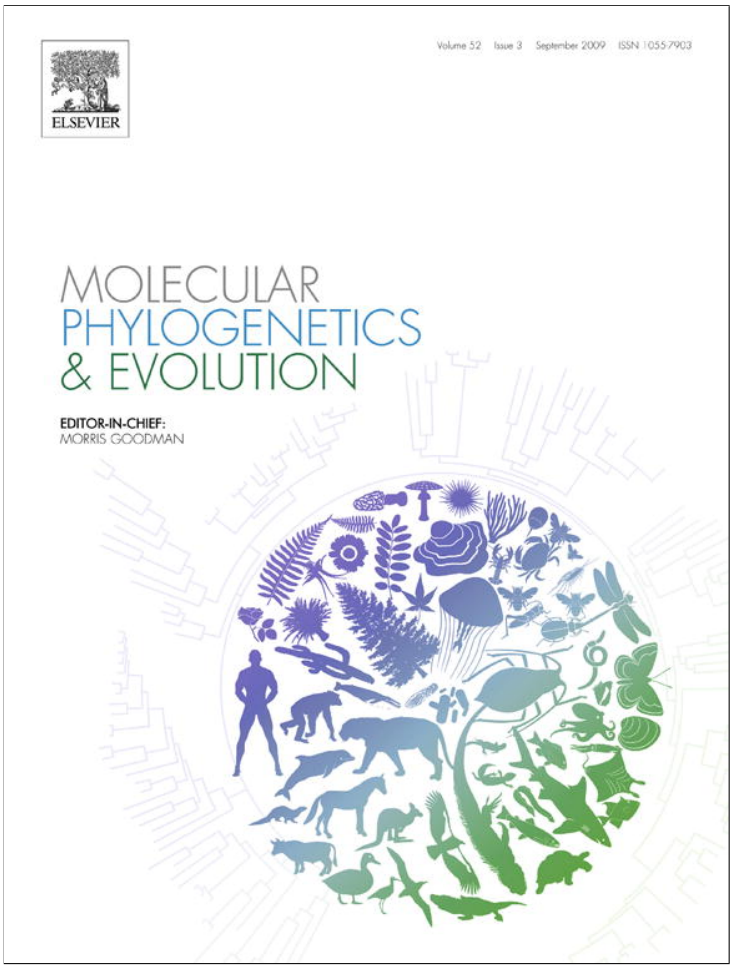

Molecular Phylogenetics and Evolution 52 (2009) 715–726

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Molecular dating and biogeography of fig-pollinating wasps

Carlos Lopez-Vaamonde [a,*], Niklas Wikström [b], Karl M. Kjer [c], George D. Weiblen [d], Jean Yves Rasplus [e], Carlos A. Machado [f], James M. Cook [g]

[a] INRA, UR 633 Zoologie Forestière, F-45075 Orléans, France
[b] Department of Systematic Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, Sweden
[c] Rutgers University, Cook College, Department of Ecology, Evolution and Natural Resources, 14 College Farm Rd., New Brunswick, New Jersey 08901, USA
[d] Bell Museum of Natural History and Department of Plant Biology, University of Minnesota, 1445 Gortner Avenue, Saint Paul, Minnesota 55108, USA
[e] Institut National de la Recherche Agronomique, Centre de Biologie et de Gestion des Populations, Campus International de Baillarguet-CS 30 016, 34988 Montferrier-sur Lez, France
[f] Department of Biology, University of Maryland, 1210 Biology-Psychology Building, College Park, MD 20742, USA
[g] School of Biological Sciences, University of Reading, Reading RG6 6AS, UK

ARTICLE INFO

Article history:
Received 14 October 2008
Revised 5 May 2009
Accepted 26 May 2009
Available online 13 June 2009

Keywords:
Agaonidae
Bayesian molecular clock
Divergence times
Dominican amber
Ficus
Plant–insect interactions

ABSTRACT

Figs and fig-pollinating wasps are obligate mutualists that have coevolved for over 60 million years. But when and where did pollinating fig wasps (Agaonidae) originate? Some studies suggest that agaonids arose in the Late Cretaceous and the current distribution of fig-wasp faunas can be explained by the break-up of the Gondwanan landmass. However, recent molecular-dating studies suggest divergence time estimates that are inconsistent with the Gondwanan vicariance hypothesis and imply that long distance oceanic dispersal could have been an important process for explaining the current distribution of both figs and fig wasps. Here, we use a combination of phylogenetic and biogeographical data to infer the age, the major period of diversification, and the geographic origin of pollinating fig wasps. Age estimates ranged widely depending on the molecular-dating method used and even when using the same method but with slightly different constraints, making it difficult to assess with certainty a Gondwanan origin of agaonids. The reconstruction of ancestral areas suggests that the most recent common ancestor of all extant fig-pollinating wasps was most likely Asian, although a southern Gondwana origin cannot be rejected. Our analysis also suggests that dispersal has played a more important role in the development of the fig-wasp biota than previously assumed.

1. Introduction

The obligate mutualistic interactions between figs (*Ficus* spp.) and their pollinating wasps, Agaonidae *sensu* (Walker, 1846), have made them classic model organisms for studying coadaptation (Jousselin et al., 2003; Weiblen, 2004) and cospeciation (Lopez-Vaamonde et al., 2001; Machado et al., 2001, 2005; Weiblen and Bush, 2002). In addition, they represent an extremely impressive co-radiation of insects and plants leading to over 800 extant fig species with associated wasps and these mutualistic systems play important ecological roles in many tropical habitats. However, the age, geographic origin and major period of diversification of figs and fig wasps remain controversial. Here, we briefly review existing hypotheses and evidence on these issues, before presenting a major new dataset and accompanying analyses in our paper.

With respect to age, molecular-dating analyses first indicated that agaonids, which form a monophyletic group (Rasplus et al., 1998), may have originated in the Late Cretaceous, 75–100 million years ago (MYA) (Machado et al., 2001). Likewise, independent analyses of the origin of genus *Ficus* suggest that figs are both monophyletic and ancient, with a minimum age estimate of 83 MYA (Datwyler and Weiblen, 2004). These results indicate that fig and fig-wasp radiations may have occurred contemporaneously, as later shown by Rønsted et al. (2005). However, divergence times from fig wasps are estimates of the crown group ages whereas fig age estimates refer to the stem lineage ages (Datwyler and Weiblen, 2004). Thus, although both figs and fig wasps appear to have a Cretaceous origin, the crown radiation of *Ficus* may have occurred more recently than suggested by Machado et al. (2001), as later suggested by Zerega et al. (2005), who conducted molecular dating using a Bayesian approach that gave a range of 40–51 MYA for the crown group age of *Ficus*. Zerega et al.'s results do not support the hypothesis of simultaneous diversification between figs and fig wasps and suggest that figs may have radiated more recently, during the Tertiary, well after the break-up of Gondwana.

* Corresponding author. Address: Institut National de la Recherche Agronomique (INRA) – Centre d'Orléans, Unité de Zoologie Forestière, 2163 Avenue de la Pomme de Pin, CS 40001 Ardon, 45075 ORLEANS Cedex 2, France. Fax: +33 2 38 41 78 79.
E-mail address: carlos.lopez-vaamonde@orleans.inra.fr (C. Lopez-Vaamonde).

1055-7903/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.05.028

The discrepancy between the estimated ages of origin of figs and their pollinators could be due to methodological biases. On the one hand, the methods used by Machado et al. (2001) required the use of ultrameric trees and thus the pruning of several taxa that violated the constant substitution rate assumption. Newer methods that allow the inclusion of lineages with different substitution rates (Sanderson, 2002) have since been developed and new analyses of the same data could produce younger divergence times. Indeed, a more recent analysis (Rønsted et al., 2005) suggests a Late Cretaceous–Early Paleocene origin of the fig-wasp mutualism (stem lineage age of figs and fig-pollinating wasps of 98–105 and 66–101 MYA, respectively), and a younger estimate of crown group age for the wasps (51–78 MYA) using the same mtDNA data. On the other hand, it is also likely that the younger estimate for the crown group age of *Ficus* is the result of the age constraints used by Zerega et al. (2005) in their calibrations, where the age of the oldest known fossil fig (55 MYA) (Collinson, 1989) was used to define the minimum age constraint for the crown age of *Ficus*.

With respect to geographical origin, it has been proposed that agaonids evolved in the southern hemisphere/Gondwana (Murray, 1985). This hypothesis is supported by phylogenetic analyses, which show that wasps from major biogeographical regions split in a chronological order that is congruent with the break-up of the Gondwanan landmass (Machado et al., 2001). The hypothesis of a Gondwanan origin of fig-pollinating wasps is further supported by the fact that most figs and fig wasps show a southern tropical distribution, and that early-diverging lineages of extant figs and pollinators are South American (West Gondwanan origin) (Machado et al., 2001; Rønsted et al., 2005). However, Rønsted et al.'s (2005) estimated date for the origin of agaonids 83.33 ± 17.61 MYA (65–100 MYA) is not consistent with the hypothesis that the fig-wasp mutualism arose before the fragmentation of the southern Gondwana landmass during the Late Cretaceous (Machado et al., 2001). Consequently, it has been suggested that long distance oceanic dispersal could also be an important process in explaining the present distribution of figs and fig wasps (Datwyler and Weiblen, 2004; Rønsted et al., 2005).

The aim of this study is to reconstruct the phylogeny and estimate the age and area of origin of agaonids using a substantially enlarged and improved molecular dataset, including nuclear sequence data for the first time. Our phylogeny is based on the largest sequence data set gathered for this group so far, including mitochondrial (COI) and nuclear (28S) DNA sequences for most of the extant genera. We use the phylogeny to estimate divergence times using three alternative methods. Relative ages of the fig wasps are converted into absolute dates using information from two amber fossils (Penalver et al., 2006). The absolute ages of the fig wasps are used to re-evaluate the hypothesis that fig-pollinating wasps have a Gondwanan origin and that their extant distribution is the result of vicariance due to the break-up of Gondwana. In addition, we infer ancestral areas using both Bayesian and maximum likelihood approaches to identify the most likely area of origin for agaonids.

2. Materials and methods

2.1. Taxon sampling

We obtained DNA sequences for 64 species from 15 out of 20 known genera of fig pollinators (see Supplementary data, Appendix 1). Of these, 57 (28S: 6; COI: 51) were published previously and 73 (28S: 59; COI: 14) sequences are reported here for the first time. *Anaphes nitens* (Mymaridae) was used as outgroup. Mymarids have long been considered to be the sister group of the rest of the Chalcidoidea (Gibson et al., 1999), however recent phylogenetic analysis failed to support monophyly of Chalcidoidea excluding Mymaridae (Gibson et al., 1999). Nevertheless, in absence of a strong phylogenetic hypothesis for Chalcidoidea, we used a mymarid as an outgroup since this family is still considered one of the most basal groups of Chalcidoidea (Campbell et al., 2000). Voucher fig-wasp specimens are deposited at INRA Orléans, INRA Montpellier, the Natural History Museum, London and the Museum of Comparative Zoology, Harvard University.

2.2. DNA sequencing

Phylogenetic relationships were estimated using sequence data from two markers, portions of the mitochondrial cytochrome oxidase I gene (COI) and the 28S nuclear large ribosomal subunit (28S).

For some species, total genomic DNA was extracted from single individuals using 50 μl of an extraction buffer containing 2% Chelex 100 Resin (Bio-Rad). For others we used aQIAamp® Tissue Kit (QIAGEN Inc.) and an extraction protocol modified as described in Weiblen (2001). PCR and sequencing procedures for 28S and COI using primers and protocols have been previously described in Lopez-Vaamonde et al. (2001) and Weiblen (2001), respectively.

2.3. Sequence alignment and phylogeny estimation

COI sequences were very similar in length and easily aligned by eye using the codon structure. On the other hand, 28S sequences showed substantial unaligned length variation and were aligned following two approaches:

Firstly, all 28S sequences were aligned with Clustal X using the default setting (open gap penalty = 10, gap extend = 5, transition weight = 0.5, delay divergent = 40). Misaligned fragments were corrected by eye by the first author using MacClade version 4.02 (Maddison and Maddison, 2001). We will refer to this manual alignment as "by eye" alignment.

Secondly, we used a manual alignment based on secondary structure. Initial 28S alignments were made using Clustal and the resulting files were then aligned manually in Microsoft Word using the structural methods described in Kjer (1995), and Kjer et al. (2007, 2009) and secondary structure models of 28S rRNA in Chalcidoid wasps, based on Gillespie et al. (2005). We will refer to this manual alignment as "structural". We excluded 368 bp across regions of rDNA alignments where positional homology could not be established using structural criteria, including regions of alignment ambiguity. Alignment ambiguous regions were defined as single stranded length heterogeneous regions flanked by hydrogen-bonded stem regions that were also lacking in sequence motifs that could be identified across taxa (Kjer et al., 2009). We used the Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa, 1999; Goldman et al., 2000), implemented in PAUP* 4.0b8a to compare topologies derived from the two manual alignments. SH tests showed that topologies derived from 28S "by eye" and structural alignments were significantly different ($P = 0.0001$). We therefore decided to use both alignments in further analyses (see dating and biogeographical sections below). All DNA alignments are available from TreeBASE (http://treebase.org/treebase/).

We used Markov Chain Monte Carlo (MCMC) methods (Larget and Simon, 1999) within a Bayesian framework to estimate the posterior probabilities of the phylogenetic trees. Adopting a Bayesian approach allowed us to account for phylogenetic uncertainty, both with respect to topology and to branch lengths, in our estimates of divergence times by conducting all subsequent analyses on the posterior distribution of trees and parameter estimates (Huelsenbeck et al., 2000; Pagel and Lutzoni, 2002). Analyses were conducted using MrBayes 3.0B5 (Huelsenbeck et al., 2001) on a two-gene combined data set including a total of 2137 nucleotide characters (824 characters from mitochondrial COI;

1313 characters from nuclear 28S). Two sets of analyses were conducted implementing alternative partitioning of the data. The first set split the data into two partitions, corresponding to the two gene regions. Among partition rate variability was modeled in these analyses by setting the rate prior to variable (*prset ratepr = variable*). The second set of analyses combined the two regions into a single partition. MrModelTest 2.0 (Nylander, 2004), a simplified version of Modeltest 3.5 (Posada and Buckley, 2004; Posada and Crandall, 1998), was used to select the model of nucleotide substitution for each partition. A general time reversible model (GTR + Γ + I) allowing for rate heterogeneity across sites, assuming a discrete gamma distribution for each partition, and for a proportion of sites to be invariable, was selected for each partition and analysis. The same model was selected using either the Akaike information criterion (AIC) or the hierarchical likelihood ratio test (hLRT) as implemented in MrModelTest 2.0.

In both sets of analyses, two independent runs were conducted, each comprising four chains for 10,000,000 generations. Trees were sampled every 1000th generation yielding a total of 10,000 trees and parameter estimates in each of the two posterior distributions. Convergence between the two independent analyses was evaluated and confirmed. Posterior probability values reported (Fig. 1) are from the unpartitioned set of analyses and were averaged over the two independent analyses.

2.4. Estimating divergence times

Node ages for the fig wasps were estimated based on the combined data set (28S "by eye" alignment + COI) using four alternative methods: (i) nonparametric rate smoothing (NPRS) (Sanderson, 1997); (ii) penalized likelihood (PL) (Sanderson, 2002), as implemented in the computer program r8s ver. 1.70 (Sanderson, 2003); (iii) a model-based Bayesian implementation of rate autocorrelation (Thorne et al., 1998) as implemented in the program Multidivtime ver. 9/25/03 (Kishino et al., 2001; Thorne and Kishino, 2002); (iv) a relaxed Bayesian molecular clock with uncorrelated rates in BEAST v1.4.8 (Drummond and Rambaut, 2007). To address the effect of alignment uncertainty on our results, we used the 28S "structural" alignment in combination with COI to calculate age estimates using BEAST. In addition, we ran dating analyses from partitioned Mrbayes runs and compared results with those based on unpartitioned analyses. Age estimates between partitioned and unpartitioned data were very similar so, for the purpose of comparing the methods, all analyses were based on the results from the unpartitioned phylogenetic analyses using the 28S "by eye" alignment combined with COI. This effectively controlled for differences resulting from the fact that Multidivtime can, whereas r8s cannot, deal with multiple partitions and their associated branch lengths. Divergence times were estimated for all nodes with at least 95% posterior probability. To account for topological and branch length uncertainties, 100 trees and parameter estimates, filtered to include all nodes supported by 95% Bayesian Posterior probability (pp) or more, were randomly drawn from the posterior distribution and used as input into all subsequent analyses. All age distributions were tested for normality. Some proved significantly non-normal (Kolmogorov-Smirnov test in SPSS for Macintosh v 10.0). Therefore, for each supported node the distribution of divergence times across the 100 trees, conditional on the phylogenetic model (GTR + Γ + I) and the calibration point, were obtained by local density estimation using the program LOCFIT (Loader, 1999), implemented in the "R" statistical package (Ihaka and Gentlemen, 1996), for a similar approach (see Lopez-Vaamonde et al., (2006). To summarize the fitted distributions, we report their modes and 95% upper and lower Highest Posterior Density (UHPD and LHPD) limits (Table 1). The mode represents the most likely divergence time under the specified model and calibration, and the HPD limits the confidence interval for this estimate. In the r8s analysis (both PL and NPRS) our estimated error bars result from the analyses of 100 trees, and the LHPD–UHPD are the 90% HPD value. Notice that the posterior distribution of trees was filtered so that all nodes with 95 or higher BPP were present in all 100 "random" trees. In the Multidivtime analysis, we also use the 90% HPD values, but here the errors are the mode of the variability across the 100 trees ± mode of SD across the 100 trees. In the BEAST analysis, the error estimate results from the posterior distribution (5000 trees), and the LHPD–UHPD are here the 95% HPD value.

2.4.1. r8s analysis: NPRS and PL

In penalized likelihood analyses using r8s the relative contribution of a parameter-rich model and a numerical penalty (introduced to avoid extensive rate variation among nearby branches) into the estimated rates and divergence times is regulated by a smoothing parameter λ. The NPRS method described by Sanderson (1997) and the maximum likelihood clock model, outlined by Langley and Fitch (1974) could be seen as logical extremes on the scale of different smoothing values used in a penalized likelihood analysis (Sanderson, 2002). At one extreme, when little smoothing is enforced ($\lambda \to 0$) the age estimates correspond to those from an NPRS analysis. At the other extreme, when we enforce considerable rate smoothing ($\lambda \to \infty$), age estimates converge towards those obtained by using a clock model (Sanderson, 2002). In penalized likelihood analysis, the optimal smoothing value for a particular tree and branch length is estimated through a cross-validation procedure (Sanderson, 2002). In our penalized likelihood analyses, we conducted such cross-validation analyses on each of our 100 randomly drawn trees to obtain optimal smoothing values for each tree. Using a fixed root age of 1, a log10 transform of the optimal values ranged from −2.5 to −0.5, and the individual optimal values were subsequently set for each tree and used in the analyses. In our NPRS analyses, the log10 value of the smoothing parameter was arbitrarily set to −4 in all analyses. This low value yields results that correspond to those from a standard NPRS analysis but at the same time allows us to use the considerably faster TN algorithm in the r8s program (Sanderson, 2003).

To prevent the r8s age estimation algorithm from converging on local optima, each analysis was started at five different initial time estimates (num_time_estimates = 5) and the local stability of solutions was checked by perturbing and restarting the analysis five times (num_restarts = 5). To obtain absolute age estimates for supported nodes, the resulting ultrameric trees were calibrated by fixing the crown group *Pegoscapus* at 30 MYA based on information from the fossil record (see below).

2.4.2. Multidivtime analysis

In the Bayesian approach for estimating divergence times, explicit models are used to introduce rate autocorrelation among closely related branches. Conceptually, rate autocorrelation is also used in the approaches developed by Sanderson (2002, 2003) but instead of being modeled, it is introduced as a numerical penalty against rapid changes in rates between ancestral branches and their descendants. In Sanderson's approaches, the relative closeness of two taxa is strictly associated with their topological placement on the tree (like in any other optimization technique), but in the model-based approach, closeness will also depend on the branch duration parameter (time since they shared a common ancestor). This seems biologically reasonable in that sister taxa recently diverged from their common ancestor are assumed to show greater similarity (in evolutionary rates) than sister taxa that diverged a long time ago.

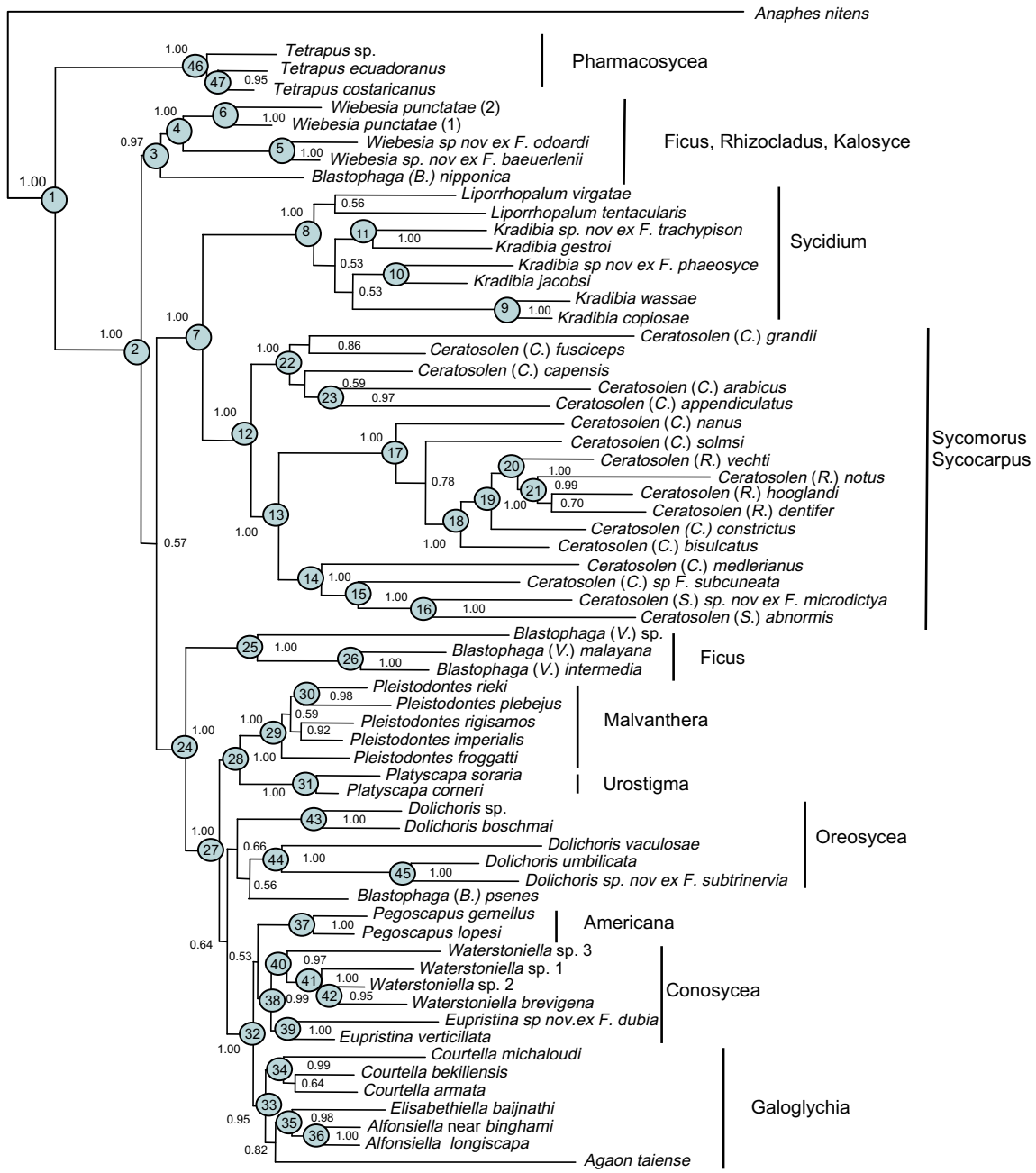

Fig. 1. Majority rule consensus of 100 randomly chosen trees reconstructed by Bayesian inference using the 28S "by eye" alignment + COI. Trees have been filtered to all include the nodes supported by 95% pp or more. Therefore, the consensus tree includes the uncertainty below 95%, but all the nodes supported by 95% or more in the overall analysis are present in 100% of the 100 random trees. Branch lengths are shown proportional to the amount of change along the branches assuming the GTR + G + I model of evolution. Posterior probability values are listed above nodes. Fig host associations are shown on the right of the cladogram. Nodes are numbered from 1 to 47 and their stem group age estimates are reported in Table 1.

Multidivtime analyses were conducted on the same 100 trees drawn at random from the posterior distribution as our r8s analyses. The prior distribution for the rate at the ingroup root node for all analyses was a gamma distribution with a mean of 0.35 and a standard deviation (SD) of 0.18 changes per site and unit time (rtrate = 0.35; rtratesd = 0.18), where one time unit represents 100 million years. The mean value (0.35 changes per site and unit time) was obtained by averaging the mean rates reported in the 100 r8s analyses. Technically, this is a violation of the definition of a prior probability, but in order to keep the prior reasonably unconstrained, a comparatively large SD value (0.18) was specified. Additional priors specified were rttm = 1 (a priori expected number of time units between tip and root); rttmsd = 1 (standard deviation of prior for time between tip and root); brownmean = 1.0 (mean of prior for brownian motion constant "nu"); brownsd = 1.0 (standard deviation of prior for brownian motion constant "nu").

Two series of analyses (both over all the 100 random trees) were done, introducing variable amounts of fossil-based information into the analyses (see Table 1): (i) the first series (multidiv1) mirrors as closely as possible the r8s analyses by more or less fixing crown group *Pegoscapus* (node 36, Fig. 1) at 30 MYA. (ii) The second series (multidiv2) of analyses are perhaps more realistic in that they, in a better way, account for the uncertainty concerning the dating of Dominican amber (see below for details). Here,

Table 1
Age estimates (MYA) using three methods: penalized likelihood and NPRS as implemented in r8s and Bayesian inference as implemented in Multidivtime. The latter is divided into two analyses using different calibration points. Multidivtime 1: *Pegoscapus* (node 36) was fixed at an age of 30 MYA; Multidivtime 2: *Pegoscapus* was constrained at a minimum age of 15 MYA and a maximum age of 45 MYA, and crown group *Tetrapus* (node 45) was constrained at a minimum age of 30 MYA; BEAST: *Pegoscapus* was fixed at an age of 30 MYA. Nodes are numbered from 1 to 46 indicated on the cladogram (see Fig. 1). UHPD and LHPD are the upper and lower Highest Posterior Densities, respectively.

		r8s (PL)		r8s (NPRS)		BEAST (by eye)		BEAST (structural)		Multidivtime 1				Multidivtime 2			
		Age (MYA)		Age (MYA)		Age (MYA)		Age (MYA)		Age (MYA)		SD (MYA)		Age (MYA)		SD (MYA)	
Node	Clade	Mode	LHPD–UHPD	Mode	LHPD–UHPD	Mode	LHPD–UHPD	Mode	LHPD–UHPD	Mode	LHPD–UHPD	Mode	LHPD–UHPD	Mode	LHPD–UHPD	Mode	LHPD–UHPD
1	Agaonidae	149	130–216	98	84–130	146	83–209	143	91–192	113	105–121	21	19–23	111	107–115	29	27–31
2		128	108–176	87	71–110	114	74–172	122	79–162	92	86–98	16	14–17	90	87–93	24	22–25
3	*Wiebesia*	95	79–134	78	60–105	66	45–138			78	75–90	14	13–16	77	75–86	21	19–24
4		76	64–111	69	53–89	49	31–112	104	63–140	69	66–81	13	12–15	69	67–77	19	18–22
5		25	18–34	22	16–33	28	7–37	34	13–66	22	21–27	7	6–8	22	21–25	8	7–9
6		44	31–59	42	31–58	29	10–61	76	23–106	44	42–51	10	9–12	44	42–49	13	12–15
7		113	94–160	84	62–99	100	63–151	97	63–136	77	71–83	14	12–15	75	72–79	20	19–22
8	Pollinators of *Ficus* section Sycidium	87	60–106	49	37–66	72	45–113	48	39–105	51	44–55	10	9–11	49	45–53	14	13–16
9		17	11–27	9	6–16	26	7–36	28	11–47	14	7–15	5	2–5	14	7–14	5	3–6
10		38	29–56	22	17–32	30	14–56	20	8–44	25	23–28	6	6–7	25	23–27	8	7–9
11		53	34–70	29	21–40	34	18–60	18	15–51	30	27–32	7	6–8	30	28–31	9	9–10
12	*Ceratosolen*	104	84–145	71	54–86	93	57–135	85	52–122	66	61–72	12	11–13	65	62–68	18	16–19
13		100	79–137	66	49–81	83	52–125	80	46–114	61	57–67	12	10–12	60	58–63	17	15–18
14		85	70–125	52	41–69	75	39–105	70	35–92	53	49–58	11	9–11	52	50–55	15	14–16
15		78	59–111	43	33–61	45	32–90	–	–	45	41–49	10	8–10	45	42–46	13	12–14
16		64	43–83	33	24–44	35	20–66	53	18–69	33	30–36	8	7–9	32	31–34	10	9–11
17		62	54–93	38	29–51	66	38–96	50	35–81	40	35–44	9	8–9	39	35–42	12	10–13
18		55	38–74	27	21–38	57	27–75	29	19–50	28	26–32	7	6–7	28	26–30	9	8–9
19		41	31–63	21	17–32	39	24–64	–	–	23	21–27	6	5–6	23	22–25	8	7–8
20		43	28–53	17	14–24	33	18–53	–	–	20	18–22	5	5–6	20	19–21	6	6–7
21		40	26–50	16	12–22	28	14–46	22	13–41	18	15–20	5	4–5	17	16–18	6	5–6
22		110	73–128	62	48–77	71	39–111	62	42–109	55	51–62	11	9–12	55	51–59	15	14–17
23		77	61–111	48	36–64	66	30–90	41	21–86	43	37–50	9	8–10	41	39–48	12	11–14
24		127	90–153	86	67–101	106	62–147	86	72–142	79	78–86	14	13–15	79	77–83	20	20–22
25	*Blastophaga (Valisia)*	89	70–121	64	48–80	73	35–113	75	44–94	64	59–66	12	11–13	61	58–64	17	16–18
26		41	29–50	26	21–36	34	13–54	52	19–63	31	28–32	8	7–9	29	28–31	10	9–11
27		97	76–123	77	60–90	73	47–109	–	–	69	68–75	12	10–12	69	66–73	18	17–20
28		72	58–94	64	48–79	66	33–87	–	–	64	54–66	10	9–11	60	53–64	17	14–17
29	*Pleistodontes*	44	34–56	42	32–54	31	17–51	43	23–70	44	37–46	9	7–9	43	37–45	13	10–13
30		34	28–45	30	25–44	23	11–41	38	18–62	33	29–39	8	7–9	33	28–38	10	9–12
31	*Platyscapa*	25	18–31	25	18–32	31	5–35	26	9–44	24	21–26	7	5–7	24	20–25	8	7–9
32		76	63–97	59	52–81	58	38–85	–	–	59	56–60	9	8–9	58	54–59	15	14–15
33	Pollinators of *Ficus* section Galoglychia	75	58–90	54	48–74	48	31–74	–	–	53	49–54	8	7–8	50	48–53	14	13–14
34	*Courtella*	58	43–70	49	39–64	29	21–63	27	20–54	44	41–47	7	7–8	44	40–46	12	11–13
35	*Elisabethiella*	52	39–67	46	39–62	36	20–56			43	40–44	7	6–7	41	39–43	11	11–12
36		25	19–38	31	21–45	20	7–35	18	9–41	27	25–28	6	5–6	26	25–28	8	8–9
37	*Pegoscapus*	30	–	30	–	30	–	30	–	30	–	–	–	28	27–30	7	7–8
38	Pollinators of *Ficus* section Conosycea	71	55–84	58	46–71	51	31–74	–	–	52	49–54	8	7–8	51	48–52	13	13–14
39		52	43–69	49	38–62	36	18–59	40	22–57	43	40–44	7	7–8	42	39–43	11	11–12
40		65	49–76	48	41–65	50	26–65	–	–	46	43–47	7	7–8	45	42–46	12	11–13
41		43	32–51	40	29–48	22	15–45	–	–	32	30–33	6	6–6	31	29–32	9	9–9
42		33	24–42	31	22–42	17	11–37	24	10–42	27	25–27	5	5–6	25	24–27	8	7–8
43	*Dolichoris* new Caledonia	42	31–55	30	25–43	21	16–57	40	15–64	33	29–34	7	6–7	31	29–33	10	9–10
44	*Dolichoris*	80	64–103	58	47–72	51	30–85	58	46–95	54	49–56	10	9–10	52	49–55	15	13–15
45		33	26–48	23	18–31	30	10–47	37	14–51	22	19–23	5	5–6	21	19–22	7	6–7
46	*Tetrapus* pollinators of *Ficus* section Pharmacosycea	38	29–49	30	23–45	22	12–47	29	15–57	47	43–49	13	12–15	48	46–48	15	13–15
47		28	21–39	25	17–36	21	8–36	25	12–47	36	35–38	12	11–13	38	37–39	13	12–13

the crown group *Pegoscapus* was constrained at a minimum age of 15 MYA and a maximum age of 45 MYA, and the crown group *Tetrapus* (node 45, Fig. 1) was constrained at a minimum age of 30 MYA.

For each supported node the mean, standard deviation, and 95% error limits of the estimated posterior distribution of divergence times, conditioned on the phylogenetic model and prior parameters, were obtained for all 100 random trees. Across the 100 random trees the estimates of both mean and standard deviation were not always normally distributed, so we report the mode for each supported node (Table 1).

2.4.3. BEAST analysis

Divergence times were estimated with phylogeny using the program BEAST, which does not assume that substitution rates are autocorrelated across the tree and estimates branch lengths,

topologies, substitution model parameters and dates simultaneously. We ran two analyses using two different data sets: 28S "by eye" alignment + COI and 28S structural alignment + COI. For each analysis, we used a constant-rate Yule (speciation process) prior, and all other priors and operators were the default settings. Four independent chains were run for 100,000,000 generations. Trees and parameters were sampled every 1000th generation during the last 10,000,000 generations yielding a total of 10,000 trees. Mean parameter estimates and 95% highest posterior densities (HPDs) were determined through analyzing the combined BEAST tree files in TreeAnnotator 1.4.8 (Drummond and Rambaut, 2007).

2.4.4. Fossil record and calibration points

There are several fig-pollinating wasp fossils reported in the literature (Penalver et al., 2006; Poinar, 1993). Numerous agaonid female specimens have been discovered in Dominican amber and they all belong to the Neotropical genera *Tetrapus* and *Pegoscapus* (Penalver et al., 2006). There is one specimen from the Florissant, Colorado, suggested to belong to the genus *Tetrapus* (Brues, 1910; Weiblen, 2002), but critical reappraisal of this specimen (Fig. 2a) leads us to conclude that it is not an agaonid (Kjellberg et al., 2005).

To convert the relative ages obtained through the r8s, Multidivtime and BEAST analyses into dates, we used fossil-based information from seven specimens preserved in Dominican amber from George Poinar's collection. Based on the adult morphology, we were able to compare these fossils to extant species and assign them to extant genera, *Pegoscapus* (Fig. 2b) and *Tetrapus* (Fig. 2c and d) and use them as calibration points in our dating analyses.

However, the dating of Dominican amber remains controversial, with the youngest proposed age of 15–20 MYA based on Foraminifera (Iturralde-Vinent and MacPhee, 1996) and the oldest of 30–45 MYA based on coccoliths (Schlee, 1990). We took into account this uncertainty using several calibration points: (i) For the r8s analysis, we used the mid value of this range (30 MYA) as a fixed calibration point for the *Pegoscapus* crown group (node 36). (ii) For the Multidivtime analysis, the fossil-based information was used in two alternative analyses: For comparison, the first series mirrors as closely as possible the r8s analyses by more or less fixing the crown group *Pegoscapus* at 30 MYA through a minimum age constraint at 29 MYA and a maximum age constraint at 31 MYA. The second series of analyses accounts for the uncertainty of the Dominican amber age by specifying less rigorous constraints. Here, the *Pegoscapus* crown group (node 36) is constrained at a minimum age of 15 MYA and a maximum age of 45 MYA, and *Tetrapus* crown group (node 45) is constrained at a minimum age of 30 MYA. (iii) For the BEAST analysis we fixed the crown group *Pegoscapus* at 30 MYA as in Multidivtime 1.

2.5. Reconstruction of ancestral area states

To infer ancestral areas where fig-pollinating wasps may have originated, we applied both Bayesian and ML analyses on the combined data sets (28S "by eye" alignment + COI and 28S structural alignment + COI); for a similar approach see Pereira et al. (2007).

The Bayesian method was implemented in SIMMAP 1.0 Beta 2.0.8 Build 10042006 (Bollback, 2006) and biogeographic regions were treated as discrete characters. We accounted for phylogenetic and branch length uncertainties using the 100 trees and their respective branch lengths included in the post-burn-in portion of the Bayesian posterior distribution of trees. Without adding additional outgroups, the analyses cannot resolve how the first branch, between *Anaphes nitens* (our outgroup) and ingroup, should be split into its two components, and the rooting point could be placed anywhere along this branch. To explore if alternative positions of the rooting point along this branch would affect the inferred ancestral areas, we conducted two separate analyses. One analysis had 10% of the length assigned to the part leading to *Anaphes nitens* and 90% to the part leading to the ingroup. The second analysis partitioned the length in the opposite way, with 10% assigned to the part leading to the ingroup and 90% to the part leading to *Anaphes nitens*.

The ML approach is implemented in Mesquite 1.1 (build h61) (Maddison and Maddison, 2007). We reconstructed ancestral area states using a stochastic Markov model of evolution (Mk1 model) and took into account phylogenetic uncertainty using all 100 random posterior trees assuming branch lengths to be scaled to divergence times, as estimated in the five different molecular-dating

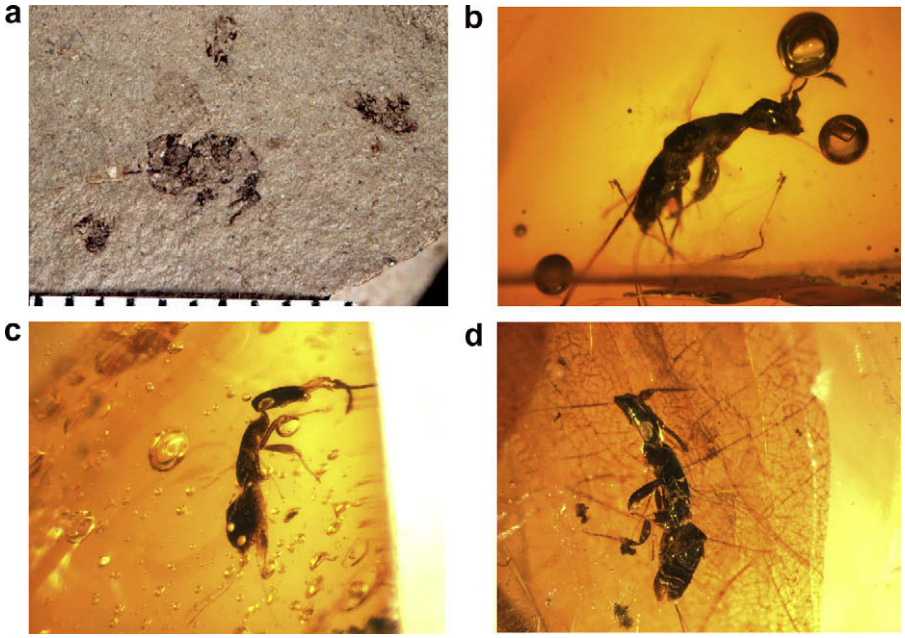

Fig. 2. Fossils of Chalcidoidea wasps. (2a) Type of *Tetrapus mayri* (MCZ No. 2067 = No. 13,976 of the S.H. Scudder Collection) from the Florissant, Colorado, (Brues, 1910; Weiblen, 2002) (photo by Herb Meyer). (2b) Fossils of the genus *Pegoscapus*, (Fig. 2b) and *Tetrapus* (Fig. 2c and d) from Dominican amber (photos by Laurent Soldati).

Table 2
The probability of each biogeographic region as the ancestral area of the Agaonidae (Fig. 1 node 1) as estimated from different dated phylogenies using the program Mesquite. Maximum likelihood values were calculated over 100 random posterior trees with branch lengths equivalent to divergence times calculated using four different molecular-dating approaches (see Table 1). Maximum likelihood values are expressed as either means ± SD or median (interquartile range) depending on whether the distribution differs significantly from a normal distribution (Kolmogorov-Smirnov test).

	Neotropical	Asian	Afrotropical	Asian + Australasia	Australasia
r8s (NPRS)	0.14 ± 0.02	0.54 ± 0.08	0.03 (0.00)	0.03 (0.00)	0.28 ± 0.09
r8s (PL)	0.14 ± 0.01	0.40 ± 0.06	0.04 (0.01)	0.04 (0.01)	0.39 ± 0.05
Multidiv1	0.15 ± 0.03	0.49 ± 0.08	0.03 (0.01)	0.03 (0.01)	0.29 ± 0.09
Multidiv2	0.16 ± 0.03	0.49 ± 0.08	0.03 (0.01)	0.03 (0.01)	0.29 ± 0.09
BEAST (by eye)	0.13 ± 0.02	0.38 ± 0.09	0.04 (0.01)	0.04 (0.01)	0.38 ± 0.08
BEAST (structural)	0.15 ± 0.02	0.40 ± 0.07	0.03 (0.01)	0.03 (0.01)	0.38 ± 0.06

approaches used (NPRS, PL, Multidivtime 1 and 2 and Beast) (see Table 2).

For both Bayesian and ML analyses we categorized the current distribution of fig-wasp genera as a multistate character based on four biogeographical areas: (1) Neotropical (Southern and Central America), (2) Afrotropical (Africa and Madagascar), (3) Australasian (including Australia, New Caledonia, New Guinea and islands east of Lydekker's line and the Molluccas), (4) Asian (Asia including 3 species which are also present in the Palearctic). However, because several taxa (i.e., *Ceratosolen* (*C.*) *appendiculatus*) occur in both Asian and Australasian regions, and both Bayesian and ML methods require unique character states, we also defined a fifth state: (5) Asian plus Australasian.

3. Results

3.1. Fig-pollinating wasp phylogenetics

Mitochondrial DNA provided 479 parsimony-informative characters out of 824 aligned nucleotide positions (58%). Nuclear ribosomal DNA (nrDNA) provided 555 informative characters out of 1313 positions (42%) based on the "by eye" alignment. On the other hand, based on the 28S structural alignment, 368 bp were excluded because positional homology could not be established using structural criteria. Out of the remaining 1790 included characters, 774 were parsimony-informative (43.2%). Only 18 positions in the aligned mtDNA were gaps compared to 544 for 28S "by eye" alignment and 408 for 28S structural alignment. GTR + I + G was the best fitting model chosen by ModelTest for each gene separately and in combination.

Results of the unpartitioned Bayesian analysis are presented as a 50% majority rule consensus tree in Fig. 1. Results of the partitioned Bayesian analysis (not shown) were completely congruent with these results, and all groups supported by 95% pp or more in the unpartitioned analyses (Fig. 1) were also supported by 95% pp or more in the partitioned analyses and showed similar branch lengths.

Comparing three separate nonparametric bootstrap analyses (not shown), the 28S ("by eye" alignment) resolved 38 clades, 28S ("structural" alignment) resolved 31 clades, while COI resolved only 18 clades with greater than 50% support. In addition, both 28S trees were better supported along the backbone, while COI resolved lower-level relationships within genera such as *Ceratosolen* and *Kradibia*. All clades in the COI bootstrap consensus were recovered in the 28S bootstrap consensus, with two exceptions. These included (1) the position of *Ceratosolen medlerianus,* that differed among separate analyses but conflicting support was less than 60%, and (2) the relative positions of *Liporrhopalum* and *Kradibia* which were in contradiction with 81% and 87% from COI and 28S ("by eye" alignment), respectively. This was the only moderately supported conflict.

Our phylogenetic reconstructions of fig-pollinating wasps showed high levels of phylogenetic conservatism (Fig. 1) as wasps that pollinate figs of the same section tend to form monophyletic groups. Thus, all wasps that pollinate figs of the *Galoglychia* section form a clade. Similarly, wasp species that pollinate sections *Pharmacosycea, Conosycea, Sycidium, Malvanthera* and *Americana*, each form clades (Fig. 1). In contrast, a few genera, such as *Blastophaga* and *Dolichoris,* were not monophyletic. This is supported by recent morphological studies which suggest that these two genera are heterogeneous and in need of taxonomic revision (Kjellberg et al., 2005).

3.2. Molecular dating

Age estimates for nodes supported by 95% pp or more differ widely between the different methods (Table 1 and Fig. 3). For instance, the crown radiation of modern fig-pollinating wasp genera (node 1) was estimated between 54 (95% confidence limit in Multidivtime 2) and 216 MYA (PL, UHPD value) depending on the method used to estimate divergence times (Table 1).

The PL analyses consistently yielded the oldest age estimates (Table 1 and Fig. 3), whereas NPRS and the two Multidivtime analyses gave reasonably similar divergence times. The Multidivtime 2 analyses generally had larger standard deviations than Multidivtime 1 (Fig. 3), which is expected given that the calibration is less fixed. Age estimates obtained with BEAST using the "by eye" alignment were very similar to those obtained with the structural alignment (Table 1 and Fig. 3).

The large differences between the PL and the NPRS analyses indicated that the PL analysis was highly sensitive to the choice of smoothing value. In our PL analyses, optimal smoothing values ranges between 0.001 and 0.032 across the 100 trees (as obtained from the cross-validation analyses) and in the NPRS analyses we arbitrarily set the smoothing to 0.0001 across all trees. Although these differences may seem small, they have a considerable effect on the resulting age estimates (Table 1).

All methods showed large error intervals, in particular at basal and intermediate nodes (Table 1 and, Figs. 3 and 4).

3.3. Reconstruction of ancestral area states

Fig-pollinating wasps occur mainly in tropical and subtropical areas of the southern hemisphere. Their diversity varies across zoogeographical regions with the Asian and Australasian regions harboring the highest species richness. Our reconstructions of ancestral areas suggest an Asian origin for the ancestors of extant fig-pollinating wasps (Fig. 5). Indeed, the Markov-ML reconstruction of ancestral areas indicated that the most recent common ancestor of all extant Agaonidae lived in Gondwana, and most likely in Asia (Table 2). The ML proportions attributed to each biogeographic region clearly indicate that the Asian region is the most likely ancestral area when using r8s, Multidivtime and BEAST (structural alignment) generated chronograms (Table 2). However, when using BEAST ("by eye" alignment) generated chronograms, both Asia and Australasia are equally likely ancestral areas. Bayes-

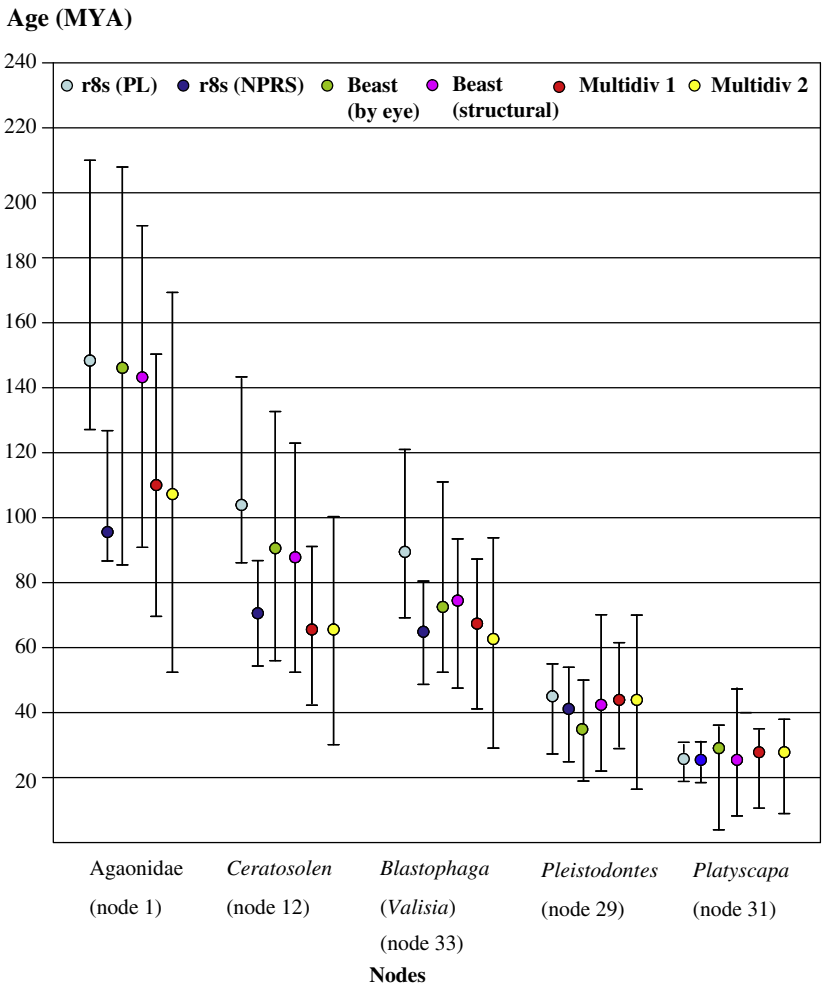

Fig. 3. Comparison of divergence date estimates across six dating analyses: mean and 95% confidence intervals for a selection of five nodes. Node numbers are the same used in Fig. 1.

ian reconstruction of ancestral areas (using both by eye and structural alignments) also favored a Gondwanan and most likely Asian origin (posterior probability using by eye and structural alignments = 0.5224; 0.6112, respectively) followed by the Neotropics (0.3123; 0.3265) and Australasia (0.1546; 0.1189) as the only other regions to receive non-zero likelihoods.

By counting the number of shifts in the estimates of ancestral areas (Fig. 5), it appears that from Asia, at least two lineages reached Africa (i.e., pollinators of figs in the section *Galoglychia* and *Platyscapa soraria*), three reached Australasia (i.e., pollinators of figs in the sections *Malvanthera*, *Rhizocladus, Oreosycea*), two lineages independently reached the American continent (pollinators of figs in the sections *Pharmacosycea* and *Americana*) and one lineage (*Courtella bekiliensis*) dispersed from Africa to Madagascar (Fig. 5).

4. Discussion

4.1. Classification and phylogenetic conservatism of host association

Our results confirm that Agaonidae (*sensu* Rasplus et al., 1998) is monophyletic (Gibson et al., 1999) but do not support the monophyly of Wiebes' (1982) subfamilies (Agaoninae and Blastophaginae) based on the basis of the shape of the mandibular appendage, which suggests that this is not a good morphological synapomorphy and probably subject to convergent evolution.

Previous phylogenetic studies have not fully resolved relationships among genera of Agaonidae because of limited sampling of taxa and characters (all based on COI sequence data) and the limited phylogenetic informativeness of mitochondrial DNA regarding ancient events (Townsend et al., 2008). Our combined nuclear and mitochondrial DNA dataset of 64 species from 15 of the 20 genera recovered a relatively well-supported at the genus level. However, the interrelationships among major clades within Agaonidae remain unresolved. This lack of resolution of deep nodes may be resolved with further genes and taxa.

In contrast to previous studies (Machado et al., 2001; Jiang et al., 2006), the genus *Ceratosolen* is monophyletic (Weiblen, 2001) and shows a strongly supported sister relationship with the genera *Liporrhopalum* and *Kradibia*. The genus *Blastophaga* (pollinators of the subgenus *Ficus*) is polyphyletic but species of the subgenus *Valisia* form a well-supported monophyletic clade. Interestingly, *Blastophaga psenes* falls within a clade formed by *Dolichoris* species (pollinators of *Oreosycea* figs) making the latter paraphyletic and suggesting the need for future taxonomic revision of this species. In addition, *Pleistodontes froggatti* comes out as the most basal species within this genus, which agrees with the placement of its host plant, *Ficus macrophylla*, in a recent phylogenetic study of *Malvanthera* figs (Rønsted et al., 2008).

Our phylogenetic analyses show a high level of phylogenetic conservatism with pollinators of several fig sections (*Pharmacosycea, Galoglychia, Conosycea, Americana, Malvanthera*) being mono-

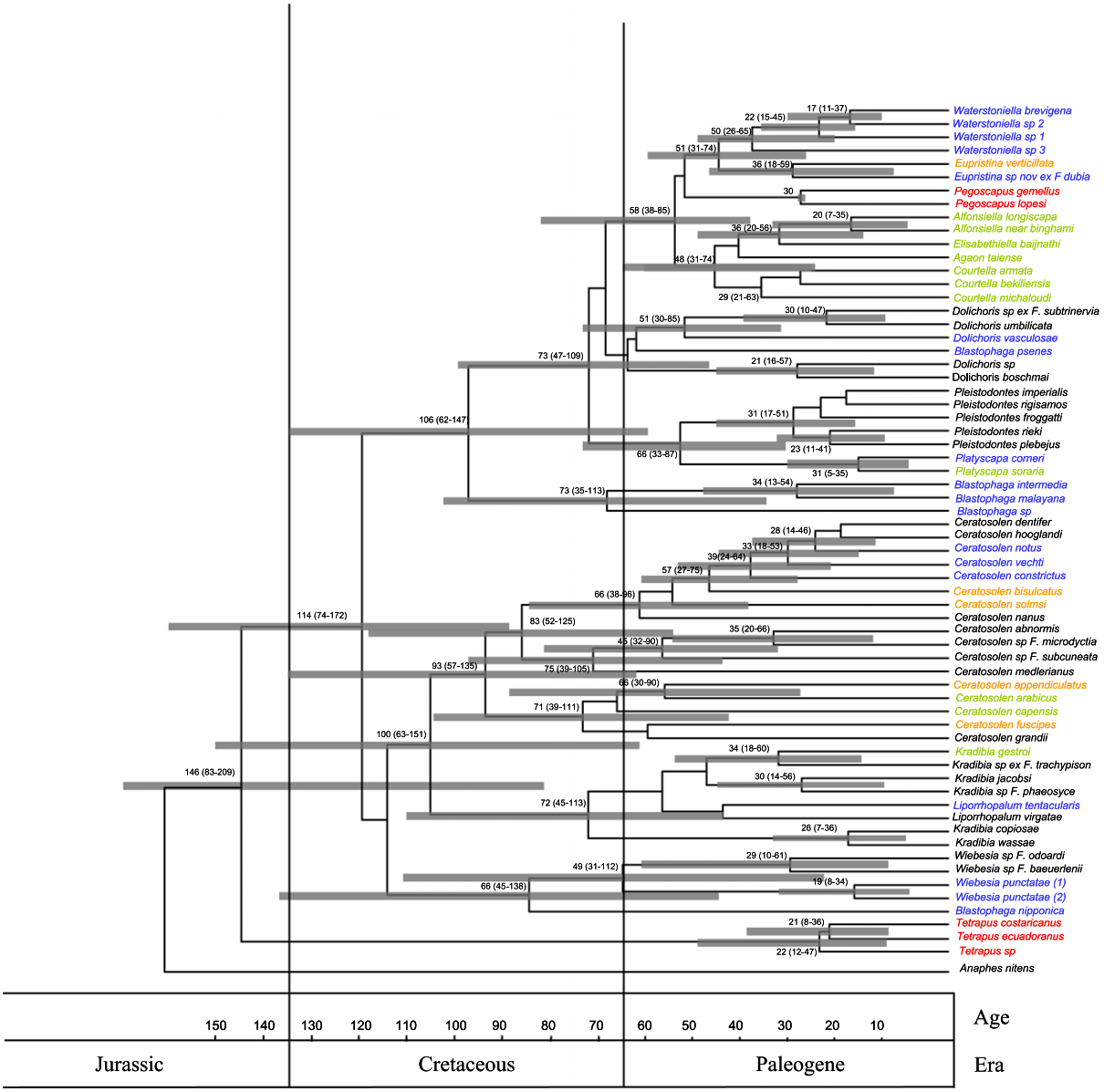

Fig. 4. Chronogram obtained for pollinating fig wasps under a Bayesian non autocorrelated rates relaxed clock model using BEAST and applied to the combined data (28S "by eye" alignment + COI) and calibrated with *Pegoscapus* (30 MYA). Age estimates with their 95% credibility intervals shown on nodes. Green, Africa (including Madagascar); red, central and south America; blue, Asia; black, Australia/New Guinea/Polinesia; orange, Asia and Australasia.

phyletic. This confirms the findings of previous studies (Machado et al., 2001; Rønsted et al., 2005).

4.2. How old are fig-pollinating wasps?

Rønsted et al. (2005) used a double dating (independent dating of two associated lineages) approach and considered fig wasps (Agaonidae) to have co-diverged with their associated hosts (*Ficus*) during the last 60 million years. However, it is clear that if we account for the uncertainties in molecular age estimates (Table 1), we cannot exclude alternative scenarios. In fact, unless we consider one or more of our analyses to yield unreasonable and incorrect age estimates, the crown group of Agaonidae (node 1) could have originated anytime between 54 and 216 MYA. This wide error margin partly originates from differences between methods. However, even if we only consider the Multidivtime 2 analyses, the origin of Agaonidae still ranges from 54 to 168 MYA. Errors associated with uncertainties in how rates have changed, in branch length estimates, in topology, and in the calibration point are best accounted for in both the BEAST and the Multidivtime 2 analysis. Unless we can eliminate or reduce these uncertainties, we are faced with very wide error margins. How could we obtain more precise age estimates? One way would be to incorporate additional information from the fossil record. Enforcing further age constraints while running the analyses would likely result in narrower error margins, but this requires that such information is available. Alternatively, we could make stronger prior assumptions about rates and how they are allowed to change across the tree (Welch et al., 2005). Although the error margins could be reduced by such an approach, we would have to justify such changes in our prior assumptions in some way.

Comparing the alternative methods reveals some notable patterns. One involves the error margins in the Multidivtime analyses and how they are distributed. We have two different error margins: one stemming from our 100 trees (i.e., the effect of topological and branch length uncertainties), and a second reported by

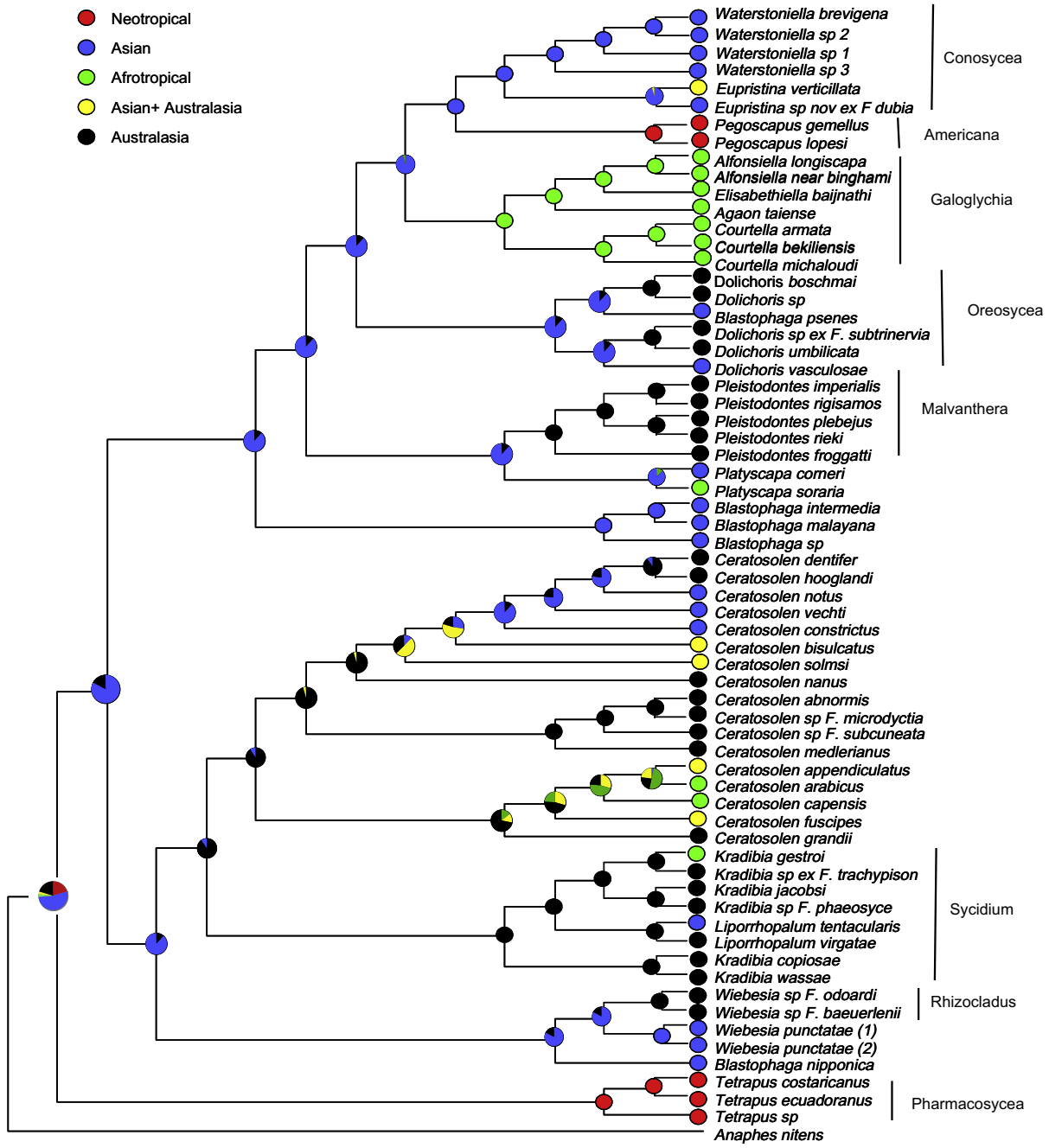

Fig. 5. One of the 100 cladograms obtained with Multidivtime of fig-pollinating wasp diversification and reconstruction of ancestral areas. Present biogeographic distribution is indicated as colored circles at the tips, and the proportion of the total likelihood received by each biogeographic region as the ancestral area of a given clade (calculated with Mesquite) is represented by pie charts at internal nodes.

Multidivtime for each point estimate. It is clear that errors for the point estimates are always larger than those associated with the 100 different trees. We conclude that topological and branch length uncertainties are of minor importance compared to uncertainty arising from the relaxed clock as implemented in Multidivtime.

Also interesting are the considerable differences between the PL and the NPRS analyses. As previously explained, the NPRS method on the one hand, and a maximum likelihood clock model on the other, can be seen as logical extremes on the scale of different smoothing values in a PL analysis (Sanderson, 2002). Moving from NPRS ($\lambda \rightarrow 0$ and where considerable rate changes are allowed) towards a clock model ($\lambda \rightarrow \infty$ and where rates are not allowed to change) we introduce an increasing amount of resistance towards rate changes. In our case, using a calibration point (*Pegoscapus*, node 36) that is nested well inside Agaonidae, this introduced resistance has the effect of pushing our root node back in time. A cross-validation approach was suggested by Sanderson (2002) to evaluate how much resistance we should impose in each analysis. However, if the estimate obtained is an average across the entire tree, we might get large errors in parts of the tree where there have been significant changes in rates.

Finally, it has recently been shown that phylogenetic uncertainty in calibration fossils can greatly increase confidence intervals of inferred divergence (Lee et al., 2009). However, uncertainty in the phylogenetic position of our calibration points

does not explain the wide error margins of our study since neither *Pegoscapus* nor *Tetrapus* show large phylogenetic uncertainty across the 100 sampled trees.

4.3. Where did fig-pollinating wasps originate?

Ancestral area analyses favor an Asian origin for fig wasps, in sharp contrast to previous studies which have suggested a South American origin for the mutualism, on the basis that the basalmost lineages of extant figs (Pharmacosycea) and associated extant fig wasps (*Tetrapus*) occur only in South America. Our phylogeny also shows that *Tetrapus* is not only sister to all other fig wasps, but also the earliest diverging lineage, so the difference in inferred area of origin (Asian instead of Neotropical) might be due to the use of different analytical methodology and/or taxa sampling. However, it is important to emphasize that ancestral character state reconstruction is subject to several sources of error (Lopez-Vaamonde et al., 2003). Firstly, it has been shown that a combination of rapid evolution and unequal probabilities of gains and losses can lead to errors in ancestral state reconstruction (Cunningham et al., 1998). Secondly, the density of taxon sampling is important for an accurate reconstruction of ancestral character states. Although we included 64 Agaonidae species covering most genera in this study, they comprise only about 6% of the estimated species total in this huge family. A third source of error is unresolved topology, although this problem is not expected to affect greatly our results since our analyses (both Bayesian and ML) take phylogenetic uncertainty into account. Fourthly, most reconstruction methods are predisposed to yield ancestral areas corresponding to the most common area, in this case, Asia (Ree and Smith, 2008). In addition, although our analyses favor an Asian origin, we did not include Antarctica as a character state, since no agaonids presently occur on this continent. Therefore, we cannot exclude the possibility that Antarctica was the ancestral area for all agaonids. This possibility arises from recent paleogeographic reconstructions showing that South America, Antarctica, Australia and India–Madagascar remained linked for most of the Late Cretaceous until at least 80 MYA (Hay et al., 1999), a time that is consistent with the estimated range for the time of origin of Agaonidae. On the other hand, Antarctica had a temperate climate during this period and therefore unlikely to support lineages adapted to humid tropical conditions (Zerega et al., 2005).

Our phylogenetic analysis confirms previous studies showing that the Neotropical fig-wasp genus *Tetrapus* is consistently placed as sister group to the rest of pollinating fig wasps (Machado et al., 2001; Rønsted et al., 2005). These studies suggested that the origin of agaonids could be in West Gondwana, but our ancestral area reconstruction shows that an eastern Gondwana origin is most likely.

The relatively young age of both the earliest diverging lineage of fig wasps, *Tetrapus*, (maximum 49 MYA, Table 1) and their associated figs (section *Pharmacosycea*) (about 60 MYA, Rønsted et al., 2005) postdates the separation of South America from Africa (about 90–100 MYA, Smith et al., 1994) during the break-up of Gondwana. If our age estimates and ancestral area reconstructions are correct, it would imply that the current presence of *Tetrapus* and its associated *Pharmacosycea* figs in South America could be the result of a long distance oceanic dispersal event (Rønsted et al., 2005) from Asia to the Neotropics. Alternatively, high levels of stem lineage extinction could explain the observed young age for extant species of this lineage of Neotropical figs and fig wasps. In addition, it is important to bear in mind that we are using sequence data from just three *Tetrapus* species. Sequencing of more *Tetrapus* species is needed to confirm both the young age of this genus and its sister relationship with the rest of agaonids.

5. Conclusions

Our age estimates range widely, depending on which molecular-dating approach is used. This makes it difficult to assess with any certainty the Cretaceous/Gondwanan origin of agaonids. However, our reconstruction of ancestral areas identifies the Asian region as the most likely area of origin of agaonids, although a southern Gondwana origin cannot be rejected. Our analyses clearly suggest a pervasive pattern of dispersal.

Acknowledgments

The authors thank the colleagues listed in the Appendix 1 for providing specimens. CLV is most grateful to Dale Dixon, and Samson Laup, for assistance in the field. We are grateful to George Poinar for kindly letting us examine his amber preserved fig-wasp collection. Thanks to Philip D. Perkins for sending the photo of *Tetrapus mayri*. Stephanie Swenson and Vanessa Bull added new COI and 28S sequences respectively. We are grateful to Alfried Vogler, David Lees and Nina Rønsted for comments on the manuscript and to Akito Kawahara for his helpful comments on BEAST analysis. JYR thanks ANR Biodiversity for funding the project NiceFigs and JMC thanks NERC (UK) for funding.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.05.028.

References

Bollback, J.P., 2006. SIMMAP: stochastic character mapping of discrete traits on phylogenies. BMC Bioinformatics 7, 88.
Brues, C., 1910. The parasitic Hymenoptera of the Tertiary of Florissant, Colorado. Bull. Mus. Comp. Zool. 54, 1–125.
Campbell, B., Heraty, J., Rasplus, J.Y., Chan, K., Steffan-Campbell, J., Babcock, C., 2000. Molecular systematic of the Chalcidoidea using 28S-rDNA. In: Austin, A.D., Dowton, M. (Eds.), The Hymenoptera: Evolution, Biodiversity and Biological Control. CSIRO, Canberra, pp. 59–73.
Collinson, M.E., 1989. The fossil history of the Moraceae, Urticaceae (including Cecropiaceae), and Cannabaceae. In: Crane, P.R., Blackmore, S. (Eds.), Evolution, Systematics, and Fossil History of the Hamemelidae. Clarendon Press, Oxford, pp. 319–339.
Cunningham, C.W., Omland, K.E., Oakley, T.H., 1998. Reconstructing ancestral character states: a critical reappraisal. Trends Ecol. Evol. 13, 361–366.
Datwyler, S.L., Weiblen, G.D., 2004. On the origin of the fig: phylogenetic relationships of Moraceae from ndhF sequences. Am. J. Bot. 91, 767–777.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Gibson, G.A.P., Heraty, J.M., Woolley, J.B., 1999. Phylogenetics and classification of Chalcidoidea and Mymarommatoidea – a review of current concepts (Hymenoptera, Apocripta). Zool. Scr. 28, 87–124.
Gillespie, J.J., Munro, J.B., Heraty, J.M., Yoder, M.J., Owen, A.K., Carmichael, A.E., 2005. A secondary structural model of the 28S rRNA expansion segments D2 and D3 for chalcidoid wasps (Hymenoptera: Chalcidoidea). Mol. Biol. Evol. 22, 1593–1608.
Goldman, N., Anderson, J.P., Rodrigo, A.G., 2000. Likelihood based tests of topologies in phylogenetics. Syst. Biol. 49, 652–670.
Hay, W.W., DeConto, R.M., Wold, C.N., Wilson, K.M., Voigt, S., Schulz, M., Wold, A.R., Dullo, W.-C., Ronov, A.B., Balukhovsky, A.N., Soeding, E., 1999. Alternative global Cretaceous paleogeography. In: Barrera, E., Johnson, C.C. (Eds.), Evolution of the Cretaceous Ocean-Climate System. Geological Society of America, Boulder, pp. 1–47.
Huelsenbeck, J.P., Rannala, B., Masly, J.P., 2000. Accommodating phylogenetic uncertainty in evolutionary studies. Science 288, 2349–2350.
Huelsenbeck, J.P., Ronquist, F., Posada, D., Crandall, K.A., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.
Ihaka, R., Gentlemen, R., 1996. R: a language for data analysis and graphics. J. Comput. Graph. Stat. 5, 299–314.
Iturralde-Vinent, M.A., MacPhee, R.D.E., 1996. Age and paleographical origin of Dominican amber. Science 273, 1850–1852.
Jiang, Z.-F., Huang, D.-W., Zhu, C.-D., Zheng, W.-Q., 2006. New insights into the phylogeny of fig pollinators using Bayesian analyses. Mol. Phyl. Evol. 38, 306–315.

Jousselin, E., Rasplus, J.-Y., Kjellberg, F., 2003. Convergence and coevolution in a mutualism: evidence from a molecular phylogeny of *Ficus*. Evolution 57, 1255–1269.

Kishino, H., Thorne, J.L., Bruno, W.J., 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol. Biol. Evol. 18, 352–361.

Kjellberg, F., Jousselin, E., Hossaert-Mckey, M., Rasplus, J.-Y., 2005. Biology, ecology and evolution of fig-pollinating wasps (Chalcidoidea, Agaonidae). In: Schaefer, C.W., Toni, M. (Eds.), Biology, Ecology and Evolution of Gall-Inducing Arthropods. Science Publishers, Enfield, pp. 539–572.

Kjer, K.M., 1995. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from the frogs. Mol. Phylogenet. Evol. 4, 314–330.

Kjer, K.M., Gillespie, J.J., Ober, K.A., 2007. Opinions on multiple sequence alignment, and an empirical comparison of repeatability and accuracy between POY and structural alignments. Syst. Biol. 56, 133–146.

Kjer, K.M., Roshan, U., Gillepie, J.G., 2009. Structural and evolutionary considerations for multiple sequence alignment of RNA, and the challenges for algorithms that ignore them. In: Perspectives on Biological Sequence Alignment: Where, How, and Why It Matters". University of California Press.

Langley, C.H., Fitch, W., 1974. An estimation of the constancy of the rate of molecular evolution. J. Mol. Evol. 3, 161–177.

Larget, B., Simon, D.L., 1999. Markov Chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Mol. Biol. Evol. 16, 750–759.

Lee, M.S.Y., Oliver, P.M., Hutchinson, M.N., 2009. Phylogenetic uncertainty and molecular clock calibrations: a case study of legless lizards (Pygopodidae, Gekkota). Mol. Phylogenet. Evol. 50, 661–666.

Loader, C., 1999. Local Regression and Likelihood. Springer Mathematics, New York, USA.

Lopez-Vaamonde, C., Godfray, H.C.J., Cook, J.M., 2003. Evolutionary dynamics of host use in a genus of leaf-mining moths. Evolution 57, 1804–1821.

Lopez-Vaamonde, C., Rasplus, J.Y., Weiblen, G.D., Cook, J.M., 2001. Molecular phylogenies of fig wasps: partial co-cladogenesis between pollinators and parasites. Mol. Phylogenet. Evol. 21, 55–71.

Lopez-Vaamonde, C., Wikstrom, N., Labandeira, C., Godfray, H.C.J., Goodman, S.J., Cook, J.M., 2006. Fossil-calibrated molecular phylogenies reveal that leaf-mining moths radiated several million years after their host plants. J. Evol. Biol. 19, 1314–1326.

Machado, C.A., Jousselin, E., Kjellberg, F., Compton, S.G., Herre, E.A., 2001. Phylogenetic relationships, historical biogeography and character evolution of fig-pollinating wasps. Proc. R. Soc. Lond. B 268, 1–10.

Machado, C.A., Robbins, N., Gilbert, M.T.P., Herre, E.A., 2005. Critical review of host specificity and its coevolutionary implications in the fig/fig-wasp mutualism. Proc. Natl. Acad. Sci. USA 102, 6558–6565.

Maddison, D.R., Maddison, W.P., 2001. MacClade 4.02. Sinauer Associates, Sunderland, MA.

Maddison, W.P., Maddison, D.R., 2007. Mesquite: a modular system for evolutionary analysis. Version 2.0.

Murray, M.G., 1985. Figs (*Ficus* spp.) and fig wasps (Chalcidoidea, Agaonidae): hypotheses for an ancient symbiosis. Biol. J. Linn. Soc. Lond. 26, 69–81.

Nylander, J., 2004. MrModeltest 2.0. Program Distributed by the Author. Uppsala University, Evolutionary Biology Centre.

Pagel, M.D., Lutzoni, F., 2002. Accounting for phylogenetic uncertainty in comparative studies of evolution and adaptation. In: Lässig, M., Valleriani, A. (Eds.), Biological Evolution and Statistical Physics. Springer-Verlag, Berlin, Heidelberg, pp. 148–161.

Penalver, E., Engel, M.S., Grimaldi, D.A., 2006. Fig wasps in Dominican amber (Hymenoptera: Agaonidae). Am. Mus. Novit. 3541, 1–16.

Pereira, S.L., Johson, K.P., Clayton, D.H., Baker, A.J., 2007. Mitochondrial and nuclear DNA sequences support a Cretaceous origin of Columbiformes and a dispersal-driven radiation in the paleogene. Syst. Biol. 56, 656–672.

Poinar, G.O., 1993. Insects in amber. Annu. Rev. Entomol. 46, 145–159.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of the Aic and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rasplus, J.Y., Kerdelhué, C., Le Clainche, I., Mondor, G., 1998. Molecular phylogeny of fig waps (Hymenoptera). Agaonidae are not monophyletic. Compte Rendu de l'Académie des Sciences de Paris 321, 517–527.

Ree, R.H., Smith, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57, 4–14.

Rønsted, N., Weiblen, G.D., Cook, J.M., Salamin, N., Machado, C.A., Savolainen, V., 2005. Sixty million years of co-divergence in the fig-wasp symbiosis. Proc. R. Soc. Lond. B 272, 2593–2599.

Rønsted, N., Weiblen, G.D., Savolainen, V., Cook, J.M., 2008. Phylogeny, biogeography, and ecology of *Ficus* section *Malvanthera* (Moraceae). Mol. Phylogenet. Evol. 48, 12–22.

Sanderson, M., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.

Sanderson, M., 2003. R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19, 301–302.

Sanderson, M.J., 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14, 1218–1231.

Schlee, D., 1990. Das Bernstein-Kabinett. Stuttg. Beitr. Naturkunde (C) 28, 1–100.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 114–116.

Smith, A.G., Smith, D.G., Funnel, B.M., 1994. Atlas of Mesozoic and Cenozoic Coastlines. Cambridge University Press, Cambridge.

Thorne, J.L., Kishino, H., 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51, 689–702.

Thorne, J.L., Kishino, H., Painter, I.S., 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15, 1647–1657.

Townsend, J.P., Lopez-Giraldez, F., Friedman, R., 2008. The phylogenetic informativeness of nucleotide and amino acid sequences for reconstructing the vertebrate tree. J. Mol. Evol. 67, 437–447.

Walker, F., 1846. Description of Chalcidites. Trans. Linn. Soc. Lond. 20 (1), 153–162.

Weiblen, G.D., 2001. Phylogenetic relationships of fig wasps pollinating functionally dioecious figs based on mitochondrial DNA sequences and morphology. Syst. Biol. 50, 243–267.

Weiblen, G.D., 2002. How to be a fig wasp. Annu. Rev. Entomol. 47, 299–330.

Weiblen, G.D., 2004. Correlated evolution in fig pollination. Syst. Biol. 53, 128–139.

Weiblen, G.D., Bush, J.L., 2002. Speciation in fig pollinators and parasites. Mol. Ecol. 11, 1573–1578.

Welch, J.J., Fontanillas, E., Bromham, L., 2005. Molecular dates for the "Cambrian Explosion": the influence of prior assumptions. Syst. Biol. 54, 13–19.

Wiebes, J.T., 1982. The phylogeny of the Agaonidae (Hymenoptera, Chalcidoidea). Neth. J. Zool. 32, 395–411.

Zerega, N.J.C., Clement, W.L., Datwyler, S.L., Weible, G.D., 2005. Biogeography and divergence times in the mulberry family (Moraceae). Mol. Phylogenet. Evol. 37, 402–416.