PROCEEDINGS ——— **THE ROYAL** **B** | **BIOLOGICAL**
—— **OF** ——— **SOCIETY** | **SCIENCES**

Tangled in a sparse spider web: single origin of orb weavers and their spinning work unravelled by denser taxonomic sampling

Dimitar Dimitrov, Lara Lopardo, Gonzalo Giribet, Miquel A. Arnedo, Fernando Álvarez-Padilla and Gustavo Hormiga

Supplementary data	"Data Supplement" http://rspb.royalsocietypublishing.org/content/suppl/2011/10/25/rspb.2011.2011.DC1.html
References	**This article cites 80 articles, 26 of which can be accessed free** http://rspb.royalsocietypublishing.org/content/279/1732/1341.full.html#ref-list-1
Subject collections	Articles on similar topics can be found in the following collections evolution (1184 articles) taxonomy and systematics (146 articles)
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Proc. R. Soc. B* go to: **http://rspb.royalsocietypublishing.org/subscriptions**

Tangled in a sparse spider web: single origin of orb weavers and their spinning work unravelled by denser taxonomic sampling

Dimitar Dimitrov[1,6,*], **Lara Lopardo**[2], **Gonzalo Giribet**[3], **Miquel A. Arnedo**[4], **Fernando Álvarez-Padilla**[5] and **Gustavo Hormiga**[6]

[1]*Center for Macroecology, Evolution and Climate, Zoological Museum, University of Copenhagen, Copenhagen, Denmark*
[2]*Zoologisches Institut und Museum, Allgemeine und Systematische Zoologie, Greifswald, Germany*
[3]*Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA*
[4]*Departament Biologia Animal, Biodiversity Research Institute UB, Universitat de Barcelona, Barcelona, Spain*
[5]*Facultad de Ciencias, Departamento de Biología Comparada, Universidad Nacional Autónoma de México, Mexico DF, Mexico*
[6]*Department of Biological Sciences, The George Washington University, Washington, DC, USA*

In order to study the tempo and the mode of spider orb web evolution and diversification, we conducted a phylogenetic analysis using six genetic markers along with a comprehensive taxon sample. The present analyses are the first to recover the monophyly of orb-weaving spiders based solely on DNA sequence data and an extensive taxon sample. We present the first dated orb weaver phylogeny. Our results suggest that orb weavers appeared by the Middle Triassic and underwent a rapid diversification during the end of the Triassic and Early Jurassic. By the second half of the Jurassic, most of the extant orb-weaving families and web designs were already present. The processes that may have given origin to this diversification of lineages and web architectures are discussed. A combination of biotic factors, such as key innovations in web design and silk composition, as well as abiotic environmental changes, may have played important roles in the diversification of orb weavers. Our analyses also show that increased taxon sampling density in both ingroups and outgroups greatly improves phylogenetic accuracy even when extensive data are missing. This effect is particularly important when addition of character data improves gene overlap.

Keywords: web evolution; spider diversification; taxon sampling; molecular dating

1. INTRODUCTION

Fossil evidence suggests that silk evolved in the common ancestor of spiders and their closest relatives, the extinct uraraneids, sometime in the Devonian [1]. Ever since silk has been inextricably entangled with the biology of spiders. From birth to death, from reproduction to foraging, much of the behaviour and lifestyle of spiders depends on silk production [2]. Nowhere is this better illustrated than in the orb-weaving spiders (Orbiculariae), with their highly geometric snare designs, often regarded as a pinnacle of animal engineering, and their ability to produce up to seven different types of silk with different functions and mechanical properties, which continue to inspire materials scientists [3].

The basic orb web architecture includes a frame holding radii that support a spirally arranged sticky thread. The web must absorb the kinetic energy of the prey and retain them long enough to give the spider time to locate and subdue them [4]. Orb-weaving spiders are of Jurassic origin; the oldest fossilised orb web is from the Lower Cretaceous [5] and the oldest orb-weaving spider fossils are from the Middle Jurassic [6–8]. Two groups of spiders make geometrically similar orb webs using different materials, but similar stereotypical behaviours: the deinopoids and the araneoids [9]. The sticky spiral of deinopoid webs is made of dry cribellate silk, which is metabolically expensive to produce. Cribellar threads are formed by thousands of fine-looped fibrils woven on a core of two axial fibres. The adhesive properties of cribellar silk are attained by van der Waals and hygroscopic forces [10,11]. In contrast, the sticky spiral thread of araneoid webs is coated with a viscid glycoprotein that coalesces as regularly spaced droplets around the axial fibres [12]. This type of composite sticky thread is produced faster and more economically, and has higher stickiness, than the dry and fuzzy deinopoid counterpart [13]. In addition, the axial fibres of the araneoid capture thread are more extensible than those of cribellar threads [14], which contribute to its increased stickiness by allowing longer spans of capture thread to contact the prey surface [15]. Cribellar silk evolved with aerial webs, at the divergence of araneomorphs ('modern' spiders) from mygalomorphs (tarantulas and their kin) and, thus it is plesiomorphic relative to the more recently evolved viscid silk.

About 30 per cent of the more than 42 000 described extant spider species belong to Orbiculariae [16]. The

* Author for correspondence (dimitard.gwu@gmail.com).

vast majority of orb-weaving species belong to Araneoidea (11 720 extant species). By comparison, their putative sister group, Deinopoidea, are very species-poor: only 322 species spin orb webs with cribellate silk [16]. This asymmetry in species diversity has been attributed to the shift in type of capture thread from dry, fuzzy cribellate silk to viscid sticky silk, combined with changes in the silk spectral reflective properties and a transition from horizontal to vertical orb webs [17–19]. Interestingly, most araneoid species build foraging webs that are no longer recognizable as orbs, such as sheet webs (e.g. Linyphiidae) or cobwebs (e.g. Theridiidae). Some lineages have abandoned the use of capture webs altogether (e.g. Mimetidae) [20]. These 'derived' araneoid webs are so highly transformed from the ancestral orbicular architecture that in the few cases in which the spider building behaviours have been studied [21–23], the stereotypical behavioural algorithms involved in orb web construction [9] are no longer identifiable as such. Furthermore, web architecture across Araneoidea is extraordinarily diverse, although still largely undocumented for the majority of species [20,24–26].

Understanding web evolution and diversification requires an empirically robust hypothesis about the phylogenetic patterns that underlie the diversity of this extraordinary group of animals. One of the oldest and most enduring questions in web evolution has been that of the origin of the orb web itself, a controversy that goes back to the late nineteenth century (see a review in [25]). In the last two decades or so, the empirical evidence has clearly tipped the balance towards the monophyletic origin of the orb web [9,20,25–29]. But because the hypothesis about the monophyly of orb weavers has been supported mainly by behavioural and spinning organ characters, it is challenging to test the possibility that orb webs are not convergent in the cribellate and ecribellate orb weavers without recourse to the building behaviours and silk products. Owing to this reliance on characters of the webs themselves, there have been several attempts to use independent sources of evidence (DNA and proteins) to test the monophyly of orb weavers, but these studies have not been designed to explicitly test orb weaver monophyly or have used small taxonomic samples [28,30,31]. In a study of spider silk proteins, Garb *et al.* [28] argued for orbicularian monophyly based on the shared presence of certain spidroins such as *Flag* and *MaSp2*. Their results, however, could also be compatible with a non-monophyletic origin of the orb web (see also a discussion in [28]). More recently, Blackledge *et al.* [29] used DNA sequences from six genetic markers and a set of morphological and behavioural characters to study the phylogenetic relationships of orbicularians. Although their work was based mostly on molecular evidence, only when morphology and behaviour were included in their analyses did they recover orbicularians as monophyletic. Their molecular data alone did not corroborate the monophyly of Orbiculariae because the retrolateral tibial apophysis (RTA) clade—which includes wolf spiders, crab spiders and their relatives—was nested within orbicularians. With over 21 000 species, the RTA clade has almost double the number of species of the Orbiculariae [16]. As Miller *et al.* pointed out, since 'most of the classical characters supporting the original monophyletic origin theory concern the orb web itself, the lack of corroboration from an independent source

of data (i.e. molecular sequence) cannot be considered supportive of the original hypothesis' (p. 12 of [32]).

One concern that all these molecular studies elicit is the potential for taxon sampling artefacts, such as recovering spurious groups or inferring erroneous character transformations. All previous molecular studies use fairly modest taxon sampling. The importance of dense taxon sampling and its positive effects on phylogenetic accuracy are well documented [33–36]. Given the enormous number of species of both spiders and orb weavers, any phylogenetic study at such higher taxonomic level will have to be based on a very small fraction of the total diversity. Rather than the number of species included in the analysis, the issue at hand is how much of the phylogenetic diversity is represented by the taxonomic sample used. Any evolutionary inferences attempting to explain the diversity of the group will be bound by how the taxon sample represents such diversity. For example, Griswold *et al.* [20] did not include any taxa not assigned to Orbiculariae, thereby failing to test that group's monophyly. Lopardo & Hormiga [37] found that a minimal addition of taxa to the analysis of Griswold *et al.* [20] produced a novel phylogenetic hypothesis for orbicularians, which required a diphyletic origin of orb webs. In Blackledge *et al.*'s study [29], 10 of the 22 orbicularian families were not included; many of them spin modified webs (such as the sheet webs of cyatholipids), the phylogenetic placement of which is critical to understand the diversification of orb weavers.

One of the principal goals of our study is to use sequence data to test the monophyly of Orbiculariae with the most thorough taxonomic sample available at this time. An expanded taxon sample should offer a fine-grained image of orbicularian evolution, and thus provide a more accurate phylogenetic context for the comparative biology of highly studied groups such as *Nephila* and its relatives. In addition, we wanted to explore the suitability of the genetic markers generally used in spider systematics to recover the monophyly of highly corroborated groups (such as Tetragnathidae or Theridiidae) and to help resolve the interrelationships among orbicularian families. A number of araneoid lineages have been difficult to place using morphological data and their exact position has varied across studies, and thus their placement in the orbicularian tree remains to be satisfactorily resolved [37]. Revealing the phylogenetic relationships of some of these groups is critical to understand the evolution and diversification of web architectures.

Information on the timeframe for the origin and the diversification of the main orbicularian lineages provides further illumination for understanding the evolution of the orb web in a broader context. Calibrated molecular phylogenies allow testing-specific hypotheses about the age of groups and the timing of evolutionary events. The expanded taxonomic and character sample of our study aims to provide a more robust test of the age of orb weavers and to test for other evolutionary events that may correlate with the diversification of orb webs. To address these questions, we have compiled a dataset that surpasses in size all prior molecular studies of spider phylogeny.

2. MATERIAL AND METHODS
(a) *Taxon sampling*
The goal of our taxon sampling strategy was to maximize the representation of orbicularian higher taxa and of their

potential close relatives. Classification and nomenclature follows Platnick [16] (but see [38]). Since the exact position of Nicodamidae is contentious [20,27,29,37], here we use the name 'Orbiculariae' to include Deinopoidea, Araneoidea and Nicodamidae, although there is no robust evidence for the exact placement of nicodamids or even their membership in orbicularians. We have included representatives from 21 extant families of Orbiculariae (see [29,37,39]), representing all except Synaphridae (no samples available) and Sinopimoidae (which we do not consider a valid family; see [40]). In total, 291 species in 222 genera, representing 50 families (see electronic supplementary material, table S1), were included in our analyses. Trees were rooted with the mygalomorph *Euagrus chisosesus* (Dipluridae). We have tried to strike a balance between familial rank representation and proportion of missing data. Our sample favours taxa well represented by slow-evolving genes, as those are most appropriate for recovering the sort of deep cladogenetic events of interest to the hypotheses being tested. Sequence data were taken from our own work or from GenBank.

(b) *Sequencing, alignment and phylogenetic analyses*

To date, with very few exceptions [41,42], inference of higher-level phylogenetic relationships in non-model organisms has been based on the combination of recurrent genes for which robust, universal primers were available. These markers would include a handful of mitochondrial genes, the nuclear ribosomal genes and, occasionally, one or two single-copy nuclear genes. While waiting for next-generation sequencing approaches to multiply the availability of sequence data information for organisms such as spiders, a rigorous and thoroughly sampled phylogeny of spiders must currently rely on a limited number of genes. The following genetic markers were used: *28S rRNA, 18S rRNA, 16S rRNA, cytochrome* c *oxidase subunit I, histone H3* and *wingless* (see electronic supplementary material, table S1). Molecular protocols and primers follow Dimitrov & Hormiga [39] and Lopardo *et al.* [38]. Static multiple sequence alignment was carried out using MAFFT v. 6 [43] with the L-INS-i or E-INS-i strategies. Character matrices resulting from statically aligned sequences were analysed using RAxML [44] and TNT [45]. In the RAxML analyses, data were partitioned by gene. For the protein-coding genes, the model parameters were optimized independently for 1st + 2nd and 3rd codon positions. In the TNT analyses, gaps were treated as 'missing data' (to make results more directly comparable to those of the maximum-likelihood analyses) and all characters were equally weighted. The combined matrix of the statically aligned sequences had a total of 6993 characters. For dynamic homology under direct optimization [46], analyses were conducted using POY v. 4.1.2.1 [47]. To investigate the effects of missing data and wild card taxa, we built several matrices (see electronic supplementary material, table S2) based on the proportion of missing data or leaf stability indices [48,49]. Leaf stability indices were calculated using PHYUTILITY v. 2.2 [50]. Additional details on the phylogenetic analyses are given in the electronic supplementary material. In total, nine different variants of the dataset were analysed. Web architecture was coded following the character states of Blackledge *et al.* [29] (plus a ninth state: 'no foraging web'), and ancestral state reconstructions were carried out in MESQUITE v. 2.74 [51]. We used both parsimony and likelihood methods (under the Mk1 model of [52]) to reconstruct the ancestral states.

(c) *Dating cladogenetic events*

We used relaxed and uncorrelated lognormal methods [53], as implemented in the program BEAST v. 1.6.1 [54], to date cladogenetic events. As a starting tree for the BEAST analyses, we used the cladogram from the maximum likelihood (ML) analysis of dataset1_v1. Six fossil calibration points, distributed within and outside Orbiculariae, were implemented as minimum ages. More details on the dating analyses and the specific calibration points are given in the electronic supplementary material.

3. RESULTS AND DISCUSSION

(a) *Phylogenetic results*

We report here on the major orbicularian clades—that is, those groups that are critical for testing the hypotheses and answering the questions outlined in §1 (for a more detailed taxonomic discussion see the electronic supplementary material, in which Table S4 summarizes the groups recovered across the nine data matrices and three analytical methods used in this study, as well as the measures of clade support). Orbiculariae monophyly is recovered by the full dataset under both ML and static parsimony analyses, as well as for most of the data matrix variations (10/14) analysed using these two methods. Araneoidea (the ecribellate orb weavers) receive high support across matrices and methods, and only in 3 of the 25 analyses was their monophyly rejected. Deinopoids (the cribellate orb weavers) and Nicodamidae are monophyletic in the ML analysis of the full matrix, but are not in many of the remaining analyses (in total Deinopoidea are monophyletic in 4 out of the 25 analyses). The monophyly of the RTA clade receives high support (with the Phyxelididae representatives nested within), but not its placement within the orbicularian lineages. The monophyly of most orbicularian families is corroborated by the analyses (see the electronic supplementary material). By contrast, the monophyly of Mimetidae and of 'Symphytognathoids' is never recovered.

We have chosen the results of the ML analyses of a nearly complete data matrix (272 taxa) to guide the discussion of the phylogenetic patterns of the data (figure 1), in order to make our results directly comparable with previous, widely recognized hypotheses of web evolution, e.g. Griswold *et al.* [20] and Blackledge *et al.* [29]. The Griswold *et al.* [20] hypothesis for the phylogeny of orbicularian families has served as the empirical and conceptual basis for much of the phylogenetic and taxonomic research in orb weavers in the decade to follow. Griswold *et al.*'s [20] study provided a comparative analysis of morphology and web-building behaviours across a taxonomic sample of the then-accepted orbicularian families. Their study was not intended to test the monophyly of Orbiculariae, which was taken as given, but to resolve relationships among araneoid families. Orbicularian monophyly had already been tested in earlier works [26]. Perhaps one of the most significant findings of Griswold *et al.* regarding the evolution and the diversification of web architectures among orb weavers was the notion that 'for most orbicularians, the orb web has been an evolutionary base camp rather than a summit' (p. 25 of [20]). The majority of subsequent studies of higher-level orbicularian phylogenetics have used Griswold *et al.*'s [20] cladogram to add outgroup taxa. In the last decade or so numerous phylogenetic
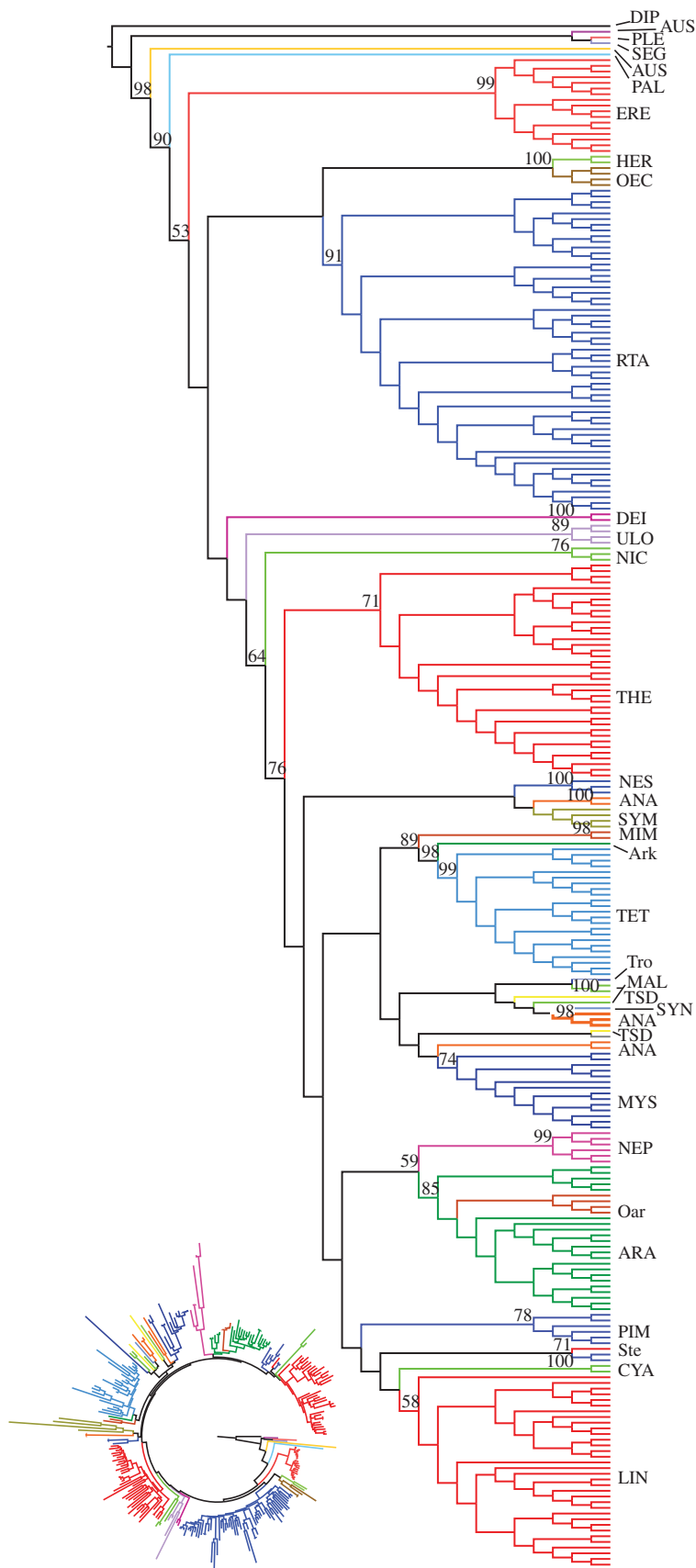
Figure 1. ML tree of data matrix dataset1_v1; miniature illustration at bottom left shows the same tree with proportional branch lengths. Bootstrap values of the major lineages are shown above branches. DIP, Dipluridae; AUS, Austrochilidae; PLE, Plectreurideae; SEG, Segestriidae; PAL, Palpimanidae; ERE, Eresidae; HER, Hersiliidae; OEC, Oecobiidae; RTA, RTA clade; DEI, Deinopidae; ULO, Uloboridae; NIC, Nicodamidae; THE, Theridiidae; NES, Nesticidae; ANA, Anapidae; SYM, Symphytognathidae; MIM, Mimetidae; Ark, *Arkys* (Araneidae); TET, Tetragnathidae; Tro, *Trogloneta* (Mysmenidae); MAL, Malkaridae; TSD, Theridiosomatidae; MYS, Mysmenidae; NEP, Nephilidae; Oar, Oarcinae; ARA, Araneidae; PIM, Pimoidae; Ste, *Stemonyphantes* (Linyphiidae); CYA, Cyatholipidae; LIN, Linyphiidae. Taxonomic groups are depicted with different colours; grey is unknown (misidentified as *Nesticus* sp.; see the electronic supplementary material).

analyses have studied interfamilial relationships [55–60], but no analysis has included an interfamilial taxon sample comparable to that of Griswold *et al.* [20]—14 families plus nephilids. From a taxon-sampling perspective, the closest study has been that of Blackledge *et al.* [29] (nine orbicularian families plus nephilids), but their taxon sampling lacked representatives of families that are critical for understanding the diversification of orb webs and orb weavers (e.g. Mysmenidae, Theridiosomatidae, Cyatholipidae and Synotaxidae). Previous molecular analyses that included orbicularian species [29–31] have failed to recover the monophyly of Orbiculariae. Only Ayoub *et al.* [61], using *elongation factor-1 γ* sequence data and four orbicularian species, found orbicularians monophyletic in their ML analyses (orbicularians were paraphyletic with respect to the RTA clade in their parsimony analyses). Our results therefore provide the first empirical support for the monophyly of Orbiculariae based exclusively on nucleotide sequence data and a broad higher-level taxon sample.

(b) Dating and diversification of orb webs

The molecular dating of orbicularian diversification (figure 2) suggests an older origin for orb weavers than that inferred from the fossil record (see also [61]), which has to be expected given that fossils and morphological data tend to underestimate divergence times, while molecular data tend to overestimate them [62]. Our analysis indicates that by the Jurassic most of the orb-weaving families had already originated. The estimated age of the most recent common ancestor (MRCA) of Orbiculariae is about 207–231 Ma (figure 2). The estimated age of the MRCA of the RTA clade is about 135–175 Ma. Such results are not entirely surprising either, since many of the Jurassic and Early Cretaceous fossils represent rather distal lineages of the modern spider families [8,63]. The Jurassic fossil record of the extinct araneoid family Juraraneidae [64] provides additional and independent support for our dating results.

Both the parsimony and ML reconstructions support a single origin of orb webs and multiple independent modifications of the ancestral orb into a great diversity of snare architectures. Most species of the large family Araneidae build two-dimensional orbs that are very similar to those of Tetragnathidae and the cribellate family Uloboridae. Based on this architectural similarity, in conjunction with the relatively basal placement of araneids and tetragnathids within Araneoidea [29,37], it has been suggested that this general type of orb web architecture is ancestral in araneoids. The sheet webs and cobwebs spun by some lineages of araneoids are generally thought to be derived from typical orb-weaving ancestors [25]. This has led to the idea of somewhat sequential evolutionary change from the more primitive and typically geometric orbs to the more derived sheet webs or cobwebs [20]. In addition, the results of Griswold *et al.* [20] also suggested a monophyletic origin for these derived araneoid webs (their 'Araneoid sheet weavers' clade). Although the monophyly of araneoid sheet weavers was refuted by the analysis of Lopardo & Hormiga [37], their results still placed araneids and tetragnathids at the base of Araneoidea, and thus implied that the typical geometric orb web is plesiomorphic. Our results suggest that there have been multiple independent transitions from orbs to other

web architectures within Araneoidea, starting early on in their diversification. In fact, one of the earliest lineage divergences, the split between nicodamids and araneoids, resulted in a deep modification of the ancestral orb into a sheet web in some species of the former group (figure 2). Given their phylogenetic placement, nicodamids may have ancestrally spun cribellate orb webs, and the cribellum may have been lost independently in some nicodamids and in araneoids. In our cladogram, most araneoid lineages that spin typical orb webs are found relatively distally placed within Araneoidea. Such a macroevolutionary pattern poorly fits the predictions of a hypothesis suggesting that the evolution of orb architecture tracks the diversification of flying insects [3]. If that were the case, one would expect a sequential change from orbs to more derived and variable web architectures as insect diversity increases (particularly after the diversification of angiosperms and their flying insect pollinators). An alternative scenario that has been proposed to explain orb web evolution (which also invokes a correlation with insect diversification, but in a different way) suggests that spider predation drove the evolution of insects [65]. According to this view, spider webs evolved in an evolutionary predator–prey arms race [3]. Although the traditional view of orbicularian relationships [20] fits well a codiversification hypothesis, as pointed out by Penney [66], the present results show that the history of web diversification is probably much more complex than previously thought.

There is ample fossil evidence that most of the modern flying insect orders were already present during the Permian–Early Triassic and that the origin of flying insects as such extends back to the Devonian ([67] and references therein). The early abundance of fossil insects has been interpreted as indirect evidence for the predator–prey arms race hypothesis [3,66]. Alternatively, one could explain the diversification of web architectures as a strategy for better niche partitioning and a more efficient use of the available trophic resources. The diversification of one of the largest spider lineages, the RTA clade, in the Late Jurassic may be interpreted as indirect evidence that trophic resources were sufficiently diverse and abundant. Sheet webs and cobwebs do not have to obey the same strict architectural constraints that govern orb webs. This allows spiders to use spaces where orbs cannot be constructed or are very inefficient in catching prey. Furthermore, the ability to build webs in different spaces would also reduce intraguild competition. Thus, prey abundance and structural complexity of the habitat may have played a more important role in web evolution than the actual diversification of their prey. Structurally complex forests were present already in the Carboniferous [68]. Most of these forests were probably affected by the Permo-Triassic extinction [69], but forests with complex species composition and stratigraphic structure were again abundant by the end of the Triassic, and even more so during the Jurassic [70,71], long before the diversification of angiosperms. Relaxed architectural constraints of sheet webs result in relaxed constraints over the behaviours needed to produce them, hence the lack of (or reduced) use of stereotypical behaviours during web construction in these groups [22]. The very conservative stereotypical behaviour of orb building in orbicularians was considered to be evolutionarily constrained, but our results show that shifts in behavioural patterns have
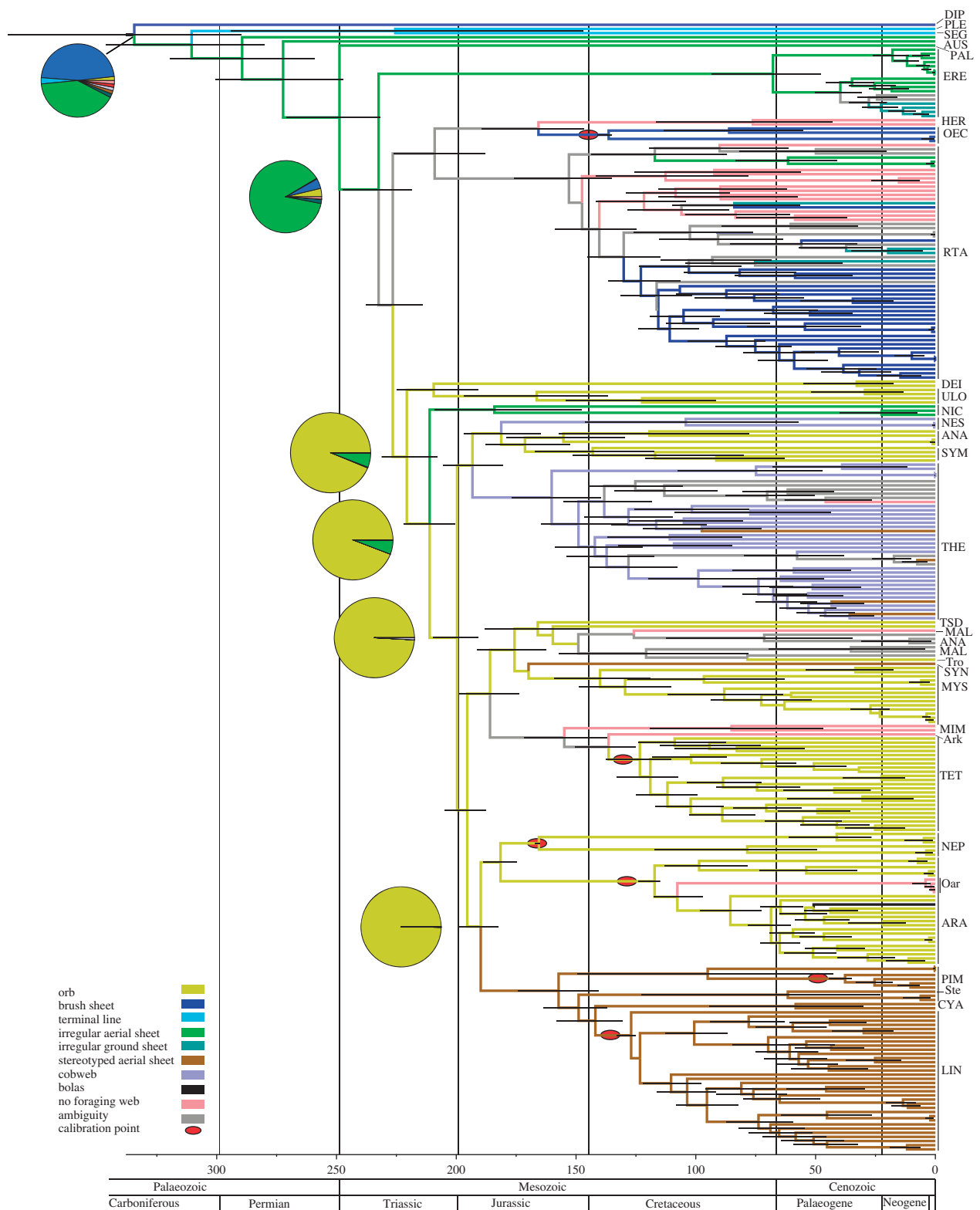
Figure 2. Chronogram of orb weavers and relatives based on dataset1_v1. Branches are coloured according to the parsimony ancestral reconstruction of web types; pie charts represent probabilities of ancestral states from the Mk1 model of [52]. Node bars show confidence intervals. Taxonomic abbreviations are as in figure 1.

happened multiple times. This finding is also in agreement with the 'imprecision mechanism' for such changes proposed by Eberhard [72,73]. Imprecision in reproducing specific behaviours occurs often, and if these changes result in somewhat more successful (or at least neutrally selective) variants they may be carried on and further modified in the descendants. Furthermore, linyphiid and theridiid sheet webs and cobwebs are cheaper to produce

as they contain less sticky silk and have a longer functional 'lifespan,' which might provide additional advantages.

Thus, the combination of structurally complex habitat, available prey and a mechanistic model that allows the origin of behavioural modifications without invoking very rare mutation events or a predator–prey arms race seems sufficient to explain both web diversity and numerous shifts in web architecture in orb weavers.

(c) Methodological implications and missing data

As stated above, the available 'standard' genetic markers support the monophyly of most family-rank groups in our analyses, but are insufficient for reconstructing relationships among families, analogous to the problem shown in studies of lepidopteran phylogenetic relationships [41]. Very few of the deeper interfamilial nodes within orbicularians receive high values of clade support and most of the internal braches are short (figure 1). This pattern is consistent with rapid diversification condensed in a relatively short period of time, as suggested by the estimated ages of the major araneoid lineages and by the fossil record [7]. This pattern may also explain the high posterior probabilities for unconventional relationships in [29] (e.g. the RTA clade embedded as an orbicularian lineage). Bayesian methods often assign high posterior probabilities to arbitrary groups owing to limitations in the tree-proposal step, which does not allow polytomies [74].

We more than doubled the number of families included in the study of Blackledge *et al.* [29], and tried to further improve resolution and support by increasing sampling density within families, improving the chance for breaking down long terminal branches [75]. This strategy also augments the probability of sampling plesiomorphic states present in the terminals that may improve the reconstruction of more basal nodes of the tree and allows for a finer-grained study of web architecture evolution. Plesiomorphic morphological character states may be a proxy for that, but we rarely recognize and use the genes that are responsible for these morphological traits. It is important to emphasize that this reasoning also holds true for taxa treated as outgroups. Outgroups are often sampled at lower densities as they are generally included merely to root the tree or to test the monophyly of the taxon of interest. Thus, in many studies, the internal relationships of the outgroup lineages are considered out of the scope and only a few representatives of the outgroup taxa are included in the analysis.

Increasing sampling density, however, often presents additional caveats for the analyses. Probably the most important is the possibility of increasing the proportion of missing entries. In our case, this is due to the fact that some taxa come from studies that used different character sampling designs (e.g. sequences deposited in GenBank) or represent taxa for which some of the targeted gene fragments could not be sequenced. In fact, missing data are often one of the most important factors when designing phylogenetic studies [76]. An increase in missing data can occur for two different reasons: (i) owing to length differences in the fragments sequenced for each genetic marker and/or (ii) as a result of missing data for the entire genetic marker(s). These two options represent the 'too many missing data cells' and 'too few characters' scenarios outlined by Wiens [77].

Blackledge *et al.*'s molecular data matrix was built entirely from sequences generated during their study and had few missing entries [29]. In constructing our full matrix, we added numerous taxa to the study of Blackledge *et al.*, most of them with a large proportion of missing data. Both missing data cells and characters were simultaneously added to the dataset, increasing the overall proportion of missing data to 50.1 per cent, with several taxa having up to 75 per cent empty cells. Nonetheless, analysis of this dataset showed improved resolution with respect to Blackledge *et al.*'s results [29] (see the electronic supplementary material); for example, the RTA clade was placed outside orbicularians and Araneidae was monophyletic, as expected based on morphological grounds. This result provides further evidence that addition of taxa may be beneficial despite increased missing data, in accordance with previous findings based mainly on simulations [76–79].

Several taxa had poor overlap of data partitions in the full dataset, as exemplified by *Clitaetra irenae*, which does not overlap in any of the data partitions with taxa sequenced only for the ribosomal genes (i.e. sequence data for this taxon are available for only two protein fragments). The exclusion of terminals with non-overlapping fragments improved resolution considerably. However, it has been shown that even in situations with little gene overlap, when using large numbers of taxa, data overlap may not be that important for taxonomic resolution [80].

A different approach for determining the effect of particular taxa is the calculation of their leaf stability index [48,49,81,82]. This measure reflects topological stability and is often used to select taxa for exclusion based on the presumption that taxa with low leaf stability will tend to affect negatively the resolution of phylogenetic analysis. We used leaf stability as an alternative approach to study the effects of increased taxon sampling density. Our results show that removing taxa with lower leaf stability did not improve resolution significantly, unlike optimization of fragment overlap, which had a positive effect on the results (electronic supplementary material, table S4). At first, this may seem counterintuitive, as decrease in overlap is often expected to result in a decrease in stability. Taxon instability, however, may be also a result of character incongruence between data partitions [83].

4. CONCLUSIONS

Increased taxon sampling helps to improve results even when missing data are present in a significant proportion. Ideally, one would aim to produce a matrix where all taxa would be complete for all genes. But despite advances in next-generation sequencing technology (e.g. Illumina), it is unlikely that such data will be available soon, especially in spiders and other megadiverse arthropod groups. And even if available, missing data will always be an issue [84]. To increase taxon sampling density, we used data already available in GenBank or other repositories, which confirms that this is an effective and cost-efficient strategy to increase the representation of taxa [85].

We have aimed to build a basic phylogenetic framework to help tackle some of the most intriguing questions regarding orb weaver evolution, such as: (i) could the diversification of insects has promoted a diversification of web architectures? And (ii) how does the diversification of spiders and their foraging techniques affect evolution of insects? Some of these questions have been touched upon elsewhere [66], but a lack of phylogenetic treatments with broad representation of orbicularians and their outgroups has limited conclusions in previous studies. Earlier discussions have centred around coradiation and evolutionary arms race scenarios [3,66]. Our results, however, suggest that factors such as biotic and abiotic environmental changes, intraguild competition, and radiation into new niches may have played important roles

for the diversification of orb weavers, although a resource (abundance of flying insects) had to be present to exploit such new niches.

REFERENCES

1 Selden, P. A., Shear, W. A. & Sutton, M. D. 2008 Fossil evidence for the origin of spider spinnerets, and a proposed arachnid order. *Proc. Natl Acad. Sci. USA* **105**, 20 781–20 785. (doi:10.1073/pnas.0809174106)

2 Brunetta, L. & Craig, C. L. 2010 *Spider silk: evolution and 400 million years of spinning, waiting, snagging, and mating.* New Haven, CT: Yale University Press.

3 Vollrath, F. & Selden, P. 2007 The role of behavior in the evolution of spiders, silks, and webs. *Annu. Rev. Ecol. Evol. Syst.* **38**, 819–846. (doi:10.1146/annurev.ecolsys. 37.091305.110221)

4 Chacon, P. & Eberhard, W. G. 1980 Factors affecting numbers and kinds of prey caught in artificial spider webs, with considerations of how orb webs trap prey. *Bull. Br. Arachnol. Soc.* **5**, 29–38.

5 Peñalver, E., Grimaldi, D. A. & Delclòs, X. 2006 Early Cretaceous spider web with its prey. *Science* **312**, 1761–1761. (doi:10.1126/science.1126628)

6 Selden, P. A. 1989 Orb-web weaving spiders in the Early Cretaceous. *Nature* **340**, 711–713. (doi:10.1038/340711a0)

7 Selden, P. A. & Penney, D. 2010 Fossil spiders. *Biol. Rev.* **85**, 171–206. (doi:10.1111/j.1469-185X.2009.00099.x)

8 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. (doi:10.1098/rsbl.2011.0228)

9 Eberhard, W. G. 1982 Behavior of arachnids. *Science* **216**, 876–877. (doi:10.1126/science.216.4548.876)

10 Hawthorn, A. C. & Opell, B. D. 2002 Evolution of adhesive mechanisms in cribellar spider prey capture thread: evidence for van der Waals and hygroscopic forces. *Biol. J. Linn. Soc.* **77**, 1–8. (doi:10.1046/j.1095-8312.2002.00099.x)

11 Hawthorn, A. C. & Opell, B. D. 2003 van der Waals and hygroscopic forces of adhesion generated by spider capture threads. *J. Exp. Biol.* **206**, 3905–3911. (doi:10.1242/jeb.00618)

12 Opell, B. D. & Bond, J. E. 2000 Capture thread extensibility of orb-weaving spiders: testing punctuated and associative explanations of character evolution. *Biol. J. Linn. Soc.* **70**, 107–120. (doi:10.1111/j.1095-8312.2000.tb00203.x)

13 Opell, B. D. 1998 Economics of spider orb-webs: the benefits of producing adhesive capture thread and of recycling silk. *Funct. Ecol.* **12**, 613–624. (doi:10.1046/j.1365-2435.1998.00222.x)

14 Blackledge, T. A. & Hayashi, C. Y. 2006 Unraveling the mechanical properties of composite silk threads spun by cribellate orb-weaving spiders. *J. Exp. Biol.* **209**, 3131–3140. (doi:10.1242/jeb.02327)

15 Opell, B. D., Markley, B. J., Hannum, C. D. & Hendricks, M. L. 2008 The contribution of axial fiber extensibility to the adhesion of viscous capture threads spun by orb-weaving spiders. *J. Exp. Biol.* **211**, 2243–2251. (doi:10.1242/jeb.016147)

16 Platnick, N. I. 2011 The world spider catalog, version 12.0. American Museum of Natural History. See http://research. amnh.org/iz/spiders/catalog/INTRO1.html. (doi:10.5531/ db.iz.0001)

17 Craig, C. L., Bernard, G. D. & Coddington, J. A. 1994 Evolutionary shifts in the spectral properties of spider silks. *Evolution* **48**, 287–296. (doi:10.2307/2410093)

18 Bond, J. E. & Opell, B. D. 1998 Testing adaptive radiation and key innovation hypotheses in spiders. *Evolution* **52**, 403–414. (doi:10.2307/2411077)

19 Opell, B. D. 1997 The material cost and stickiness of capture threads and the evolution of orb-weaving spiders. *Biol. J. Linn. Soc.* **62**, 443–458. (doi:10.1111/j.1095-8312.1997.tb01635.x)

20 Griswold, C. E., Coddington, J. A., Hormiga, G. & Scharff, N. 1998 Phylogeny of the orb-web building spiders (Araneae, Orbiculariae: Deinopoidea, Araneoidea). *Zool. J. Linn. Soc.* **123**, 1–99. (doi:10.1111/j. 1096-3642.1998.tb01290.x)

21 Benjamin, S. P. & Zschokke, S. 2003 Webs of theridiid spiders: construction, structure and evolution. *Biol. J. Linn. Soc.* **78**, 293–305. (doi:10.1046/j.1095-8312.2003.00110.x)

22 Benjamin, S. P. & Zschokke, S. 2004 Homology, behaviour and spider webs: web construction behaviour of *Linyphia hortensis* and *L. triangularis* (Araneae: Linyphiidae) and its evolutionary significance. *J. Evol. Biol.* **17**, 120–130. (doi:10.1046/j.1420-9101.2004. 00667.x)

23 Eberhard, W. G. 1977 Aggressive chemical mimicry by a bolas spider. *Science* **198**, 1173–1175. (doi:10.1126/ science.198.4322.1173)

24 Eberhard, W. G., Agnarsson, I. & Levi, H. W. 2008 Web forms and the phylogeny and theridiid spiders (Araneae: Theridiidae): chaos from order. *Syst. Biodivers.* **6**, 415–475. (doi:10.1017/S1477200008002855)

25 Coddington, J. A. 1986 The genera of the spider family Theridiosomatidae. *Smith. Contrib. Zool.* **422**, 1–96. (doi:10.5479/si.00810282.422)

26 Coddington, J. A. 1990 Cladistics and spider classification: Araneomorph phylogeny and the monophyly of orbweavers (Araneae: Araneomorphae, Orbiculariae). *Acta Zool. Fenn.* **190**, 75–87.

27 Griswold, C. E., Ramírez, M. J., Coddington, J. A. & Platnick, N. I. 2005 Atlas of phylogenetic data for entelegyne spiders (Araneae: Araneomorphae: Entelegynae) with comments on their phylogeny. *Proc. Calif. Acad. Sci.* **56**(Suppl. 2), 324.

28 Garb, J. E., Dimauro, T., Vo, V. & Hayashi, C. Y. 2006 Silk genes support the single origin of orb webs. *Science* **312**, 1762. (doi:10.1126/science.1127946)

29 Blackledge, T. A., Scharff, N., Coddington, J. A., Szuts, T., Wenzel, J. W., Hayashi, C. Y. & Agnarsson, I. 2009 Reconstructing web evolution and spider diversification in the molecular era. *Proc. Natl Acad. Sci. USA* **106**, 5229–5234. (doi:10.1073/pnas.0901377106)

30 Hausdorf, B. 1999 Molecular phylogeny of araneomorph spiders. *J. Evol. Biol.* **12**, 980–985. (doi:10.1046/j.1420-9101.1999.00104.x)

31 Vink, C. J., Hedin, M., Bodner, M. R., Maddison, W. P., Hayashi, C. Y. & Garb, J. E. 2008 Actin 5C, a promising

nuclear gene for spider phylogenetics. *Mol. Phylogenet. Evol.* **48**, 377–382. (doi:10.1016/j.ympev.2008.03.003)

32 Miller, J. A. *et al.* 2010 Phylogeny of entelegyne spiders: affinities of the family Penestomidae (NEW RANK), generic phylogeny of Eresidae, and asymmetric rates of change in spinning organ evolution (Araneae, Araneoidea, Entelegynae). *Mol. Phylogenet. Evol.* **55**, 786–804. (doi:10.1016/j.ympev.2010.02.021)

33 Pollock, D. D., Zwickl, D. J., McGuire, J. A. & Hillis, D. M. 2002 Increased taxon sampling is advantageous for phylogenetic inference. *Syst. Biol.* **51**, 664–671. (doi:10.1080/10635150290102357)

34 Hillis, D. M., Pollock, D. D., McGuire, J. A. & Zwickl, D. J. 2003 Is sparse taxon sampling a problem for phylogenetic inference? *Syst. Biol.* **52**, 124–126. (doi:10.1080/10635150390132911)

35 Zwickl, D. J. & Hillis, D. M. 2002 Increased taxon sampling greatly reduces phylogenetic error. *Syst. Biol.* **51**, 588–598. (doi:10.1080/10635150290102339)

36 DeBry, R. W. 2005 The systematic component of phylogenetic error as a function of taxonomic sampling under parsimony. *Syst. Biol.* **54**, 432–440. (doi:10.1080/10635150590946745)

37 Lopardo, L. & Hormiga, G. 2008 Phylogenetic placement of the Tasmanian spider *Acrobleps hygrophilus* (Araneae, Anapidae) with comments on the evolution of the capture web in Araneoidea. *Cladistics* **24**, 1–33. (doi:10.1111/j.1096-0031.2007.00173.x)

38 Lopardo, L., Giribet, G. & Hormiga, G. 2011 Morphology to the rescue: molecular data and the signal of morphological characters in combined phylogenetic analyses—a case study from mysmenid spiders (Araneae, Mysmenidae), with comments on the evolution of web architecture. *Cladistics* **27**, 278–330. (doi:10.1111/j.1096-0031.2010.00332.x)

39 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. (doi:10.1111/j.1096-3642.2010.00662.x)

40 Hormiga, G. 2008 On the spider genus *Weintrauboa* (Araneae, Pimoidae), with a description of a new species from China and comments on its phylogenetic relationships. *Zootaxa* **1814**, 1–20.

41 Regier, J. C. *et al.* 2009 Toward reconstructing the evolution of advanced moths and butterflies (Lepidoptera: Ditrysia): an initial molecular study. *BMC Evol. Biol.* **9**, 280. (doi:10.1186/1471-2148-9-280)

42 Zahiri, R., Kitching, I. J., Lafontaine, J. D., Mutanen, M., Kaila, L., LHolloway, J. D. & Wahlberg, N. 2011 A new molecular phylogeny offers hope for a stable family level classification of the Noctuoidea (Lepidoptera). *Zool. Scr.* **40**, 158–173.

43 Katoh, K., Kuma, K., Toh, H. & Miyata, T. 2005 MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518. (doi:10.1093/nar/gki198)

44 Stamatakis, A. P., Meier, H. & Ludwig, T. 2008. RAxML: A parallel program for phylogenetic tree inference. See http://sco.h-its.org/exelixis/software.html.

45 Goloboff, P. A., Farris, J. S. & Nixon, K. C. 2008 TNT, a free program for phylogenetic analysis. *Cladistics* **24**, 774–786. (doi:10.1111/j.1096-0031.2008.00217.x)

46 Wheeler, W. C. 1996 Optimization alignment: the end of multiple sequence alignment in phylogenetics? *Cladistics* **12**, 1–9. (doi:10.1111/j.1096-0031.1996.tb00189.x)

47 Varón, A., Sy Vinh, L. & Wheeler, W. C. 2010 POY version 4: phylogenetic analysis using dynamic homologies. *Cladistics* **26**, 72–85. (doi:10.1111/j.1096-0031.2009.00282.x)

48 Thorley, J. L. & Wilkinson, M. 1999 Testing the phylogenetic stability of early tetrapods. *J. Theor. Biol.* **200**, 343–344. (doi:10.1006/jtbi.1999.0999)

49 Wilkinson, M. 2006 Identifying stable reference taxa for phylogenetic nomenclature. *Zool. Scr.* **35**, 109–112. (doi:10.1111/j.1463-6409.2005.00213.x)

50 Smith, S. A. & Dunn, C. W. 2008 Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics* **24**, 715–716. (doi:10.1093/bioinformatics/btm619)

51 Maddison, W. P. & Maddison, D. R. 2010 Mesquite: a modular system for evolutionary analysis. Version 2.73. See http://mesquiteproject.org.

52 Lewis, P. O. 2001 A likelihood approach to estimating phylogeny from discrete morphological characters data. *Syst. Biol.* **50**, 913–925. (doi:10.1080/106351501753462876)

53 Drummond, A. J., Ho, S. Y., Phillips, M. J. & Rambaut, A. 2006 Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88. (doi:10.1371/journal.pbio.0040088)

54 Drummond, A. J. & Rambaut, A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214. (doi:10.1186/1471-2148-7-214)

55 Agnarsson, I. 2004 Morphological phylogeny of cobweb spiders and their relatives (Araneae, Araneoidea, Theridiidae). *Zool. J. Linn. Soc.* **141**, 447–626. (doi:10.1111/j.1096-3642.2004.00120.x)

56 Arnedo, M. A., Coddington, J., Agnarsson, I. & Gillespie, R. G. 2004 From a comb to a tree: phylogenetic relationships of the comb-footed spiders (Araneae, Theridiidae) inferred from nuclear and mitochondrial genes. *Mol. Phylogenet. Evol.* **31**, 225–245. (doi:10.1016/S1055-7903(03)00261-6)

57 Kuntner, M., Coddington, J. A. & Hormiga, G. 2008 Phylogeny of extant nephilid orb-weaving spiders (Araneae, Nephilidae): testing morphological and ethological homologies. *Cladistics* **24**, 147–217. (doi:10.1111/j.1096-0031.2007.00176.x)

58 Schütt, K. 2003 Phylogeny of Symphytognathidae *s.l.* (Araneae, Araneoidea). *Zool. Scri.* **32**, 129–151. (doi:10.1046/j.1463-6409.2003.00103.x)

59 Rix, M. G., Harvey, M. S. & Roberts, J. D. 2008 Molecular phylogenetics of the spider family Micropholcommatidae (Arachnida: Araneae) using nuclear rRNA genes (18S and 28S). *Mol. Phylogenet. Evol.* **46**, 1031–1048. (doi:10.1016/j.ympev.2007.11.001)

60 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. (doi:10.1111/j.1096-0031.2008.00242.x)

61 Ayoub, N. A., Garb, J. E., Hedin, M. & Hayashi, C. Y. 2007 Utility of the nuclear protein-coding gene, elongation factor-1 gamma (EF-1γ), for spider systematics, emphasizing family level relationships of tarantulas and their kin (Araneae: Mygalomorphae). *Mol. Phylogenet. Evol.* **42**, 394–409. (doi:10.1016/j.ympev.2006.07.018)

62 Steiper, M. E. & Young, N. M. 2008 Timing primate evolution: lessons from the discordance between molecular and paleontological estimates. *Evol. Anthropol.* **17**, 179–188. (doi:10.1002/Evan.20177)

63 Penney, D. & Selden, P. A. 2002 The oldest linyphiid spider, in Lower Cretaceous Lebanese amber (Araneae, Linyphiidae, Linyphiinae). *J. Arachnol.* **30**, 487–493. (doi:10.1636/0161-8202(2002)030[0487:TOLSIL]2.0.CO;2)

64 Eskov, K. 1984 A new fossil spider family from the Jurassic of Transbaikalia (Araneae: Chelicerata). *Neues Jahrb. Geol. Palaeontol. Mon. hefte* **1984**, 645–653.

65 Eisner, T., Alsop, R. & Ettershank, G. 1964 Adhesiveness of spider silk. *Science* **146**, 1058–1061. (doi:10.1126/science.146.3647.1058)

66 Penney, D. 2004 (for 2003 Does the fossil record of spiders track that of their principal prey, the insects? *Trans. R. Soc. Edinb. Earth Sci.* **94**, 275–281.

67 Grimaldi, D. A. & Engel, M. S. 2005 *Evolution of the insects.* Cambridge, UK: Cambridge University Press.

68 Gastaldo, R. A., Stevanovic-Walls, I. M., Ware, W. N. & Greb, S. F. 2004 Community heterogeneity of early Pennsylvanian peat mires. *Geology* **32**, 693–696. (doi:10.1130/G20515.1)

69 Looy, C. V., Twitchett, R. J., Dilcher, D. L., Konijnenburg-Van Cittert, J. H. A. & Visscher, H. 2001 Life in the end-Permian dead zone. *Proc. Natl Acad. Sci. USA* **98**, 7879–7883. (doi:10.1073/pnas.131218098)

70 Hinz, J. K., Smith, I., Pfretzschner, H.-U., Wings, O. & Sun, G. 2010 A high-resolution three-dimensional reconstruction of a fossil forest (Upper Jurassic Shishugou Formation, Junggar Basin, Northwest China). *Palaeobiodivers. Palaeoenviron.* **90**, 215–240. (doi:10.1007/s12549-010-0036-y)

71 Thorn, V. 2005 A Middle Jurassic fossil forest from New Zealand. *Palaeontology* **48**, 1021–1039. (doi:10.1111/j.1475-4983.2005.00490.x)

72 Eberhard, W. G. 1990 Imprecision in the behavior of *Leptomorphus* sp. (Diptera, Mycetophilidae) and the evolutionary origin of new behavior patterns. *J. Insect Behav.* **3**, 327–357. (doi:10.1007/BF01052113)

73 Eberhard, W. G. 2000 Breaking the mold: behavioral variation and evolutionary innovation in *Wendilgarda* spiders (Araneae Theridiosomatidae). *Ethol. Ecol. Evol.* **12**, 223–235. (doi:10.1080/08927014.2000.9522797)

74 Lewis, P. O., Holder, M. T. & Holsinger, K. E. 2005 Polytomies and Bayesian phylogenetic inference. *Syst. Biol.* **54**, 241–253. (doi:10.1080/10635150590924208)

75 Hillis, D. M. 1996 Inferring complex phylogenies. *Nature* **383**, 130–131. (doi:10.1038/383130a0)

76 Wiens, J. J. 2006 Missing data and the design of phylogenetic analyses. *J. Biomed. Inform.* **39**, 34–42. (doi:10.1016/j.jbi.2005.04.001)

77 Wiens, J. J. 2003 Missing data, incomplete taxa, and phylogenetic accuracy. *Syst. Biol.* **52**, 528–538. (doi:10.1080/10635150390218330)

78 Wiens, J. J. 2005 Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? *Syst. Biol.* **54**, 731–742. (doi:10.1080/10635150500234583)

79 Wiens, J. J. & Moen, D. S. 2008 Missing data and the accuracy of Bayesian phylogenetics. *J. Syst. Evol.* **46**, 307–314.

80 Goloboff, P. A., Catalano, S. A., Mirande, J. M., Szumik, C. A., Arias, J. S., Källersjö, M. & Farris, J. S. 2009 Phylogenetic analysis of 73 060 taxa corroborates major eukaryotic groups. *Cladistics* **25**, 211–230. (doi:10.1111/j.1096-0031.2009.00255.x)

81 Dunn, C. W. *et al.* 2008 Broad taxon sampling improves resolution of the animal tree of life. *Nature* **452**, 745–749. (doi:10.1038/nature06614)

82 Lee, M. S. Y. 2005 Choosing reference taxa in phylogenetic nomenclature. *Zool. Scr.* **34**, 313–328. (doi:10.1111/j.1463-6409.2005.00196.x)

83 Giribet, G. & Edgecombe, G. D. 2006 Conflict between data sets and phylogeny of centipedes: an analysis based on seven genes and morphology. *Proc. R. Soc. B* **273**, 531–538. (doi:10.1098/rspb.2005.3365)

84 Hejnol, A. *et al.* 2009 Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc. R. Soc. B* **276**, 4261–4270. (doi:10.1098/rspb.2009.0896)

85 Driskell, A. C., Ane, C., Burleigh, J. G., McMahon, M. M., O'Meara, B. C. & Sanderson, M. J. 2004 Prospects for building the tree of life from large sequence databases. *Science* **306**, 1172–1174. (doi:10.1126/science.1102036)