).

167. Sixth metasomal sternum (females):

- (0) similar to other segments, except for troughlike vertical side walls.
- (1) elongate, forming an exposed tapering tube through which sting is exerted.

168. Apex of female sternum VI medially:

- (0) simple (truncate, slightly rounded or emarginate), more or less continuous with lateral portions of the apex.
- (1) forming a medial lobe (Fig. 79 in Melo 1999).
- (2) with two pointed projections separated by a deep V-emargination (Fig. 80 in Melo 1999).
- (3) denticulate (Fig. 81 in Melo 1999).

169. Seventh metasomal tergum of female I:

- (0) partly exposed and evenly sclerotized.
- (1) hidden under tergum VI and considerably desclerotized.

170. Seventh metasomal tergum of female II:

- (0) two broad, lateral plates connected anteriorly by a sclerotized bridge.
- (1) as (0), but bridge displaced toward posterior margin of tergum.
- (2) lateral plates narrow (not sclerotized dorsad to spiracles), connected by a narrow, but strongly sclerotized bridge.
- (3) as (2), but lateral segments of bridge forming a 90° angle with dorsal segment.
- (4) forming two separate lateral plates (hemitergites), connected by membrane only.

This character applies only to taxa assigned state (1) in the preceding character.

171. Female hemitergites VIII:

- (0) narrowly connected dorsally by a sclerotized bridge.
- (1) connected by membrane only.

172. Female gonapophyses VIII in lateral view:

- (0) strongly curving downward.
- (1) gently curving downward.
- (2) straight.
- (3) curving upward.

Melo (1999): *Ctenocolletes* is assigned (?) because its sting is very reduced (gonapophyses widely separated at their bases).

173. Articulation within gonocoxite IX of female:

- (0) absent.
- (1) present.

174. Male tergum VII:

- (0) entire.
- (1) with lateral lobes.

175. Male cerci (tergum VIII):

- (0) present.
- (1) absent.

176. Male sternum VII:

- (0) partly exposed.
- (1) completely hidden under sternum VI.

177. Apex of male sternum VIII:

- (0) continuous with disk or forming a broad and short medial lobe (broader than or as broad as long).
- (1) forming a medial lobe or projection longer than broad, but less than 3x longer than wide, relatively broad.
- (2) medial projection more than 3x longer than wide, broad, sides diverging distally.
- (3) as (2), but narrow and sides parallel or converging distally.
- (4) with two long, lateral spiniform projections.
- (5) forming three long, spiniform projections.

178. Lateral margin of projection of male sternum VIII:

- (0) entire.
- (1) serrate (Fig. 82 in Melo 1999).
- (2) with short, thick bristles.

179. Apical margin of male sternum VIII:

- (0) entire.
- (1) denticulate.

180. Apical margin of male sternum VIII – part II:

- (0) not bidentate.
- (1) bidentate (Apêndice 1 Figura 3e).

181. Basal margin of male sternum VIII:

- (0) with a medial and elongated projection (Apêndice 1 Figura 3^a e 3c).
- (1) with U-shape.
- (2) with a large medial projection (Apêndice 1 Figura 3b, 3d e 3e).
- (3) with a truncate base (Apêndice 1 Figura 3f).

- (4) forming a unique projection.

The condition described in state (4) is found only in Alyssontini (Bembicidae).

182. Posterior edge of gonobase foramen (ventrally):

- (0) close to bases of gonocoxites.
- (1) widely separated from bases of gonocoxites.

183. Volsellae:

- (0) clearly differentiated from gonocoxites.
- (1) largely fused to gonocoxites, usually small or sometimes apparently absent.

184. Volsellae II:

- (0) cusps and digitus not fused.
- (1) cusps and digitus fused.

185. Gonapophyses of male genitalia dorsally:

- (0) connected by membrane only or by a sclerotized bridge, but not forming a distinct tube.
- (1) completely fused, forming a tube.

Melo (1999): Various forms of sclerotized bridges are found among the taxa analyzed, making difficult the recognition of discrete states. This character is used to recognize the distinct condition found among several Crabroninae.

186. Apicoventral edge of gonapophyses of male genitalia:

- (0) without teeth.
- (1) with numerous short teeth.

Sometimes, these teeth are very inconspicuous, like those in *Mellinus* [see illustrations in Menke (1996)].

187. Number of ovarioles:

- (0) 3:3.
- (1) 4:4.

This character was codified using the following references: Alexander and Rozen (1987), Iwata and Sakagami (1966), Ohl and Linde (2003), Rozen and Kamel (2007), Rozen and Ozbek (2003), Rozen (1986), Rozen (2003) and Serrão and Martins (2004).

188. Body pilosity:

- (0) pelos simples.
- (1) pelos ramificados.

189. Metallic coloration:

- (0) absent.
- (1) present.

The metallic coloration could be in the whole body like some *Ampulex* or only in a tagma like *Stangella*.

190. Larval integument:

- (0) with minute spicules or smooth.
- (1) with dense, short spicules.
- (2) with dense, conspicuous seta-like acanthae.

Melo (1999): In the descriptions of the larva of *Mimesa bicolor*, Janvier (1956) makes no reference to the integument; *Mimesa* was coded as (?).

191. Position of larval anus:

- (0) terminal, directed caudad.
- (1) ventral, preapical, directed ventrad.

192. Larval antennal papillae:

- (0) absent (sometimes the orbits are protuberant, but there are no papillae).
- (1) present, usually well developed and conspicuous.

193. Larval maxillae:

- (0) directed mesad apically, closely associated with labium and hypopharynx.
- (1) projecting apically as large, free lobes.

194. Larval galea:

- (0) large, subequal in size to maxillary palpus.
- (1) small, less than half the size of maxillary palpus.
- (2) absent.

195. Larval spinneret:

- (0) a transverse slit.
- (1) with paired openings, each at the end of a projection.
- (2) absent.

196. Provisions for larvae:

- (0) Orthopteroids (Blattodea, Mantodea, Phasmatodea and Orthoptera).
- (1) Thysanoptera.
- (2) Hemiptera (Heteroptera and Homoptera)
- (3) immature Holometabola.
- (4) adult Holometabola.
- (5) pollen.
- (6) Araneae.

In the fossils *Succinapis* and *Cretotrigona* the corbículas are present. For this reason it was attributed (5).

Melo (1999): Besides preying on aphids, *Nitela* is also known to prey on Psocoptera, but this state was not included because, among the exemplar taxa, it would be present only in this genus. *Lindenius* and *Arpactophilus* are assigned more than one state. *Nysson* and *Eusapyga* are coded(?) since they are cleptoparasites.

197. Relocation of larval food:

- (0) absent.
- (1) present.

Melo (1999): *Chlorion* is assigned both states. *Epyris* females are known to relocate their prey, but this is an exception for bethylids; it is assigned (0).

198. Construction of a nest before obtaining larval food:

- (0) absent.
- (0) present.

Apêndice 3 Tabela 1. Matriz de caracteres e estados atribuídos para cada terminal. Polimorfismos: a (01), b (02) c (10), d (12), e (24).

95

Apêndice 3 Tabela 1. Continuação (2/7). Polimorfismos: a (01), b (02) c (10), d (12), e (24).

Terminais	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4
	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	0
<i>Epyris</i>	1	2	0	1	0	0	0	0	0	1	0	0	1	0	0	0	5	0	0	0	0	1	0	0	0	0
<i>Clystospenella</i>	1	2	0	1	1	0	0	0	0	0	-	0	1	0	0	0	0	0	0	0	0	4	0	0	0	0
<i>Notocyphus</i>	0	0	-	0	0	0	0	0	0	0	-	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Rhopalosoma</i>	0	0	-	0	0	0	0	0	0	0	-	1	0	0	0	0	0	3	1	0	0	1	0	0	0	
<i>Trimeria</i>	0	2	0	0	0	0	0	0	0	0	-	1	1	0	0	0	5	1	0	0	1	0	0	0	0	
<i>Trielis</i>	0	1	-	0	0	0	0	1	0	0	-	1	1	0	0	0	0	1	0	0	0	1	0	0	0	
<i>Anthosila</i>	1	2	0	0	0	0	0	1	0	1	0	1	1	0	0	0	2	0	0	0	1	0	0	0	0	
<i>Eusapyga</i>	0	1	-	0	0	0	0	1	0	0	-	1	1	0	0	0	5	1	0	0	0	1	0	0	0	
<i>Sierolomorpha</i>	0	1	-	0	0	0	0	0	0	0	-	1	1	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Pachycondyla</i>	1	0	-	0	0	0	0	0	0	0	-	1	1	0	0	0	0	0	0	0	0	0	0	0	0	
<i>?Haidomyrmex</i>	? 0	-	-	0	0	0	?	0	0	0	-	1	1	?	0	0	0	0	0	0	0	0	?	0	?	
<i>?Burnasphex</i>	? 2	0	0	0	0	?	0	0	1	?	1	?	?	?	0	0	?	0	0	?	1	0	?	?	0	
<i>?Fóssil A</i>	? 0	-	-	0	0	0	?	0	0	0	-	1	?	?	0	0	?	0	0	?	1	0	?	?	0	
<i>Ampulex</i>	0	2	0	0	0	1	0	1	1	0	-	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Apheiotoma</i>	0	2	0	0	0	1	0	1	0	0	-	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Dolichurus</i>	0	1	-	0	0	0	1	0	1	0	-	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
<i>?Dolichurus fóssil</i>	0	2	-	1	0	1	?	1	1	0	-	1	0	0	0	0	0	0	0	0	0	1	0	?	0	
<i>?Cretampulex</i>	0	2	0	0	0	?	?	1	0	0	-	1	0	0	0	0	0	0	0	0	0	0	9	?	0	
<i>Heterogyna</i>	0	2	0	0	0	0	0	0	0	0	-	1	0	0	0	0	2	0	0	0	0	1	0	0	0	
<i>?Fóssil G</i>	? ?	? ?	? ?	? ?	? ?	? ?	?	0	0	0	-	1	0	0	?	0	0	?	0	?	?	?	?	?	?	
<i>?Apodolichurus</i>	0	0	-	0	0	0	?	0	0	0	-	1	0	0	0	0	0	0	0	0	0	0	?	?	0	
<i>?Fóssil B</i>	? 0	-	-	0	1	0	?	0	0	0	-	1	?	?	0	0	?	0	0	?	1	?	?	?	0	
<i>?Fóssil C</i>	? 0	-	-	0	0	0	?	0	0	0	-	1	?	?	0	0	?	0	0	?	1	?	?	?	0	
<i>Chlorion</i>	0	1	-	0	1	0	0	0	0	0	-	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Palmodes</i>	0	0	-	0	1	0	0	0	0	0	-	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Podalonia</i>	0	0	-	0	0	0	0	0	0	0	-	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Penepodium</i>	0	1	-	0	1	0	0	0	0	0	-	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Stangeella</i>	0	0	-	0	1	0	0	0	0	0	-	1	0	0	0	0	0	0	0	0	0	1	0	0	0	
<i>?Fóssil D</i>	0	0	-	0	0	0	?	0	0	1	?	1	1	0	0	0	1	0	0	0	0	1	0	?	0	
<i>?Angarosphex</i>	? 1	-	?	?	?	?	?	0	0	?	?	1	1	?	?	?	0	0	0	?	?	?	?	?	0	
<i>Mellinus</i>	0	0	-	0	0	0	1	0	0	0	-	1	1	0	0	0	0	0	0	0	0	1	0	0	0	
<i>?Fóssil H</i>	? 2	0	0	0	0	0	?	0	0	1	?	1	1	0	?	0	0	0	0	0	1	0	?	?	0	
<i>Dinetus</i>	1	2	0	1	1	0	0	0	0	1	1	1	1	0	1	0	2	0	0	0	1	1	0	1	0	
<i>Laphyragopus</i>	0	0	-	0	1	0	?	0	0	0	-	1	1	0	0	1	1	0	1	1	1	1	0	?	0	
<i>Xenosphex</i>	0	1	-	0	0	1	?	0	0	1	0	1	1	1	1	1	4	0	0	0	1	1	0	1	0	
<i>Anacrabro</i>	0	2	0	0	0	1	?	0	0	1	0	1	1	1	1	1	0	4	0	0	1	1	0	1	0	
<i>Lindenus</i>	0	2	0	0	0	1	0	0	0	0	-	1	1	1	1	1	0	4	0	0	1	1	0	1	0	
<i>?Lindocerus</i>	0	2	0	0	0	0	?	0	0	0	-	1	1	1	1	1	0	4	0	0	1	1	0	?	0	
<i>?Crococerus fóssil</i>	0	2	0	0	0	0	?	0	0	0	-	1	1	1	1	0	4	0	0	0	0	1	0	?	0	
<i>Oxybelus</i>	1	2	1	1	0	0	0	0	0	0	1	0	1	1	1	1	0	0	0	0	1	1	0	1	0	
<i>Palarus</i>	0	0	-	0	1	0	0	0	0	0	-	1	1	0	0	1	2	0	1	0	1	0	0	0	0	
<i>Gastrosericus</i>	1	2	0	0	0	0	1	0	0	0	-	1	1	0	0	1	2	0	0	0	0	2	0	0	0	
<i>Tachytes</i>	0	1	-	0	0	0	1	0	0	0	-	1	1	0	0	1	2	0	0	0	0	2	0	0	0	
<i>Pisonopsis</i>	0	2	0	0	0	0	0	0	0	0	-	1	1	0	0	0	2	1	0	0	0	2	0	0	0	
<i>Entomopison</i>	0	2	0	0	0	0	0	0	0	0	-	1	1	0	0	0	2	1	0	0	0	2	0	0	0	
<i>?Trypoxylon fóssil</i>	0	0	-	0	0	0	?	0	0	0	-	1	1	0	0	0	2	1	0	0	0	2	0	?	0	
<i>Willinkiella</i>	0	2	0	0	0	0	0	0	0	0	-	1	1	0	0	0	0	0	0	0	0	1	0	1	1	
<i>Scapheutes</i>	1	2	0	1	0	0	0	0	0	0	-	1	1	1	1	0	4	0	0	0	0	1	0	0	1	
<i>Lyroda</i>	0	2	0	0	0	0	0	0	0	0	-	1	1	0	0	0	0	0	0	0	0	1	0	1	0	
<i>Miscophus</i>	0	2	0	0	0	0	0	0	0	1	0	1	1	0	0	0	1	0	0	0	0	1	0	0	0	
<i>Nitela</i>	1	2	2	1	0	0	1	0	0	1	0	1	1	0	0	0	2	0	0	0	0	1	0	1	0	
<i>Paranysson</i>	1	1	-	0	1	0	0	0	0	1	0	1	1	0	0	0	2	0	0	0	0	2	0	1	0	
<i>Plenoculus</i>	0	2	0	0	0	0	0	0	0	1	0	1	1	1	1	0	2	0	0	0	1	1	0	1	0	
<i>Sericophurus</i>	1	2	0	1	0	0	0	0	0	0	-	1	1	0	0	0	2	0	0	0	1	2	0	0	0	
<i>Solierella</i>	0	2	0	0	0	0	0	0	0	0	-	1	1	1	1	1	2	0	0	0	0	2	0	0	0	
<i>?Fóssil M</i>	0	?	?	?	0	?	?	0	0	0	-	1	1	0	?	?	0	0	?	?	?	0	0	?	?	
<i>?Fóssil N</i>	? 1	-	-	0	0	0	?	0	0	0	-	1	?	?	?	0	?	0	0	?	?	0	0	?	0	
<i>?Fóssil O</i>	0	?	?	?	0	0	?	0	0	0	-	1	1	0	?	?	1	0	?	?	?	?	?	?	?	
<i>?Palanga</i>	0	1	-	?	0	0	?	0	0	0	-	1	1	0	?	?	0	0	?	?	?	?	?	?	?	
<i>Alysson</i>	0	1	-	0	0	0	0	0	0	0	-	1	1	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Didineis</i>	0	1	-	0	0	0	0	0	0	0	-	1	1	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Nysson</i>	0	1	-	0	0	0	0	1	0	0	-	1	1	0	0	0	5	0	0	0	0	1	0	0	0	
<i>Zanysson</i>	0	1	-	0	0	0	0	1	0	0	-	1	1	0	0	0	6	0	0	0	0	2	0	0	0	
<i>Heliocausus</i>	0	2	0	0	0	0	0	0	0	1	1	1	1	1	1	1	3	0	1	1	1	0	0	0	0	
<i>Tigupa</i>	0	2	0	0	0	0	0	0	0	1	1	1	1	1	1	1	3	0	1	1	1	0	0	0	0	
<i>Bembecinus</i>	0	0	-	0	0	0	0	0	0	1	0	1	1	0	0	0	4	0	0	0	0	1	0	0	0	
<i>Stictia</i>	0	1	-	0	0	0	0	0	0	0	1	0	1	1	1	1	2	0	0	0	0	1	0	0	0	
<i>Glenostictia</i>	0	1	-	0	0	0	0	0	0	0	1	0	1	?	?	1	0	0	0	?	?	1	0	0	0	
<i>Ammatomus</i>	0	0	-	0	0	1	0	0	0	1	0	1	1	1	1	0	4	0	0	0	0	2	0	0	0	
<i>Argogorytes</i>	0	1	-	0	0	0	0	0	0	0	-	1	1	0	0	0	0	0	0	0	0	1	0	0	0	
<i>Austrogorytes</i>	0	?	-	0	0	1	0	0	0	1	0	1	1	1	?	?	4	0	0	0	0	1	0	1	0	
<i>Gorytes</i>	0	0	-	0	0	1	0	0	0	1	0	1	1	1	1	1	4	0	0	0	0	1	0	1	0	
<i>Hoplisodes</i>	0	1	-	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	1	0	1	0	
<i>Megist</i>																										

Apêndice 3 Tabela 1. Continuação (3/7). Polimorfismos: a (01), b (02) c (10), d (12), e (24).

Terminais	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6	7	7	7	7	7	7	
	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5
Epyris	0	0	0	0	0	-	0	0	0	0	0	1	0	-	0	0	0	1	0	0	0	0	0	0	0	0
Clystospenella	0	0	0	0	0	-	0	0	0	0	0	0	0	-	0	0	0	1	0	0	0	1	0	0	0	0
Notocyphus	0	0	0	0	0	-	0	0	0	0	0	1	0	-	0	0	2	0	0	0	0	0	0	2	-	?
Rhopalosoma	0	0	0	0	0	-	0	0	0	0	0	1	0	-	0	0	0	0	0	0	0	0	0	2	-	?
Trimeria	0	0	0	0	0	-	0	0	0	0	0	0	0	-	0	1	2	0	0	0	0	0	0	1	-	0
Trielis	0	0	0	0	0	-	0	0	0	0	0	1	0	-	0	0	2	0	0	0	0	1	0	1	-	0
Anthosila	0	0	0	0	0	-	0	0	0	0	0	0	1	0	0	1	2	0	0	0	0	0	0	0	0	0
Eusapyga	0	0	0	0	0	-	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	-	0
Sierolomorpha	0	0	0	0	0	-	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Pachycondyla	0	0	0	0	0	-	0	0	0	0	0	1	1	0	-	0	2	0	0	0	0	0	0	0	1	0
†Haidomyrmex	0	0	0	0	0	-	0	0	0	0	0	1	?	?	?	?	?	?	?	?	?	0	0	0	0	?
†Burnasphex	1	0	1	0	0	-	1	0	0	0	0	1	?	?	?	?	?	?	?	?	?	0	0	0	0	?
†Fóssil A	1	1	1	0	0	-	1	0	0	0	0	1	?	?	?	?	?	?	?	?	?	0	0	0	0	?
Amplex	1	1	1	1	0	0	1	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	1
Aphelotoma	1	1	1	1	0	0	1	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	1
Dolichurus	1	1	1	1	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	0	0	0	0	0	1
†Dolichurus fóssil	1	1	1	1	0	0	0	0	0	0	0	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
†Cretampulex	1	1	1	1	0	0	1	0	0	0	0	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Heterogyna	1	1	1	1	0	0	1	0	0	0	0	1	0	-	0	0	0	0	0	0	0	1	0	1	-	1
†Fóssil G	1	1	1	1	0	0	1	0	0	0	0	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
†Apodolichurus	1	1	1	1	0	0	1	0	0	0	0	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
†Fóssil B	1	1	1	1	0	0	1	0	0	0	0	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
†Fóssil C	1	1	1	1	0	0	1	0	0	0	0	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Chlorion	1	1	1	1	0	0	1	0	0	0	0	1	0	-	0	1	0	0	0	0	0	0	0	1	-	1
Palmodes	1	1	1	1	0	0	1	0	0	0	0	1	0	-	0	1	0	0	0	0	0	0	0	1	-	1
Podalonia	1	1	1	1	0	0	1	0	0	0	0	1	0	-	0	1	0	0	0	0	0	1	0	2	-	?
Penepodium	1	1	1	1	0	0	1	0	0	0	0	1	0	-	0	1	0	0	0	0	0	0	0	1	-	1
Stangeella	1	1	1	1	0	0	1	0	0	0	0	1	0	-	0	1	0	0	0	0	0	0	0	1	-	1
†Fóssil D	1	1	1	1	0	0	1	2	0	0	1	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
†Angarosphex	1	1	1	1	0	0	2	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Mellinus	1	1	1	1	0	0	0	2	0	0	1	0	1	0	0	2	0	0	0	0	0	1	0	1	-	1
†Fóssil H	1	1	1	1	0	0	1	1	0	0	1	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Dinetus	1	1	1	1	0	0	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	1	0	1	-	1
Laphyragogus	1	1	1	1	0	0	1	0	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Xenosphex	1	1	1	1	0	0	1	2	0	0	1	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Anacrabro	1	1	1	1	1	0	0	2	0	0	1	1	1	0	0	2	0	0	0	0	1	1	0	1	-	1
Lindenius	1	1	1	1	2	0	0	2	0	0	1	1	1	0	0	2	0	0	0	0	0	1	0	1	-	1
†Lindocerus	1	1	1	1	2	0	1	2	0	0	1	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
†Crassocerus fóssil	1	1	1	1	2	0	1	2	0	0	1	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Oxybelus	1	1	1	1	1	0	1	2	0	0	1	0	1	0	0	2	0	0	0	0	1	1	0	1	-	1
Palarus	1	1	1	1	0	0	1	1	0	0	1	1	1	0	0	1	0	0	0	1	1	1	0	1	-	1
Gastrosericus	1	1	1	1	0	0	1	2	0	0	1	0	1	0	0	1	0	0	0	1	1	1	0	3	-	?
Tachytes	1	1	1	1	0	0	0	2	0	0	1	1	1	0	0	1	0	0	0	1	1	1	0	1	-	1
Pisonopsis	1	1	1	1	0	0	0	2	0	0	1	0	1	0	0	0	0	0	0	1	1	1	0	1	-	1
Entomopison	1	1	1	1	0	0	0	2	0	0	1	0	1	0	0	0	0	0	0	1	1	1	0	1	-	1
†Trypoxylon fóssil	1	1	1	1	0	0	0	2	0	0	1	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Willinkiella	1	1	1	1	0	0	0	2	0	0	1	0	1	0	0	2	0	0	0	1	1	1	0	1	-	1
Scapheutes	1	1	1	1	0	0	0	2	0	0	1	0	1	0	0	2	0	0	0	1	1	1	0	1	-	1
Lyroda	1	1	1	1	0	0	0	2	0	0	1	0	1	0	0	2	0	0	0	1	1	1	0	2	-	1
Miscophus	1	1	1	1	0	0	0	2	0	0	1	1	1	0	0	1	0	0	0	1	1	1	0	2	-	1
Nitela	1	1	1	1	0	0	0	2	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	3	-	?
Paranysson	1	1	1	1	0	0	0	1	0	0	1	0	1	0	0	1	1	0	0	1	1	1	0	1	-	1
Plenoculus	1	1	1	1	0	0	0	2	0	0	1	1	1	0	0	1	1	0	0	1	1	1	0	1	-	1
Sericophurus	1	1	1	1	0	0	1	2	0	0	1	0	1	0	0	1	0	0	0	1	1	1	0	1	-	1
Solierella	1	1	1	1	0	0	0	2	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	1	-	1
†Fóssil M	1	1	1	1	0	0	1	2	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
†Fóssil N	1	1	1	1	0	0	1	2	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
†Fóssil O	1	1	1	1	0	0	?	2	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
†Palanga	1	1	1	1	0	0	?	2	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Alysson	1	1	1	1	0	0	0	0	0	0	2	1	1	0	-	0	0	1	0	1	0	0	1	1	2	-
Didineis	1	1	1	1	0	0	0	0	0	0	2	1	1	0	-	0	0	1	0	1	0	0	1	1	3	-
Nysson	1	1	1	1	0	0	0	0	0	0	2	1	1	1	1	0	0	0	0	0	0	0	1	0	3	-
Zanysson	1	1	1	1	0	0	0	0	0	0	2	1	1	1	1	0	0	0	0	0	0	0	1	0	3	-
Heliocausus	1	1	1	1	0	0	1	2	1	2	1	1	1	1	1	0	1	0	0	0	0	0	1	0	3	-
Tigupa	1	1	1	1	0	0	1	2	1	2	1	1	1	1	1	0	1	0	0	0	0	0	1	0	3	-
Bembecinus	1	1	1	1	0	0	0	0	0	2	1	1	1	1	1	0	1	0	0	0	0	0	1	0	3	-
Stictia	1	1	1	1	0	0	1	0	0	2	1	1	1	1	1	0	1	0	0	0	0	0	1	0	3	-
Glenostictia	1	1	1	1	0	0	1	0	0	2	1	1	1	1	1	0	1	0	0	0	0	0	1	0	3	-
Anmatomus	1	1	1	1	0	0	1	2	0	2	1	1	1	1	1	0	2	0	0	1	0	0	1	0	3	-
Argogorytes	1	1	1	1	0	0	1	2	0	2	1	1	1	1	1	0	0	0	0	0	0	0	1	0	2	-
Austrogorytes	1	1	1	1	0	0	0	0	0	0	2	1	1	1	1	0	1	0	0	1	0	0	1	0	3	-
Gorytes	1	1	1	1	0	0	0	0	0	0	2	1	1	1	1	0	1	0	0	1	0	0	1	0	1	-
Hoplosides	1	1	1	1	0	0	0	0	0	0	2	1	1	1	1	0	1	0								

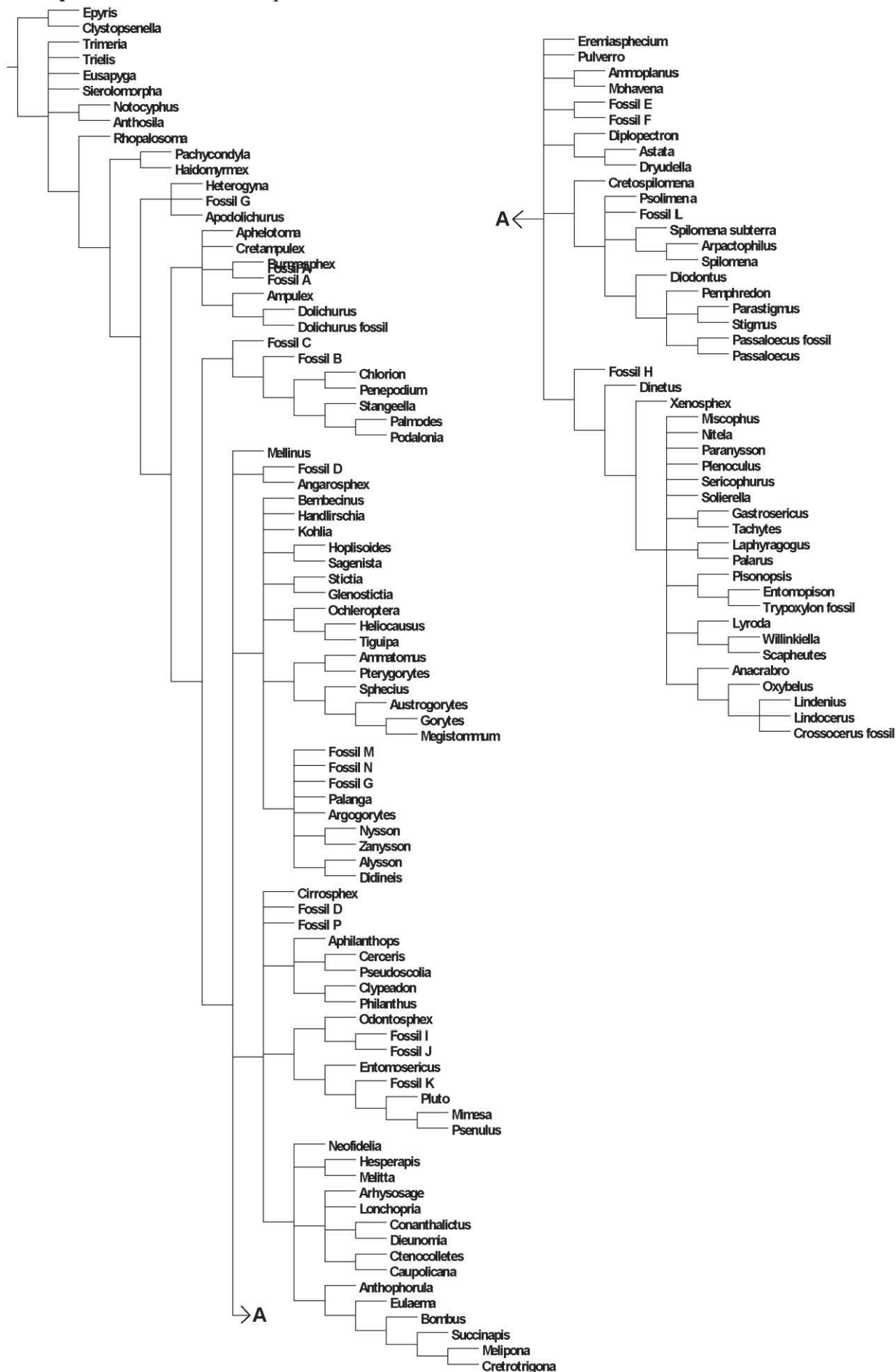
	7	7	8	7	8	8	8	8	8	8	9	9	9	9	9	9	9	9	9								
Terminais	6	7	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9		
Epyris	1	0	0	0	0	0	0	0	0	5	4	0	0	-	0	0	1	-	0	0	0	-	-	0	-		
Clystospenella	0	0	0	0	0	0	0	0	0	5	4	0	0	-	0	0	1	-	1	0	0	1	-	-	0	-	
Notocyphus	-	0	0	0	1	0	0	0	1	5	4	0	0	-	2	0	1	-	1	0	0	-	-	0	-		
Rhopalosoma	-	0	0	0	0	0	0	0	0	5	4	0	0	-	0	0	1	-	1	0	0	-	-	0	-		
Trimeria	0	2	0	0	0	0	0	0	0	5	4	0	0	-	4	0	1	-	1	0	0	-	-	0	-		
Trielis	0	2	0	0	0	0	0	0	2	5	4	0	0	-	2	1	1	-	1	0	0	-	0	0	-		
Anthosila	1	2	0	0	0	0	0	0	1	5	4	0	0	-	0	0	1	-	1	0	0	-	-	0	-		
Eusapyga	0	2	0	0	0	0	0	0	1	5	4	0	0	-	0	0	1	-	1	0	0	-	-	0	-		
Sierolomorpha	1	2	0	0	0	0	0	0	0	5	4	0	0	-	0	0	1	-	1	0	0	-	-	0	-		
Pachycandyla	2	0	0	0	0	0	0	0	2	5	4	0	0	-	2	1	1	-	1	0	0	-	-	0	-		
†Haidomyrmex	2	0	0	0	0	0	?	?	0	2	5	4	0	0	-	2	1	?	?	?	?	0	0	-	0	-	
†Burmasphex	1	1	0	0	0	?	?	?	0	2	5	4	0	0	-	1	1	?	?	?	?	0	1	?	3	0	-
†Fossil A	1	1	0	0	0	0	?	?	0	2	5	4	0	0	-	1	1	?	?	?	?	0	1	?	3	0	-
Ampulex	1	1	0	0	0	0	0	0	2	5	4	0	0	1	0	0	1	1	-	0	0	0	-	0	0	-	
Apheletoma	1	0	0	1	0	1	0	0	2	5	4	0	0	1	0	0	1	1	-	0	0	0	-	0	0	-	
Dolichurus	1	0	1	0	0	0	0	0	2	5	4	0	0	1	0	0	1	1	-	0	0	0	-	0	0	-	
†Dolichurus fossil	1	0	1	0	0	?	?	?	0	2	5	4	0	0	1	0	0	1	?	?	?	0	0	-	0	0	-
†Cretatmplex	1	0	0	0	0	?	?	?	0	2	5	4	0	0	1	0	0	1	?	?	?	0	0	-	0	0	-
Heterogyna	0	0	1	1	0	1	0	0	2	5	4	0	0	1	0	0	1	?	-	1	0	0	-	1	0	-	
†Fossil G	?	?	?	?	0	?	?	?	0	2	5	4	0	0	1	0	1	?	?	?	?	0	0	-	?	0	-
†Apodolichurus	0	0	0	1	1	0	?	?	0	2	5	4	0	1	0	1	?	?	?	?	?	0	0	-	?	0	-
†Fossil B	0	0	0	?	?	?	?	?	0	2	1	1	0	0	-	3	1	?	?	?	?	1	2	0	2	0	-
†Fossil C	1	0	0	1	1	0	?	?	0	2	1	1	0	0	-	3	1	?	?	?	?	1	2	0	2	0	-
Chlorion	0	0	1	1	0	1	0	0	2	4	1	0															

[illegible]

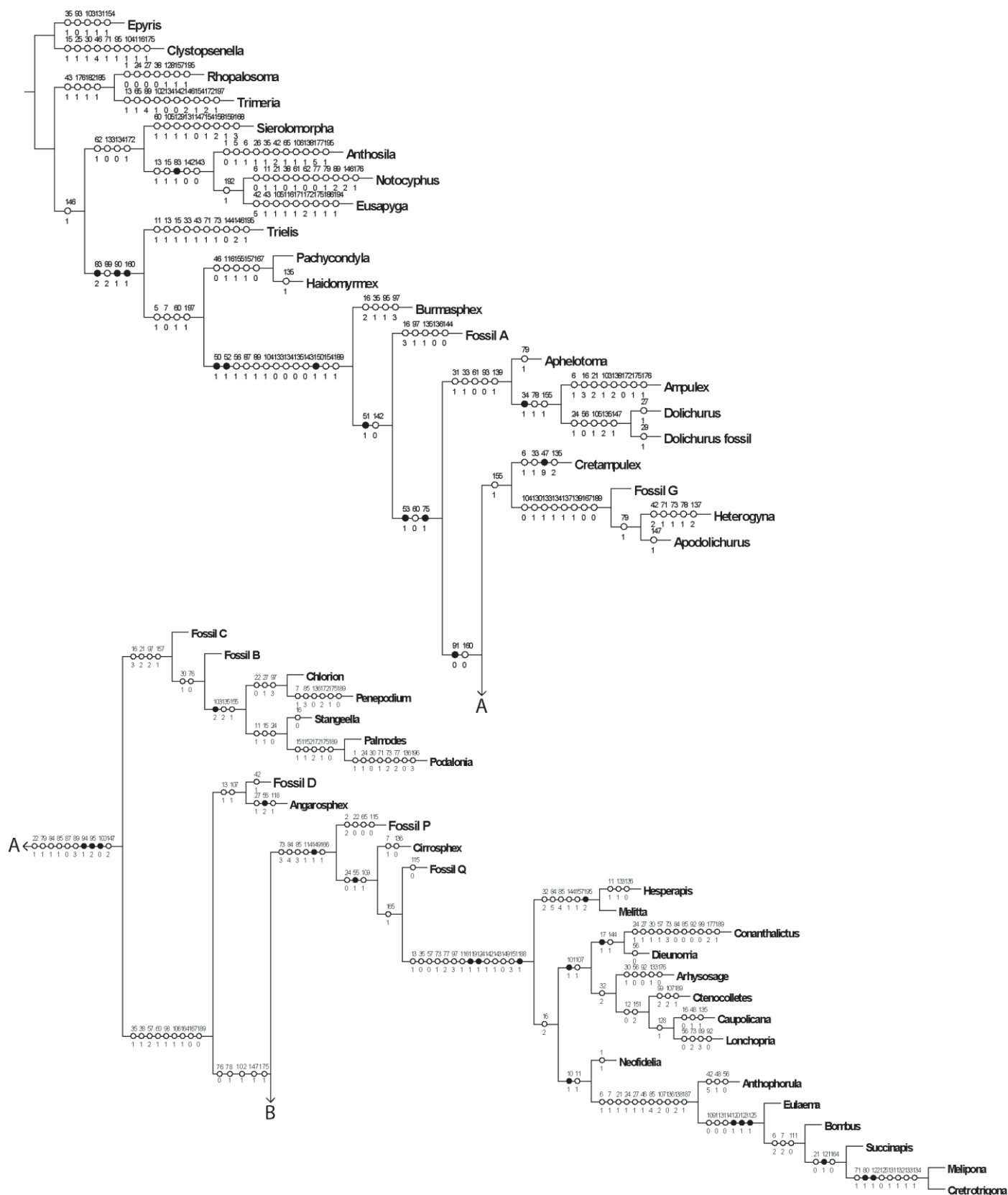
[illegible]

[illegible]

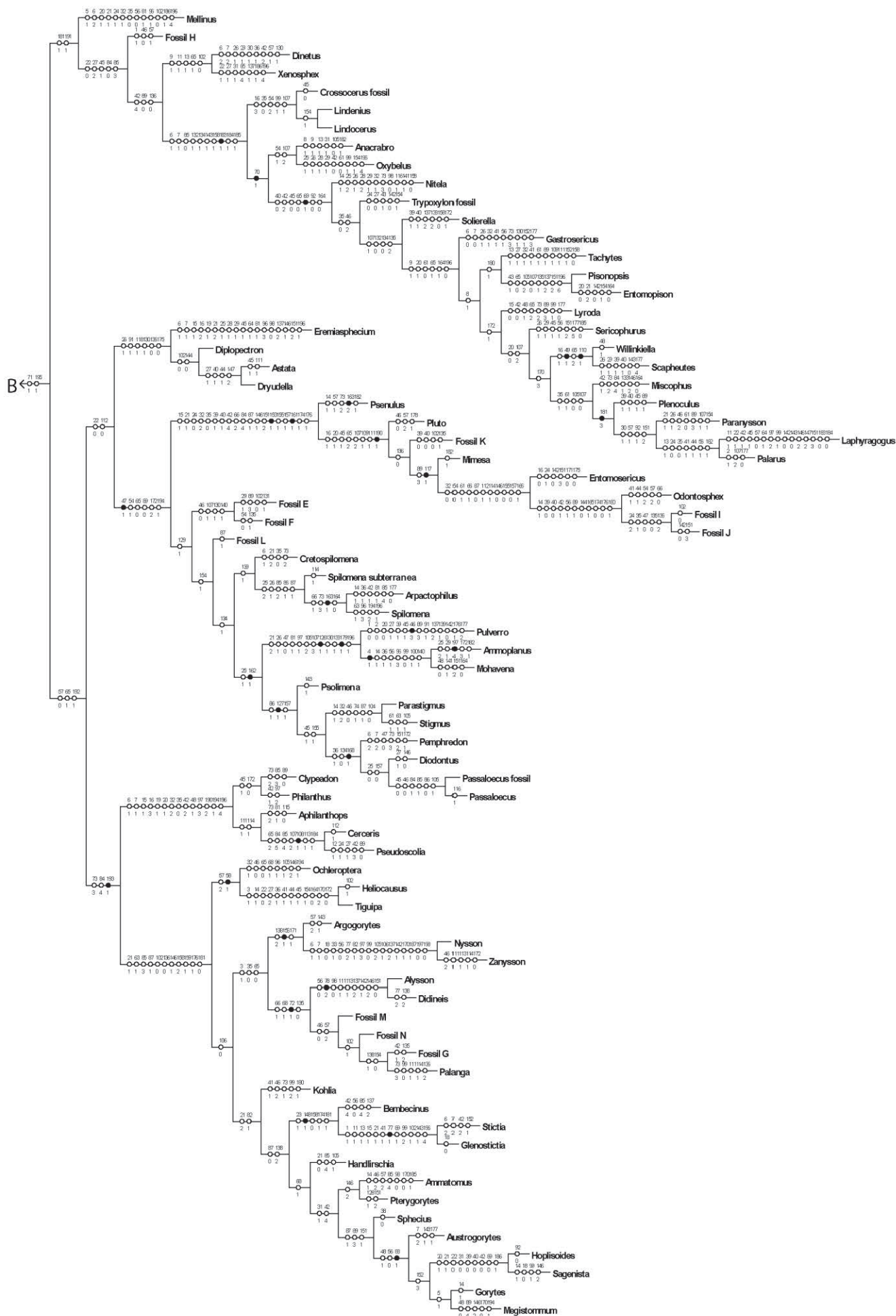
7.4 Apêndice 4. Árvores de parcimônia.



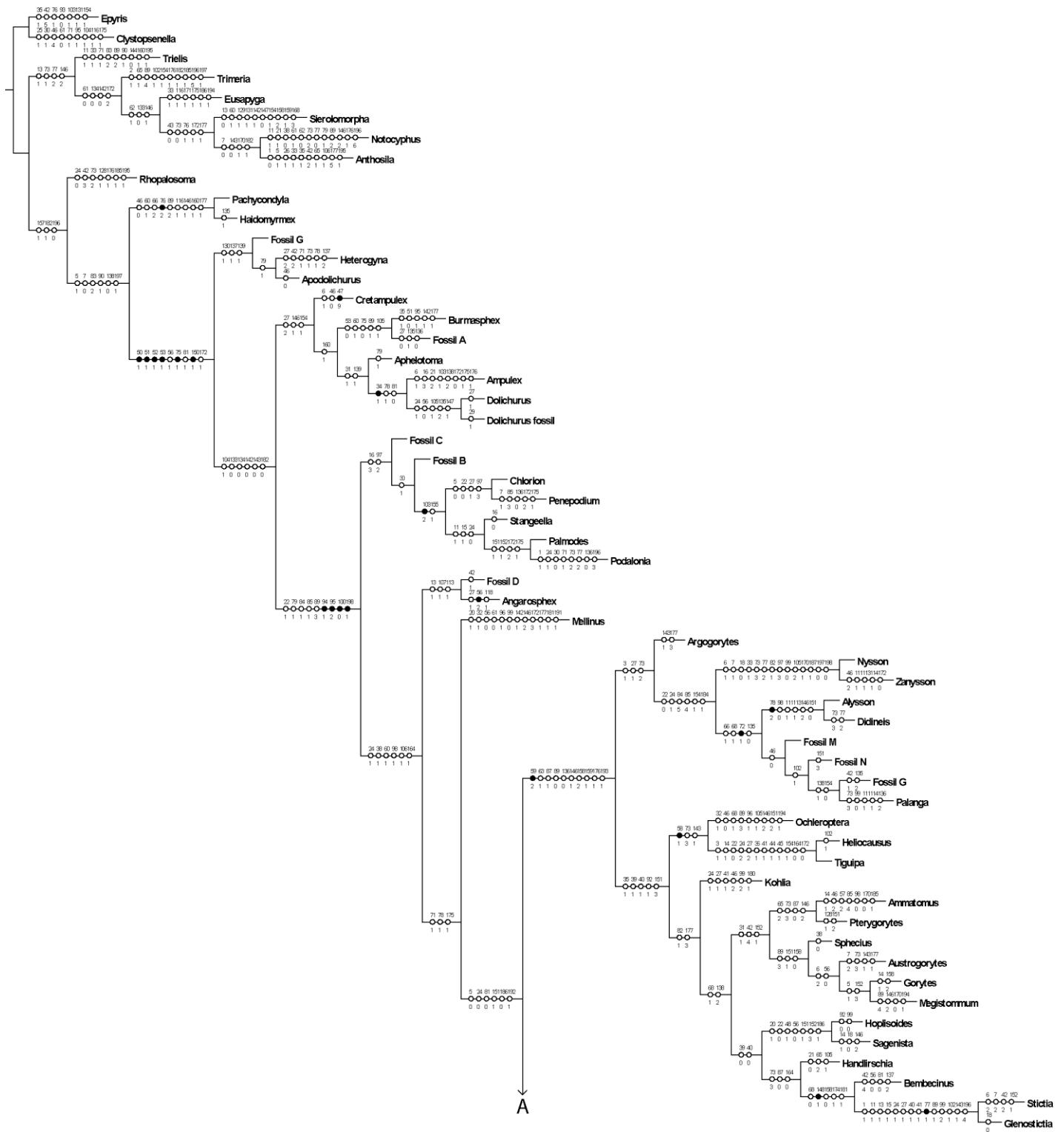
Apêndice 4 Figura 1. Consenso estrito de 380 árvores igualmente parcimoniosas com pesos iguais (2214 passos, CI = 13, RI = 66).



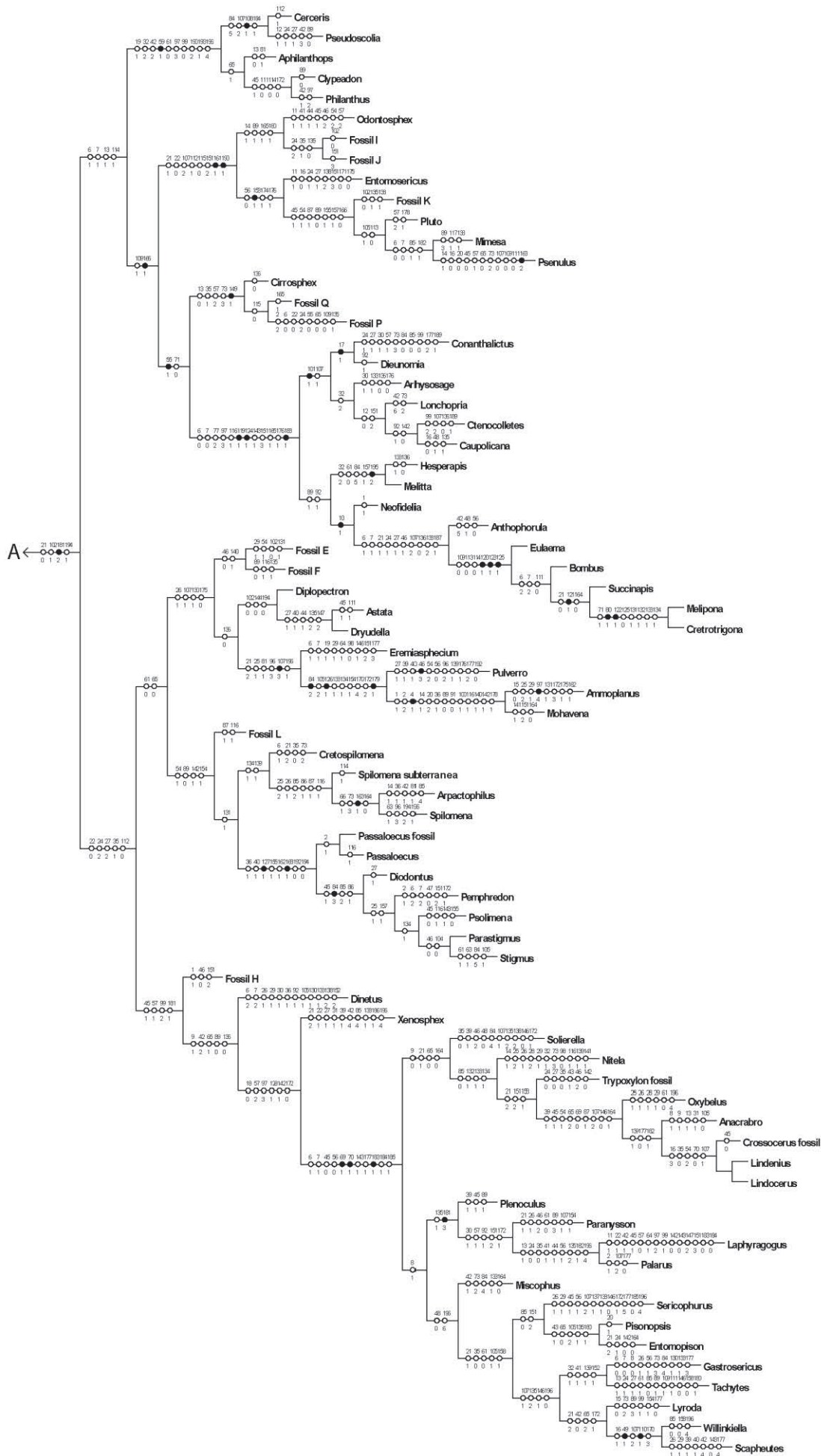
Apêndice 4 Figura 2. Árvore mais parcimoniosa obtida de uma análise de parcimônia com pesagem implícita com valor de $K = 3$ (default) (2089 passos, $CI = 14$, $RI = 68$). Otimização não ambígua dos caracteres.



Apêndice 4 Figura 2. Continuação. (2/2). Árvore mais parcimoniosa obtida de uma análise de parcimônia com pesagem implícita com valor de $K = 3$ (default) (2089 passos, $CI = 14$, $RI = 68$). Otimização não ambígua dos caracteres.

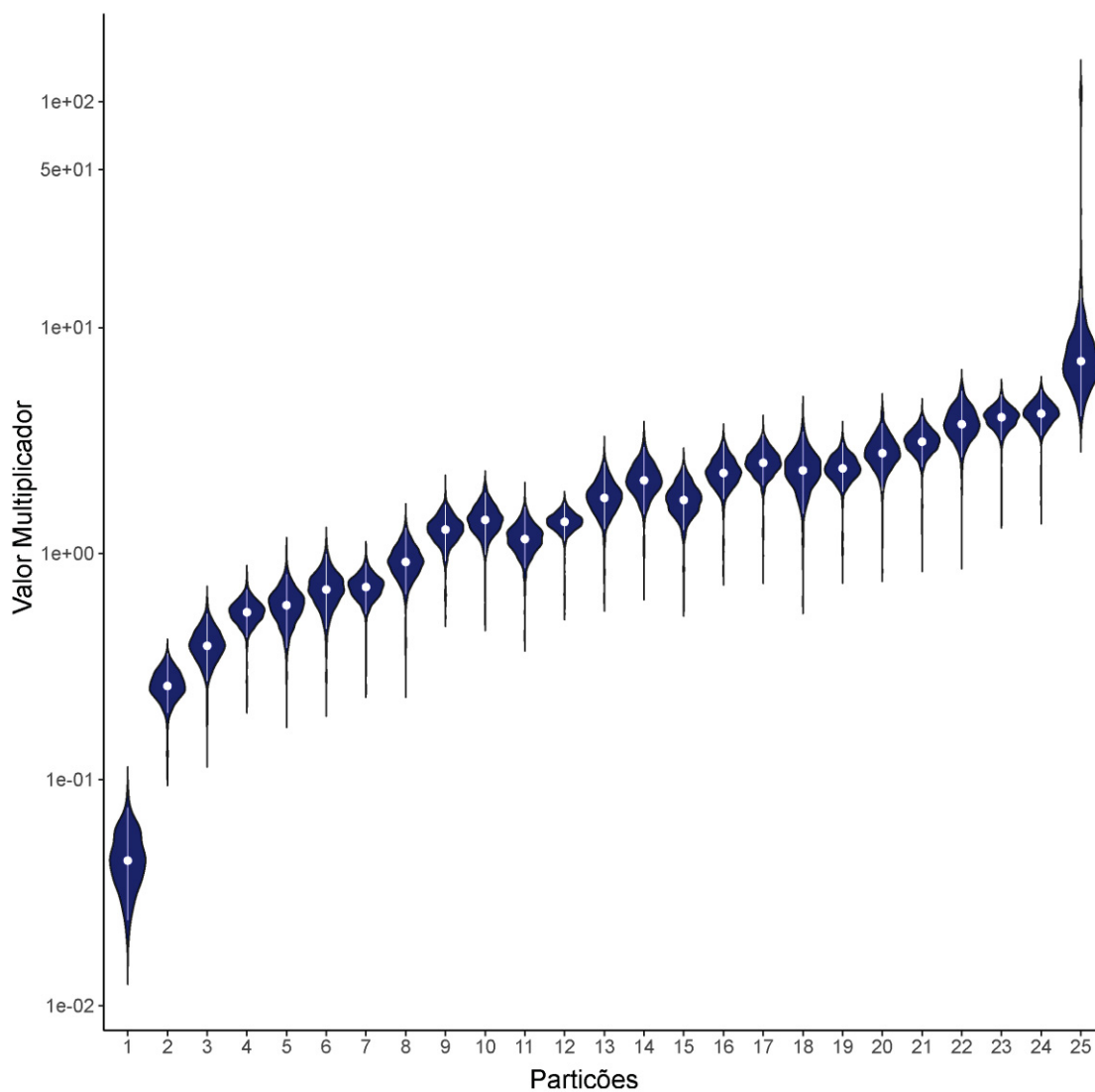


Apêndice 4 Figura 3. Árvore mais parcimoniosa obtida de uma análise de parcimônia com pesagem implícita com valor de $K = 27.226563$ (default) (1995 passos, $CI = 14$, $RI = 70$). Otimização não ambígua dos caracteres.



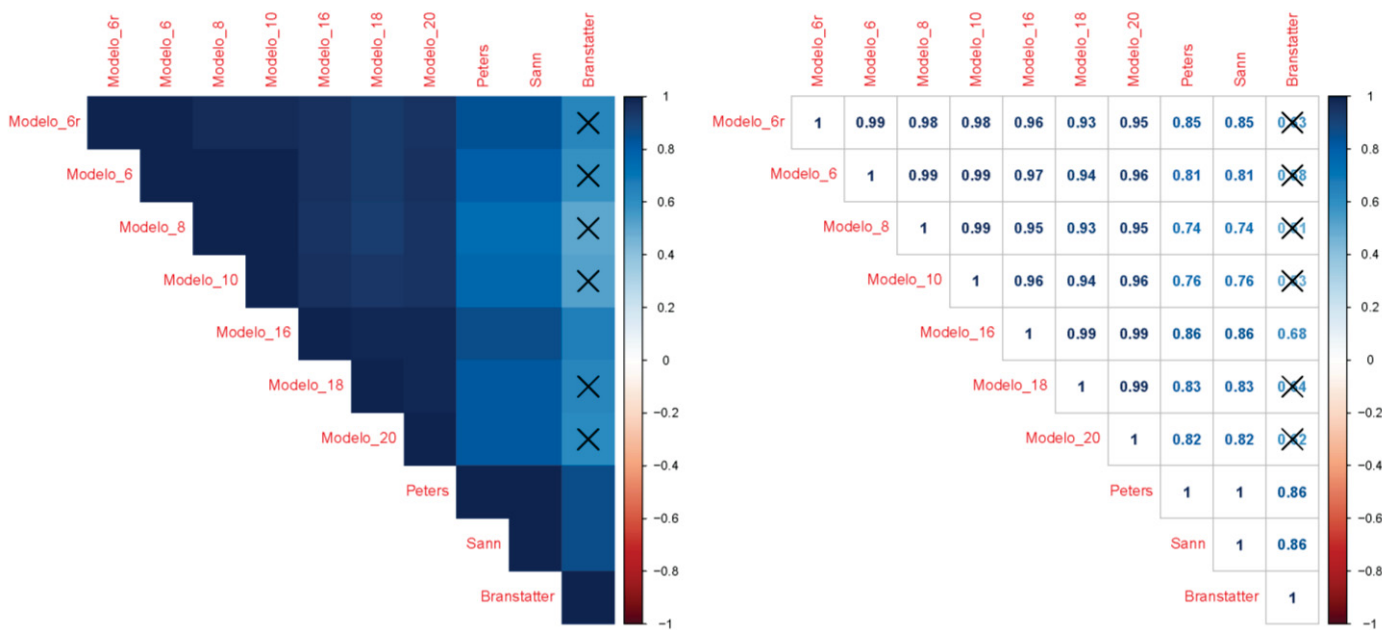
Apêndice 4 Figura 3. Continuação. (2/2). Árvore mais parcimoniosa obtida de uma análise de parcimônia com pesagem implícita com valor de K = 27.226563 (default) (1995 passos, CI = 14, RI = 70). Otimização não ambígua dos caracteres.

7.5 Apêndice 5. Valores multiplicadores das partições morfológicas.



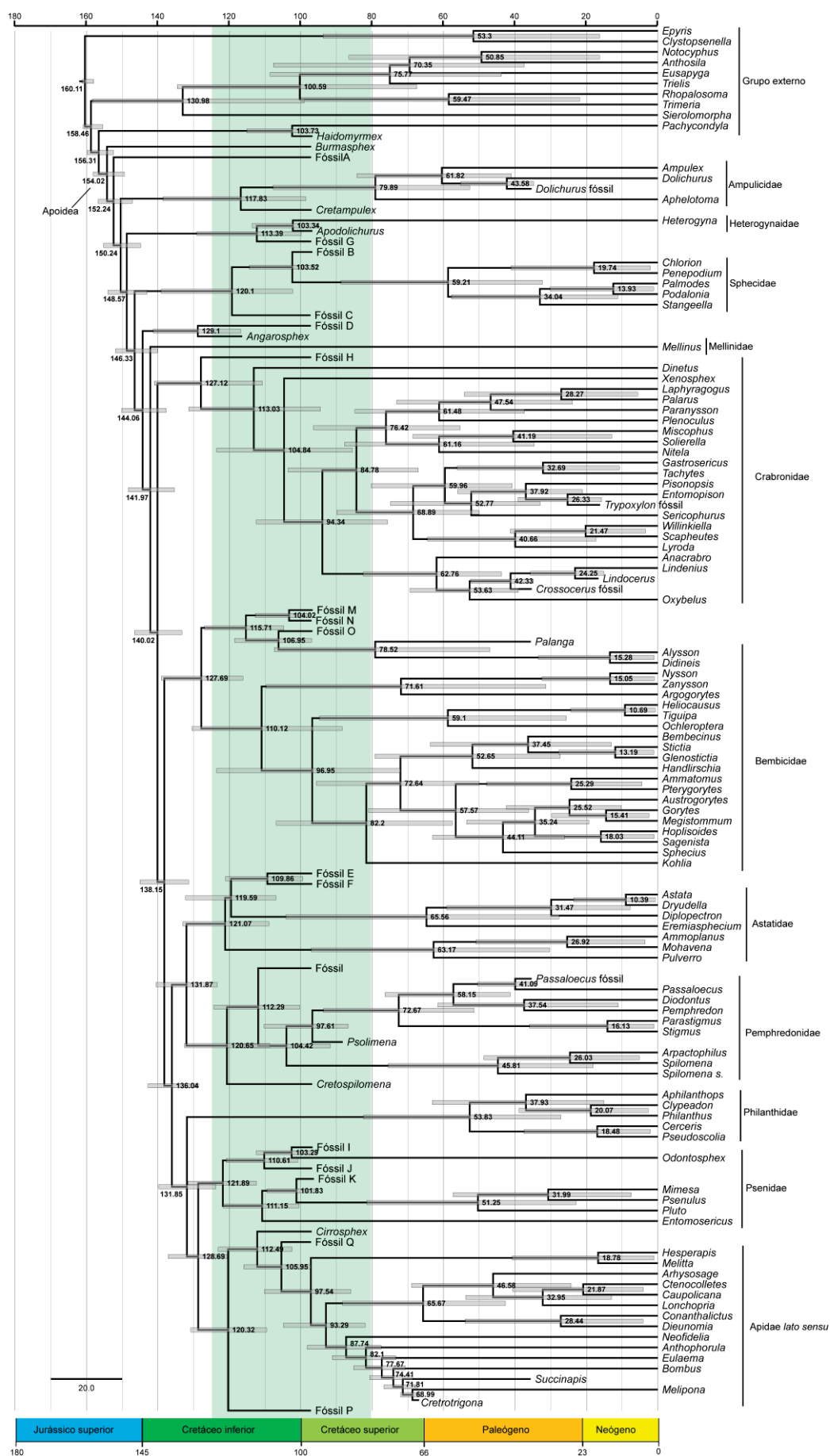
Apêndice 5 Figura 1. Valor multiplicador para cada partição (*rate multiplier per partition*) em um particionamento máximo de dados morfológicos.

7.6 Apêndice 6. Correlações entre as médias das idades estimadas para diferentes modelos de relógio.



Apêndice 6 Figura 1. Correlações entre as médias das idades estimadas de 10 hipóteses avaliadas (os sete modelos morfológicos e as três hipóteses moleculares) para os 16 clados de Apoidea. Branstatter: Branstatter et al. (2017). Peters: Peters et al. (2017). Sann: Sann et al. (2018). Valores de $R^2 > 0.00$ apresentam correlações positivas. Comparações marcadas com “X” apresentam correlações não significativas ($p > 0.05$).

7.7 Apêndice 7. Cronograma das linhagens de Apoidea contraposta com o período KTR.



Apêndice 7 Figura 1. Cronograma com base em dados morfológicos em abordagem Bayesiana. Relógio morfológico do Modelo 6. Nove partições. Todos os grupos compatíveis adicionados (*Contype=allcompat*). Escala em milhões de anos. Em verde o intervalo de 120-80 Ma (KTR).

7.8. Apêndice 8. Restrições do cronograma da Figura 5.

O cronograma da Figura 5 possui 20 restrições. Essas restrições foram atribuídas seguindo os resultados de testes topológicos descritos no materiais e método (2.4 testes de modelos) e apresentados nos resultados (2.1.4 Testes de hipóteses). Também foram feitas restrições para simular as idades dos clados presentes na Figura 1.

A anotações das restrições são as seguintes:

constraint 1 = 13-.;
constraint 2 = 14-.;
constraint 3 = 19-.;
constraint 4 = 22-.;
constraint 5 = 31-.;
constraint 6 = 32-.;
constraint 7 = 56-.;
constraint 8 = 81-.;
constraint 9 = 81-101;
constraint 10 = 102-.;
constraint 11 = 107-.;
constraint 12 = 115-.;
constraint 13 = 33-55;
constraint 14 = 34-55;
constraint 15 = 128-132;
constraint 16 = 129-132;
constraint 17 = 131-132;
constraint 18 = 126-132;
constraint 19 = 120-125;
constraint 20 = 120-132;

- 1) Restrições para o Fóssil A e *†Burmasphex*. A relação desses dois táxons foi fortemente rejeitada (KRS=34.66). Correspondem as restrições 1 e 2.
- 2) Posicionamento de Heterogynaidae. Seguiu-se a topologia da Figura 1. Correspondem as restrições 3 e 4.
- 3) Restrições para *Mellinus* como linhagem independente. A relação dessa linhagem juntamente com os demais Crabronidae *stricto sensu* foi altamente rejeitada (KRS= 13.94). Correspondem as restrições 5 e 6.
- 4) Restrições para o clado 8. Esse clado apresentou baixo suporte, porém seguiu-se a topologia da Figura 1. Corresponde a restrição 7.
- 5) Restrições para o clado 9. Esse clado apresentou baixo suporte, porém seguiu-se a topologia da Figura 1. Corresponde a restrição 8.
- 6) Restrições para o clado 10. O posicionamento de Philanthidae no clado 10 apresentou baixo suporte. Esse posicionamento de baseia na topologia da Figura 1. O posicionamento de Psenidae por outro lado foi fortemente suportado em teste de hipótese (KRS=10.74). Correspondem as restrições 10, 11 e 12.
- 7) Restrições internas de Crabronidae *stricto sensu*. O posicionamento independente das linhagens *Dinetus* e *Xenosphex* apresentaram alto suporte de ramos, PP=1 e PP=0.88 respectivamente. Restrições 13, 14 e 15.
- 8) As demais restrições foram para as relações internas em Apinae e são explicadas nos materiais de método. Correspondem as restrições 16 a 20.

8. **Anexo.** Homoplasy-based partitioning outperforms alternatives in Bayesian analysis of discrete morphological data.

Disponível em: <https://doi.org/10.1093/sysbio/syz001>

ACCEPTED MANUSCRIPT

Homoplasy-based partitioning outperforms alternatives in Bayesian analysis of discrete morphological data

Brunno B Rosa ✉, Gabriel A R Melo, Marcos S Barbeitos ✉

Systematic Biology, syz001, <https://doi.org/10.1093/sysbio/syz001>

Published: 11 January 2019 **Article history** ▼

“ Cite 🔑 Permissions ➦ Share ▼

Homoplasy-Based Partitioning Outperforms Alternatives in Bayesian Analysis of Discrete Morphological Data

BRUNNO B. ROSA¹, GABRIEL A.R. MELO¹, AND MARCOS S. BARBEITOS^{2,*}

¹Laboratório de Biologia Comparada de Hymenoptera, Departamento de Zoologia, Universidade Federal do Paraná, Caixa Postal 19020, Curitiba 81530-980, Brazil; and ²Laboratório de Evolução dos Organismos Marinhos, Departamento de Zoologia, Universidade Federal do Paraná, Curitiba 81530-980, Brazil

*Correspondence to be sent to: Laboratório de Evolução dos Organismos Marinhos, Departamento de Zoologia, Universidade Federal do Paraná, Curitiba 81530-980, Brazil;
 E-mail: msbarbeitos@gmail.com.

Received 24 July 2017; reviews returned 20 December 2018; accepted 4 January 2019
 Associate Editor: Luke Harmon

Abstract.—Bayesian analysis of morphological data is becoming increasingly popular mainly (but not only) because it allows for time-calibrated phylogenetic inference using relaxed morphological clocks and tip dating whenever fossils are available. As with molecular data, recent studies have shown that modeling among-character rate variation (ACRV) in morphological matrices greatly improves phylogenetic inference. In a likelihood framework this may be accomplished, for instance, by employing a hidden Markov model to assign characters to rate categories drawn from a (discretized) Γ distribution and/or by partitioning data sets according to rate heterogeneity and estimating per-partition branch lengths, conditioned on a single topology. While the first approach is available in many phylogenetic analysis software, there is still no clear consensus on how to partition data, except perhaps in the simplest cases (e.g., “by codon” partitioning of coding sequences). Additionally, there is a trade-off between improvement in likelihood scores and the number of free parameters in the analysis, which rises quickly with the number of partitions. This trade-off may be dealt with by employing statistics that penalize overfitting of complex models, such as Akaike or Bayesian information criteria, or the more recently introduced stepping-stone method for marginal likelihood approximation. We applied the latter to three distinct matrices of discrete morphological data and demonstrated that sorting characters by homoplasy scores (obtained from implied weighting parsimony analysis) outperformed other partitioning strategies (anatomically-based and PartitionFinder2). The method was in fact so efficient in segregating characters by rates of evolution that no within-partition ACRV modeling was necessary, while among-partition rate variation was adequately accommodated by rate multipliers. We conclude that partitioning by homoplasy is a powerful and easy-to-implement strategy to address ACRV in complex data sets. We provide some guidelines focusing on morphological matrices, although this approach may be also applicable to molecular data sets. [Aves; fossil; homoplasy; Hymenoptera; implied weighting parsimony; MrBayes; stepping-stone; TNT.]

Although DNA-sequence data have come to dominate phylogenetic inference, Bayesian analysis of morphological data is gaining increasing attention in recent years, especially for estimation of divergence time using both molecular and morphological characters in a total evidence approach (Quental and Marshall 2010; Wiens et al. 2010; Pyron 2011; Ronquist et al. 2012a; Lee et al. 2013, 2014; Dávalos et al. 2014; Giribet 2015; Lee and Palci 2015; Lee 2016; King et al. 2017). Bayesian inference based on morphology-only data sets is expected to continuously increase with possibility of novel studies investigating morphological evolution (Clarke and Middleton 2008; Klopstein et al. 2015) and of inferring dated phylogenies under morphological clocks (Ronquist et al. 2012a; Lee et al. 2013, 2014; Lee and Palci 2015; Lee 2016).

Bayesian analysis of morphological data has been based on Lewis’ (2001) Mk model. It was developed as a generalization of the Jukes–Cantor model (Jukes and Cantor 1969), with the single estimated parameter being the instantaneous rate of transition between character states (Lewis 2001). More recent developments on the analysis of discrete morphological data involve choice of the best distribution for among-character rate variation (ACRV) (Wagner 2012; Harrison and Larsson 2015), nonstationary models (Klopstein et al. 2015), and choice of the best hyperprior for modeling state frequencies

(Wright 2015). Overall the main concern has been in modeling heterogeneity in rates of evolution across characters.

An alternative (but not mutually exclusive) approach is to segregate characters into separate subsets or partitions according to their rates of evolution (Harrison and Larsson 2015; Wright 2015). Data partitioning of molecular data has become commonplace and it has been shown to strongly influence tree topology, branch lengths, clade support, and final likelihoods (Lemmon and Moriarty 2004; Brandley et al. 2005; Marshall et al. 2006; Brown and Lemmon 2007; Kainer and Lanfear 2015). Considering that genes and gene regions (e.g., exons versus introns, codon positions, paired versus unpaired regions in ribosomal genes) may evolve under distinct rates, sorting them into subsets allows for accommodation of heterogeneity on evolutionary rate. Partitioning of molecular data is currently carried out from a functional perspective, that is, sorting by gene and gene regions, or under an automated site-rate approach, such as that implemented in PartitionFinder (Lanfear et al. 2017). These approaches have been called, respectively, “ad hoc” and “algorithmic” by Kainer and Lanfear (2015).

Methods aimed at proposing partitions for morphological data are still incipient. Clarke and Middleton (2008) were the first to analyze morphological

data broken up in subsets based on morpho-functional knowledge of the studied characters. In their case, the subsets corresponded to characters derived from specific body regions, such as pelvic and pectoral bones, whose set elements are assumed to evolve independently from those of other sets. More recently, Clarke and Middleton's principles were followed by Tarasov and Génier (2015) for inferring relationships among dung beetles and by Lee (2016) to infer a dated phylogeny for mammals applying total-evidence multiple clocks chosen through grouping of partitions with similar relative branch lengths. This approach is comparable to the functional approach used for molecular data and is referred here as application of an anatomical criterion for partitioning of morphological data.

In parallel to the automated site-rate approach applied to molecular data, the current version of PartitionFinder (2.1.1; Lanfear et al. 2017) implemented the ability of analyzing morphological data sets. Using a preliminary version of PartitionFinder, Wright (2015) evaluated a large number of morphology-only data sets, including that of Clarke and Middleton (2008), and found support for the use of partitions in most of them. As in both Clarke and Middleton (2008) and Tarasov and Génier (2015), data partitioning considerably outperformed the results of the unpartitioned data sets. However, the partition schemes returned by PartitionFinder for Clarke and Middleton's data set in Wright (2015) resulted in lower marginal likelihood scores compared with the original anatomical partitions used by Clarke and Middleton (2008).

Here, we introduce a new method based on homoplasy as a criterion for partitioning of discrete morphological characters. We rely on the assumption that levels of homoplasy (that can be efficiently estimated from maximum parsimony trees) can be a good proxy for the evolutionary rate of the characters, because fast evolving characters will also be highly homoplastic. Therefore, partitions containing characters with similar homoplasy levels are expected to be evolving at similar rates. An analogous approach developed by Kjer and Honeycutt (2007) was applied to molecular data sets, with further modifications better suited to phylogenomic matrices proposed by Misof et al. (2014).

Bayesian analyses under models spanning a wide range of complexity were carried out using unpartitioned data sets and matrices partitioned by different "algorithmic" (including homoplasy partitioning) and "ad hoc" strategies. Results were compared using Bayes factors (BF) computed from

marginal likelihoods (MgL) obtained by stepping-stone (SS) importance sampling (Xie et al. 2011).

MgL is computed via SS as the weighted average of posterior likelihoods (the probability of the data given the model), being the weights provided by the prior. MgL scores are proportional to the degree of overlap between prior and posterior densities. Such as Akaike's or Bayesian information criteria (AIC and BIC), SS sampling also penalizes overparameterization. Complex models tend to have flatter prior landscapes due to the higher dimensionality of the parameter space and hence a poorer match between joint prior and posterior distributions. Thus, MgL scores obtained via SS tend to be worse for parameter-rich models than for simpler ones, unless the latter fail to adequately capture the complexity of the data.

In this study, we show that homoplasy partitioning was the most efficient way to segregate characters according to their rates of evolution. Under this scheme, there was no need for modeling rate variation among characters within partitions, whereas among-partition variation was adequately accommodated by rate multipliers. Character sorting using alternative strategies was not nearly as accurate, generating partitions with higher rate heterogeneity. Therefore, modeling rate variation under such schemes requires more complex models and priors, resulting in MgL scores significantly smaller than those obtained under homoplasy partitioning.

MATERIALS AND METHODS

Data Sets

Three matrices of discrete morphological data were used in the analyses (Table 1). Clarke and Middleton (2008) were the first to establish a partition criterion for discrete morphological data on birds, using the matrix from Clarke's et al. (2006). For this reason, we chose this matrix (henceforth called CEA) for comparisons with Clarke and Middleton's partition proposals. The second matrix (OZL) was obtained from Lee et al. (2014), who used an expanded and slightly modified version of O'Connor and Zhou's (2012) data set in explorative analyses of bird diversification employing morphological clocks (see Lee et al. 2014 for details). The third matrix contains morphological data of a rare family of aculeate wasps, Scolebythidae (SCO), chosen for the small number of terminals and a large set of fossil taxa.

TABLE 1. Summary of the analyzed data sets and corresponding references

Taxa	Matrix acronym	Number of terminals	Number of characters	References
Aves	CEA	25 (5 extant; 20 extinct)	205	Clarke's et al. (2006), Clarke and Middleton (2008)
Aves	OZL	65 (4 extant; 61 extinct)	247	O'Connor and Zhou (2012), Lee et al. (2014)
Hexapoda, Scolebythidae	SCO	18 (6 extant; 11 extinct)	28	Carpenter (1999), Engel and Grimaldi (2007), Engel et al. (2013), this study ^a

^aSee Supplementary File SM1 available on Dryad for characters and matrix.

The scolebythid matrix was assembled originally by [Carpenter \(1999\)](#) and has been modified, with addition of characters and fossil taxa, by [Engel and Grimaldi \(2007\)](#) and [Engel et al. \(2013\)](#). We also added data from *Clystosphenella mirabilis* Engel (extinct species from Dominican amber) and *Pristapenesia asiatica* Azevedo et al. (an extant species from Asia), as well as an additional character (Supplementary File SM1 available on Dryad at <http://dx.doi.org/10.5061/dryad.1hh515h>).

Partitioning

Two partitioning criteria were applied to all three data sets in order to compare with the unpartitioned scheme. First, we used anatomical partitioning according to [Clarke and Middleton \(2008\)](#). Under this approach, characters are allocated to partitions using morpho-functional criteria, further detailed in the following sections on each data set. Second, characters were partitioned according to their homoplasy levels, that is, they were mapped onto partitions corresponding to their homoplasy scores. [Farris \(1969\)](#) introduced the consistency index (CI) as a measure of character homoplasy on a given tree. CI is defined as the ratio between the minimum number of state changes in a matrix and the actual number of steps on a given tree. When $CI = 1$, the number of steps on the tree is the same as the minimum number of changes in the matrix. Therefore, the character has 0 homoplasy. [Goloboff \(1993\)](#) realized that CI penalized multistate over binary characters because the former would, by definition, have more changes on any given matrix if more than two states were found in the terminals. He improved on this measure by introducing a function that corrects for the difference in state numbers.

We used Goloboff's unbiased measure of homoplasy (f) implemented in the software TNT ([Goloboff et al. 2008](#)), with the default concavity parameter ($k=3$). The values returned by the implied weight analysis are normalized between 0 and 1, with the lowest value corresponding to no homoplasy. No values are assigned to non-informative characters, and these were assigned to their own partition. Table 2 provides a summary of the main partition schemes and the abbreviation adopted for each scheme by criterion. Additionally, each model has an individual sequential number available in [Supplementary Tables S1–S5](#) available on Dryad. The partition schemes for each data set are detailed below.

We also used the newly released version of PartitionFinder 2.1.1 ([Lanfear et al. 2017](#)), which now includes a script for partitioning of morphological characters. Data sets were analyzed under all possibilities of available parameters (using the option `–min-subset-size=1`), as follows: (1) tree topology: neighbor joining (default), parsimony (equal and implied weighting), and maximum likelihood; (2) branch lengths: linked and unlinked; (3) models of evolution: multistate+G+A; and (4) model selection: AIC, AICc, and BIC. We tested 24 parameter combinations for each data set, totaling 72 tested models ([Supplementary Table S6](#) available on Dryad).

Single-partition solutions and redundant schemes returned under different settings were not considered. PartitionFinder (PF) returned up to seven distinct schemes for a single data set ([Supplementary Table S6](#) available on Dryad). Given the large number of models to be tested in downstream phylogenetic analyses of each scheme (described below), we attempted to reduce the workload by setting a dissimilarity threshold below which schemes were considered too similar to

TABLE 2. Criteria, number of partitions, acronyms, and references for each scheme

Matrix	Criterion	Number of partitions	Acronyms	References
CEA	Anatomy	2	C-Ana1	Clarke and Middleton (2008)
	Anatomy	3	C-Ana2	Clarke and Middleton (2008)
	Anatomy	4	C-Ana3	Clarke and Middleton (2008)
	Homoplasy	8	C-Hom	This study
	Partition Finder	2	C-ParFin1	This study
	Partition Finder	4	C-ParFin2	This study
	Partition Finder	8	C-ParFin3	This study
	Partition Finder	18	C-ParFin4	This study
	Partition Finder	19	C-ParFin5	This study
	Partition Finder	22	C-ParFin6	This study
OZL	Partition Finder	25	C-ParFin7	This study
	Anatomy	7	O-Ana	This study ^a
	Homoplasy	14	O-Hom	This study
	Partition Finder	3	O-ParFin1	This study
	Partition Finder	5	O-ParFin2	This study
	Partition Finder	23	O-ParFin3	This study
	Partition Finder	25	O-ParFin4	This study
	Partition Finder	26	O-ParFin5	This study
	Partition Finder	27	O-ParFin6	This study
SCO	Anatomy	3	S-Ana	This study
	Homoplasy	6	S-Hom	This study
	Partition Finder	2	S-ParFin1	This study
	Partition Finder	8	S-ParFin2	This study

^aBased on anatomical regions in [O'Connor and Zhou \(2012\)](#) (see text for details).

deserve further evaluation. For all criteria (homoplasy, anatomical, and PF) each partitioning scheme was represented as a square matrix M , defined by $M_{ij}=1$ if the i th and j th characters were assigned to the same partition and $M_{ij}=0$ if otherwise. This is analogous to the representation of phylogenetic trees as adjacency matrices, whose cells are set to 1 if corresponding taxa belong to the same clade or to 0 if they belong to different clades. We then computed pairwise Hamming (or editing) distances between all matrices within each data set. Distances were rescaled between 0 and 1, where 1 is the maximum possible distance between matrices (i.e., when no characters shared the same partition between them) and represented as dendrograms using single linkage as the clustering method (Supplementary Fig. S1 available on Dryad). Dissimilarity among schemes had a clear bimodal distribution for CEA and OZL data sets, which had a number of PF schemes well below the arbitrarily established dissimilarity threshold of 1% (red line Supplementary Fig. S1 available on Dryad). In such cases, only the PF scheme with the smallest number of partitions within its cluster was further evaluated (highlighted in boldface in Supplementary Fig. S1 available on Dryad).

Aves: Clarke et al. data set.—Under the anatomical criterion, we used three partition schemes previously established by Clarke and Middleton (2008). The first scheme contains two partition sets (Supplementary Table S7 available on Dryad): pelvic plus cranial plus axial characters ($n=122$) and pectoral characters ($n=82$). The second partition scheme has three sets (Supplementary Table S8 available on Dryad): cranial plus axial ($n=71$), pectoral ($n=82$), and pelvic characters ($n=51$). The third scheme has four partition sets (Supplementary Table S9 available on Dryad): cranial ($n=52$), axial ($n=19$), pectoral ($n=82$), and pelvic characters ($n=51$). Under the homoplasy criterion eight different values of homoplasy were returned (Supplementary Tables S10 and S11 available on Dryad) and characters were sorted in eight corresponding partitions. Under the PF criterion, we selected four schemes that exceeded the dissimilarity threshold explained above, with 2, 4, 8, and 18 partitions, respectively (Supplementary Fig. S1 available on Dryad). Also, see the Supplementary Tables S12 to S18 available on Dryad for the partition sets returned by PF.

Aves: O'Connor and Zhou's data set modified by Lee et al.—Under the anatomical criteria, we used the different body sections of the character list in O'Connor and Zhou (2012) (Supplementary Table S19 available on Dryad): skull and mandible ($n=48$), vertebral column and ribs ($n=33$), thoracic girdle and sternum ($n=39$), thoracic limb ($n=57$), pelvic girdle ($n=21$), pelvic limb ($n=47$), and integument ($n=2$). Under the homoplasy criterion 14 distinct values of homoplasy scores were obtained (Supplementary Tables S20 and S21 available on Dryad) and characters were also partitioned according to these

values. We also evaluated three PF schemes, with 3, 5, and 23 partitions (Supplementary Fig. S1 available on Dryad). Also, see the Supplementary Tables S22 to S27 available on Dryad for the partition sets returned by PF.

Scolecbythidae data set.—Under the anatomical criterion the matrix was divided into three partitions (Supplementary Table S28 available on Dryad): head ($n=7$), mesosomal plus one metasomal character ($n=12$), and wing characters ($n=9$). According to the homoplasy criterion the Scolecbythidae matrix exhibited five different classes of homoplasy values and hence five partitions (Supplementary Tables S29 and S30 available on Dryad). We also tested the two schemes returned by PF, with 2 and 8 partitions, respectively (Supplementary Fig. S1 available on Dryad). Also, see the Supplementary Tables S31 and S32 available on Dryad for the partition sets returned by PF.

Parameters and Priors

All possible permutations of the tested parameters and priors were analyzed for each data set, resulting in a total of 1624 models. For CEA, we tested 1016 different models, 304 for OZL and 304 for the SCO data set, respectively. Details of the tested parameters are provided below.

Ascertainment bias (coding).—The implementation of the Mk model (Lewis 2001) in MrBayes allows for correction of the ascertainment bias exhibited by morphological data sets, in which only variable characters are coded. We tested two coding corrections: (1) *lset coding=variable*, which corrects for lack of non-variable characters in the matrix; and (2) *lset coding=informative*, which in addition to lack of non-variable characters also assumes that only parsimony-informative characters have been scored (Ronquist et al. 2011). We evaluated the effect of coding on unpartitioned, homoplasy- and anatomically-partitioned data sets. Informative coding resulted in much better MgL in several analyses (as obtained by Tarasov and Génier 2015), whereas others crashed repeatedly before completion. Conversely, mixed-coding trial runs in which non-informative characters (i.e., characters with a single state change) were appropriately scored as variable, and the remainder as informative, returned much lower MgL scores than observed for single-coding runs (data not shown). This pattern was consistent across all data sets. The problems arising when characters were scored as informative led us to focus only on variable-coding data sets because (in all completed analyses) the best models were the same for either ascertainment bias correction. We reported all comparisons between variable and informative coding using BF in Supplementary Table S5 available on Dryad, but data using informative coding were omitted from figures and tables in the Results section.

Among-partition rate variation (APRV) on branch lengths.—Following Marshall et al. (2006) and Clarke and Middleton (2008), we also evaluated distinct priors of rate variation among partitions implemented in MrBayes: (1) *prset ratepr=fixed*, which is fixed to the average state change rate across partitions; (2) *prset ratepr=variable*, which allows for variation among partitions, while constraining the average rate of change across partitions to 1. Branch lengths remain linked among partitions through rate multipliers which are, by default, drawn from a flat Dirichlet prior; and (3) *unlink brlens=(all)*, which allows for branch lengths to be independently estimated across partitions.

Among-character rate variation (ACRV).—In addition to the rate variation among partitions it is also possible to model within-partition rate variation, assuming an underlying Γ distribution whose shape parameter (α) is sampled in MrBayes from a uniform prior in the (0.05,50) interval. MrBayes allows the following settings: (1) *lset rates=equal* and *link shape=(all)*, which assumes no rate variation across characters; (2) *lset rates=gamma* and *link shape=(all)*, which assumes a single Γ distribution for all partitions, also referred to as shared Γ ; and (3) *lset rates=gamma* and *unlink shape=(all)*, in which each partition has its own Γ distribution, also referred as per-partition Γ .

Branch length prior.—We also tested the effect of different exponential priors for branch length using the inverse scale parameter (λ) suggested by Marshall et al. (2006) for morphological characters and previously evaluated by Clarke and Middleton (2008), that is, $\lambda = 5$, $\lambda = 10$, $\lambda = 20$, and $\lambda = 40$.

Ordered versus unordered characters

Restrictions to changes between states of multistate characters are sometimes applied in analysis of morphological data, resulting in a preset ordering of the state changes. Here the characters were left unordered in the analyses of all data sets. Considering that Clarke and Middleton (2008) ordered 38 of the 205 characters, we also reanalyzed CEA implementing these same restrictions. For model comparison, however, we calculated MgL using SS (see below) and not harmonic means as done by these authors. The same parameters described above were applied.

Analysis and Model Testing

The analyses were conducted in MrBayes 3.2.6 (Ronquist et al. 2012b) through the CIPRES Science Gateway portal (Miller et al. 2010) and the Ohio Supercomputer Center (OSC, 1987). The different models were compared using BFs, interpreted according to Kass and Raftery (1995). Because MrBayes reports the natural log of MgL, BF are calculated not as the ratio, but the

difference between the $\ln(\text{MgL})$ of competing models, hence they are also reported as the natural log of the ratio, or $\ln(\text{BF})$. Kass and Raftery's statistic (KRS) is $2 \cdot \ln(\text{BF})$, or twice the pairwise differences between the MgLs of the most (H_1) and least (H_0) likely hypotheses (or the models with the highest and lowest MgL). Under Kass and Raftery's criteria, if this statistic is between 0 and 2, evidence against H_0 is "not worth more than a bare mention". Values between 2 and 6 indicate "positive evidence" against H_0 , from 6 to 10 should be interpreted as "strong evidence" and over 10, "very strong evidence" against the model with lowest MgL.

MgL were computed following the SS method (Xie et al. 2011). We employed 50 steps in MgL estimation via SS, while the number of generations and sampling frequencies varied according across data sets, as follows. A total of 50×10^6 generations and a sampling frequency of 1000 were employed for both CEA and OZL data sets, while 20×10^6 generations and a sampling frequency of 100 were used for SCO. All models were compared against a standard analysis of the non-partitioned data following the default parameters (variable coding, $\lambda = 10$ and equal rates) indicated in the MrBayes 3.2 manual for morphological characters. Burnin was set to the first 25% of all samples.

In order to evaluate the tree topologies, additional MCMC analyses were carried out for the best models (those with KRS under 2) selected in the previous phase. The number of generations applied for each matrix were: 5×10^6 for CEA; 5 to 15×10^6 for OZL; and 2×10^6 for SCO. Four independent runs (four chains each) and a burnin of 25% were applied on all matrices. Convergence was checked in Tracer 1.6 (Rambaut et al. 2013), AWTY (Wilgenbush et al. 2004) with recommendations from Nylander et al. (2008), and in MrBayes. Chain temperatures were also empirically adjusted for each matrix in order to ensure proper mixing: *temp*=0.025 for OZL matrix; and *temp*=0.100 for both CEA and SCO matrices. The trees shown are majority-rule consensus trees (*Contype*=*Halfcompat*), with the resulting posterior probabilities as branch support.

We also evaluated the number and support of recovered clades in resulting majority-rule consensus tree obtained from each data set under the best models. We scored the number of clades with support above 0.95 and above 0.99. Our results were also qualitatively compared with previous works by evaluating changes in positions of major clades (see Supplementary File SM3 available on Dryad for results).

RESULTS

Partitioning Criteria and Choice of Parameters and Priors

Best models for each data set are compiled in the Table 3 and full results are detailed in the Supplementary Tables S1–S4 available on Dryad, in which all tested models are listed together with the respective values of

TABLE 3. Best models by criterion for all data sets analyzed with variable coding and unordered characters.

CEA							
Model	Partitions	APRV ^b	ACRV	λ	MgL	KRS ^c	KRS ^d
C-Unp-6	1	N/A	Shared Γ	10	-1625.19	126.36	0.00
C-Unp-7	1	N/A	Shared Γ	20	-1625.46	126.9	0.54
C-Unp-2 ^a	1	N/A	Equal rates	10	-1626.05	128.08	1.72
C-Ana1-34	2	Unlinked branch lengths	Shared Γ	10	-1612.56	101.1	0.00
C-Ana1-26	2	Unlinked branch lengths	Equal rates	10	-1613.01	102	0.90
C-Ana2-26	3	Unlinked branch lengths	Equal rates	10	-1613.28	102.54	1.44
C-Hom-14	8	Linked branch lengths	Equal rates	10	-1562.01	0.00	0.00
C-ParFin2-14	4	Linked branch lengths	Equal rates	10	-1600.60	77.18	0.00
C-ParFin2-32	4	Unlinked branch lengths	Per-partition Γ	40	-1600.77	77.52	0.34
C-ParFin1-14	2	Linked branch lengths	Equal rates	10	-1601.08	78.14	0.96
C-ParFin1-22	2	Linked branch lengths	Shared Γ	10	-1601.19	78.36	1.18
C-ParFin1-18	2	Linked branch lengths	Per-partition Γ	10	-1601.29	78.56	1.38
C-ParFin2-22	4	Linked branch lengths	Shared Γ	10	-1601.55	79.08	1.90
OZL							
Model	Partitions	APRV ^b	ACRV	λ	MgL	KRS ^c	KRS ^d
O-Unp-6	1	N/A	Shared Γ	10	-3770.20	192.92	0.00
O-Unp-2 ^a	1	N/A	Equal rates	10	-3812.92	280.36	85.44
O-Ana-10	7	Fixed branch lengths	Shared Γ	10	-3770.19	194.90	0.00
O-Hom-14	14	Linked branch lengths	Equal rates	10	-3672.74	0.00	0.00
O-ParFin2-18	5	Linked branch lengths	Per-partition Γ	10	-3713.71	81.94	0.00
SCO							
Model	Partitions	APRV ^b	ACRV	λ	MgL	KRS ^c	KRS ^d
S-Unp-2 ^a	1	N/A	Equal rates	10	-284.44	23.58	0.00
S-Unp-6	1	N/A	Shared Γ	10	-284.97	24.64	1.06
S-Ana-2	3	Fixed branch lengths	Equal rates	10	-284.50	23.70	0.00
S-Ana-10	3	Fixed branch lengths	Shared Γ	10	-285.00	24.70	1.00
S-Ana-6	3	Fixed branch lengths	Per-partition Γ	10	-285.42	25.54	1.84
S-Hom-13	6	Linked branch lengths	Equal rates	5	-271.65	0.00	0.00
S-ParFin2-6	8	Fixed branch lengths	Per-partition Γ	10	-284.22	23.14	0.00
S-ParFin1-10	2	Fixed branch lengths	Shared Γ	10	-284.35	23.40	0.08
S-ParFin2-10	8	Fixed branch lengths	Shared Γ	10	-284.39	23.48	0.10
S-ParFin1-6	2	Fixed branch lengths	Per-partition Γ	10	-284.40	23.50	0.24
S-ParFin1-2	2	Fixed branch lengths	Equal rates	10	-284.47	23.64	0.24

Best model overall under each scheme in boldface.

Part. = Partitions; APRV = among-partition rate variation; ACRV = among-character rate variation; λ = inverse scale parameter of the exponential branch length prior; MgL = marginal likelihood; KRS = Kass and Raftery's statistic; N/A = not applicable.

^aDefault model.

^b"Fixed branch lengths": likelihood estimation was conditional on a single topology and set of branch lengths across partitions; "Linked branch lengths": branch lengths were allowed to vary across partitions, but remained linked by rate multipliers; "Unlinked branch lengths": branch lengths corresponding to each partition were independently estimated.

^cKass and Raftery's statistic computed with respect to the best model overall.

^dKass and Raftery's statistic computed with respect to the best model within each criterion.

(ln) MgL and Kass and Raftery's (1995) statistic (KRS). We detail results for each data set in the sections below.

CEA

Reanalysis of CEA shows that the best overall model was obtained for homoplasy partitioning, linked branch lengths, $\lambda = 10$ and no modeling of ACRV within partitions (i.e., equal rates; C-Hom-14; Table 3). The best model for the unpartitioned data set (C-Unp-6) also employed $\lambda = 10$, but ACRV was modeled via Γ distribution. This model was strongly rejected in favor of C-Hom-14 (KRS = 126.36; Table 3), despite its much smaller number of free parameters. The three best models using anatomical partitioning were significantly more complex because all of them employed unlinked branch lengths. The best model (C-Ana1-34) used a shared Γ distribution to accommodate ACRV, but

applying a single rate to each partition (C-Ana1-26) did not significantly worsen the marginal likelihood (KRS = 0.90; Table 3). Both models correspond to Clarke and Middleton (2008) first scheme (two partitions). Interestingly, the third best model (C-Ana2-26, KRS = 1.44) had the same settings as C-Ana1-26, but it was selected under the second scheme (3 partitions; Table 3). Even though anatomical schemes clustered together, C-Ana2 was closer to C-Ana3 than to C-Ana1 (Supplementary Fig. S1 available on Dryad). All anatomical models were also significantly worse than C-Hom-14 (KRS > 100; Table 3). However, anatomical partitioning did generate a very strong improvement in model fit over analyses of unpartitioned data (KRS > 26 in all pairwise comparisons, calculations not shown).

Four PartitionFinder2 (PF) schemes were analyzed: C-ParFin-1, C-ParFin-2, C-ParFin-3, and C-ParFin-4, with 2, 4, 8, and 18 partitions, respectively. There were actually seven possible schemes, but similarity among

C-ParFin-4, C-ParFin-5, C-ParFin-6, and CParFin-7 (18, 19, 22, and 25 partitions, respectively) exceeded 99%, hence only C-ParFin-4 was analyzed. All PF schemes were clustered in the single-linkage dendrogram, with the exception of C-ParFin1, which was the most dissimilar of them all, most likely because it had the smallest number of partitions (Supplementary Fig. S1 available on Dryad). C-ParFin-1 was returned by PF when using unlinked branch lengths and either AICc or BIC, regardless of starting tree choice (default NJ, ML, MP, or Implied-Weights MP). C-ParFin-2 was selected for any given tree by AICc when branches were linked, whereas C-ParFin-4 was chosen by AIC, also using linked branch lengths, for all starting trees, except implied-weights MP. Overall, linking branch lengths when the selection criterion was either AIC or BIC resulted in schemes with larger numbers of partitions (Supplementary Table S6 available on Dryad).

There were six best models under PF: three of them were recovered for scheme C-ParFin-1 and three for C-ParFin-2. The model with the best marginal likelihood (C-ParFin2-14, $MgL = -1600.60$) had the same parameters as C-Hom-14, but its partitioning scheme was very strongly rejected in favor of the latter's ($KRS = 77.18$; Table 3). There was Bayesian preference for

linked branch lengths and $\lambda = 10$ in the case of all other top PF models, with the exception of C-ParFin2-32 (unlinked branch lengths and $\lambda = 40$). Modeling ACRV using equal rates, shared or per-partition Γ distributions had negligible effect on the MgL of the six top models (Table 3). In every case, rejection of unpartitioned and anatomically partitioned models in favor of PF was also very strong ($KRS > 44$ and $KRS > 22$, respectively, calculations not shown). A graphical summary of the results for CEA is shown in Supplementary Fig. S2 available on Dryad. There was either strong or very strong Bayesian preference for C-Hom-14 (highlighted in red) over all alternative models.

As stated above, characters were so appropriately segregated by homoplasy partitioning that rate variation in CEA was fully captured by rate multipliers. Multipliers varied more than 800-fold across partitions, from 0.1 (or a tenth of the average rate) to ~82, although the 95% highest posterior density (HPD) ranged from 0.1 to ~32. Figure 1a shows the per-partition posterior distributions of these multipliers as violin plots, arranged in growing order of unbiased homoplasy values (f), with the exception of partition 8, to which non-informative characters were assigned. Dots correspond to medians and whiskers to 95% HPDs. Due to the

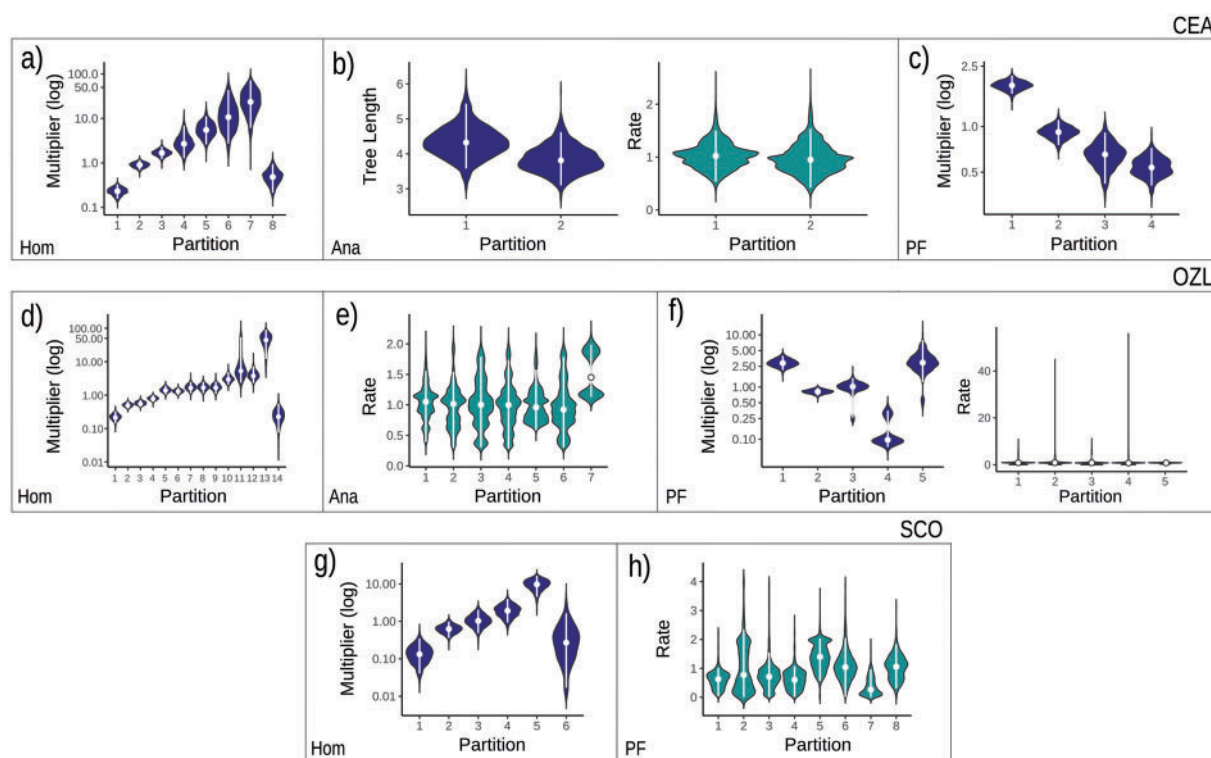


FIGURE 1. Violin plots representing per-partition Bayesian posterior distributions of tree lengths (model employed unlinked branch lengths to accommodate APRV), rate multipliers (linked branch lengths), and rates of character evolution (ACRV was approximated by a shared or per-partition Γ distribution). Multipliers and tree lengths are represented by blue and rate posteriors by green violins. Distributions are shown only for the top models within each partitioning strategy, listed in Table 3 and indicated in each panel (Hom = homoplasy; Ana = anatomy; PF = PartitionFinder2). Panels are arranged in rows corresponding to each data set, indicated on the right end of the row (CEA, OZL, SCO). Rate multipliers are represented in multiplicative ($\log_{10}$) scale. Note that violin widths are scaled with reference to their own partitions. a) C-Hom-14, b) C-Ana1-34, c) C-Parfin2-14, d) O-Hom-14, e) O-Ana-10, f) O-ParFin-18, g) S-Hom-13, h) O-ParFin8-18.

multiplicative nature of the parameter, posteriors are shown in a logarithmic scale. When computed on a per-partition basis, distributions were unimodal and symmetrical, suggesting that variation in rate multipliers results simply from the stochastic behavior of the Markov chain. Figure 1a suggests that the median of multipliers adjusted by MrBayes increased exponentially with f (hence the linear relationship in log-scale), while the median multiplier of the uninformative partition (8) was, as expected, considerably lower.

All best models for the anatomical schemes had unlinked branch lengths and the top model did require ACRV accommodation (shared Γ ; Table 3). This model was obtained under scheme C-Anat1 (two partitions), although character segregation was poor, as indicated by the strong overlap between tree length (blue) and rate (green) posteriors (Fig. 1b). In standard phylogenetic analysis, the shape (α) and rate (β) parameters of Γ are set to the same value (Yang 1993). As a consequence, the mean of the distribution (α/β) is always 1, leading to posteriors concentrated around this value (Fig. 1b).

PartitionFinder2's top model (C-ParFin2-14) was identical to C-Hom-14, that is, it used linked branch lengths and equal rates. However, modeling ACRV using either shared (C-ParFin2-22) or per-partition Γ distributions (C-ParFin2-18) did not significantly worsened the MgL (Table 3). Although best model partitions did vary with respect to rate multipliers, densities of partitions 2, 3, and 4 overlap and the latter resembles a mixture of distributions of similar masses (Fig. 1c).

OZL

In the analysis of OZL, a single best model was also selected for both unpartitioned and anatomically partitioned data sets. In both cases, $\lambda = 10$ and ACRV was accommodated by Γ distributions. SS analysis preferred fixed branch lengths under the top anatomical scheme (O-Ana-10). Fixing branch lengths means that a single set of branch lengths will be fit in all partitions, which is tantamount to suppressing the effect of partitioning in the analysis. It rendered O-Ana-10 and the top unpartitioned model (O-Unp-6) identical and their MgLs virtually indistinguishable (-3770.19 vs. -3770.20 ; Table 3). In other words, anatomical partitioning of OZL was so inadequate that the top model for this data set was no better than the best model using the unpartitioned matrix. Once more, the best homoplasy scheme outperformed the unpartitioned or anatomical alternatives ($KRS > 192$; Table 3). OZL's overall best model (O-Hom-14) had the same parameters as that for CEA (C-Hom-14). Likewise, the second overall best model (O-Hom-24) was also obtained when shared Γ and $\lambda = 40$ were set, but there was positive evidence against it ($KRS = 6.44$; Table 3; [Supplementary Table S2](#) available on Dryad).

When using linked branch lengths, OZL's PF schemes selected by either AIC or BIC had large numbers of partitions (23 or more). Unlinking branch lengths and/or

using the corrected version of AIC (AICc) returned schemes with 1, 3, or 5 five partitions ([Supplementary Table S6](#) available on Dryad). Although the number of partitions in the first group ranged from 23 to 27, they formed a cluster whose normalized dissimilarity was lower than 1% ([Supplementary Fig. S1](#) available on Dryad). Hence only the least partitioned scheme (O-ParFin3) was analyzed together with O-ParFin1 and O-ParFin2 (3 and 5 partitions, respectively). O-ParFin1 selection was independent of the starting tree or branch length modeling when either BIC or AICc were used. O-ParFin2 was always selected under AIC and unlinked branch lengths, regardless of the starting tree and O-ParFin3 was returned under AIC for linked branch lengths and only for the (default) neighbor-joining tree ([Supplementary Table S6](#) available on Dryad). Scheme O-ParFin2 yielded a single top model (O-ParFin2-18) employing linked branch lengths, shared Γ distribution and $\lambda = 10$. Once again, it was very strongly rejected ($KRS = 81.94$) in favor of the overall best model (O-Hom-14; Table 3) but also represented a very strong improvement over unpartitioned and anatomically partitioned data ($KRS > 114$). All SS results for OZL are summarized in [Supplementary Fig. S2](#) available on Dryad. Overall patterns were similar to CEA (Table 3).

OZL was the largest and, rate-wise, the most complex matrix analyzed in this study. There were 14 partitions under homoplasy, whose median rate multipliers again increased with f , although there was significant overlap in mid-range bins. The top anatomical model for OZL failed to segregate characters with respect to rate variation because there was Bayesian preference for fixed branch lengths. ACRV was accommodated by a shared Γ and it was extensive within each partition, as shown by the evident multi-modality of the rate posteriors (Fig. 1e).

The top PF model for OZL used linked branch lengths and per-partition Γ . Multiplier posteriors showed significant overlap between partitions 1 and 5 and partitions 2 and 3 and, on a multiplicative scale, partitions 3–5 distributions were slightly bimodal. Partition 5 also had long tails, although they were off the 95% HPD (in blue, Fig. 1f). Rate densities within partitions were, as expected, concentrated around 1, although the chain sporadically sampled extremely high values from the prior, leading to heavily skewed distributions (in green, Fig. 1f).

SCO

The smallest matrix (Scolebythidae) had the second largest number of competing best models. There were three when no partitions were applied to the data set, three under anatomic partitioning, one top model under homoplasy and five under PF. The latter strategy returned two schemes, with 2 and 8 partitions ([Supplementary Table S6](#) available on Dryad) and both schemes yielded best models, as listed in Table 3. There was Bayesian preference for $\lambda = 10$ in all models and for fixed branch lengths in all partitioned analyses; but

because fixing branch lengths actually eliminates the effect of partitioning, models varied only with respect to ACRV modeling. The only exception was homoplasy, whose best model (S-Hom-13) again employed linked branch lengths and $\lambda=5$ instead of $\lambda=10$. Excluding S-Hom-13, all other models were equivalent to S-Unp-2, if they employed fixed rates, or to S-Unp-6 if they employed shared or per-partition Γ . Marginal log-likelihoods varied from -285.42 to -284.22 (Table 3) thus KRS for the maximum possible $\ln(\text{BF})$ was 2.20 units, which is borderline evidence against the worse-fit model. It is interesting to notice that the largest Bayes factor was observed for models, which are supposed to be the same (S-Ana-6 and S-ParFin-6). This suggests that the “positive evidence” threshold of 1 log-likelihood unit proposed by Kass and Raftery (1995) may be too liberal when MgL are computed via SS. We conclude that all models aside from S-Hom-13 were roughly equivalent, in spite of differences in ACRV modeling, but once more strongly rejected in favor of homoplasy partitioning ($\text{KRS} > 23$, Table 3).

Per-partition posterior distributions of rate multipliers obtained from S-Hom-13 (Fig. 1g) resembled the ones obtained for CEA (Fig. 1a) although median dispersion was smaller. A single character (25) was assigned to the uninformative partition, hence the large variance of the corresponding distribution. The top PF model (S-ParFin2-6) employed fixed branch lengths and per-partition Γ . Rate posteriors were mostly bimodal with long tails (Fig. 1h), although not as extreme as found in OZL (Fig. 1e). Because the top model under anatomical partitioning used fixed branch lengths and equal rates, there were no comparable results to be shown.

Tree Topologies and Clade Support

A summary of the trees resulting from MCMC analyses implementing the default model (unpartitioned data) and the best models, by partitioning criteria, is shown in [Supplementary File SM3](#) available on Dryad. In general, no striking differences in number of recovered clades are observed among the results from all data sets. The homoplasy partitioning schemes, however, tend to increment the number of clades that are recovered in 50% or more trees in the posterior distributions (and hence not collapsed in the majority rule consensus tree, improving its resolution), in particular for OZL and SCO data sets. For CEA, the homoplasy schemes increase the number of clades with 0.99 posterior probability when applied in both unordered and ordered characters (see [Supplementary File SM2](#) available on Dryad).

DISCUSSION

Homoplasy Partitioning of Discrete Morphological Characters

Data partitioning in conjunction with the choice of adequate priors are decisive in Bayesian modeling of

morphological character evolution. Here we showed that implementing homoplasy partitioning in all analyzed data sets had a large effect on MgL when compared with anatomically partitioned or unpartitioned data sets (Table 3 and [Supplementary File SM2](#) available on Dryad). Additionally, we also found improved resolution of consensus both in OZL and SCO data sets ([Supplementary File SM3](#) available on Dryad). Our results agree with previous empirical studies that have also demonstrated higher performance of partitioned data sets (Clarke and Middleton 2008; Tarasov and Génier 2015; Wright 2015; Gavryushkina et al. 2016; King et al. 2017). Appropriate sorting of characters into partition sets by the homoplasy approach is evidenced by the Bayesian preference for linked branch lengths and equal rates in analyses of all data sets (Tables 3 and 4). The choice of no ACRV means that evolution across characters within homoplasy partitions was so homogeneous that fitting a single rate to each one was sufficient. On the other hand, there was considerable rate heterogeneity among partitions, as evidenced in Figure 1. In homoplasy partitioned matrices, among-partition rate variation (APRV) was efficiently accommodated by linking branch lengths, which adds a modest number of parameters to the model as the number of partitions increases. This result contrasts with what was obtained for Nylander et al. (2004) in analysis of combined morphological and molecular data sets. Although the authors did not employ any likelihood penalization method, they found an increase in 600–900+ log-likelihood units when comparing partitioned to unpartitioned data sets, whereas allowing ACRV increased log-likelihood in roughly 3000 units. Those authors partitioned the molecular data by gene, but not by codon, and did not attempt to partition the morphological matrix, so ACRV was probably necessary to accommodate the rate variation not captured by their schemes. In our study, there was positive to very strong Bayesian support for no ACRV modeling in homoplasy partitioned data (Table 3).

Homoplasy based criterion of data partitioning has been previously applied to molecular data sets by Kjer and Honeycutt (2007). They used similar reasoning,

TABLE 4. Comparisons between stepping-stone marginal likelihoods of models by partitioning criteria for Clarke and Middleton (2008) data set (here referred as Clarke's et al. data set). Highest MgL values in boldface.

Character ordering	Partitioning criteria	MgL by SS ^a
Unordered	Unpartitioned	–1639.44 to –1637.94
Unordered	Anatomy	–1642.41 to –1628.27
Unordered	Homoplasy	–1633.15 to –1574.67
Ordered	Unpartitioned	–1615.13 to –1608.08
Ordered	Anatomy	–1610.93 to –1612.16
Ordered	Homoplasy	–1602.92 to –1548.56
Ordered	PartitionFinder2	–1734.4 to –1623.5 ^b

^aThe following set of parameters was implemented in the analyses: coding variable, APRV unlinked and linked, ACRV per partition gamma, and $\lambda=5$.

^bMgL values taken from Wright (2015).

that is, partitioning of sites according to their levels of homoplasy. They calculated Farris' CI for each site based on a set of trees generated from 1000 bootstrap runs and sorted the returned CI's into six discrete rate classes, each corresponding to a partition. Considering the penalization of multistate characters imposed by the CI (see the section *Partitioning* in Methods and Materials), the method proposed here, based on Goloboff's unbiased measure of homoplasy, represents an improvement over that of Kjer and Honeycutt. A thorough evaluation of the application of this approach to molecular data sets is beyond the scope of this study, but should be explored in future work, including both nucleotide and aminoacid data sets.

The adequacy of homoplasy levels as proxy for rates of evolution is supported by the evident ordinal correlation between the per-partition median rate multipliers and Goloboff's unbiased homoplasy scores (Fig. 1a, d and g). If the posterior conforms to a Gaussian (normal) distribution, its mean (μ) is an unbiased estimator of any given parameter's expectation. Variation around the mean arises from the stochastic nature of the chain after reaching stationarity, hence expressing phylogenetic uncertainty or noise. A suitable measure for noise is the standard deviation (σ) of the posterior and the signal-to-noise ratio (SNR) of the Markov process is given by μ/σ (i.e., the reciprocal of the coefficient of variation). In signal processing analysis, SNR is usually converted to a log-ratio scale by the formula $10 \cdot \log_{10}(\mu/\sigma)$. When taken in this form, SNR is expressed in decibels (dB). In this scale, $\text{SNR} > 0$ means more signal than noise in the data and the opposite if $\text{SNR} < 0$ (Lockhart and Cheetham 1985). In the context of phylogenetic analysis, SNR thus calculated could be viewed as a measure of the phylogenetic information contained in the data partition. For instance, if the expected rate multiplier of a partition is 0.1 (i.e., $\mu = 0.1$) and $\sigma = 0.2$, noise is twice as high as signal. This means that the chain sampled much more freely from the prior than it would have if $\mu = 1.0$, in which case the standard deviation would only be 20% of the mean. In the first case, $\text{SNR} = -3.01$ dB and in the

second $\text{SNR} = 6.98$ dB. The difference in SNR between the first and second cases is a function of how much the data in the partition updated the prior; thus, SNR is effectively quantifying phylogenetic signal over noise. The number of characters also tends to decrease with the f score (see Supplementary Tables S11, S21 and S30 available on Dryad) thus exacerbating the drop in the partition's phylogenetic content.

Figure 1a, d and g show that there is indeed increase in within-partition rate multipliers variances with homoplasy scores. Because more homoplastic characters are less informative, the chain samples more freely from the prior, generating posteriors with wider HPDs. Therefore, their higher medians could be a mere byproduct of decreasing phylogenetic signal. To evaluate the normality of the multiplier posteriors we visually inspected corresponding probability-probability plots (Supplementary Fig. S3a–c available on Dryad). Most distributions were very close to normality, being skewed to the left in some cases (blue curve is below the confidence envelope), with the noticeable exception of OZL's partition 11, which is highly skewed (Supplementary Fig. S3b available on Dryad). We computed SNR for unordered CEA, SCO, and OZL and graphed it against f scores arranged in increasing order (Fig. 2). Results showed an overall trend of decreasing phylogenetic signal with increasing homoplasy, but noise never overcame signal except for partition 11 in OZL (Fig. 2b) and the non-informative partition (labeled “#” in Fig. 2c) in SCO.

OZL's partition 11 is made up by a single character, the relative size between the alular (“first wing”) digit and the metatarsal (numbered 168 in O'Connor and Zhou, 2012). Its corresponding posterior is highly skewed to the left, as it can be clearly seen in Figure 1d. Because it deviates strongly from normality (probably due to lack of adequate convergence), SNR is not accurately estimated for this partition because its mean is lower than expected, hence its departure from the overall pattern. It is interesting to notice that partition 13 also includes a single limb-related character, the distal

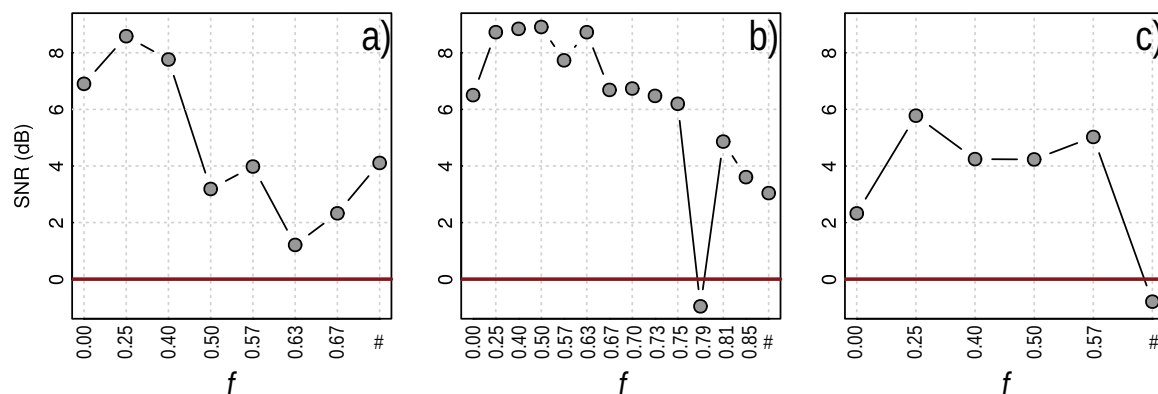


FIGURE 2. SNR plots correlating the amount of phylogenetic information in each partition to the corresponding homoplasy scores (f). Scores are in ordinal and not in ratio scale because no numeric values are assigned by TNT to the parsimony non-informative partition (labeled #). Negative SNRs (below the red line) denote more noise than signal in the partition. a) CEA, b) OZL, and c) SCO.

extent of metatarsal II, relative to metatarsal IV of the pelvic limb (236 in O'Connor and Zhou, 2012). However, its posterior met normality (Supplementary Fig. S3b available on Dryad) and its SNR exceeded 2 dB, although its *f* score is greater than that of partition 11 (0.85 vs. 0.79, Fig. 2b). Thus, low SNR does not seem to be attributable only to partition size. The same pattern was observed for SCO. Character 28 (shape of margin of pterostigma inside the marginal cell) was assigned to partition 5, whose *f* score was 0.57, while character 25 (shape of prosternum) is parsimony non-informative (partition 6, see Supplementary File SM1 available on Dryad for the full character list). Again, the multiplier posterior of the former conforms well to a normal distribution and has the second highest SNR, whereas the latter has negligible phylogenetic signal, although this estimate may not be very accurate for, as character 168 in OZL, it also departs from normality (Fig. 2c and Supplementary Fig. S3c available on Dryad).

This also has implications in deciding when partitions exhibiting similar behavior could be fused into a single partition. TNT's implementation of implied weighting analysis groups all characters with a single state change (singletons or autapomorphies, in parsimony context) into the same category. The corresponding posterior distributions of rate multipliers (the last ones in Fig. 1a, d and g) have medians which are closer to the ones obtained for the least homoplastic partitions (the first ones in each graph) than to any other partition in the same data set. Fusing the last and first partitions together resulted in even better MgLs than the best models, although evidence against them was not positive in every case (CAE KRS = 1.88, OZL KRS = 4.40, and SCO KRS = 2.82). At worst, these results suggest that these two partitions may be reduced to one with no significant worsening in model likelihoods.

We conclude that, although rising APRV is in part a byproduct of the reduction in phylogenetic information, there is some signal left even in the most homoplastic or parsimony uninformative partitions. In every case where the posterior significantly departed from normality, it was skewed to the left. In these cases, μ is lower than expected, so corresponding SNRs were likely underestimates, thus corroborating the prevalence of signal. It is been long known in molecular phylogenetic analysis (e.g., Philippe et al. 1994) that saturated characters, such as the third codon positions, evolve at higher rates. Thus, rate of evolution and phylogenetic signal are inherently intertwined and this was effectively captured by homoplasy partitioning, as evidenced in Figure 2. In the following sections, we discuss the performance of this strategy against its competing alternatives.

Anatomical Partitioning

The best models under anatomical partitioning of CEA employed unlinked branch lengths and resulted in considerable improvement over unpartitioned data

(see Results and Table 3). However, the posterior distributions of the top model's (C-Ana-34) tree lengths and rates showed considerable overlap (Fig. 1b and Supplementary File SM4 available on Dryad). All anatomical schemes were strongly rejected in favor of homoplasy (Table 3). Results were so insensitive to the choice of anatomical partitions proposed by Clarke and Middleton (2008) that despite a normalized dissimilarity of over 10% between C-Ana1 and C-Ana2 (Supplementary Fig. S1 available on Dryad), the three best models were based on either scheme (Table 3). In the two other data sets (OZL and SCO), tentative anatomical partitioning was even more inefficient because there was Bayesian preference for fixed branch lengths in all cases, thus results were no better than those obtained from the analysis of unpartitioned data (Table 3).

Besides model performance, we must also take into consideration the differences in implementation when comparing anatomical and homoplasy criteria. Use of an anatomical criterion, as implemented by Clarke and Middleton (2008), presupposes *a priori* functional units or anatomical modules, which have been hypothesized to evolve independently from one another. In the words of Clarke and Middleton (2008), these modules are expected to show "disjunct patterns of evolution". A similar reasoning was given by Tarasov and Génier (2015), who assumed "that characters of the same anatomical region undergo similar evolutionary dynamics". Mapping of homoplasy scores onto the longitudinally arranged characters (Fig. 3) suggests that, both in Aves and Scolebythidae, patterning rates of evolution is not necessarily correlated with anatomy. One could still argue that the major point in such partitioning strategy is that "disjunct evolution" means that branch lengths ought to be independently estimated across modules. Such scheme would still be superior to the unpartitioned analyses, thus simply mapping homoplasy to individual characters is not a fair test. Table 3 shows this to be true in the case of CEA for which there was an improvement of 12–13 log-likelihood units with respect to the unpartitioned analyses when unlinked branch lengths were employed. However, this was not the case of OZL and SCO: top models for these data sets employed fixed branch lengths, which is equivalent to no data partitioning. If on the one hand OZL and SCO matrices were not explicitly partitioned according to functional modules, on the other, Clarke and Middleton's (CEA) schemes vary substantially among themselves and yet resulted in remarkably similar MgL, as noted in the beginning of this section. Therefore, an additional problem confronting someone trying to implement this criterion derives from the *ad hoc* hypotheses invoked to support the delimitation of the functional modules. We expect that different groups of organisms will require their own specific hypotheses and that different researchers will eventually reach non-coincident partitions due to the profusion of functional hypotheses available. One could always resort to developmental data or attempt a more objective

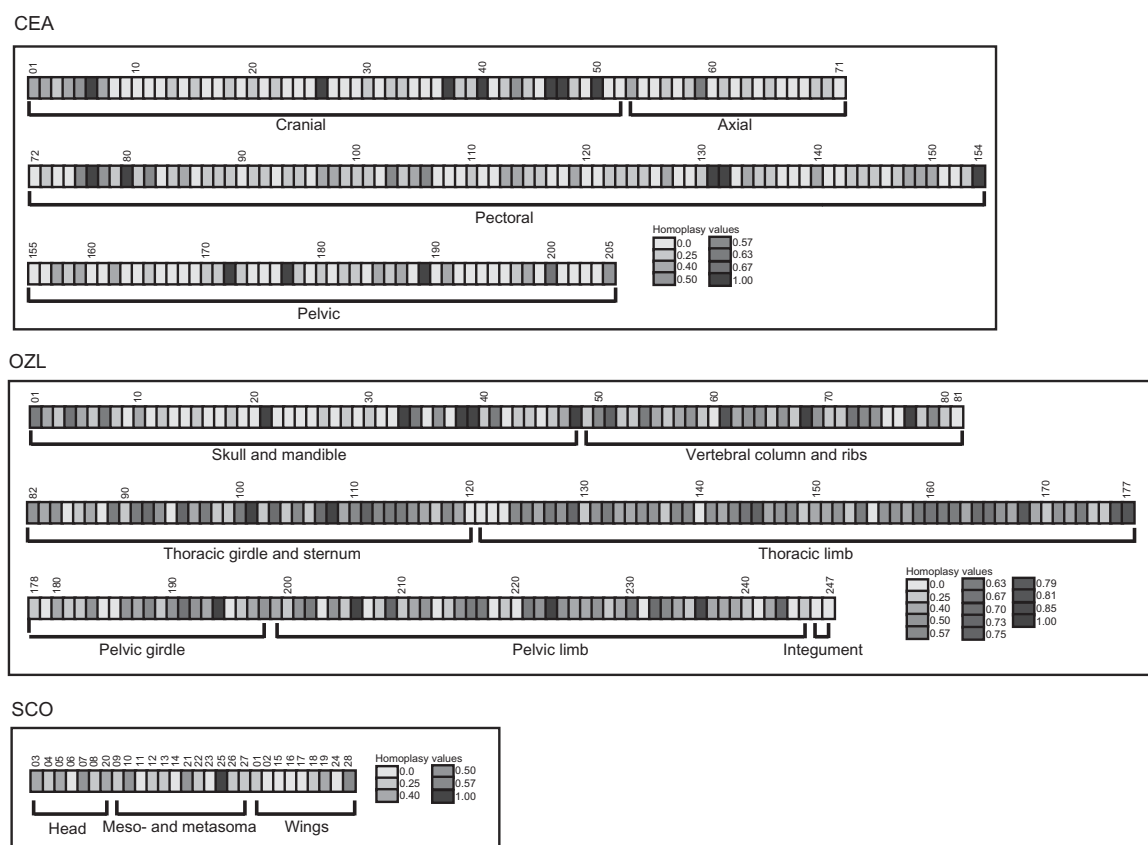


FIGURE 3. Homoplasy scores obtained from implied weighting parsimony mapped onto the corresponding anatomic partitions for the three analyzed data matrices. See text and [Supplementary Tables S4–S14](#) available on Dryad for details.

approach to module identification (e.g., [Magwene 2001](#)). At any rate, the implementation of anatomical data partitioning requires more complex premises than the homoplasy criterion.

Homoplasy versus PartitionFinder

Unlike the other strategies, PartitionFinder 2.1.1 returned several schemes for each analyzed data set. Scheme selection was highly sensitive to modeling of rate variation among partitions and the criterion used to penalize overparameterization (AIC, AICc, and BIC), but much more invariant with respect to the guide tree. Overall, we found a uniform pattern across all data sets: linking branch lengths reduced the number of free parameters, but led to oversplitting under AIC and especially under BIC, whereas starting trees had little influence on scheme choice ([Supplementary Table S6](#) available on Dryad). This is rather surprising given that BIC is supposed to be the most conservative criterion in model selection ([Brewer et al. 2016](#)). Our results also differed from tests with an earlier version of PartitionFinder2, in which [Wright \(2015\)](#) analyzed 333 data sets selecting linked branch lengths and found the average number of partitions supported by AICc to be 10, 15 for BIC, and 43 for BIC. She does not inform in her

thesis the minimum subset size that was allowed, so her results are perhaps not comparable to ours.

The number of partitions in C-ParFin4, 5, 6 and 7, O-ParFin3, 4, 5 and 6, and S-ParFin2 were at least 60% greater than the corresponding homoplasy scheme, or 8 partitions for the small *Scolecophoridae* matrix (28 characters) and over 18 for the two *Aves* data sets ([Supplementary Fig. S1](#) available on Dryad). These partitions invariably formed clusters of high similarity, indicating differential allocation of just a few characters among them. The only exception was O-ParFin6 that was very different from other schemes ([Supplementary Fig. S1b](#) available on Dryad). Oversplitting could probably be reduced by increasing the value of the *–min-subset-size* option, which was set to 1. Indeed, all highly partitioned schemes had partitions with a single character. There are no clear guidelines on how to select the minimum subset size, but increasing it may not necessarily lead to better partitioning schemes: all homoplasy schemes employed in this study also had at least one single-character partition and top models under this strategy strongly outperformed the best PF models (Table 3). Top PF models were based on schemes that had 5 partitions or less in the case of CEA and OZL, whereas homoplasy partitioning returned 8 and 14 partitions, respectively. Differences between homoplasy and PF schemes go beyond the number of partitions.

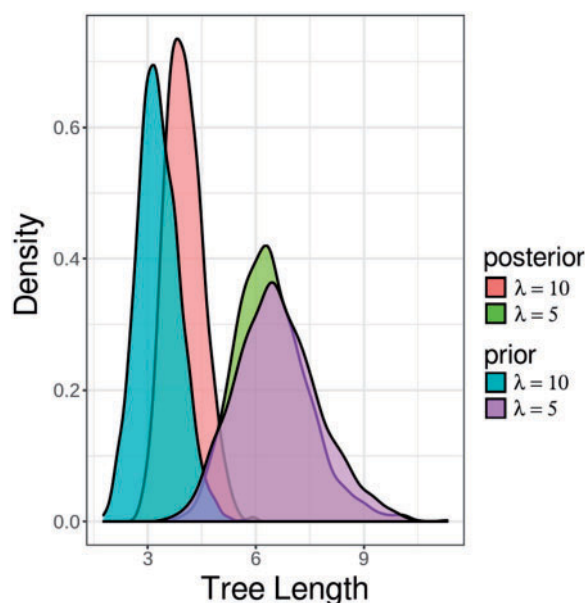


FIGURE 4. Prior and posterior tree length densities sampled via MCMC using the best (S-Hom-13, $\lambda=5$) and second best (S-Hom-14, $\lambda=10$) phylogenetic models, applied to the homoplasy partitioned *Scolecbythidae* (SCO) data set. The analysis employing S-Hom-14 shows that the prior distribution (cyan) was strongly updated by the data, for most of the posterior's density (pink) is shifted towards the prior's right tail. Setting the inverse scale parameter to five doubles the prior's average tree length (purple distribution), greatly improving its fit to the posterior (in green).

For instance C-ParFin3 and the scheme C-Hom (from which the best overall model resulted) both had 8 partitions but were very dissimilar (Supplementary Fig. S1a available on Dryad), demonstrating that these strategies are fundamentally different.

In the case of both CEA and OZL, PF partitioning very strongly improved model likelihood if compared with either unpartitioned or anatomically partitioned data sets (see Results). Our results contrast with those obtained using the preliminary version of PF2 by Wright (2015). When applied to Clarke and Middleton (2008) data set (CEA) with ordered characters, PF2 suggested various partition schemes, with the number of subsets varying from 2 to 55. Wright's comparisons between the partition schemes returned by PF2 with those of Clarke and Middleton showed improved MgL associated with the latter models. Wright (2015) then concluded that the scheme chosen by Clarke and Middleton was superior to those of PF2 and that all available biological information should be taken into account for partitioning of morphological data sets. The author used variable coding, unlinked and linked branch lengths, ACRV approximated by per-partition Γ distributions and $\lambda = 5$ in her analysis of Clarke and Middleton's data. Our MgL values obtained under these same set of parameters (C-Hom(Ord)-41 and C-Hom(Ord)-29) were compared with those provided by Wright (2015) (see Table 4). Independently of the MgL estimation method, Clarke and Middleton's anatomical schemes still have improved MgL values when compared with those of the

preliminary version of PF2. Perhaps an increase in the MgL associated with her PF2 partition schemes could result from a broader exploration of the parameter space. For example, our results show that for CAE better MgL values are associated with $\lambda = 10$ –20, instead of $\lambda = 5$.

Unlike the Aves data sets, PartitionFinder2 did not improve model fit to the smallest matrix (SCO, 28 characters) over the analysis of unpartitioned data and neither did the anatomical scheme (Table 3). Interestingly, this was the only data set in which the number of PF partitions exceeded homoplasy. These results suggest that model fit may be particularly sensitive to partitioning in smaller data sets.

Choice of Parameters

According to our results, effective partitioning strategies should result in very strong Bayesian preference for linked branch lengths and equal rates. This was the case in all top models, always recovered for homoplasy partitioned data sets. Our results also show that MrBayes default's value for the inverse scale parameter of the exponential prior distribution ($\lambda = 10$) resulted in best models in most analyses, regardless of the partitioning strategy (Table 3). An important exception was the top model for *Scolecbythidae* (S-Hom-13) that showed Bayesian preference for $\lambda = 5$. The exponential distribution mean is the reciprocal of λ , so when $\lambda = 10$, the mean is 0.1 and twice as large (0.2) when $\lambda = 5$ (Ronquist et al. 2009), which also makes the prior flatter and less informative. Thus, there may be two reasons that explain the preference for $\lambda = 5$ in the case of the smallest data matrix: either the data are informative and pull the chain towards longer branches or the data are simply not informative, and there is Bayesian preference for a vaguer prior.

To distinguish among these two hypotheses, we sampled the prior distribution of tree lengths using MrBayes' option *mcmc data=no*, and compared with the posteriors obtained using $\lambda = 5$ and $\lambda = 10$. Results show that the data are actually informative (Fig. 4). Most of the posterior mass for $\lambda = 10$ (cyan) is shifted to the right with respect to the prior, indicating positive preference for longer branches (Table 3). This is reinforced by the much better fit between prior and posterior when $\lambda = 5$, which explains the improved marginal likelihood.

We suspected that the longer branches in this and all other analyses employing the same strategy (see Results and Supplementary File SM4 available on Dryad) may be due to the segregation of the most homoplastic characters in its singleton partitions (see Supplementary Tables S11, S21 and S30 available on Dryad). The influence of these higher multipliers on branch lengths was likely exacerbated in the small *Scolecbythidae* matrix and hence the difference in λ choice.

We ran analyses with $\lambda=5$ and $\lambda=10$, after excluding partition 5 (which has a single character but is informative according to Fig. 2c) and fusing partitions 1 and 6 (respectively, the homoplasy-free and

parsimony-uninformative partitions; see first section of Discussion), and comparison of resulting MgL indicated positive Bayesian preference for $\lambda=10$ (KRS = 2.30), corroborating our initial hypothesis. This suggests that smaller values of λ may provide a better model fit in smaller matrices with high homoplastic characters. There is, however, no reason why any particular rate prior should be appropriate for all data sets. The preference for $\lambda=10$ may be peculiar to the tested matrices or simply because it makes the prior vaguer and hence more conservative. Evaluating alternative branch length priors is beyond the scope of this study, but one should also consider this possibility when applying homoplasy partitioning to his/her data set.

CONCLUSIONS

Use of homoplasy as a criterion for data partitioning is an efficient approach to explore discrete morphological characters under Bayesian inference. The models implementing homoplasy partitioning schemes showed significantly better fit for the three data sets analyzed than all other alternatives. They also resulted in increased resolution of the tree topologies, in particular for OZL, the largest and most complex matrix analyzed in this study. In the case of the smallest matrix (SCO) homoplasy was indeed the only strategy that effectively partitioned the data, resulting in superior model likelihood with improved tree resolution and branch support. This criterion is easily implemented, requiring no *a priori* premises, and can be applied to data sets irrespective of their numbers of characters. It is also more straightforward than an algorithmic alternative (PartitionFinder2) that has a much larger parameter space that must be surveyed by the user and whose schemes were outperformed in every comparison with homoplasy partitioning. Morphological data sets frequently contain less than 100 characters and our results show that a full analysis of a matrix twice as large (CEA) may be accomplished in a matter of hours. Thus, we believe homoplasy partitioning to be a promising approach in model-based phylogenetic analyses.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.1hh515h>.

FUNDING

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico [309641/2016-0; 140263/2015-2].

ACKNOWLEDGMENTS

We would like to thank Eduardo Almeida, Rodrigo Feitosa, Peter Foster, Luke Harmon, Mike Lee, April Wright, and an anonymous reviewer for comments on an earlier draft of the manuscript.

REFERENCES

- Brandley M.C., Schmitz A., Reeder T.W. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Syst. Biol.* 54:373–390.
- Brewer M.J., Butler A., Cooksley S.L. 2016. The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity. *Methods Ecol. Evol.* 7:679–692.
- Brown J.M., Lemmon A.R. 2007. The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. *Syst. Biol.* 56:643–655.
- Carpenter J.M. 1999. What do we know about chrysoid (Hymenoptera) relationships? *Zool. Scr.* 28:215–231.
- Clarke J.A., Middleton K.M. 2008. Mosaicism, modules, and the evolution of birds: results from a Bayesian approach to the study of morphological evolution using discrete character data. *Syst. Biol.* 57:185–201.
- Clarke J.A., Zhou Z., Zhang F. 2006. Insight into the evolution of avian flight from a new clade of Early Cretaceous ornithurines from China and the morphology of *Yixianornis grabaui*. *J. Anat.* 208:287–308.
- Dávalos L.M., Velazco P.M., Warsi O.M., Smits P.D., Simmons N.B. 2014. Integrating incomplete fossils by isolating conflicting signal in saturated and non-independent morphological characters. *Syst. Biol.* 63:582–600.
- Engel M.S., Grimaldi D. 2007. Cretaceous Scolebythidae (Hymenoptera) and phylogeny of the family. *Am. Mus. Novit.* 3568.
- Engel M.S., Ortega-Blanco J., McKellar R.C. 2013. New scolebythid wasps in Cretaceous amber from Spain and Canada, with implications for the phylogeny of the family (Hymenoptera: Scolebythidae). *Cretac. Res.* 46:31–42.
- Farris J.S. 1969. A successive approximations approach to character weighting. *Syst. Zool.* 18:374–385.
- Gavryushkina A., Heath T.A., Ksepka D.T., Stadler T., Welch D., Drummond A.J. 2017. Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* 66:57–73.
- Giribet G. 2015. Morphology should not be forgotten in the era of genomics—a phylogenetic perspective. *Zool. Anz.* 256:96–103.
- Goloboff P. 1993. Estimating character weights during tree search. *Cladistics.* 9:83–91.
- Goloboff P., Farris S., Nixon K. 2008. TNT, a free program for phylogenetic analysis. *Cladistics.* 24:774–786.
- Harrison L.B., Larsson H.C.E. 2015. Among-character rate variation distributions in phylogenetic analysis of discrete morphological characters. *Syst. Biol.* 64:307–324.
- Jukes T.H., Cantor C.R. 1969. Evolution of protein molecules. In: Munro H.N., editor. *Mammalian protein metabolism*. New York: Academic Press, p. 21–132.
- Kainer D., Lanfear R. 2015. The effects of partitioning on phylogenetic inference. *Mol. Biol. Evol.* 32:1611–1627.
- Kass R.E., Raftery A.E. 1995. Bayes factors. *J. Am. Stat. Assoc.* 18:773–795.
- King B., Qiao T., Lee M.S.Y., Zhu M., Long J.A. 2017. Bayesian morphological clock methods resurrect placoderm monophyly and reveal rapid early evolution in jawed vertebrates. *Syst. Biol.* 66:499–516.
- Kjer K.M., Honeycutt R.L. 2007. Site specific rates of mitochondrial genomes and the phylogeny of Eutheria. *BMC Evol. Biol.* 7:8. doi: 10.1186/1471-2148-7-8.
- Klopfstein S., Vilhelmsen L., Ronquist F. 2015. A nonstationary Markov model detects directional evolution in hymenopteran morphology. *Syst. Biol.* 64:1089–1103.
- Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2017. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34:772–773.
- Lee M.S.Y. 2016. Multiple morphological clocks and total-evidence tip-dating in mammals. *Biol. Lett.* 12:20160033.
- Lee M.S.Y., Palci A. 2015. Morphological phylogenetics in the genomic age. *Curr. Biol.* 25:R922–R929.
- Lee M.S.Y., Soubrier J., Edgecombe G.D. 2013. Rates of phenotypic and genomic evolution during the Cambrian explosion. *Curr. Biol.* 23:1889–1895.

- Lee M.S.Y., Cau A., Naish D., Dyke G.J. 2014. Morphological clocks in paleontology, and a mid-Cretaceous origin of crown Aves. *Syst. Biol.* 63:442–449.
- Lemmon A., Moriarty E. 2004. The importance of proper model assumption in Bayesian phylogenetics. *Syst. Biol.* 53:265–277.
- Lewis P.O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* 50:913–925.
- Lockhart G.B., Cheetham B.M.G. 1985. Basic digital signal processing: Butterworths basic series. London: Butterworth-Heinemann.
- Magwene, P. L. 2001. New tools for studying integration and modularity. *Evolution*. 55(9):1734–1745.
- Marshall D., Simon C., Buckley T. 2006. Accurate branch length estimation in partitioned Bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors. *Syst. Biol.* 55:993–1003.
- Miller M.A., Pfeiffer W., Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE)*, 14 Nov. 2010, New Orleans, LA pp 1–8.
- Misof B., Liu S., Meusemann K., Peters R.S., Donath A., Mayer C., Frandsen P.B., Ware J., Flouri T., Beutel R.G., Niehuis O., Petersen M., Izquierdo-Carrasco F., Wappler T., Rust J., Aberer A.J., Aspöck U., Aspöck H., Bartel D., Blanke A., Berger S., Böhm A., Buckley T.R., Calcott B., Chen J., Friedrich F., Fukui M., Fujita M., Greve C., Grobe P., Gu S., Huang Y., Jermini L.S., Kawahara A.Y., Krogmann L., Kubiak M., Lanfear R., Letsch H., Li Yiyuan Li Z., Li J., Lu H., Machida R., Mashimo Y., Kapli P., McKenna D.D., Meng G., Nakagaki Y., Navarrete-Heredia J.L., Ott M., Ou Y., Pass G., Podsiadlowski L., Pohl H., von Reumont B.M., Schütte K., Sekiya K., Shimizu S., Slipinski A., Stamatakis A., Song W., Su X., Szucsich N.U., Tan M., Tan X., Tang M., Tang J., Timelthaler G., Tomizuka S., Trautwein J., Tong X., Uchifune T., Walz M.G., Wiegmann B.M., Wilbrandt J., Wipfler B., Wong T.K.F., Wu Q., Wu G., Xie Y., Yang S., Yang Q., Yeates D.K., Yoshizawa K., Zhang Q., Zhang R., Zhang W., Zhang Yunhui Zhao J., Zhou C., Zhou L., Ziesmann T., Zou S., Li Y., Xu X., Zhang Y., Yang H., Wang J., Wang J., Kjer K.M., Zhou X. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science*. 346:763–767.
- Nylander J.A.A., Ronquist F., Huelsenbeck J.H., Nieves-Aldrey J.L. 2004. Bayesian phylogenetic analysis of combined data. *Syst. Biol.* 53:47–67.
- Nylander J.A.A., Wilgenbusch J.C., Warren D.L., Swofford D.L. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*. 24:581–583.
- O'Connor J.K., Zhou Z. 2012. A redescription of *Chaoyangia beishanensis* (Aves) and a comprehensive phylogeny of Mesozoic birds. *J. Syst. Palaeontol.* 2019:1–18.
- Ohio Supercomputer Center. 1987. Columbus OH: Ohio Supercomputer Center. Available from: <http://osc.edu/ark:/19495/f5s1ph73/>.
- Philippe H., Söhrhann U., Baroin A., Perasso R., Gasse F., Adoutte A. 1994. Comparison of molecular and paleontological data in diatoms suggests a major gap in the fossil record. *J. Evolution Biol.* 7:247–65.
- Pyron R.A. 2011. Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Syst. Biol.* 60:466–481.
- Quental T.B., Marshall C.R. 2010. Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol. Evol.* 25:435–441.
- Rambaut A., Suchard M., Drummond A. 2013. Tracer 1.6. Available from: <http://tree.bio.ed.ac.uk/software/tracer/>.
- Ronquist F., Huelsenbeck J., Teslenko M. 2011. MrBayes Version 3.2 Manual: tutorials and model summaries. Available from: http://mrbayes.sourceforge.net/mb3.2_manual.pdf.
- Ronquist F., Klopfstein S., Vilhelmsen L., Schulmeister S., Murray D.L., Rasnitsyn A.P. 2012a. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Syst. Biol.* 61:973–999.
- Ronquist, F., Van Der Mark P., Huelsenbeck J. 2009. Bayesian phylogenetic analysis using MrBayes. In: Lemey P., Salemi M., Vandamme A.M., editors. *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. New York: Cambridge University Press, p. 210–66.
- Ronquist F., Teslenko M., Van Der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012b. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539–542.
- Tarasov S., Génier F. 2015. Innovative Bayesian and parsimony phylogeny of dung beetles (Coleoptera, Scarabaeidae, Scarabaeinae) enhanced by ontology-based partitioning of morphological characters. *PLoS One*. 10(3): e0116671.
- Wagner P.J. 2012. Modelling rate distributions using character compatibility: implications for morphological evolution among fossil invertebrates. *Biol. Lett.* 8:143–146.
- Wiens J.J., Kuczyński C.A., Townsend T., Reeder T.W., Mulcahy D.G., Sites J.W. 2010. Combining phylogenomics and fossils in higher-level squamate reptile phylogeny: molecular data change the placement of fossil taxa. *Syst. Biol.* 59:674–688.
- Wilgenbusch J.C., Warren D.L., Swofford D.L. 2004. AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. Available from: <http://ceb.csit.fsu.edu/awty>.
- Wright A.M. 2015. Estimating phylogenetic trees from discrete morphological data [PhD Dissertation]. University of Texas at Austin.
- Wright A.M., Lloyd G.T., Hillis D.M. 2015. Modeling character change heterogeneity in phylogenetic analyses of morphology through the use of priors. *Syst. Biol.* 65:602–611.
- Xie W., Lewis P.O., Fan Y., Kuo L., Chen M.H. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Syst. Biol.* 60:150–160.
- Yang Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol. Biol. Evol.* 10:1396–1401.