WILEY-BLACKWELL

Cladistics 29 (2013) 132–146

Cladistics

10.1111/j.1096-0031.2012.00419.x

Pholcid spider molecular systematics revisited, with new insights into the biogeography and the evolution of the group

Dimitar Dimitrov[a,b,*], Jonas J. Astrin[c] and Bernhard A. Huber[c]

[a]*Center for Macroecology, Evolution and Climate, Zoological Museum, University of Copenhagen, Copenhagen, Denmark;* [b]*Department of Biological Sciences, The George Washington University, Washington, DC, USA;* [c]*Forschungsmuseum Alexander Koenig, Adenauerallee 160, D-53113 Bonn, Germany*

Accepted 5 June 2012

Abstract

We analysed seven genetic markers sampled from 165 pholcids and 34 outgroups in order to test and improve the recently revised classification of the family. Our results are based on the largest and most comprehensive set of molecular data so far to study pholcid relationships. The data were analysed using parsimony, maximum-likelihood and Bayesian methods for phylogenetic reconstruction. We show that in several previously problematic cases molecular and morphological data are converging towards a single hypothesis. This is also the first study that explicitly addresses the age of pholcid diversification and intends to shed light on the factors that have shaped species diversity and distributions. Results from relaxed uncorrelated lognormal clock analyses suggest that the family is much older than revealed by the fossil record alone. The first pholcids appeared and diversified in the early Mesozoic about 207 Ma ago (185–228 Ma) before the breakup of the supercontinent Pangea. Vicariance events coupled with niche conservatism seem to have played an important role in setting distributional patterns of pholcids. Finally, our data provide further support for multiple convergent shifts in microhabitat preferences in several pholcid lineages. Our findings suggest that both adaptive and non-adaptive speciation may have played an important role in the diversification of pholcid lineages.

Pholcids (Araneae: Pholcidae) have been singled out (e.g. Huber, 2000, 2011a; Bruvo-Madaric et al., 2005) for their often high abundance in tropical forests (e.g. Huber, 2000, 2003a), unusually high number of synanthropic species (e.g. Huber, 2000, 2011a), and the extensive knowledge that exists on their sexual behaviour and on the functional morphology of their copulatory organs (e.g. Huber, 1994, 1995, 1996a,b, 2002; Uhl et al., 1995; Uhl, 1998). Commonly known as daddy-longlegs spiders, they may be among the arthropods with the longest legs relative to their body size. Some synanthropic species such as *Pholcus phalangioides* have probably shared a roof with all of us on at least one occasion, yet many aspects of the biology and diversity of pholcids remain unknown. With more than 1200 described species (Huber, 2011a; Platnick, 2012) pholcids are among the most species-rich spider families. Generic revisions, however, often double the number of described taxa in their focal genera, suggesting that the actual number of species is much higher (e.g. Huber, 2005a, 2011a). Since the first cladistic analysis of Pholcidae (Huber, 2000), there have been numerous phylogenetic studies focusing on this group of haplogyne spiders (e.g. Huber, 2000, 2003a, 2005a,b, 2011a; Bruvo-Madaric et al., 2005; Astrin et al., 2007; Dimitrov and Ribera, 2007; Huber and Astrin, 2009) some of which primarily rely on molecular data (e.g. Bruvo-Madaric et al., 2005; Astrin et al., 2007; Dimitrov et al., 2008; Huber and Astrin, 2009; Huber et al., 2010). The evidence presented in these studies has resulted in a recent revision of the classification of the family (Huber, 2011b). Despite a few earlier attempts (Petrunkevitch, 1928; Mello-Leitão, 1946) to revise the original classification proposed by Simon (1893), his system has been

*Corresponding author.
E-mail address: dimitard.gwu@gmail.com

used for more than a century with only a few amendments. Although a big step forward, Huber's revised classification was explicitly proposed as "… a working hypothesis …" (Huber, 2011b, p. 221) as there are still numerous genera that have not been studied in a phylogenetic context or because cladistic analyses based on morphological data alone have not been able to unambiguously resolve their relationships. There are also several instances where generic limits continue to be contentious (e.g. *Spermophora*, *Leptopholcus*, *Belisana*, etc.) and several nodes within and among subfamilies remain ambiguous.

The lack of a rigorous phylogenetic hypothesis has largely prevented higher level biogeographical and diversification research in pholcids. As a result, just a couple of island radiations have been studied in some detail (Dimitrov et al., 2008; Huber et al., 2010). Despite some fossils, most of which can be placed in extant genera (e.g. Huber and Wunderlich, 2006; Penney, 2006; Wunderlich, 2008a; Dunlop et al., 2012), nothing is known about the origins of the family and its early history except that Pholcidae may be much older than the oldest known fossils (see comments on haplogyne spiders in Ayoub et al., 2007).

A robust phylogenetic framework is therefore essential to test the limits and relationships of most pholcid subfamilies and many genera. To provide and improve such a framework, we analyse seven molecular markers sequenced from a large sample of pholcid species representing all major lineages within the family and 36 of the 85 extant genera (Platnick, 2012). In addition, we used information from the fossil record to calibrate the pholcid family tree and estimate the age of the main lineages within the family. The availability of a dated phylogeny also allows us to contextualize evolutionary events and address biogeographical and macroevolutionary questions.

Materials and methods

Taxon sampling and gene selection

The in-group taxon sampling focused on maximizing the representation of pholcid lineages at the generic level. In a few cases in which monophyly of genera was uncertain, several species representing potentially non-monophyletic groups were included. Particular effort has been made to include distinct *Pholcus* lineages, by far the largest pholcid genus, and closely related taxa in order to test the monophyly of *Pholcus* and its composition. Numerous trips to collect DNA-ready material of different pholcid lineages have been carried out as part of the present research. We have specifically targeted undersampled lineages and regions. In a few cases (e.g. *Carapoia*, *Spermophora*) additional effort to achieve high coverage at the species level has been made to address specific biogeography and/or conservation questions. Outgroup taxa from all major spider lineages were added, with emphasis on haplogynes. Denser outgroup sampling has been shown to increase accuracy of the results (Dimitrov et al., 2012). In addition, some of the outgroups (e.g. Araneidae) were selected to facilitate the use of additional fossil calibration points for the molecular dating analyses.

We have targeted seven molecular markers that represent both nuclear and mitochondrial ribosomal and protein-coding genes: *12S rRNA*, *28S rRNA, 18S rRNA, 16S rRNA, cytochrome* c *oxidase subunit I, histone H3* and *wingless*. All of them have successfully been used in previous studies of phylogenetic relationships at different taxonomic levels in pholcids and other spider groups (e.g. Bruvo-Madaric et al., 2005; Astrin et al., 2006; Dimitrov et al., 2008, 2012; Álvarez-Padilla et al., 2009; Arnedo et al., 2009; Dimitrov and Hormiga, 2011). The bulk of character data used in the present study were newly generated as a result of current sampling efforts. In addition, both taxon and character sampling were further expanded by including additional sequence data from GenBank. The complete list of taxa and the sequence accession numbers are given in Supporting Information, Table S1.

DNA extraction, amplification and sequencing

Total genomic DNA was extracted from the prosoma (usually in females), opisthosoma (usually in males) or single whole individuals (in small specimens, for example juveniles) using the BioSprint96 magnetic bead extractor and corresponding extraction kits by Qiagen (Hilden, Germany). Extracted DNA is deposited at the Biobank of the ZFMK (Zoologisches Forschungsmuseum Alexander Koenig, Bonn, Germany).

PCR was carried out in total reaction mixes of 20 μL, including 1.2–2 μL of undiluted DNA template, 1.6 μL of each primer (10 pmol/μL), 2 μL of "Q-Solution" and 9.5 μL of "Multiplex PCR Master Mix", containing hot start Taq DNA polymerase and buffers. The latter components are available in the "Multiplex PCR" kit from Qiagen. In spite of the kit's name, the PCR reactions were not multiplexed, but were run individually. The kit helped in minimizing PCR optimization. A list of primers used for amplification and sequencing is given in Table S2. Thermal cycling was performed on a GeneAmp PCR System 2700 (Applied Biosystems, Foster City, CA, USA), using the same "touch-down" PCR protocol for all genes: hot start Taq activation: 15 min at 95 °C; first cycle set (15 repeats): 35 s denaturation at 94 °C, 90 s annealing at 60 °C (−1 °C per cycle) and 90 s extension at 72 °C; second cycle set (25 repeats): 35 s denaturation at 94 °C, 90 s annealing at 50 °C, and 90 s extension at 72 °C. After enzymatic

cleanup (Exo/SAP), PCR products were sent to a sequencing facility (Macrogen, Amsterdam, Netherlands) for double stranded sequencing.

Phylogenetic analyses

Alignments, datasets and data partitions. Multiple sequence alignments were performed with Muscle ver. 3.8.31 (Edgar, 2004) under default settings. Alignments of protein coding genes were translated to amino acids and checked for unexpected stop codons. Each aligned matrix was used directly for the single-gene analyses. Single-gene analyses were preformed only under maximum likelihood and further analyses focused on the combined datasets. To investigate the effect of ambiguously aligned regions, alignments of genes with considerable sequence length variation (12S, 16S, 18S and 28S) were processed with the program Gblocks ver. 0.91 (Castresana, 2000). Gblocks was run with the default parameters except for the *min blocks* setting which was changed from 10 to 5 relaxing the default constraint on the minimum length of a block. Individual gene alignments, both treated and not treated with Gblocks, were also combined into a single matrix and analysed together under the following partitioning schemes: by gene (protein coding genes form a single partition); by gene/codon (1st and 2nd and position of protein coding genes are treated separately from the 3rd position). *Euagrus chisoseus* (Dipluridae) was used to root the trees.

Parsimony analyses. Parsimony analyses where performed with TNT ver. 1.1 (Goloboff et al., 2008). Given the size of the dataset we have opted for a driven search combining few of the new technology algorithms. This strategy has been shown to be the most efficient when dealing with large datasets (Goloboff, 1999). The specific algorithms used were: tree drifting, mixed sectorial searches and tree fusing (ratchet was not active) with the following settings: 50 initial addition sequences, initial level 10, 10 cycles of drifting, stabilizing strict consensus five times with default factor of 75. The search was executed from a script file with the following commands: *hold 80000; piwe-; const-; rseed1; xm: noverb nokeep; rat : it 0 up 4 down 4 au 0 num 36 give 99 equa; dri: it 10 fit 1.00 rfi 0.20 aut 0 num 36 give 99 xfa 3.00 equa; sec: mins 45 maxs 45 self 43 incr 75 minf 10 god 75 drift 6 glob 5 dglob 10 rou 3 xss 10- 14 + 2 noxev noeq; tf: rou 5 minf 3 best ke nochoo swap; xm : level 10 nochk rep 50 fuse 3 dri 10 rss css noxss mult nodump conse 5 conf 75 nogive notarg upda autoc 3 xmix; xm; xmult:;*

For further comments on the commands and arguments used see Dimitrov et al. (2012). Detailed explanation on command usage is available in the program manual. In addition to the parsimony analyses under equal weights we performed an additional round of analyses for the full dataset using implied weighting. Implied weighting uses the k constant to modulate the concavity of the concave function of homoplasy (Goloboff, 1993) in such way that k has a negative correlation with the degree to which homoplasy is down-weighted (i.e. low k results in stronger weighting against homoplasy). As DNA sequences are prone to higher degrees of homoplasy, larger k values (e.g. > 10) are preferable (Arnedo et al., 2009). Here we use two relatively low values of k (6 and 10) and two larger values (20 and 100). For the searches under implied weights we have used the same search strategy changing the argument of the *piwe* command in the above script to set up the k value.

Clade support was assessed by the mean of jackknife frequencies calculated in TNT. Due to the size of the dataset and to speed up the analyses, a traditional search strategy was used for the jackknife resampling. We performed 1000 pseudoreplicates with heuristic searches consisting of ten random addition sequences, followed by ten iterations of tree bisection and reconnection and holding one tree. Clade support under implied weighting was calculated with the same strategy but using symmetric resampling instead of traditional jackknife (Goloboff et al., 2003). In all parsimony analyses gaps were treated as missing data, consistent with the standard gap treatment in the likelihood, Bayesian and BEAST analyses. Treating gaps as missing data, however, could be statistically inconsistent in likelihood analyses (Warnow, 2012). Here we have used Gblocks to remove ambiguously aligned positions, which mostly correspond to sites where insertions or deletions occur. The resulting dataset was analysed in a separate set of analyses and results were compared with results from the full dataset.

Maximum-likelihood. Maximum-likelihood (ML) trees were inferred with RaxML-HPC ver. 7.2.8 (Stamatakis, 2006) as implemented on the CIPRES science gateway (Miller et al., 2010). For the ML searches a GTR + Γ model of sequence evolution was applied for each partition following the program recommendations. Clade support was estimated with the fast bootstrap algorithm using the GTRCAT model (Stamatakis et al., 2008). The following are the specific command line options used to run RaxML (note that some flags may be specific for the CIPRES implementation): *-T 6 -N autoMRE -s* infile *-n* result *-x 12345 -p 12345 -f a -m GTRCAT -q part -o* outgroup.

Bayesian inference. Bayesian analyses were run in MrBayes ver. 3. 2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2004) on the Biocluster at the Copenhagen Zoological Museum and the Pyramid cluster at the George Washington University. To select

best fitting models of sequence evolution we used the program jMODELTEST v1.0.1 (Posada, 2008). Models selected under the Akaike and Bayesian information criterion are listed in Table S3. In each analysis two independent runs each with four Markov chain Monte Carlo chains (three heated and one cold) were run for a total of 50 million generations. Trees were sampled each 1000th generation. Results were examined for convergence with Tracer v1.5 (Rambaut and Drummond, 2007) and by monitoring the deviation of split frequencies. Trees sampled before stationarity were discarded using the burnin command.

Molecular dating. The use of a strict molecular clock was rejected by the likelihood ratio test (Huelsenbeck and Crandall, 1997). To accommodate for rate heterogeneity among lineages we have used the relaxed uncorrelated lognormal clock model implemented in the program BEAST ver. 1.6.1 (Drummond and Rambaut, 2007). Fossil constraints were used to calibrate the relaxed clock and estimate lineage divergence times and their 95% highest posterior density interval (HPD). To prevent introduction of additional biases and/or sources of error fossil constraint had to satisfy two conditions: a suitable fossil must share at least one synapomorphy with a lineage within a monophyletic crown group in our preferred tree; and when several fossil species belonging to the same lineage were available only the oldest one was used. After reviewing the literature and sorting out fossil information in accordance with the aforementioned criteria seven fossil calibration points were selected for the molecular clock analyses: Nephilidae (min. age 165 Ma) (Selden et al., 2011), Araneidae (min. age 115–121 Ma) (Penney and Ortuño, 2006), Segestriidae (min. age 125–135 Ma) (Wunderlich, 2008a), Dysderidae (min. age 35–40 Ma) (Wunderlich, 2004), New World *Leptopholcus* sp. (min. age 20–45 Ma) (Huber and Wunderlich, 2006), *Quamtana* sp. (min. age 48–55 Ma) (Penney, 2006) and *Pimoa* sp. (min. age 35–40 Ma) (Wunderlich, 2008b). In addition to the fossil information, in the case of Nephilidae biogeographical patterns have been used to hypothesize a similar minimum age (> 160 Ma) for that lineage (Kuntner, 2006) a few years before the actual fossil was discovered.

Dimitrov et al. (2008) studied the tempo and mode of the *Pholcus* radiation in Macaronesian archipelagoes. Their results together with information on the geological age of the islands (Anguita and Hernán, 1975; Cantagrel et al., 1993) were used to set a constraint for the age of the Macaronesian clade in the present analyses. This constraint was applied using lognormal distribution where the highest probability density overlapped with the age estimates of Dimitrov et al. (2008), while allowing the maximum and minimum age to extend beyond these limits. Prior distribution parameters (see Table S7) were chosen in such way that the probability of the clade being younger than 5 Ma was decreasing sharply while there was a gradual decrease in probability of ages older than 13.5 Ma [2.5% of the prior probability distribution was in the range 16.24 Ma and up, allowing ages that would be closer to that of the older island in the archipelagos—Fuerteventura ca. 21–23 Ma (Cantagrel et al., 1993)].

All fossil constraints were implemented as a minimum age of the stem of the lineage to which the fossil was assigned (Renner, 2005; Donoghue and Benton, 2007). To accommodate the uncertainty in the position of the fossils along the stems and the age of their divergences from the extant species we have used probability distributions for the age priors. Both logarithmic and exponential distributions were explored. In addition, analyses without the Dysderidae calibration were run because the minimum age of 35–40 Ma may be an underestimate of the actual age of this family as most known fossils are placed in extant genera. Best trees from the ML analyses of the full and the gblocks-processed datasets were used as starting trees. In addition, a final round of dating analyses was run, strictly enforcing the topology of the ML best tree in order to compare the estimates from analyses that account for uncertainty in the relationships and analyses that do not consider topological uncertainties.

Results

Phylogenetic results

All analyses resulted in very similar topologies and here we use the tree from the ML analyses of the gblocked dataset to summarize the results (Figs 1 and 2). Differences from this tree are listed below and discussed in the next section; in addition, all remaining results are available in the supplementary material. Analyses of the full and the gblocked dataset converged to practically identical results with no major incongruence, which indicates that treatment of gaps as missing data has not biased the results of the analyses of the full dataset. An overview of support for the main groups from the different analytical treatments and data partitions is given in Table S4. The following discussion is centred on the results from analyses of the datasets derived from the concatenated individual gene matrices: the full dataset and the combined matrix resulting from the gblocks-processed gene matrices.

Our results confirm the monophyly of Pholcidae; when the full dataset is analysed, the family receives a high bootstrap support under ML (83) and a posterior probability of 1 (analyses of the gblocked dataset support the monophyly but do not provide high support). Only when low values of k are used in the

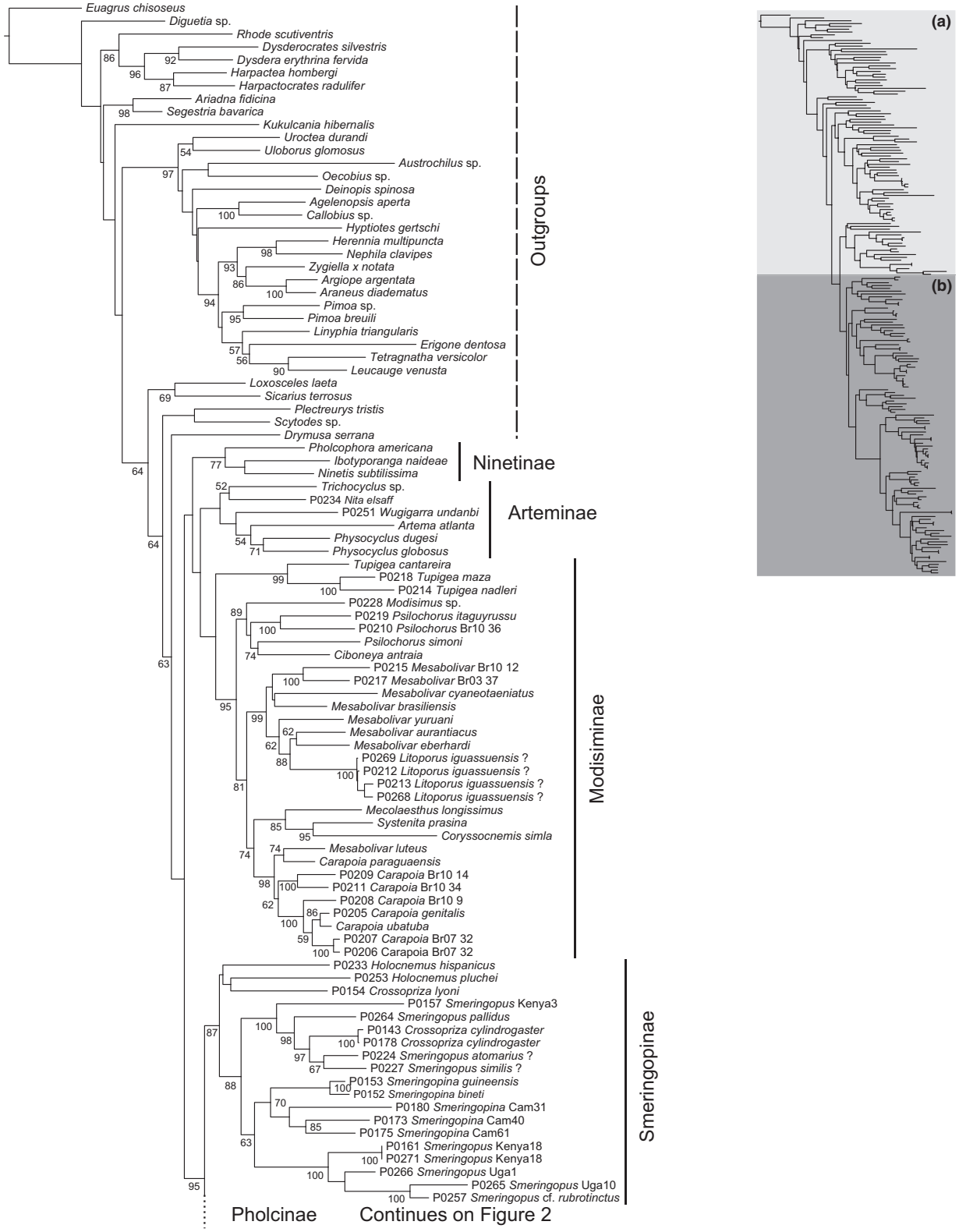

Fig. 1. ML tree from the analyses of the gblocked dataset; numbers at nodes are levels of bootstrap support. Top right miniature represents an overview of the full tree: (a) section shown here (lighter background); (b) section shown in Fig. 2 (Pholcinae). Unique identifiers for the specimens sequenced as part of this study are shown before the taxon name and start with the letter P followed by four digits. Specimens that were not identified to species level or new species include a code that follows the taxon name and refers to the morphospecies.

implied weights parsimony and under equal weights parsimony analyses of the gblocked dataset is the monophyly of pholcids compromised. The formerly proposed close relationship of diguetids and pholcids (e.g. Platnick et al., 1991; Starrett and Waters, 2007) is not recovered; diguetids are placed close to dysderids instead. The closest relatives of pholcids according to our results are drymusids, but this node is not well supported (did not receive a bootstrap support > 70 and/or posterior probability > 0.95). Plecteurids, which were also hypothesized as potentially closely related (Platnick et al., 1991), are placed in a clade together with scytodids and this clade is sister to the (Drymusidae + Pholcidae) in the Bayesian and ML trees. Parsimony analyses under equal weights differ in the position of plecteurids but this topology does not receive support, and resolution under implied weights converges to the ML and Bayesian results. All formally described pholcid subfamilies Ninetinae, Arteminae, Modisiminae, Smeringopinae and Pholcinae (Huber, 2011b) are recovered as monophyletic, albeit with a few differences in generic composition. Ninetinae is either sister to all remaining pholcids (in the Bayesian results of the full dataset and parsimony under implied weights), a member of a large clade together with Arteminae and Modisiminae or in a lineage comprising Ninetinae, Arteminae, Smeringopinae and Pholcinae (gblocked data under equal weights). None of these resolutions is well supported. The Middle Eastern genus *Nita*, which is currently in Ninetinae, is sister to *Trichocyclus* and is therefore a member of Arteminae. Another addition to Arteminae is the Australian genus *Wugigarra* [currently in Modisiminae (Huber, 2011b)]. Bayesian analyses included *Artema* in Modisiminae (probably an artefact, see Discussion). Modisiminae is monophyletic with the exclusion of *Wugigarra* and well supported. Relationships of the Modisiminae taxa included in our analyses are well resolved and highly supported but some genera change in composition. The North and South American *Psilochorus* are not monophyletic; *Mesabolivar luteus* is nested within *Carapoia* and *Litoporus iguassuensis* is placed within *Mesabolivar*. The monophyly of Smeringopinae and sister-group relationship with Pholcinae are well supported. Our analyses suggest that *Smeringopus* includes *Crossopriza cylindrogaster*. *Smeringopus* appears paraphyletic with respect to *Smeringopina*. *Holocnemus* remains problematic and the two taxa included do not form a clade.

Pholcinae is monophyletic but deeper intergeneric relationships within the subfamily receive erratic support. *Nyikoa* and *Zatavua* are not sister to the remaining pholcines as previously suggested (Huber, 2003b, 2007) but members of a larger clade that includes *Anansus* and possibly also *Khorata* and *Metagonia*. *Spermophora* has long been recognized as polyphyletic (e.g. Huber, 2003a, 2005c, 2011b) and our results confirm this. Several "*Spermophora*" lineages receive high support: Central African species are monophyletic and apparently closely related to *Spermophorides*; East African species are sister to the East African genus *Buitinga;* and the remaining African taxa are in a clade that contains *Belisana*, *Buitinga,* other "*Spermophora*" lineages and *Paramicromerys*. Within that group, in the results of the Bayesian analyses, *Belisana* and the "*Spermophora*" from Cameroon and Uganda form a poorly supported clade. *Quamtana* species are grouped together except for *Quamtana filmeri* whose placement is not resolved. *Micropholcus*, *Leptopholcus* and *Pehrforsskalia* form a monophyletic lineage that is sister to *Pholcus*. All analytical treatments find high support for a sister-group relationship of *Micropholcus* and the New World species of *Leptopholcus*, while *Pehrforsskalia* is closely related to the Old World *Leptopholcus*. *Pholcus* is monophyletic but relationships of lineages within the genus remain mostly unresolved or poorly supported and under implied weighting with low k values it includes *Leptopholcus*, *Pehrforsskalia* and *Micropholcus*. Despite our still limited taxon sampling, many of the species groups hypothesized based on morphological grounds (see Huber, 2011a) are recovered, some with high support. The *Pholcus bamboutos* group is not monophyletic as already suggested in its original description (Huber, 2011a); *Pholcus kribi* is closely related to *Pholcus nkoetye* (originally assigned to the *circularis* group) and *Pholcus kakum* is closely related to the *chappuisi* group. *Pholcus opilionoides* is nested within the *crypticolens* group. All Asian species except *Ph. atrigularis* and those in the *crypticolens* group form a well-supported monophyletic lineage which is sister to the remaining *Pholcus* (only the ML analysis of the full dataset finds a different but poorly supported topology). In all analytical treatments the Macaronesian and North American species form a fairly well-supported clade that is closely related to the East African *chappuisi* group and the West African *Ph. kakum*.

Molecular dating

Results from the molecular clock analyses under exponential and lognormal error distributions for the *tmrca* (time of the most recent common ancestor) prior were similar but using exponential distribution (Fig. S10) resulted in age estimates that were slightly (< 10%), but consistently, younger (95% confidence intervals were overlapping). Analyses under exponential prior were also more sensitive to the exclusion of the Dysderidae calibration. Use of lognormal priors resulted in slightly older age estimates and results were not affected by the Dysderidae calibration. Lognormal distributions better accommodate the error associated with the underestimation of the age of the fossil used to establish the dating constraint. For these reasons we

present here the results of the molecular clock analyses based on lognormal priors. BEAST analyses using as a starting tree the results of the ML analyses of the full and the gblock-trimmed dataset resulted in very similar age estimates and topologies. The gblocked data contain fewer ambiguities and missing data (in BEAST, gaps are treated as missing data in the same way as in ML) and allow for better estimation of branch lengths, therefore the result of BEAST using as starting tree the result of the ML analyses form the gblocked dataset is chosen to summarize the molecular dating results and discuss the tempo of evolution in pholcids (results of the analyses with exponential priors and the analyses of the full dataset are available as supplementary material).

Pholcids start diversifying some 207 Ma ago (95% HPD: 185–228 Ma) and by the mid Jurassic all of the modern subfamilies have emerged (Figs 3 and 4). Many genera appear to be old with origins in the early Palaeogene or late Cretaceous, and most of the major diversification events within subfamilies have happened during the Jurassic or early Cretaceous. Repeating the same analyses but enforcing the topology from the ML gblocked tree resulted in overlapping age estimates.

Discussion

Relationships

The present analyses are not tailored to resolve the phylogenetic relationships of haplogyne spiders because they do not include representatives of all haplogyne lineages. However, we have included a large number of outgroup taxa and our results suggest that current hypotheses for the closest relatives of pholcids need further investigation. None of our phylogenetic reconstructions placed diguetids close to pholcids; instead, drymusids are consistently recovered as the closest relatives. To test this relationship more stringently, additional haplogyne lineages need to be added in future analyses.

Despite the considerable increase in taxon and character sampling in comparison with previous molecular analyses with similar scope in pholcids (Bruvo-Madaric et al., 2005; Astrin et al., 2007), relationships among subfamilies are still not robustly resolved. The only exception is the well-supported sister-group relationship of Pholcinae and Smeringopinae. Both subfamilies have a comb on the fourth tarsus where comb hairs are spread over most of the tarsus length (comb hairs are limited to distal patches in other comb-wearing lineages) and this morphological feature was suggested as a potential synapomorphy of this group (Huber and Fleckenstein, 2008). The strong support coming from the molecular analyses presented here further strengthens this hypothesis.

The monophyly of Ninetinae is well supported by current and previous molecular analyses (Bruvo-Madaric et al., 2005; Astrin et al., 2007) as well as by several morphological synapomorphies (Huber, 2000) but the scope of the subfamily is uncertain. Huber (2011b) included 16 genera in Ninetinae but most of those have never been included in phylogenetic analyses. Huber and Brescovit (2003) have suggested that Ninetinae may be polyphyletic as currently defined. This is confirmed by the strong evidence in our results that *Nita* is a member of Arteminae. The relatively huge male palpal femur of *Nita* (Huber and El Hennawy, 2007) supports its inclusion in Arteminae and also points in that direction. Further changes in the composition of Ninetinae are likely to occur with inclusion of more taxa in future analyses.

Previous studies have always found support for Arteminae (Huber, 2001; Bruvo-Madaric et al., 2005; Astrin et al., 2007) and this is confirmed here. In addition to *Nita*, the Australian genus *Wugigarra* (previously in Modisiminae) is also newly transferred to Arteminae. The shape of the procursus in *Wugigarra* is similar to that of Arteminae (Huber, 2001) and its Australian distribution was in conflict with the former placement in Modisiminae—an otherwise exclusively New World clade. The transfer to Arteminae resolves both inconsistencies. *Trichocyclus* and *Physocyclus* have been treated as sister taxa based on the reduction of epiandrous spigots and some early molecular evidence (Huber, 2001; Bruvo-Madaric et al., 2005). Our expanded analyses suggest that *Physocyclus* is more closely related to *Artema* than to *Trichocyclus*.

Modisiminae is monophyletic except in the Bayesian analyses where *Artema* is sister to *Tupigea*. There is no morphological evidence that supports such a placement for *Artema* and this is probably an artefact of the analyses [it has been shown that high posterior probabilities can result for arbitrary clades (Lewis et al., 2005)]. Modisiminae now includes only New World taxa and in most of the analyses it is sister to Arteminae. Within Modisiminae, the Brazilian species currently placed in *Psilochorus* are close to but not sister of *Psilochorus sensu stricto*, which is restricted to North America. Other Modisiminae genera also need further attention; *Litoporus iguassuensis*, a species currently classified as *Tupigea*, is nested within *Mesabolivar;* it is a biologically very aberrant species in which males have been discovered only recently (A. Pérez, R. Baptista and B.A. Huber, unpublished data). *Mesabolivar luteus* seems to belong to *Carapoia*. Female morphology and preliminary molecular data have already pointed in this direction (Huber, 2005b; Astrin et al., 2007) but male characters rather pointed towards *Mesabolivar*. The recent discovery of a new species that shares with *M. luteus* a highly derived (inverted) resting position and whose male characters also point towards *Carapoia*

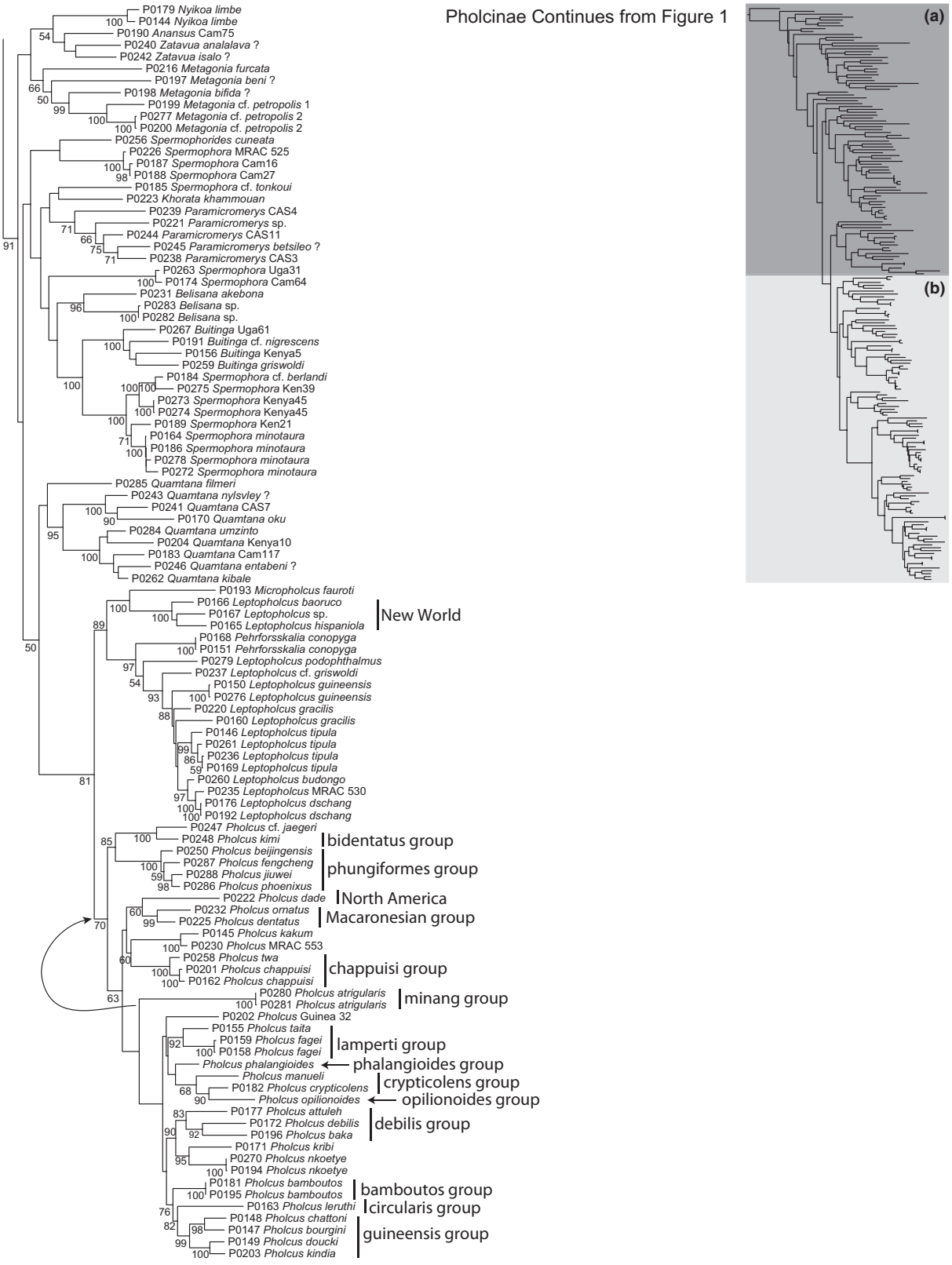

Fig. 2. ML tree from the analyses of the gblocked dataset; numbers at nodes are levels of bootstrap support. Arrow shows alternative placement for *Ph. atrigularis* when analysing the full dataset. Top right miniature represents an overview of the full tree: (a) section shown in Fig. 1; (b) Pholcinae shown here (lighter background). For explanation of codes accompanying species names see legend to Fig. 1.

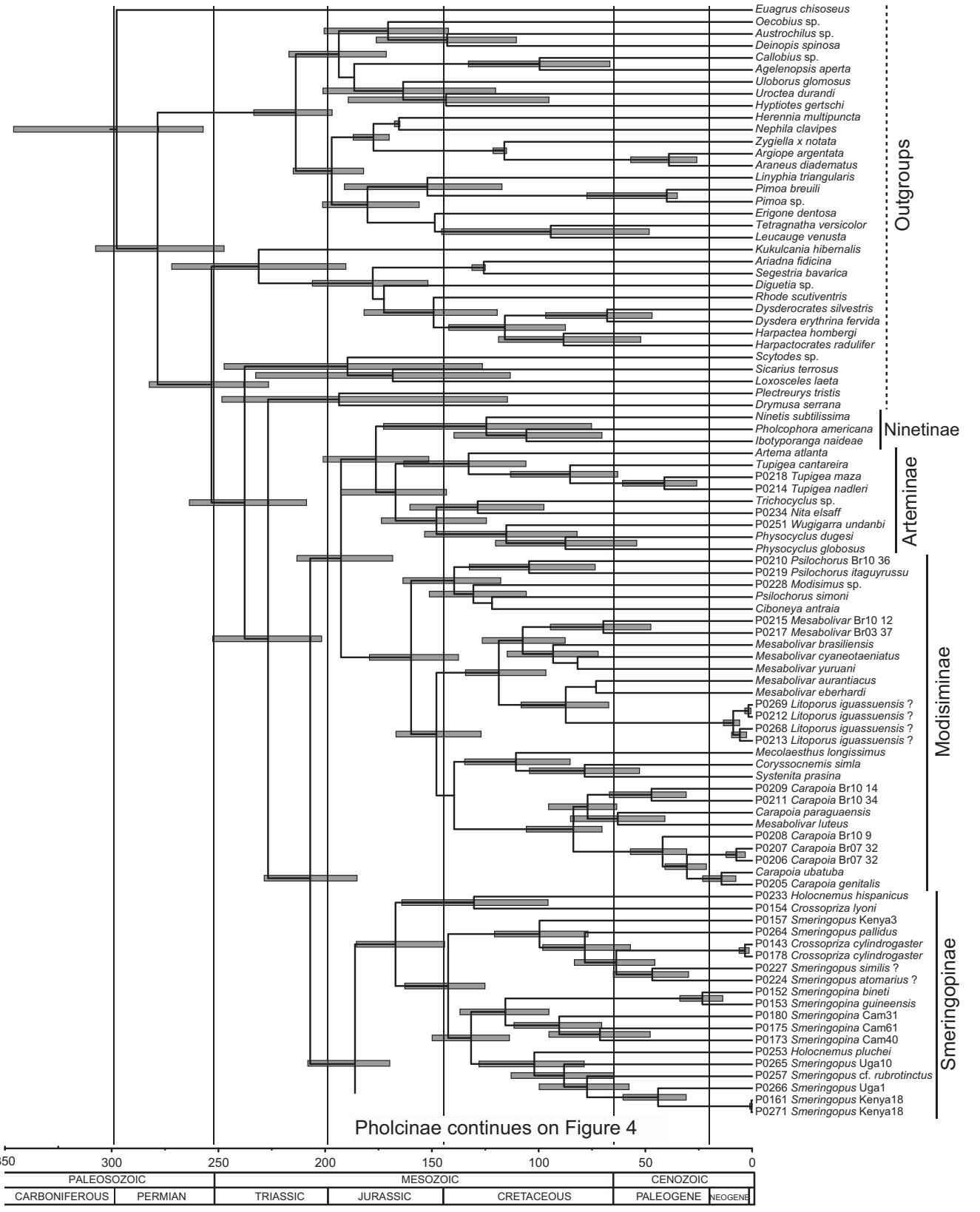

Fig. 3. Chronogram of pholcids based on the gblocked dataset and lognormal prior distributions for the fossil constraints. Bars at nodes are 95% confidence intervals. For explanation of codes accompanying species names see legend to Fig. 1.

(B.A. Huber, personal observation) strengthen this view. This is interesting because *M. luteus* and the new species are not dark leaf-litter dwellers as "typical" *Carapoia* but light leaf-dwellers; this is an indication that a shift between microhabitats has occurred in *Carapoia* but the direction remains unknown.

Smeringopinae monophyly is well supported and the subfamily is sister to Pholcinae. However, of the four genera from this subfamily included in the present analysis only *Smeringopina* is recovered as monophyletic. This is the first study to analyse molecular characters in *Smeringopina* and the results suggest that *Smeringopina* is closely related to or even nested within *Smeringopus*; the two genera also share unique morphological characters (Huber, in press). *Crossopriza cylindrogaster* is also nested within *Smeringopus* and an ongoing revision of *Smeringopus* (Huber, in press) supports this position by several morphological characters. Again, this is a case of habitat shift and the unusual general appearance of the light, leaf-dwelling *C. cylindrogaster* has masked its relationship with the darker, near-ground-dwelling closest relatives. *Holocnemus* continues to be problematic (cf. Astrin et al., 2007), and our results support a preliminary cladistic analysis of morphological characters suggesting that *Holocnemus* is not monophyletic (B.A. Huber, unpublished data).

Pholcinae deep relationships remain elusive but some structure within the subfamily is starting to emerge. This is the first phylogenetic analysis to include molecular data of the genus *Anansus*. It belongs to a Pholcinae lineage that also includes *Nyikoa* and *Zatavua*; the latter two were previously considered sister to the remaining pholcines (Huber, 2003b, 2007) and here, together with *Anansus*, *Khorata* and *Metagonia*, also appear as sister to the other members of this subfamily. One of the main goals of the current study was to shed light on the limits and relationships of two large Pholcinae lineages: *Pholcus* and *Spermophora*. Although we could not obtain samples from any "true" East Asian *Spermophora* [i.e. from one that shares putative morphological synapomorphies with *Spermophora senoculata*, the type species of the genus; see Huber (2005c)], many of the Asian and African lineages have been sampled and the present results represent a large step towards building a robust phylogenetic hypothesis for the species currently grouped in this genus. In some cases, geographical closeness seems to be a better predictor of phylogenetic relationships than morphology. For example, East African "*Spermophora*" seem to be sister of the East African endemic genus *Buitinga*. Morphology alone failed to resolve this close relationship (Huber, 2003a).

Pholcus is by far the largest genus of pholcids and currently includes 294 species (Huber, 2011a). Together with ten other genera that share specific morphological synapomorphies (*Leptopholcus*, *Micropholcus*, *Pehrforsskalia*, etc.) it has been treated as the "*Pholcus* group of genera" (Huber, 2011a; b). Some species groups currently assigned to *Pholcus* may in fact be more closely related to one of the other genera in this clade than to the type species of *Pholcus* (Huber, 2011a). Unfortunately, the present dataset is missing some of the closely related Asian and Indo-Australian genera (*Calapnita*, *Micromerys*, *Panjange*, *Sihala* and *Uthina*) and barely includes South-East Asian *Pholcus* species due to the lack of specimens suitable for DNA extraction; rigorous tests of the relationships of all "*Pholcus* group genera" will be possible only when these taxa are included in future analyses. With the current taxon sampling we find *Pholcus* to be monophyletic in all analytical treatments but this is well supported only in the results of the gblocked dataset. Our limited sampling supports several of the species groups that have been proposed (*bidentatus*-group, *phungiformes*-group, *chappuisi*-group, *lamperti*-group, *debilis*-group, *guineensis*-group, etc; see Fig. 2), but not the relationships among groups previously suggested (Huber, 2011a). In our results, East Asian species are sister to the remaining lineages while most African species appear to be more derived. One particularly interesting finding is the close relationship of the North American *kingi*-group (represented by *Ph. dade*) and the Macaronesian species group (represented by *Ph. dentatus* and *Ph. ornatus*; see Fig. 2). This clade is in turn related to a lineage including only African species. Most autochthonous North American *Pholcus* species were described very recently (Huber, 2011a) and not included in previous phylogenetic studies of Macaronesian *Pholcus* (Dimitrov and Ribera, 2007; Dimitrov et al., 2008). Therefore, the question regarding the origin of the Macaronesian *Pholcus* remains open. Geographically, Africa is the closest continent but colonization of North America from some of the Macaronesian archipelagoes or *vice versa*, although rare, has been documented (e.g. Panero et al., 1999; Stech et al., 2011).

The polyphyly of *Leptopholcus* has been suggested for some time (e.g. Huber, 2000, 2011a) but the placement of *Micropholcus* as sister to the New World "*Leptopholcus*" is new and surprising. The two lineages have a highly different general morphology and in fact this dissimilarity is so strong that a potential synapomorphy [a highly modified hair at the tip of the male palpal trochanter apophysis; see Huber (2000, figs 105 and 106)] has never been seriously considered to be homologous. The African *Pehrforsskalia* is sister to the remaining Old World *Leptopholcus*.

Biogeographical considerations

Pholcids are an old lineage that—according to our analysis—originated some 207 Ma ago (95% HPD: 185–228 Ma). This age is much older than the oldest known

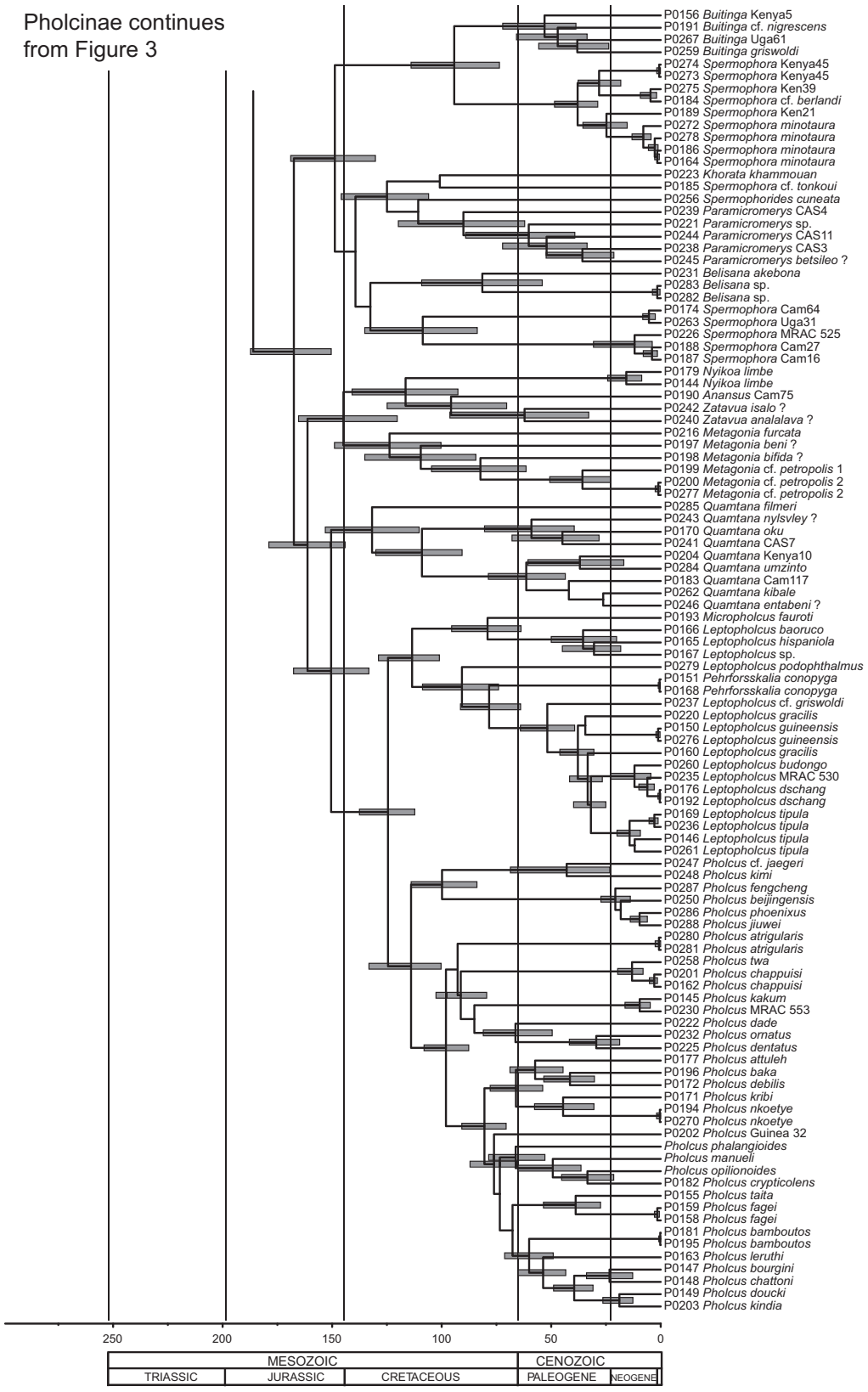

Fig. 4. Chronogram of Pholcinae based on the gblocked dataset and lognormal prior distributions for the fossil constraints (continued from Fig. 3). Bars at nodes show 95% confidence intervals. For explanation of codes accompanying species names see legend to Fig. 1.

fossils of pholcids but in line with recent findings based on molecular dating, which show that many of the spider families may be much older than suggested by the fossil record alone (e.g. Ayoub et al., 2007; Dimitrov et al., 2012). The oldest known pholcid fossils are < 60 Ma old and have been described from Palaeogene (Baltic, Le Quesnoy and Fu Shun) and Neogene (Dominican) ambers (Dunlop et al., 2012). Ayoub et al. (2007) included several haplogyne representatives in their analyses and concluded that they were about 240 Ma old, suggesting the haplogyne origin "… possibly dating as far back as the Carboniferous". We included several haplogyne lineages in our analyses and results point in the same direction. Nonetheless, it is important to note that while fossils may often underestimate the time of divergence among lineages, molecular data are generally prone to overestimate it (Steiper and Young, 2008). Hence, slight overestimation of some divergence ages in our results cannot be completely ruled out. For example, the estimated age of the Macaronesian lineage in our results appears to be older than the age of the oldest islands in the archipelagos and also exceeds by a few million years previous estimates for the age of the main Macaronesian *Pholcus* lineage (Dimitrov et al., 2008). Dimitrov et al. (2008), however, suggested that the Macaronesain *Pholcus* radiation may be older as they were not able to access the age of the early branching events in that group.

The early pholcid diversification inferred by our analyses also has important biogeographical implications. At that time all continental landmasses were still accreted into the supercontinent Pangea, and hence some of the early diversification events might have happened before the continents started drifting apart. This could explain the wide and partly fragmented distribution of some old pholcid lineages such as Ninetinae, Arteminae and Pholcinae. Early diversification might have taken place in Pangea and as a result of its fragmentations different lineages are now found on different continents. This may also explain the current distributional patterns at the subfamily level. At present all subfamilies except Smeringopinae are found on more than one major landmass that is not restricted to Laurasian or Gondwanian origin. For example, Modisiminae is found in both South and North America and lineages in North America are too old to be a result of northward expansion with the formation of the Panama isthmus or *vice versa*. Smeringopinae is exceptional in being restricted to landmasses of predominantly Gondwanian origin; the vast majority of species are restricted to Africa. Pangea fragmentation may also have been the trigger for some of the deeper diversification events within subfamilies (e.g. separation of New World *Metagonia* and Malagasy *Zatavua* from other Pholcinae).

Pholcids are sedentary and often live in very cryptic habitats (e.g. under logs or stones, under leaves, in caves, and in holes on tree bark). There is no evidence that they do disperse by ballooning (Schäfer et al., 2001), and certain species groups of *Modisimus* on Hispaniola seem to have retained their distributions basically unchanged for about 20 Myr (Huber et al., 2010). Therefore, it seems logical to invoke vicariance as an explanation of pholcid biogeographical patterns but the role of dispersal, especially when studying more recent events, should not be completely overrun. There are a few cases in which pholcids have been able to disperse to oceanic islands, some of which are highly isolated (Beatty et al., 2008; Dimitrov et al., 2008). However, the wide distribution of *Ph. ancoralis* in the Pacific may be a result of anthropogenic influence (Huber, 2011a), and the monophyly of the Macaronesian species also suggests limited dispersal abilities. Another factor that may have played a role in setting the large-scale distribution limits of the family may be phylogenetic niche conservatism [i.e. the propensity of species to retain fundamental niche preferences similar or identical to those of their ancestors (Wiens and Donoghue, 2004; Wiens and Graham, 2005)]. A circumtropical distribution of old lineages and low clade and species diversity in temperate or frigid climates would support this hypothesis with respect to the thermal dimension of the niche. One interesting observation is the apparently low number of dispersals, including human-mediated dispersals, in lineages adapted to very humid climates. It is possible that adaptation to drier conditions has allowed some species to tolerate more open environments and climates with more variable humidity (some *Pholcus* species, *Artema atlanta*, *Holocnemus pluchei*, *Smeringopus pallidus*, etc.), which has facilitated range expansion and dispersal.

Diversification of pholcids and evolution of microhabitat preferences

Even though historical biogeographical events have probably played an important role in shaping current distribution patterns of pholcids, they cannot account for the exceptional species diversity within biogeographical regions. Two main mechanisms that may explain species diversity at that scale have been proposed: speciation through adaptive radiation (Schluter, 2000) and speciation through non-adaptive processes (e.g. Kozak et al., 2006). In cases of adaptive radiations different species may, for example, inhabit different environments and exhibit differences in traits that are directly related to the way they interact with their surroundings. In contrast, when speciation has been driven by non-adaptive processes, different species live in the same or very similar environments and morphological differences, if present, are not related to the way they interact with the environment. Previous studies on pholcids and *Pholcus* in particular have revealed the

importance of non-adaptive speciation and the role of sexual selection in shaping species diversity and morphological disparity in this group (e.g. Huber, 1999, 2005d; Dimitrov et al., 2008). The role of adaptive speciation in the generation of pholcid diversity has been rather poorly explored. The present results suggest that there have been numerous cases of shifts in microhabitat preferences where leaf-dwelling species have litter-dwelling ancestors or *vice versa*. For example, the leaf-litter species *Pholcus kribi* and *Metagonia petropolis* each have leaf-dwelling ancestors, while the leaf-dwelling "*Crossopriza*" *cylindrogaster* is derived from near-ground-dwelling ancestors. This suggests that in at least some genera, speciation through adaptive radiation has occurred and has contributed to morphological diversity in these groups. Often such changes have led to convergent phenotypes as some microhabitats favour development of particular morphological characteristics [e.g. short legs, globular abdomen and dark coloration in litter dwellers vs. long legs, long abdomen and pale greenish coloration in leaf dwellers (Huber, 2003a, 2005a, 2009, 2011a; Huber et al., 2010)].

Future perspectives

One of the main goals of the present study was to provide an improved framework for the design of future research in pholcid systematics, macroecology and macroevolution; or in other words to identify some of the gaps in current knowledge and set up future priorities. A positive result in this respect is that it seems we are already on the right track by increasing taxon and gene sampling. Table S4, which summarizes levels of support from different data partitions, shows how important it is to use multiple genes in the analyses; there is not a single gene that reliably recovers relationships across pholcids, yet combining different genes that provide information at different levels results in a fairly well-resolved phylogeny. Further efforts to increase taxon and gene sampling are essential to gain a better understanding of the drivers of diversification in pholcids and also to identify monophyletic lineages in problematic genera. Last but not least we need to invest significantly more effort in obtaining better distributional data and information on the natural history of different species. Information on distribution and species traits, together with dated phylogenies, is essential and will provide us with a powerful tool to study the relative importance and interrelation of adaptive and non-adaptive speciation processes in this group.

Acknowledgements

We thank the Associate Editor, L. Prendini, M.A. Arnedo and one anonymous reviewer for comments on an earlier version of the manuscript. B.A.H. thanks the numerous people who have either assisted during collection trips or have kindly provided DNA-ready material: D. Court, B. Cutler, R. Duncan, J. Godwin, C.E. Griswold, M. Harvey, P. Jäger, R. Jocqué, R. Kamga, P. LeGall, F.J. Muafor, R. Mwakodi, A. Pérez G., R. Raven, M. Sidibe and C. Warui. We thank Hannah Wood for commenting on an earlier draft of the manuscript. We thank Christina Blume for lab work. The German Research Foundation provided financial support (DFG, HU 980/9-1). The Danish National Research Foundation provided additional support to D.D.

References

Álvarez-Padilla, F., Dimitrov, D., Giribet, G., Hormiga, G., 2009. Phylogenetic relationships of the spider family Tetragnathidae (Araneae, Araneoidea) based on morphological and DNA sequence data. Cladistics 25, 109–146.

Anguita, F., Hernán, F., 1975. A propagating fracture model versus a hot spot origin for the Canary Islands. Earth Planet. Sci. Lett. 27, 11–19.

Arnedo, M.A., Hormiga, G., Scharff, N., 2009. Higher-level phylogenetics of linyphiid spiders (Araneae, Linyphiidae) based on morphological and molecular evidence. Cladistics 25, 231–262.

Astrin, J.J., Huber, B.A., Misof, B., Klutsch, C.F.C., 2006. Molecular taxonomy in pholcid spiders (Pholcidae, Araneae): evaluation of species identification methods using CO1 and 16S rRNA. Zool. Scr. 35, 441–457.

Astrin, J.J., Misof, B., Huber, B.A., 2007. The pitfalls of exaggeration: molecular and morphological evidence suggests *Kaliana* is a synonym of *Mesabolivar* (Araneae: Pholcidae). Zootaxa 1646, 17–30.

Ayoub, N.A., Garb, J.E., Hedin, M., Hayashi, C.Y., 2007. Utility of the nuclear protein-coding gene, elongation factor-1 gamma (EF-1 [gamma]), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae). Mol. Phylogenet. Evol. 42, 394–409.

Beatty, J.A., Berry, J.W., Huber, B.A., 2008. The pholcid spiders of Micronesia and Polynesia (Araneae, Pholcidae). J. Arachnol. 36, 1–25.

Bruvo-Madaric, B., Huber, B.A., Steinacher, A., Pass, G., 2005. Phylogeny of pholcid spiders (Araneae: Pholcidae): combined analysis using morphology and molecules. Mol. Phylogenet. Evol. 37, 661–673.

Cantagrel, J.M., Fuster, J.M., Pin, C., Renaud, U., Ibarrola, E., 1993. Age Miocène inférieur des carbonatites de Fuerteventura (23 Ma: U-Pb zircon) et le magmatisme précoce d'une île océanique (îles Canaries). CR Acad. Sci. Paris 316, 1147–1153.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540.

Dimitrov, D., Hormiga, G., 2011. An extraordinary new genus of spiders from Western Australia with an expanded hypothesis on the phylogeny of Tetragnathidae (Araneae). Zool. J. Linn. Soc. 161, 735–768.

Dimitrov, D., Ribera, C., 2007. The genus *Pholcus* (Araneae, Pholcidae) in the Canary Islands. Zool. J. Linn. Soc. 151, 59–114.

Dimitrov, D., Arnedo, M.A., Ribera, C., 2008. Colonization and diversification of the spider genus *Pholcus* Walckenaer, 1805 (Araneae, Pholcidae) in the Macaronesian archipelagos: evidence

for long-term occupancy yet rapid recent speciation. Mol. Phylogenet. Evol. 48, 596–614.
Dimitrov, D., Lopardo, L., Giribet, G., Arnedo, M.A., Álvarez-Padilla, F., Hormiga, G., 2012. Tangled in a sparse spider web: single origin of orb weavers and their spinning work unravelled by denser taxonomic sampling. Proc. R. Soc. B 279, 1341–1350.
Donoghue, P.C.J., Benton, M.J., 2007. Rocks and clocks: calibrating the tree of life using fossils and molecules. Trends Ecol. Evol. 22, 424–431.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Dunlop, J.A., Penney, D., Jekel, D., 2012. A summary list of fossil spiders and their relatives. In: Platnick, N.I. (Ed.), The World Spider Catalog, Version 12.5. American Museum of Natural History, available at: http://research.amnh.org/entomology/spiders/catalog/index.html (accessed 25 May 2012).
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
Goloboff, P.A., 1993. Estimating character weights during tree search. Cladistics 9, 83–91.
Goloboff, P.A., 1999. Analyzing large data sets in reasonable times: solutions for composite optima. Cladistics 15, 415–428.
Goloboff, P.A., Farris, J.S., Källersjö, M., Oxelman, B., Ramírez, M.J., Szumik, C.A., 2003. Improvements to resampling measures of group support. Cladistics 19, 324–332.
Goloboff, P.A., Farris, J.S., Nixon, K.C., 2008. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.
Huber, B.A., 1994. Genital morphology, copulatory mechanism and reproductive biology in *Psilochorus simoni* (Berland, 1911) (Pholcidae; Araneae). Netherl. J. Zool. 44, 85–99.
Huber, B.A., 1995. Copulatory mechanism in *Holocnemus pluchei* and *Pholcus opilionoides*, with notes on male cheliceral apophyses and stridulatory organs in Pholcidae (Araneae). Acta Zool. 76, 291–300.
Huber, B.A., 1996a. Genitalia, fluctuating asymmetry, and patterns of sexual selection in *Physocyclus globosus* (Araneae: Pholcidae). Rev. Suisse. Zool. 1996, 289–294.
Huber, B.A., 1996b. Evidence for gustatorial courtship in a haplogyne spider (*Hedypsilus culicinus*: Pholcidae: Araneae). Netherl. J. Zool. 47, 95–98.
Huber, B.A., 1999. Sexual selection in pholcid spiders (Araneae, Pholcidae): artful chelicerae and forceful genitalia. J. Arachnol. 13, 5–141.
Huber, B.A., 2000. New World pholcid spiders (Araneae: Pholcidae): a revision at generic level. Bull. Am. Mus. Nat. Hist. 254, 1–348.
Huber, B.A., 2001. The pholcids of Australia (Araneae; Pholcidae): taxonomy, biogeography, and relationships. Bull. Am. Mus. Nat. Hist. 260, 1–144.
Huber, B.A., 2002. Functional morphology of the genitalia in the spider *Spermophora senoculata* (Pholcidae, Araneae). Zool. Anz. 241, 105–116.
Huber, B.A., 2003a. High species diversity in one of the dominant groups of spiders in East African montane forests (Araneae: Pholcidae: *Buitinga* n. gen., *Spermophora* Hentz). Zool. J. Linn. Soc. 137, 555–619.
Huber, B.A., 2003b. Cladistic analysis of Malagasy pholcid spiders reveals generic level endemism: Revision of *Zatavua* n. gen. and *Paramicromerys* Millot (Pholcidae, Araneae). Zool. J. Linn. Soc. 137, 261–318.
Huber, B.A., 2005a. High species diversity, male-female coevolution, and metaphyly in Southeast Asian pholcid spiders: the case of *Belisana* Thorell 1898 (Araneae, Pholcidae). Zoologica 155, 1–126.
Huber, B.A., 2005b. Revision and cladistic analysis of the spider genus *Carapoia* González-Sponga (Araneae : Pholcidae), with descriptions of new species from Brazil's Atlantic forest. Invertebr. Syst. 19, 541–556.
Huber, B.A., 2005c. Revision of the genus *Spermophora* Hentz in Southeast Asia and on the Pacific Islands, with descriptions of three new genera (Araneae: Pholcidae). Zool. Med. 79-2, 61–172.
Huber, B.A., 2005d. Sexual selection research on spiders: progress and biases. Biol. Rev. 80, 363–385.
Huber, B.A., 2007. Two new genera of small, six-eyed pholcid spiders from West Africa, and first record of *Spermophorides* for mainland Africa (Araneae: Pholcidae). Zootaxa 1635, 23–43.
Huber, B.A., 2009. Life on leaves: leaf-dwelling pholcids of Guinea, with emphasis on *Crossopriza cylindrogaster* Simon, a spider with inverted resting position, pseudo-eyes, lampshade web, and tetrahedral egg-sac (Araneae: Pholcidae). J. Nat. Hist. 43, 2491–2523.
Huber, B.A., 2011a. Revision and cladistic analysis of *Pholcus* and closely related taxa (Araneae, Pholcidae). Bonn. Zool. Monogr. 58, 1–509.
Huber, B.A., 2011b. Phylogeny and classification of Pholcidae (Araneae): an update. J. Arachnol. 39, 211–222.
Huber, B.A., In press. Revision and cladistic analysis of the Afrotropical endemic genus *Smeringopus* Simon, 1890 (Araneae: Pholcidae). Zootaxa.
Huber, B.A., Astrin, J.J., 2009. Increased sampling blurs morphological and molecular species limits: revision of the Hispaniolan endemic spider genus *Tainonia* (Araneae: Pholcidae). Invertebr. Syst. 23, 281–300.
Huber, B.A., Brescovit, A.D., 2003. *Ibotyporanga* Mello-Leitão: tropical spiders in Brazilian semi-arid habitats (Araneae: Pholcidae). Insect Syst. Evol. 34, 15–20.
Huber, B.A., El Hennawy, H., 2007. On Old World ninetine spiders (Araneae: Pholcidae), with a new genus and species and the first record for Madagascar. Zootaxa 1635, 45–53.
Huber, B.A., Fleckenstein, N., 2008. Comb-hairs on the fourth tarsi in pholcid spiders (Araneae, Pholcidae). J. Arachnol. 19, 232–240.
Huber, B.A., Wunderlich, J., 2006. Fossil and extant species of the genus *Leptopholcus* in the Dominican Republic, with the first cases of egg-parasitism in pholcid spiders (Araneae: Pholcidae). J. Nat. Hist. 40, 2341–2360.
Huber, B.A., Fischer, N., Astrin, J.J., 2010. High level of endemism in Haiti's last remaining forests: a revision of *Modisimus* (Araneae: Pholcidae) on Hispaniola, using morphology and molecules. Zool. J. Linn. Soc. 158, 244–299.
Huelsenbeck, J.P., Crandall, K.A., 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. Annu. Rev. Ecol. Syst. 28, 437–466.
Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.
Kozak, K.H., Weisrock, D.W., Larson, A., 2006. Rapid lineage accumulation in a non-adaptive radiation: phylogenetic analysis of diversification rates in eastern North American woodland salamanders (Plethodontidae: Plethodon). Proc. Biol. Sci. 273, 539–546.
Kuntner, M., 2006. Phylogenetic systematics of the Gondwanan nephilid spider lineage Clitaetrinae (Araneae, Nephilidae). Zool. Scr. 35, 19–62.
Lewis, P.O., Holder, M.T., Holsinger, K.E., 2005. Polytomies and Bayesian phylogenetic inference. Syst. Biol. 54, 241–253.
Mello-Leitão, C.F., 1946. Notas sobre os Filistatidae e Pholcidae. An. Acad. Bras. Ciênc. 18, 39–83.
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). New Orleans, LA, pp. 1–8.
Panero, J.L., Francisco-Ortega, J., Jansen, R.K., Santos-Guerra, A., 1999. Molecular evidence for multiple origins of woodiness and a New World biogeographic connection of the Macaronesian Island endemic *Pericallis* (Asteraceae: Senecioneae). Proc. Natl Acad. Sci. USA 96, 13886–13891.

Penney, D., 2006. The oldest fossil pholcid and selenopid spiders (Araneae) in Lowermost Eocene amber from the Paris Basin, France. J. Arachnol. 59, 2–598.
Penney, D., Ortuño, V.M., 2006. Oldest true orb-weaving spider (Araneae: Araneidae). Biol. Lett. 2, 447.
Petrunkevitch, A., 1928. Systema aranearum. Trans. Conn. Acad. Arts Sci. 29, 1–270.
Platnick, N.I., 2012. The World Spider Catalog, Version 12.5. American Museum of Natural History, New York, NY, available at: http://research.amnh.org/iz/spiders/catalog (accessed 25 May 2012).
Platnick, N.I., Coddington, J.A., Forster, R.R., Griswold, C.E., 1991. Spinneret morphology and the phylogeny of haplogyne spiders (Araneae, Araneomorphae). Am. Mus. Novit. 3016, 1–73.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
Rambaut, A., Drummond, A.J., 2007. Tracer v1.4, available at: http://beast.bio.ed.ac.uk/Tracer (accessed 25 May 2012).
Renner, S.S., 2005. Relaxed molecular clocks for dating historical plant dispersal events. Trends Plant Sci. 10, 550–558.
Ronquist, F., Huelsenbeck, J.P., 2004. MrBayes 3: A Tutorial. http://132.229. 93.11/taskforcemolecular/PDF/Ronquist% 20% Huelsenbeck% 20% 20% Mr Bayes% 203% 20 tutorial. pdf (accessed 8 March 2002).
Schäfer, M.A., Hille, A., Uhl, G.B., 2001. Geographical patterns of genetic subdivision in the cellar spider *Pholcus phalangioides* (Araneae). Heredity 86, 94–102.
Schluter, D., 2000. The Ecology of Adaptive Radiation. Oxford University Press, USA.
Selden, P.A., Shih, C., Ren, D., 2011. A golden orb-weaver spider (Araneae: Nephilidae: *Nephila*) from the Middle Jurassic of China. Biol. Lett. 7, 775–778.
Simon, E., 1893. Histoire naturelle des araignées, 2nd edn. Librairie Encyclopédique de Roret, Paris.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688.
Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57, 758.
Starrett, J., Waters, E.R., 2007. Positive natural selection has driven the evolution of the Hsp70s in *Diguetia* spiders. Biol. Lett. 3, 439–444.
Stech, M., Werner, O., Gonzalez-Mancebo, J.M., Patino, J., Sim-Sim, M., Fontinha, S., Hildebrandt, I., Ros, R.M., 2011. Phylogenetic inference in *Leucodon* Schwagr. subg. Leucodon (Leucodontaceae, Bryophyta) in the North Atlantic region. Taxon 60, 79–88.
Steiper, M.E., Young, N.M., 2008. Timing primate evolution: lessons from the discordance between molecular and paleontological estimates. Evol. Anthropol. 17, 179–188.
Uhl, G., 1998. Mating behaviour in the cellar spider, *Pholcus phalangioides*, indicates sperm mixing. Anim. Behav. 56, 1155–1159.
Uhl, G., Huber, B.A., Rose, W., 1995. Male pedipalp morphology and copulatory mechanism in *Pholcus phalangioides* (Fuesslin, 1775) (Araneae, Pholcidae). Bull. Br. Arachnol. Soc. 10, 1–9.
Warnow, T., 2012. Standard maximum likelihood analyses of alignments with gaps can be statistically inconsistent. PLoS Curr. 4, RRN1308.
Wiens, J.J., Donoghue, M.J., 2004. Historical biogeography, ecology and species richness. Trends Ecol. Evol. 19, 639–644.
Wiens, J.J., Graham, C.H., 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. Annu. Rev. Ecol. Evol. Syst. 36, 519–539.
Wunderlich, J., 2004. Fossil spiders (Araneae) of the superfamily Dysderoidea in Baltic and Dominican amber, with revised family diagnoses. Beitr. Araneol. 3, 633–746.
Wunderlich, J., 2008a. The dominance of ancient spider families of the Araneae: *Haplogyne* in the Cretaceous, and the late diversification of advanced ecribellate spiders of the Entelegynae after the Cretaceous–Tertiary boundary extinction events, with descriptions of new families. Beitr. Araneol. 5, 524–675.
Wunderlich, J., 2008b. Descriptions of fossil spider (Araneae) taxa mainly in Baltic amber, as well as certain related extant taxa. Beitr. Araneol. 5, 44–139.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. ML tree from the analyses of the full dataset; branch values represent bootstraps.

Figure S2. All compatible partitions tree from the Bayesian analyses of the full dataset; branch values represent posterior probabilities.

Figure S3. All compatible partitions tree from the Bayesian analyses of the gblocked dataset; branch values represent posterior probabilities.

Figure S4. Strict consensus of 214 equally parsimonious trees found under equal weights parsimony analyses of the full dataset; branch values represent jackknife proportions.

Figure S5. Strict consensus of 89 equally parsimonious trees found under equal weights parsimony analyses of the gblocked dataset; branch values represent jackknife proportions.

Figure S6. Strict consensus of four equally parsimonious trees found under implied weights parsimony analyses of the full dataset with $k = 6$; branch values represent symmetric resampling proportions.

Figure S7. Strict consensus of four equally parsimonious trees found under implied weights parsimony analyses of the full dataset with $k = 10$; branch values represent symmetric resampling proportions.

Figure S8. Strict consensus of four equally parsimonious trees found under implied weights parsimony analyses of the full dataset with $k = 20$; branch values represent symmetric resampling proportions.

Figure S9. Strict consensus of four equally parsimonious trees found under implied weights parsimony analyses of the full dataset with $k = 100$; branch values represent symmetric resampling proportions.

Figure S10. Chronogram of pholcids based on the gblocked dataset and exponential prior distributions for the fossil constrains.

Table S1. Species, family and GenBank accession numbers for the sequences used in the present analyses. n.a.—no sequences available (=missing data).

Tables S2. List of primers used in the present study.

Table S3. Best models selected by jModelTest.

Table S4. Support for the monophyly of the major clades in the different analytical treatments.

Table S5. Species names, museum collection codes and locality and collector data for the specimens used in this study.

Table S6. Number of characters, level of missing data and number of characters removed with Gblocks.

Table S7. Settings used to implement the dating constrains in the molecular clock analyses.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.