**CSIRO** PUBLISHING
www.publish.csiro.au/journals/is *Invertebrate Systematics*, 2010, **24**, 238–257

# Patterns of habitat affinity and Austral/Holarctic parallelism in dictynoid spiders (Araneae : Entelegynae)

*Joseph C. Spagna*[A,D], *Sarah C. Crews*[B,C] *and Rosemary G. Gillespie*[B]

[A]Department of Biology, William Paterson University, 300 Pompton Road, Wayne, NJ 07470, USA.
[B]Division of Organisms and Environment, University of California, Berkeley, 137 Mulford Hall, Berkeley, CA 94720-3114, USA.
[C]Present address: Berkeley City College, Department of Sciences, 2050 Center Street, Berkeley, CA 94709, USA.
[D]Corresponding author. Email: SpagnaJ@wpunj.edu

**Abstract.** The ability to survive in a terrestrial environment was a major evolutionary hurdle for animals that, once passed, allowed the diversification of most arthropod and vertebrate lineages. Return to a truly aquatic lifestyle has occurred only rarely among terrestrial lineages, and is generally associated with modifications of the respiratory system to conserve oxygen and allow extended periods of apnea. Among chelicerates, in particular spiders, where the circulatory system also serves as a hydrostatic skeleton, very few taxa have exploited aquatic environments, though these environments are abundant and range from freshwater ponds to the marine intertidal and relictual (salt) lakes. The traditional systematic positions of the taxa inhabiting these environments are controversial. Partitioned Bayesian analysis using a doublet model for stems in the nearly complete 18S rRNA gene (~1800 nt) and in the D2 and D3 regions of the 28S rRNA gene (~690 nt), and standard models for loops and full protein-coding histone H3 (349 nt) partitions (totalling 3133 bp when aligned) of dictynoid spiders and related lineages revealed that the only truly aquatic spider species, *Argyroneta aquatica* (Clerck, 1767) (Cybaeidae Banks, 1892), belongs in a clade containing other taxa with unusual habitat affinities related to an aquatic existence, including occupation of semi-aquatic (intertidal) areas (Desidae Pocock, 1985: *Paratheuma* spp.) and highly alkaline salt-crusts (Dictynidae O. Pickard-Cambridge, 1871: *Saltonia incerta* (Banks, 1898)). In a contrasting pattern, other spiders that also occupy intertidal zones, including some other members of the family Desidae (*Desis* spp., *Badumna longinqua* (L. Koch, 1867)), are an independently derived clade found primarily in the southern hemisphere. Use of the doublet model reduced some branch-support values in the single-gene trees for rRNA data, but resulted in a robust combined-data phylogeny from 18S rRNA, 28S rRNA, and histone H3. This combination of results – reduction in support in single-gene trees and gain in support in combined-data trees –is consistent with use of the doublet model reducing problematic signal from non-independent base pairs in individual data partitions, resulting in improved resolution in the combined-data analyses.

**Additional keywords:** *Argyroneta*, Cybaeidae, Desidae, Dictynoidea, doublet modelling, intertidal habitats, *Paratheuma*, partitioned Bayesian analysis, *Saltonia*, secondary structure.

## Introduction

The ability to exploit dry, terrestrial environments has represented a large obstacle in the evolutionary history of complex multicellular organisms, and has occurred only once in each of vascular plants, vertebrates, molluscs, oligochaetes, and nematodes since the Palaeozoic (Thomas 1972; Jeram *et al*. 1990; Selden 1990; MacNaughton *et al*. 2002; Pisani *et al*. 2004; Poinar *et al*. 2007). Among arthropods, the most successful terrestrial lineages have been the insects, myriapods and arachnids (Coddington *et al*. 2004). However, the return to a truly aquatic lifestyle requires a suite of adaptations related to obtaining oxygen and to osmoregulation, and has occurred only rarely, with extensive diversification in a single group – the hydrachnids or water mites – and is associated with a parasitic lifestyle (Thorp and Kovich 1991; Coddington *et al*. 2004). The spider taxa that are now found in aquatic environments tend to be limited in diversity, although they occur in highly disparate environments, including marine intertidal, salt marshes, and in one case, submerged in fresh water (Foelix 1996). They also occur in areas that were wet in the past, but are now dry salt lakes. Current taxonomy scatters these unusually distributed taxa across different families (Platnick 2010), leading to the supposition that the ability to exploit such habitats has evolved independently multiple times.

In this study, we focus on taxa in the superfamily Dictynoidea, within which the greatest range of aquatic lifestyles is found in three different families: Desidae, Cybaeidae, and Dictynidae. Within the Desidae, two genera characteristically inhabit the intertidal zone: (1) the genus *Desis* (Walckenaer, 1837), which is distributed across the southern hemisphere and contains species that hide in kelp holdfasts during daylight and high tide (McQueen and McLay 1983; Wise 1993; Forster and Forster 1999); and (2) the genus *Paratheuma* (Bryant, 1940), which occupies intertidal zones from the South Pacific to Mexico,

 10.1071/IS10001 1445-5226/10/030238

Florida and the Caribbean, and hides in barnacle shells, rock or coral rubble during daylight and high tide (Roth 1967; Roth and Brown 1975*a*; Beatty and Berry 1988*a*, 1988*b*). Within the Cybaeidae, one species is well known for its freshwater aquatic lifestyle: *Argyroneta aquatica*. The sole exemplar of its genus, it lives under water, one of the rarest environments for spiders. It is the only spider to live and hunt fully immersed, using a silken bubble attached to an aquatic plant to hold a breathable air supply (Foelix 1996; Selden 2002). It is found in slow-moving freshwater streams throughout Europe and Asia. Finally, a most unusual species, *Saltonia incerta* (Dictynidae) (Banks, 1898), inhabits an environment that is the historical remnant of an aquatic habitat, the dry alkaline salt-flats in the south-western United States of America and Northern Mexico (Roth and Brown 1975*b*).

Each of these lineages of spiders (representatives of *Paratheuma* and *Desis*, and the single species *S. incerta* and *A. aquatica*) has been difficult to assign accurately to the taxonomic rank of family based on morphology because each shares some special morphological features that may or may not be associated with its unusual lifestyle. For example, *A. aquatica* has modifications to the respiratory system, including fewer lung lamellae, but larger tracheal tubes than most other spiders (Selden 2002), which may provide some advantage when breathing underwater. *Saltonia incerta* has similar book-lungs, but not the obvious association with aquatic habitats (Roth and Brown 1975*b*). Likewise, *S. incerta* shows some morphological similarity to desids, such as having an oversized colulus (a non-functional vestige of a pair of spinnerets) (Roth and Brown 1975*b*). The uncertainty about whether these similarities are homologies or homoplasies led us to use molecular evidence to explore these relationships.

Inferring the evolutionary history of organisms in extreme or unusual environments based solely on morphological characters can be difficult, as adaptation is often accompanied by substantial modifications that may be convergent due to similar environmental pressures (Wake 1991; Givnish and Systma 1997; Lee 1998). The advent of molecular phylogenetics has greatly improved understanding of the role of convergent adaptation to a wide variety of environments (Martins 2000). Here we build on previous work on the aquatic, semi-aquatic, and salt-flat living spider lineages that has shown that these taxa belong to a small number of related families nested within an established spider group, the 'RTA' clade (Coddington and Levi 1991). The monophyly of the 'RTA' clade, so-called because of the characteristic presence of a retrolateral tibial apophysis (RTA) on the male intromittent organ, is supported by recent morphological and molecular analyses (Griswold *et al*. 1999, 2005; Spagna and Gillespie 2008), which provide a framework for rational sampling.

To address the questions of evolutionary history, we generate DNA sequence data and make use of improved methods of analysis of rRNA genes, which form the basis of many of the molecular phylogenetic analyses of both deep and recent divergences in arthropods, including crustaceans, insects, and spiders (Dixon and Hillis 1993; Hedin and Maddison 2001; Wheeler *et al*. 2001; Mallatt *et al*. 2004; Spagna and Gillespie 2006; Crews *et al*. 2010). It has long been known that the secondary structure of these genes influences their evolution, and that constraints differ between the paired bases in self-complementary helices (stems) and unpaired structures (loops) (Schöniger and von Haeseler 1994). Recently, it has become possible to account for the non-independence of stem bases using a doublet model, which has now been applied in several studies on arthropods (Kim *et al*. 2003; Kjer 2004; Angelini and Jockusch 2008; Crews *et al*. 2010). Such partitioned modelling has been shown to improve upon traditional modelling regimes (Erpenbeck *et al*. 2007) and is facilitated here by spider-specific secondary structure models, which exist for the commonly used nuclear ribosomal-RNA genes for the small subunit (18S) and large subunit (28S) rRNAs (Hendriks *et al*. 1988; Hedin and Maddison 2001; Spagna and Gillespie 2006) and can be evaluated and improved upon for use in phylogenetic analyses.

Here, we sample the RTA clade broadly, focussing on dictynoid lineages, including the fully-aquatic, semi-aquatic, and salt-flat endemic species, to improve understanding of the evolutionary history of taxa that occur in habitats not typically inhabited by spiders. To this end, we develop a secondary structure model for 28S rRNA for RTA-clade spiders using thermodynamic algorithms to modify and extend existing models. Secondary structure models (the new RTA-clade-specific 28S rRNA models, and the pre-existing spider 18S rRNA model) are used to identify the self-complementary stem positions for doublet modelling in a partitioned Bayesian analysis framework (Huelsenbeck and Ronquist 2001). With this, we estimate a phylogenetic tree that can serve as a template for further comparative ecological, physiological, and gene-expression studies.

## Methods

### *Taxon sampling*

We chose taxa from throughout the RTA clade (Coddington and Levi 1991), sampling multiple exemplars of the hypothesised relatives of the focal taxa from within the Dictynoidea, in particular Cybaeidae for *Argyroneta*, Dictynidae for *Saltonia*, and both Austral and Holarctic genera of Desidae and their putative relatives (Spagna and Gillespie 2008). We also sampled from families that have been closely allied to the Dictynoidea, in particular Agelenidae C. L. Koch, 1837 (including representatives of the agelenine, tegenariine, and textricine subfamilies) and 'Coelotines' (Amaurobiidae Thorell, 1870) that are associated with riparian habitats (Glesne 1998; Ayoub and Riechert 2004).

Though representative taxon sampling for a highly diverse group such as the RTA clade (>18 000 species; Platnick 2010) is difficult, and for practical reasons imperfect, our choice of taxa is based on both traditional sources, including revisionary and taxonomic literature (Lehtinen 1967; Coddington and Levi 1991; Platnick 2010), and more recent molecular work inclusive of many of the same taxa (Spagna and Gillespie 2008; Miller *et al*. 2010). We follow the latter two papers in our general regime of including exemplars from as many of Lehtinen's (1967) 'subfamilies' as possible. Though Lycosoidea and Dionycha are relatively undersampled here, the purpose of their inclusion is to allow broad sampling across the RTA-clade ingroup, rather than use them to test hypotheses about specific

relationships within the Dictynidae or other 3-clawed taxa, for which deep sampling of lycosoids and dionycha would have little utility. Taxon deposition and GenBank accession information can be found in Appendix 1.

Three non-RTA clade taxa were included as potential outgroups, with exemplars of Nicodamidae Simon, 1897 (*Megadictyna thilenii* Dahl, 1906), Eresidae C. L. Koch, 1950 (*Stegodyphus* sp.), and Araneidae Clerck, 1757 (*Zygiella x-notata* Clerck, 1757) used to represent the groups thought to be the nearest relatives, including the sister-group to the RTA clade (Griswold *et al*. 1999, 2005; Spagna and Gillespie 2008).

### *Molecular markers and amplification conditions*

Markers chosen for this study were nuclear sequences from the D-loop section of the 28S rRNA gene, two large sections covering >90% of the 18S rRNA gene, plus the full sequence for the protein-coding histone H3 gene. These markers, which are among the best characterised in spiders, were specifically chosen for slow accumulation of mutations, making them useful for deciphering phylogenetic relationships in arachnids at relatively deep levels (Giribet *et al*. 1996; Wheeler and Hayashi 1998; Maddison and Hedin 2003; Murphy *et al*. 2006). Faster-evolving markers such as mitochondrial cytochrome *c* oxidase subunit I (COI) and 16S rRNA were not used as they have been shown to contribute little unsaturated signal to analysis of divergences at this depth (Spagna and Gillespie 2008).

The primer pairs 28S O and 28S C (Hedin and Maddison 2001) were used to amplify a section of 28S rRNA sequence (~690 bp) in length. Because some RTA-clade spiders appear to have 28S rRNA paralogues (Murphy *et al*. 2006; J. C. Spagna and R. G. Gillespie, unpubl. data), care was taken to make sure that no double or multiple banding was seen in PCR amplification products, resulting in clean, interpretable chromatograms without double peaks for all sequences used in this project. Two primer pairs, 18S 1F and 18S 5R and 18S 4F and 18S 9R (Giribet *et al*. 1996), were used to amplify the first ~850 and second ~950 base pairs of the 18S rRNA gene, respectively. A 349 base pair section of histone H3 was amplified using the primer pairs histone H3aF and histone H3aR (Colgan *et al*. 1998). For 28S rRNA, Failsafe PCR Optimization Kit (Epicentre Technologies, Madison, WI) buffers were used to improve PCR reaction yields.

The PCR mixtures used were 12 µL total final volume, with a basic mix of 2.5 µL dNTPs, 1.0 µL buffer, 0.2 µL Taq, 2.5 µL per primer, 1.5 µL of template DNA, and 1.8 µL $H_2O$. The buffer used for each rRNA locus was determined using a FailSafe optimization kit, while standard buffer was used for the histone H3 amplifications. For PCR cycles, a basic protocol of 95°C melting temperature for 30 s, 55°C annealing temperature for 40 s, and 72°C extension temperature for 45 s, was repeated for 35 cycles. Minor modifications were made for individual taxa in terms of temperatures and cycle segment lengths (PCR buffers used and cycle optimisations are available upon request from the authors).

### *Alignment, secondary structure assessment, and incorporation in partitioned models*

Alignments of the 18S and 28S rRNA datasets were done using MUSCLE; MUSCLE was chosen for alignment for its relative accuracy, speed, and ease of use (Edgar 2004). Alignment of 28S rRNA was also conducted using secondary structure models to produce a second set of aligned datasets. Alignment using secondary structure has been demonstrated to be more appropriate than traditional alignment methods (Kjer 1995; Angelini and Jockusch 2008). Also, finding the precise location of stems and loops allows for them to be modelled separately in phylogenetic analysis, and also allows for the implementation of a doublet model (Kjer 2004). To account for secondary structure in 18S rRNA, we used the pre-existing model of Hendriks *et al*. (1988) based on sequences from a theraphosid spider, which has been shown to be generally applicable across a broad range of arachnids (Spagna and Gillespie 2006). There was no conflict between the MUSCLE-aligned 18S rRNA data and the secondary structure model used that required post hoc realignment, so bases could simply be assigned to 'stems' or 'loops' based on the Hendriks *et al*. (1988) model.

Modelling of the secondary structure of 28S rRNA was more complex because the interspecific variability in this marker made folding and self-complementary pairing more ambiguous. Existing models for arthropods (Hedin and Maddison 2001) and the mfold web server (Zuker 2003) were used to find appropriate structures for the amplified sections (~690 bp) of the D2 and D3 segments of 28S rRNA. Mfold uses a free energy algorithm following the thermodynamic rules of SantaLucia (1998), and the default parameters were used.

Assessment of behaviour of the mfold algorithm demonstrated that lower folding temperatures resulted in consistently greater levels of pairing within each arm than at the default temperature of 37°C (mean increase in the number of bases forming pairs was 48% at 21°C, 49% at 4°C). Though this default temperature setting for the mfold algorithm might appear high for an ectothermic animal, the spiders sampled live at a wide range of temperatures, often exceeding 37°C in the case of *Saltonia*, so it falls within realistic physiological limits of the animals. Because we lack corroborating crystallographic data for our sequences' secondary structures, we used 37°C as a conservative approach to base-pair modelling, as any inappropriate assignment of bases to stem partitions would result in underweighting these, effectively reducing the size of the matrix.

Each helix of the D2 region for each taxon was input singly into the server, thus producing a possible structure, or structures, for each helix of each specimen. The entire 'core + D3' region was input for each specimen, producing a possible structure. Gaps were inserted and adjusted by eye to produce the final alignment, in such a way to minimise discontinuities in the stem structures shared across taxa, assuming that unpaired loop regions are more likely to have undergone the indel events represented by the gaps. In the final alignment, it was ambiguous whether certain bases were stems or loops, and in these cases, the ambiguities were assigned to loops.

The protein coding gene histone H3 (349 base pairs) was aligned by eye after the sequences were translated to codons, which also allowed us to determine that there were no stop codons in any of the sequences.

With the data matrix including lineages with known long branches, inconsistent molecular clocks (Spagna and Gillespie 2008), and deep divergences (>100 million years) (Penney 2003),

we used a Bayesian likelihood approach to analyse the data in order to avoid possible long-branch attraction artefacts (Felsenstein 1978; Bergsten 2005). Additionally, secondary structure information for rRNA genes, using both pre-existing models and new models developed herein, was included in the Bayesian analyses, in order to compensate for the non-independence of base-pairing in the self-complementary helices of the 'stem regions' of rRNA.

Although generally considered superior to parsimony methods for dealing with long branches, Bayesian methods have also been shown to be potentially susceptible to such artefacts (Kolaczkowski and Thornton 2009). For this reason, we performed accelerated maximum likelihood analysis using RAxML-HPC 7.0.4 (Stamatakis 2006*a*) on the CIPRES Portal v 2.1 (Miller *et al*. 2009) for comparison with Bayesian outcomes. The matrix was partitioned into stems and loops for the rRNA genes and by codon position for histone H3 using the GTRCAT model (Stamatakis 2006*b*) for each partition, with the number of rate categories set to 25. Bootstrapping was conducted using the rapid bootstrap algorithm for webservers (Stamatakis *et al*. 2008), and these values were calculated for nodes on the best-scoring tree.

Bayesian analysis was conducted on each gene separately and in one concatenated dataset for both the MUSCLE-aligned datasets and the structurally aligned datasets using mixed-model partitioned analyses (Nylander 2004; Brandley *et al*. 2005) in MrBayes 3.1 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). The rRNA genes were partitioned by stems, half-stems and loops, using a doublet model of nucleotide substitution (Kjer 2004; Angelini and Jockusch 2008) for full stems to account for the non-independence of bases for which we had explicit information about complementary base positions, while histone H3 was partitioned by codon position. The doublet model as implemented in MrBayes 3.1 is highly parameterised (16 parameters total), but previous studies have shown it to be superior to models that assume independence (Tsagkogeorga *et al*. 2009). The primary risk from over-parameterisation is an increase in variance of parameters (e.g. branch lengths), rather than decreased accuracy in consensus tree topology (Huelsenbeck and Rannala 2004; Kelchner and Thomas 2007). For all non-stem partitions, models were chosen using the AIC in MrModeltest 2.2 (Nylander 2004; Posada and Buckley 2004). Chosen models are shown in Table 1. Gaps were treated as 'missing data' in all analyses. The 28S rRNA and H3 datasets were run for 10 million generations, the 18S rRNA dataset was run for 9 million generations, and the concatenated dataset was run for 12 million generations. The *cumulative* and *compare* commands were used in AWTY (Wilgenbusch *et al*. 2004) to determine burn-in and compare the two results of each MrBayes run.

**Table 1. Partitions and models used in partitioned Bayesian analyses**

| Partition | Selected model |
|---|---|
| 18S rRNA stems | HKY+I+Γ (doublet) |
| 18S rRNA loops | GTR+I+Γ |
| Histone H3 position 1 | GTR+I+Γ |
| Histone H3 position 2 | HKY |
| Histone H3 position 3 | GTR+I+Γ |
| 28S rRNA stems | GTR+Γ(doublet) |
| 28S rRNA loops | GTR+Γ |
| 28S rRNA half-stems | GTR+Γ |

## Results

### *Modelling of secondary structure*

The 18S rRNA model for spiders of Hendriks *et al*. (1988) was found to be sufficient for assignment of strings of bases to loop partitions, and pairs of bases to complementary stem partitions, for model and parameter partitioning before analysis. The 28S rRNA data were considerably more variable in terms of both base and length polymorphisms, and the secondary structure models developed for the region sequenced are seen in Figs 1–5.

Arm 1 is the most variable of the 28S rRNA structures modelled, and a representation of the variation can be seen in Fig. 2. Variation occurs both between and within clades (see phylogenetic results), in terms of total transcript length, and number and size of loops and stems. Within-taxon variation occurs even in species of the same genus (such as *Paratheuma insulana* (Banks, 1902) and *P. armata* (Marples, 1964); Fig. 2, Group1). Of groups represented in Fig. 1, the Cybaeidae *s.l.* and dictynoids + *Desis* appear the most structurally uniform, with the Agelenidae (including the subfamily Coelotinae) appearing quite variable.

Structures for Arm 2 are much more uniform than Arm 1, with most variation occurring in the form of presence or absence of a single loop (represented by an asterisk in Fig. 3) in the terminal helix of the arm (Fig. 3, panels 1–4).

While longer than both Arm 1 and Arm 2, across the spider taxa modelled, the Arm 3s (Fig. 4) are similar in uniformity to Arm 2, with a set of 7–8 positionally homologisable loop structures. Though not all Arm 3s have identical numbers of loops, the task of homologising individual loops was made simpler by the amount of asymmetry, possibly resulting from slip-strand mispairing events (Gillespie *et al*. 2004), between the 5′ and 3′ sides of the intervening loops, as these were often remarkably unequal, and this asymmetry was fairly consistent across taxa.

The primary sequences for the D3 + core structures were, in most cases, highly similar and thus alignment was largely trivial, simplifying the modelling of stems and loops across multiple taxa, with some exceptions. The structure for *Zygiella x-notata* (Fig. 5) represents the structure for conforming taxa, while the three taxa with notable variations from the primary sequence length (*Paratheuma*, *Saltonia*, and *Dictyna*) were modelled individually. Structurally, *Paratheuma* and *Dictyna* appear to share a positionally homologous, self-complementary four-base-pair motif, which makes the centre of the D-loop core appear 'pinched' into two smaller loops (top 2 panels of (F 5), which is not seen in the other study taxa.

### *Combined data phylogeny*

The majority-rule consensus tree for 6000 post-burn-in combined-data trees (representing 6 million generations, after 4 million generations were discarded as burn-in) is shown in Fig. 6. The monophyly of the ingroup, representing all sampled members of the RTA clade, was strongly supported (posterior probability (p.p.) = 1.0). The focal taxa of this study (those with

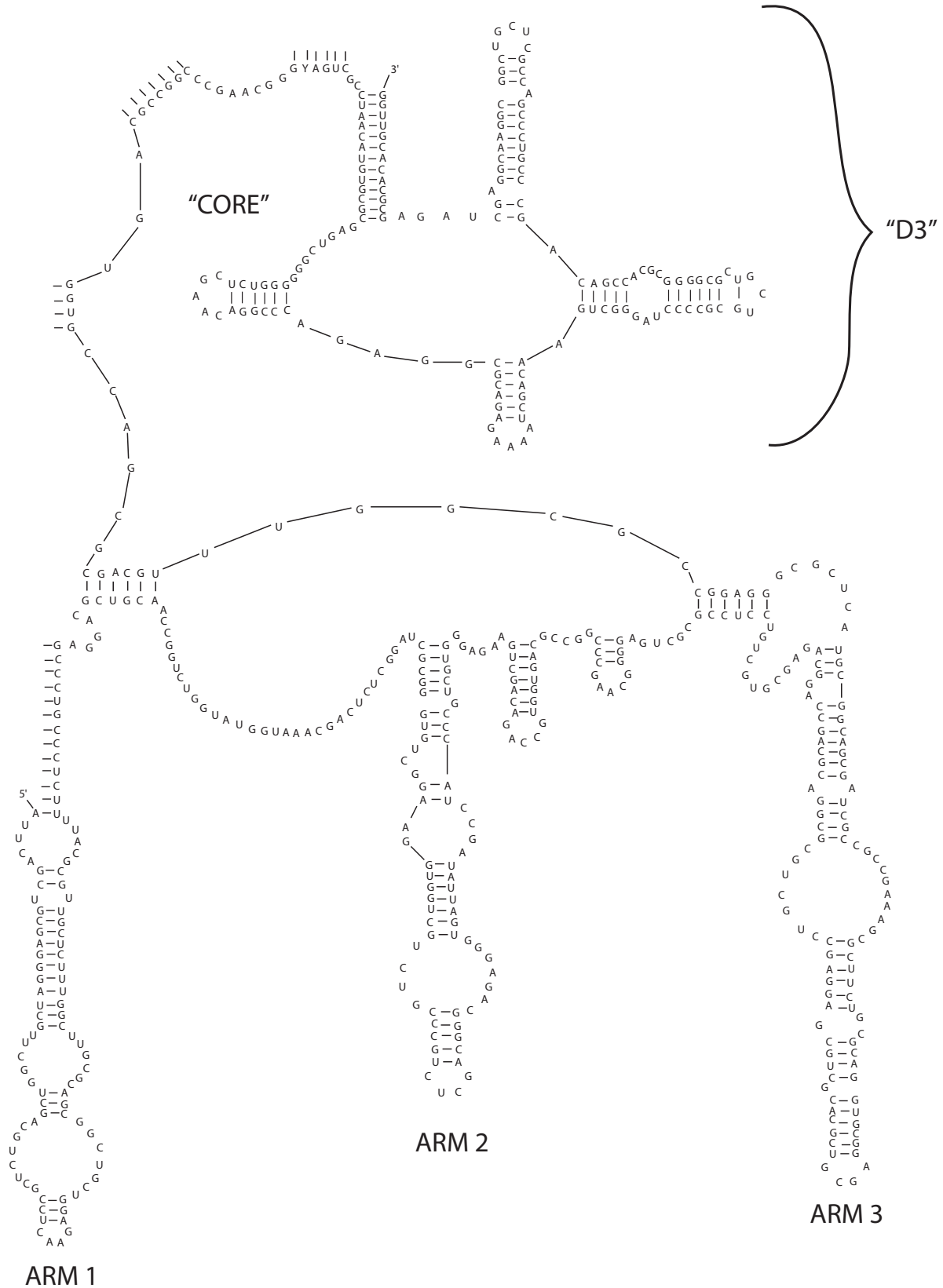


**Fig. 1.** Hypothesis of secondary structure for the D2 and D3 regions and the core region connecting them for *Saltonia incerta*, showing relative positions of the Core, D3 and Arms 1–3 in the 5′ (lower left) to 3′ (upper right). Arm and core orientations are maintained in Figs 2–5 for comparison.

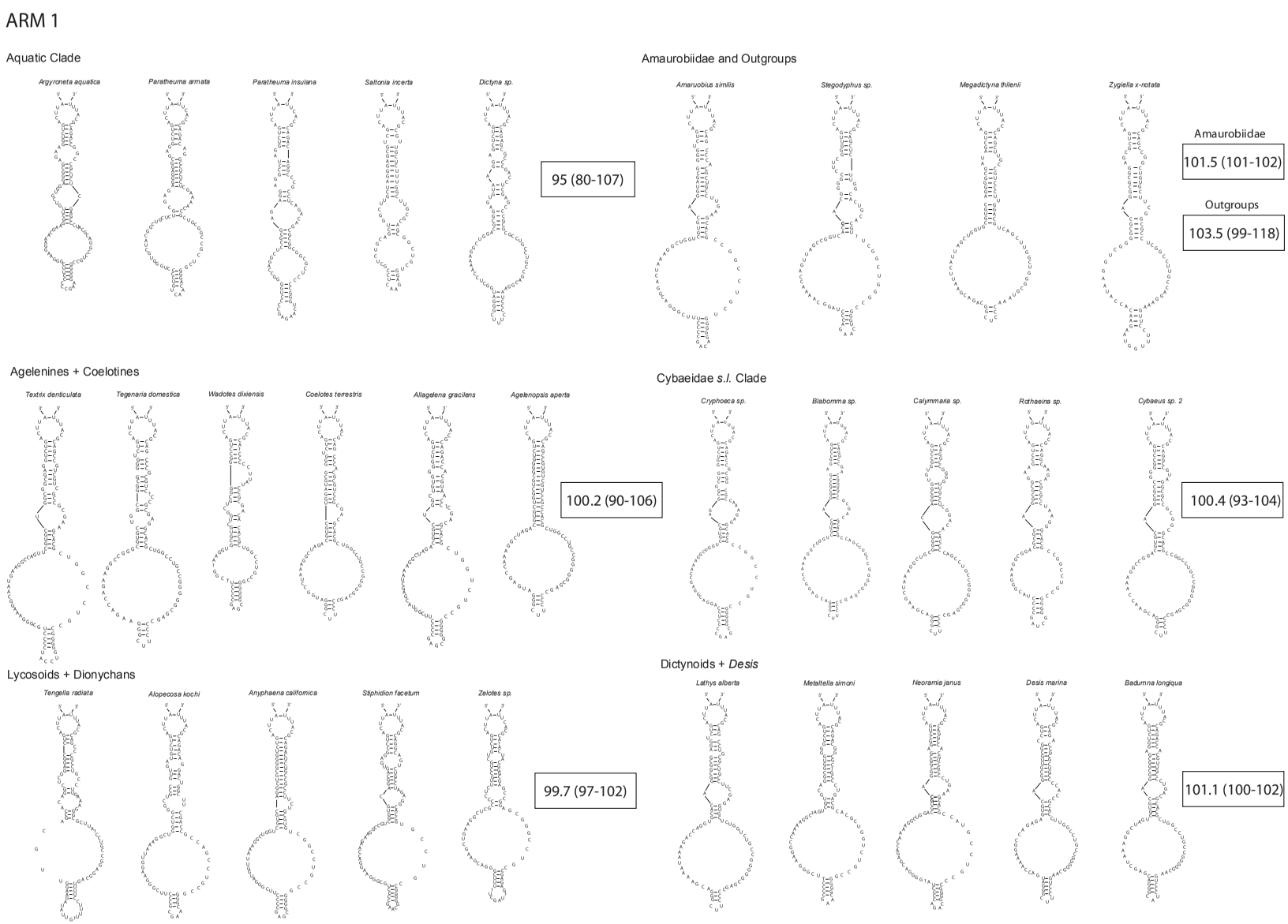


**Fig. 2.** Groups 1 and 2: the secondary structures of the 'Arm 1' of the D2 region of 28S rRNA from members of the 'aquatic clade', the amaurobiid clade and three outgroup taxa from the combined-data tree. Arm 1 is highly variable, even within genera, as demonstrated by the two species of *Paratheuma*. Groups 3 nad 4: Arm 1s from members of the Agelenidae and Coelotinae (Group 3) and the cybaeid *sensu lato* clade (Group 4). In general, the latter appear more structurally conserved than the former. Groups 5 and 6: Arm 1s from members of the *Lathys* (Desidae) and Austral clade spiders (row 5) and for Lycosoids and Dionychans (Group 6). Of the former, the genus *Desis* has a particularly conserved secondary structure for this feature. The arms shown in the secondary structure figures were chosen to represent the range of variation. Boxes to the right of each group represent mean length and minimum-maximum range (in parentheses) to capture full variation in figured and unfigured arms.

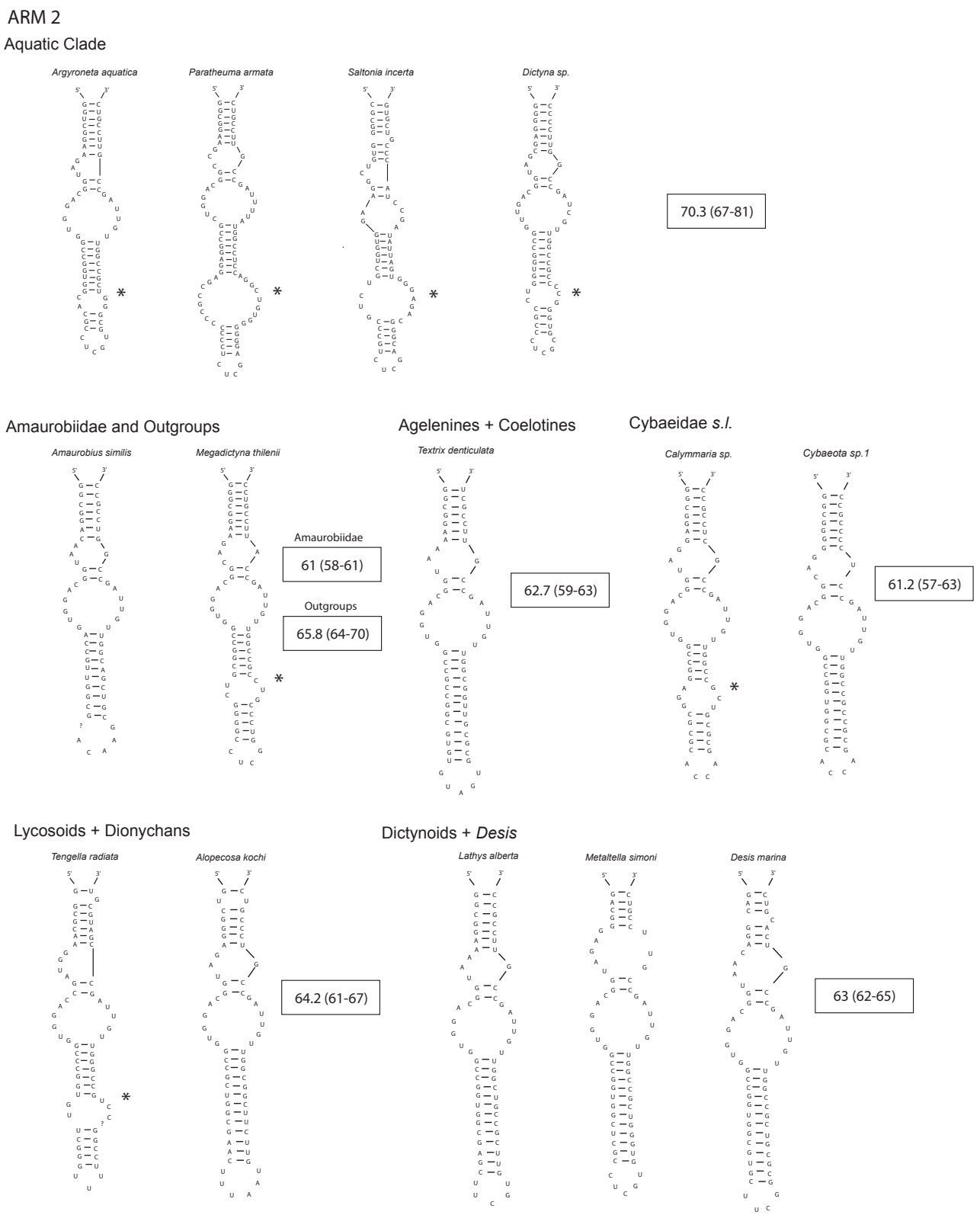


**Fig. 3.** Groups 1 and 2: the secondary structures of the 'second arm' of the D2 region of 28S rRNA from members of the 'aquatic clade' + *Dictyna*, (Group 1) an Amaurobiid, an outgroup taxon, a representative from the Agelenidae, and the *Cybaeidae s.l.* Clade (Group 2). Groups 3 and 4: Arm 2s from Lycosoids, Dictynoids and *Desis*. There is little length variation in this region overall. Arm 2 secondary structure appears to be quite conserved across taxa, with most variation seen in the presence of a single loop structure (marked with an asterisk) in some taxa. Boxes as in Fig. 2.

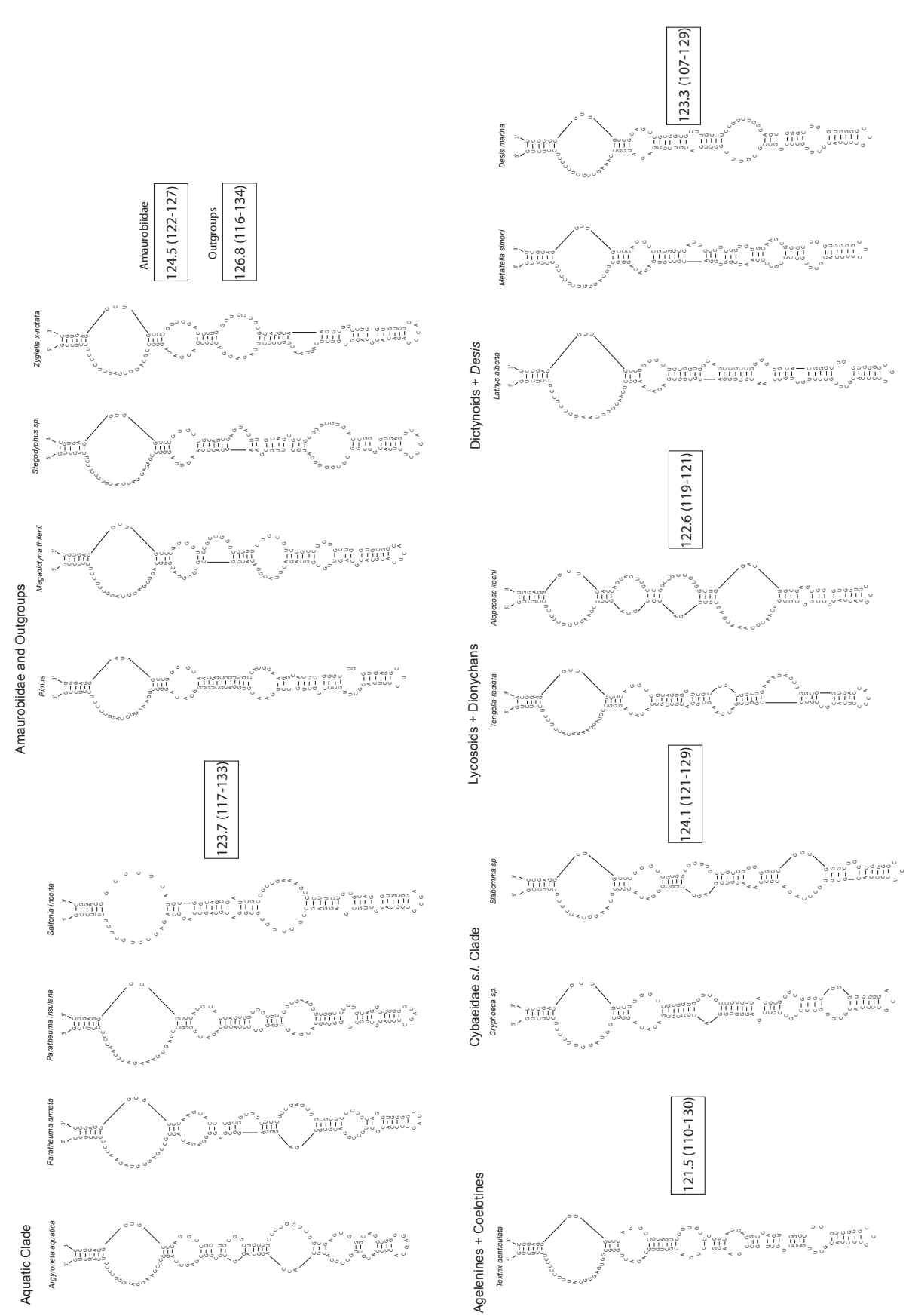


**Fig. 4.** Groups 1 and 2: the secondary structures of the 'third arm' of the D2 region of 28S rRNA from members of the aquatic clade (Group 1), the amaurobiids and outgroup taxa (Group 2). Groups 3 and 4: Arm 3s from a representative of the Agelenidae, the Cybaeidae *s.l.* clade, a lycosoids and Dictynoids plus *Desis*. Boxes as in Fig. 2.

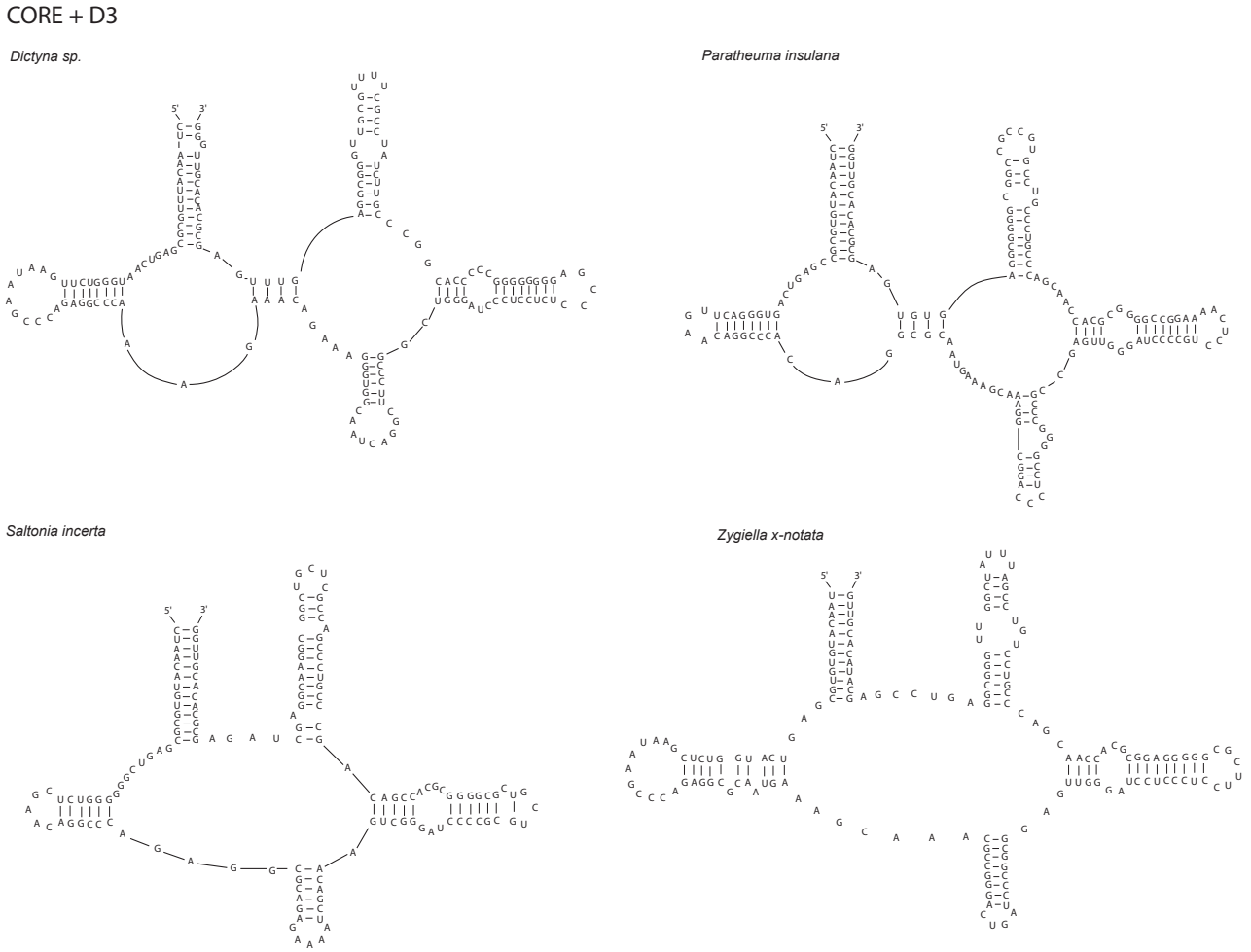


**Fig. 5.** The secondary structures for the 'core' plus 'D3' regions of 28S rRNA. This region was highly conserved across taxa, thus only a few divergent representatives are shown. These include *Dictyna*, *Saltonia* and *Paratheuma* from the 'aquatic' clade and two outgroup taxa. *S. incerta* is fairly divergent in this region of 28S rRNA, which is typically highly conserved. Several deletions appear to have taken place in the *Saltonia* 28S rRNA gene.

extreme microhabitat affinities) are found in two distinct clades, consistent with geography, but not with current taxonomy. These are represented by the thick green, blue and red branches in Fig. 6. The first clade includes two species (*Desis formidabilis* (O. Pickard-Cambridge, 1890) and *D. marina* (Hector, 1877)) that are nested within a group including the Austral cribellate spiders of New Zealand, while the second includes all exemplars of *Argyroneta*, *Saltonia*, and *Paratheuma*, as well as *Dictyna* sp. (p.p. = 1.0).

Additional, larger clades approximating taxa at the 'family' level comprise a cybaeid group including *Cybaeus* L. Koch, 1868, *Calymmaria* Chamberlin & Ivie, 1937, *Cybaeina* Chamberlin & Ivie, 1932, *Cybaeota* Chamberlin & Ivie, 1933, New genus #4 (Bennett 1991: p. 156; Ubick *et al.* 2005), *Blabomma* Chamberlin & Ivie, 1937, *Yorima* Chamberlin & Ivie, 1942, New genus #1 (Bennett 1991: p. 168; Ubick *et al.* 2005) and *Cryphoeca* Thorell, 1870 and an agelenid group including all Holarctic taxa, plus monophyletic Agelenidae (including Coelotines) (both supported with p.p. > 0.95). Other nodes with strong statistical support include a clade consisting of all cribellate Amaurobiidae sampled (p.p. = 1.0), and a sister relationship between the aforementioned Holarctic agelenids and Cybaeidae *s.l.* (p.p. = 0.96).

Comparing the topology from the combined-data, structurally modelled analysis with the topology from the MUSCLE-aligned analysis shows 26 nodes statistically supported (p.p. > 0.95; red asterisks in Fig. 6) in both analyses, while 10 additional nodes (marked with black asterisks) are supported at this level only in the structural analysis. Two nodes (grey asterisks) receive statistical support only in the analyses of the MUSCLE-aligned data, while one node (*Dictyna* Sundevall, 1833 as the sister taxon to *Paratheuma* + *Argyroneta*) receives support in the MUSCLE analysis but remains unresolved at the 0.50 level in the structurally modelled tree, and thus does not appear in the figure.

The tree topology yielded by the combined-data RAxML analysis was largely congruent with the Bayesian trees, with a major difference in the placement of the genera *Pimus* Chamberlin, 1947, *Amaurobius* C. L. Koch, 1837 and *Callobius* Chamberlin, 1947. These cribellate Amaurobiidae, which emerge as a clade nested well within the RTA clade in the Bayesian tree (see node marked with dagger in Fig. 6), appear

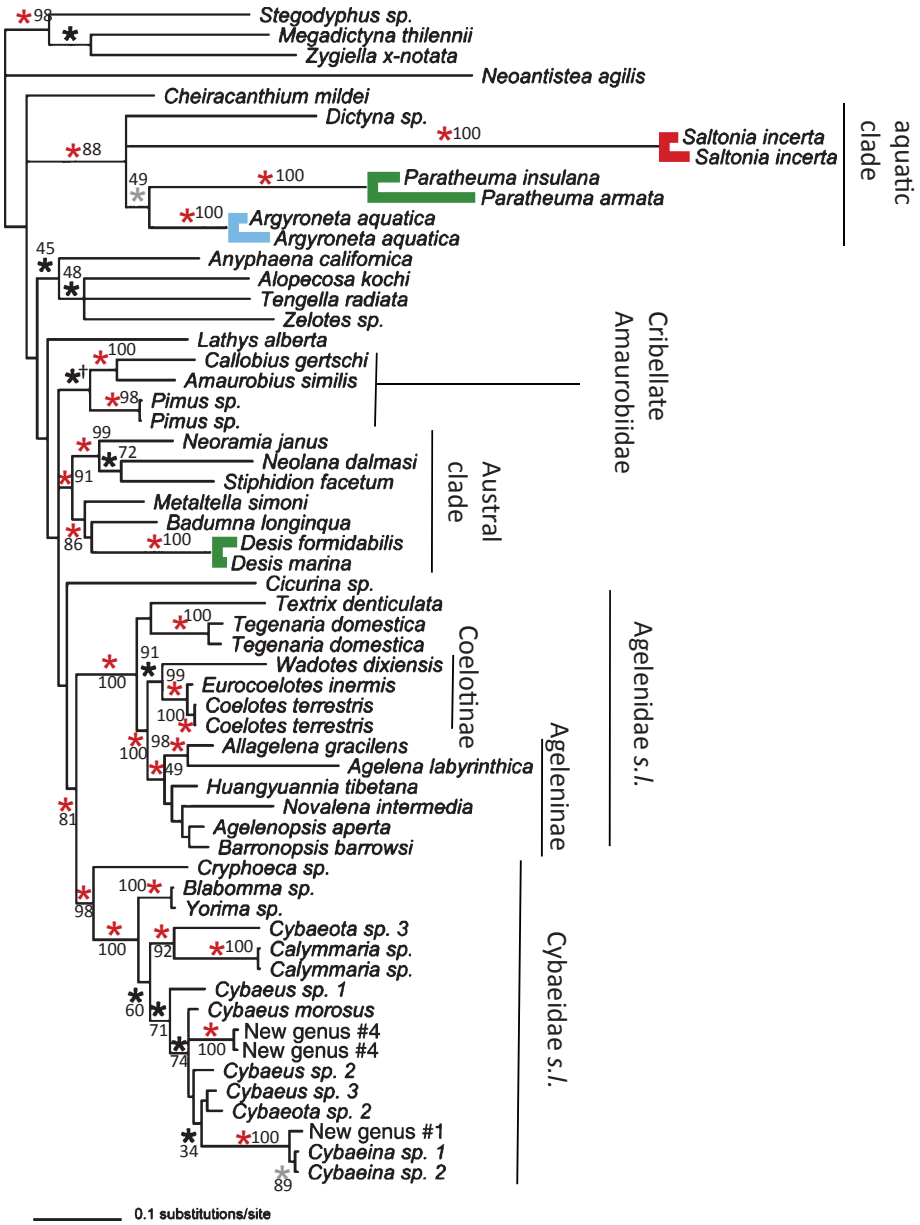


**Fig. 6.** Combined-data tree. Majority-rule consensus tree of 6 million generations after stationarity was achieved in Bayesian analysis of the combined-data matrix. Asterisks (*) denote branches receiving posterior probability support of ≥95% in the MUSCLE-alignment only (grey asterisks), structurally aligned and doublet-modelled only (black asterisks), or in both alignments (red asterisks). Coloured branches denote habitat affinity as follows: blue = aquatic, green = intertidal, and red = salt-flat/salt crusts. Large monophyletic groups supported in combined analysis and discussed in text are labelled with clade names. Numbers represent bootstrap values from the RAxML analysis on nearest nodes receiving statistical support in Bayesian analysis, with the dagger (†) representing the single supported node not supported in the best ML tree (see text for details).

as a basal clade (bootstrap value 27% for the node separating the *Pimus* species from *Amaurobius* + *Callobius*) in the RAxML analysis. Bootstrap values for the RAxML analysis can be seen in Fig. 6.

All combined-data trees show a branch to *Saltonia* that is longer than the rest of the branches in the tree, though this effect is much more extreme in the RAxML tree than in the Bayesian doublet-modelled tree.

### *Single-gene trees*

#### *18S rRNA*

The single-gene tree from the 18S rRNA data (Fig. 7) largely reflects the relationships seen in the combined-data tree with a few notable exceptions. The split between the Holarctic and Austral intertidal desids is present in the 18S rRNA tree (thick, green branches); however, within the Holarctic clade, the intertidal taxa (*Paratheuma*) form a tritomy with *Argyroneta* and (*Saltonia* + *Dictyna*) instead of being sister to *Argyroneta*. The Holarctic agelenid clade is present, but with poor resolution between the coelotine, agelenine, tegenariine and textricine taxa. Most (21 of 22) nodes receiving statistical support in the structurally partitioned tree were also supported in the MUSCLE-aligned 18S rRNA tree, while five nodes were supported in the MUSCLE-aligned tree only. Two of these five nodes do not appear in Fig. 5 – (*Cicurina* Menge, 1871 + *Cheiracanthium* C. L. Koch, 1839 + *Amaurobius* + *Callobius* + 'aquatic clade' and *Saltonia* + *Paratheuma*), as the tree in the figure represents the majority-rule consensus of structurally modelled trees. The

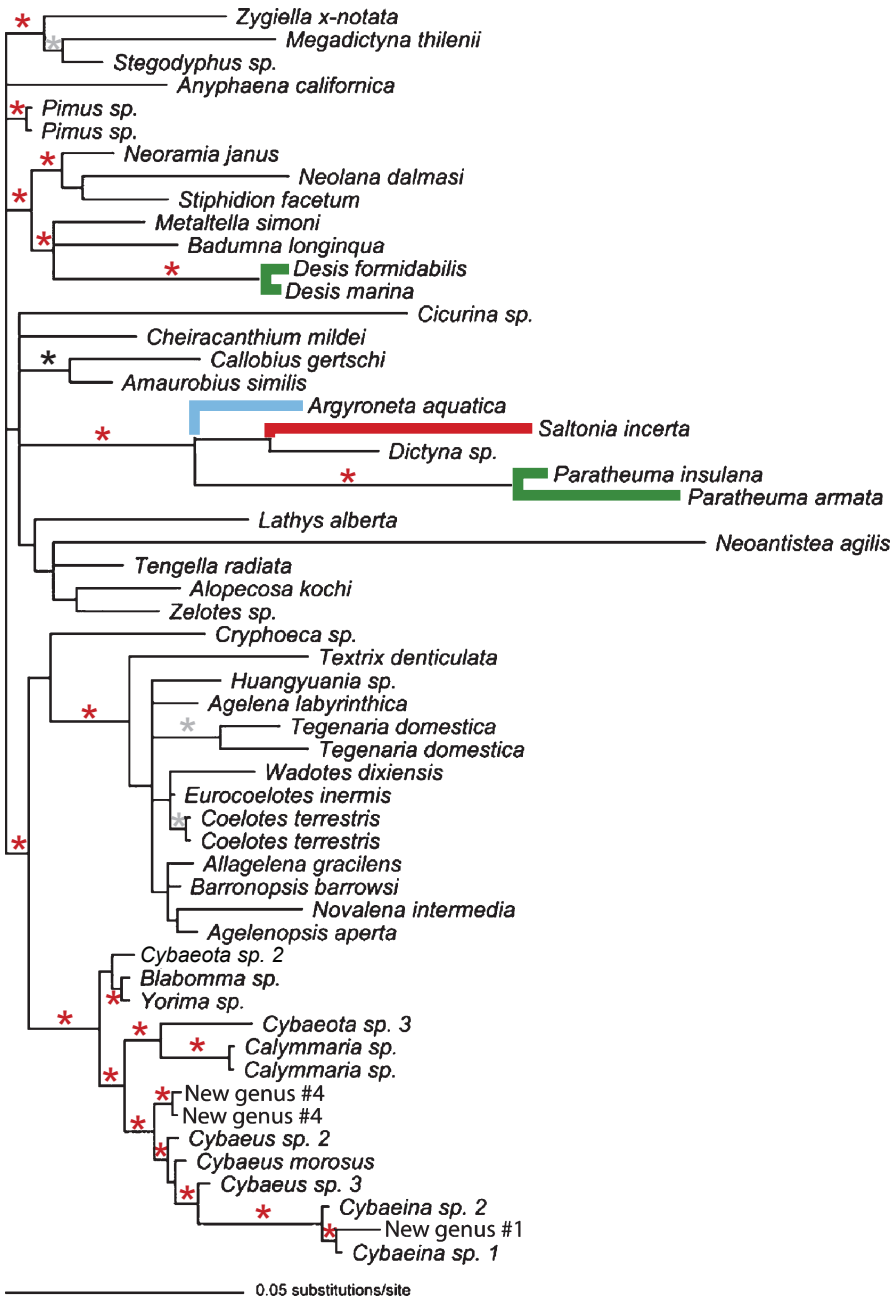


**Fig. 7.** 18S rRNA gene tree. Majority-rule consensus of 4 million post-stationarity generations of the 18S rRNA matrix. Labelling and colour conventions as in Fig. 6.

branch leading to the *Saltonia* taxa is the longest in the tree, as in the combined-data tree.

### *28S rRNA*

The majority-rule consensus 28S rRNA tree (Fig. 8) reflects many of the same general relationships as the combined-data and 18S rRNA gene trees, but shows less support for some key nodes of interest, and does not provide statistically significant (95% p.p.) support for some of the nodes in the combined-data tree that are the focus of this study. Individual differences include paraphyly of the RTA clade ingroup, dissolution of the Austral clade, and the presence of the *Saltonia* clade in an unsupported sister-relationship with *Neoantistea* Gertsch, 1934 (Hahniidae Bertkau, 1878). The relationships of the focal taxa – *Saltonia*, *Argyroneta*, *Paratheuma* and *Desis* – are similar to the combined-data tree, including the geographic split between intertidal species; however, neither the node supporting those inhabiting the northern hemisphere (*Paratheuma* spp.), nor the node supporting the Austral association is statistically supported at the 95% (p.p. = 0.91 and 0.88, respectively).

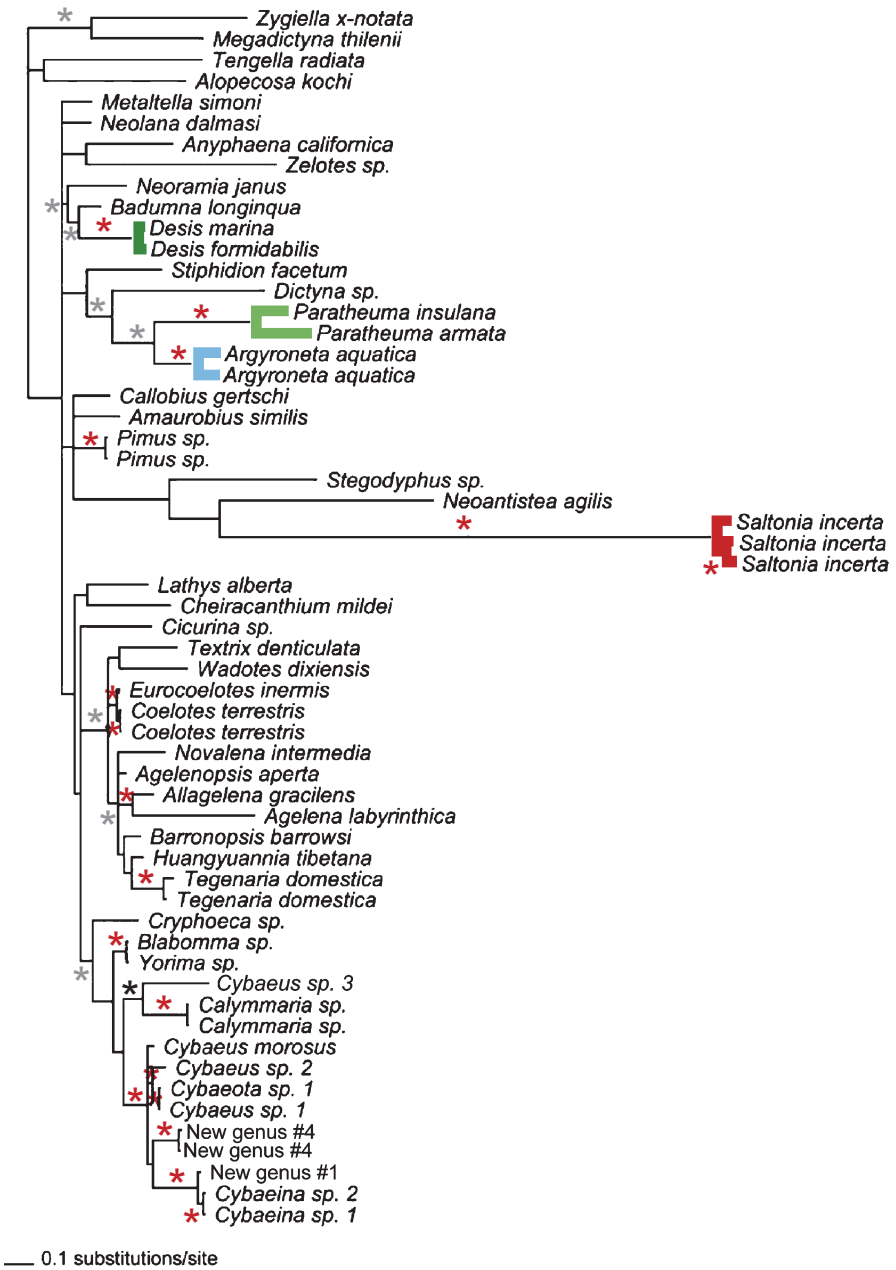


**Fig. 8.** 28S rRNA gene tree, D2 + D3 and core regions. Majority-rule consensus of 11 thousand trees sampled from 11 million post-stationarity generations of the 28S rRNA matrix. Labelling and colour conventions as in Fig. 6.

Neither the Holarctic Agelenidae nor the cybaeid group (Fig. 6) achieve statistical significance (p.p. = 0.89 and 0.91 respectively) in the 28S rRNA gene tree, and both contain unresolved and poorly supported internal nodes. Eighteen of 19 supported nodes in the structurally modelled 28S rRNA gene tree were also supported in the MUSCLE-aligned analysis, while 10 (including 2 not in the majority-rule consensus structural tree- *Stegodyphus* + *Megadictyna* Dahl, 1906 + *Zygiella* F. O. P.-Cambridge, 1902 and an Austral clade including *Stiphidion* – not figured) fell below the 0.95 p.p. threshold when the structural partitions were used.

### *Histone H3*

Contrasting with the gene trees from rRNA matrices, the protein-coding histone H3 gene tree (not included as a figure), based on 349 bp of data, offers little resolution at the 95% p.p. level, with the exception of some non-controversial groupings of closely related taxa. These stable groupings include monophyly of three focal genera (as genera, not as a single inclusive group): *Desis*, *Paratheuma* and *Saltonia*, and monophyly of all agelenine exemplars (*Agelena* Walckenaer, 1805 + *Agelenopsis* Geibel, 1869 + *Barronopsis* Chamberlin & Ivie, 1941) and coelotine exemplars (*Eurocoelotes* Wang, 2002 + *Coelotes* Blackwall, 1841).

## Discussion

### *Secondary structure patterns in the RTA clade*

The Arm 1 structures from 28S rRNA (Fig. 2) for members of the aquatic clade (Fig. 6) appear reduced in length relative to other exemplars, but there is no clear trend towards increase or reduction in size across taxa. Of the large clades represented in Fig. 1, the Cybaeidae *s.l.* and Dictynidae appear the most structurally uniform, while the Agelenidae and Coelotinae structures appear quite variable. For Arm 2 (Figs 1, 3), the aquatic-clade exemplars vary little in general secondary structure, with visually homologisable loops and stems in all taxa.

Overall, the structural conservation of the loops and stems of all three arms of the D2-loop sequenced suggest that our methods produced reasonable hypotheses of primary homology for modelling stems and loops, and can be used as a basis for similar work at a variety of levels across the RTA clade, and possibly in related spider taxa. Though modelling genes and arms individually was tractable for this study, for a larger dataset (either more sequence data, or more taxa), a more automated approach might be desirable. In such a case, a variety of methods for determination of secondary structure based on two or more RNA sequences are available, such as Dynalign (Mathews and Turner 2002, as implemented in the software package RNA Structure ver. 5.03, Mathews 2010) and RNAalifold (Hofacker *et al*. 2002; Bernhart *et al*. 2008).

### *Doublet modelling*

While doublet models have been shown to perform well with single genes (Kim *et al*. 2003; Angelini and Jockusch 2008), there are few published studies addressing the performance of doublet models when used with multiple gene partitions (16S and 28S rRNA genes – Deans *et al*. 2006; 18S and 28S rRNA genes – Rix *et al*. 2008), and none that deal with divergences at the temporal scale addressed here, in which the divergence point between the ingroup and outgroup taxa has been estimated to be mid-Jurassic (Penney and Ortuño 2006); this deep divergence is reflected in the high degree of variation in the arms of the D-loop region of 28S rRNA. Despite this variation both within and between taxa, particularly for Arm 1 of the D-loop region of 28S rRNA, there was sufficient similarity to allow primary homology assessment and assignment of stretches of primary sequence to partitions.

It might appear that the application of the doublet results in a 'worse' outcome, if one focusses only on support values and resolution of individual clades in single-gene trees. The 28S rRNA tree (Fig. 8) has considerably less support (10 of 28 shared nodes fall below the statistical threshold of 95% posterior probability) when the stem–loop partition and doublet model are applied than when the computer-generated alignment is used. The effect is less stark in the 18S rRNA tree, with only 3 of 25 shared nodes losing statistical support when the doublet is applied. However, to understand whether the loss of support in single-gene trees really represents a loss of analytical power, it is important to consider the changes in the topology and support of the combined-data trees (Fig. 6). The tree remains well resolved and well supported when the doublet model is applied, and in fact the difference between the MUSCLE-aligned trees and the doublet-modelled trees shows the opposite effect than when applying the model to the single-gene datasets: instead of losing statistical support, 10 additional nodes gain support through application of the doublet model, leading to a consensus tree in which 36 nodes are supported by statistically significant posterior probabilities.

This effect – loss of support in single-gene trees, but increase in support in combined-data trees – has been explained by Kjer (2004) in his study using 18S data of holometabolous insects: reduction in support may result when pairs of non-independent characters are effectively down-weighted when recognised and treated as non-independent. With pairs of bases as the unit character, the total number of stem characters is reduced by half, and fewer characters can reduce resolution. Even if the outcome of paired-sites modelling is not higher support or an increase in some other measure of confidence in the phylogenetic hypotheses generated, it does correct the violations of assumed independence inherent in most models used in likelihood analysis of self-complementary genes; reduced support may in fact be an improvement in the analytical outcome, if the alternative is high support for an incorrect clade.

In the context of a multi-gene analysis such as this study, the support lost in the single-gene trees can be made up for by strong signal from other partitions, but this pattern can only be seen when multiple genes are used. This is not always the case, though, as a comparable study of a different clade of spiders (Rix *et al*. 2008) showed decreased support in a combined-data 18S rRNA + 28S rRNA tree. This might be due to a smaller dataset (720 bp in the former study versus ~1800 used here, and no protein-coding data in that study), and underlines the importance of alternate lines of evidence (multiple genes, with a variety of genomic sources) in a combined analysis.

The length of the *Saltonia* branch in the combined-data analyses, which appears to be contributed primarily by the 18S rRNA data (Fig. 6), may result from either rapid evolution of the 18S rRNA gene, or a disproportionate amount of speciation/

extinction in the *Saltonia* lineage (Bergsten 2005). Though gene-duplication and functional divergence in paralogues has been shown to result in rapid genetic change leading to long branches (Fares *et al*. 2006), examination of secondary structure relative to models from other spiders shows no particular violations of stem–loop expectations, which one would expect to find in a pseudogene. Post hoc examination of base composition reveals no significant deviation in the *Saltonia* 18S rRNA gene (A = 24%, T = 26%, C = 22%, G = 28%) from the percentages found in other taxa (A = 24%, T = 25%, C = 23%, G = 28%; $\chi^2 = 0.89$, d.f. = 3, $P > 0.8$). We remain confident that the *Saltonia* 18S rRNA genes sequenced are orthologues of those from other spiders, both their adherence to secondary-structure rules for eukaryotes generally and spiders specifically (Hendriks *et al*. 1988; Spagna and Gillespie 2006), and the fact that we sequenced three individuals from distant populations and found them to be quite similar. The branch's phylogenetic position was robust to changes in method (Bayesian with/without doublet, and RAxML), and it showed no tendency to group with other long branches, such as the outgroup branches, or the hahniid exemplar *Neoantistea*. Use of the doublet model reduced the relative length of the branch, probably due to down-weighting complementary changes in paired stem bases, which further supports maintenance of function in the gene (stems would not be under selective pressure to maintain pairing integrity in a pseudogene), although the forces driving the relatively rapid accumulation of base differences remain unknown.

### Taxonomic affinities of target taxa

An unexpected outcome of the study was the close phylogenetic affinity among three of the target taxa, the arachnid extremophiles *Argyroneta*, *Paratheuma* and *Saltonia* (Fig. 6). The combined analysis placed these taxa together in a clade with *Dictyna*. Though the grouping of *Paratheuma* and *Argyroneta* has been predicted based on previous molecular analyses (Spagna and Gillespie 2008; Miller *et al*. 2010), the close relationship between these two taxa and *Saltonia*, the position of which has long been enigmatic in spider systematics (Roth and Brown 1975*b*; Foelix 1996), is unexpected. Indeed, the combined-data tree supports none of the current family-rank groupings of spiders with extreme microhabitat affinities: *Argyroneta* does not fall with others in the Cybaeidae, as has long been classified based on morphology and fossil evidence (Simon 1892; Lehtinen 1967; Selden 2002), but rather must be derived from a dictynid-like ancestor, while probably being sister to *Paratheuma*. The new position of the latter taxon – traditionally and currently considered a desid (Platnick 2010) – would make Desidae polyphyletic. Dictynidae is also polyphyletic in these analyses, with four representatives (*Dictyna*, *Lathys* Simon, 1884, *Cicurina* and *Saltonia*) never forming a clade in the combined or single-gene trees.

Despite conflicts with morphological taxonomy, the results of these analyses are robust, and are largely consistent with other recent studies including RTA-clade taxa (Griswold *et al*. 1999, 2005; Spagna and Gillespie 2008; Miller *et al*. 2010). The increase in support when the data are modelled to account for base-pairing, and data from all three genes are included (Fig. 6), suggests that we have developed a hypothesis that is more reliable than the consensus result for the matrix when secondary structure is not considered. Thus we consider our approach to be a major methodological improvement for the study of ancient divergences in spiders, although future work should benefit not only from use of secondary structure, but also from the development of newer, possibly better-parameterised, likelihood models incorporating doublet information. Such models include the RNA7a and RNA7b models as implemented in the program PHASE (Gowri-Shankar and Jow 2006) and their incorporation into partitioned Bayesian analyses. Comparison of Bayesian likelihood and accelerated likelihood underlines the consistency of the well supported clades here, with the exception of the 'Cribellate Amaurobiidae', which are placed in a basal position, and are not monophyletic within the RTA clade, only in the RAxML analysis. Though this difference does little to settle the difficult problems of the phylogenetic composition and relationships of a monophyletic Amaurobiidae, it has little effect on the evolution of the dictynoid extremophiles, and parallel evolution between Austral and worldwide species, which are the foci for this study.

### Geographic convergence

It has been suggested that the spider fauna of New Zealand and Australia represents an assortment of relict taxa, with families such as Gradungulidae Forster, 1955 and Austrochilidae Zapfe, 1955 maintaining plesiomorphic features not common in more widely distributed spider taxa (Forster 1987; Forster and Forster 1999). This idea has affected studies of spider evolution by influencing taxon sampling in datasets assembled to address deep phylogenetic questions. For example, based on the premise that Austral taxa often represent basal lineages of worldwide families, the taxon *Neoramia* has been used as the sole 'plesiomorphic' exemplar of the family Agelenidae in a variety of studies (Forster and Wilton 1973; Coddington and Levi 1991; Griswold *et al*. 1999; Blackledge *et al*. 2009). Previously, this approach has been shown to be flawed, as the Austral and worldwide Agelenidae are not monophyletic (Spagna and Gillespie 2008; Miller *et al*. 2010). The current study also renders the family Desidae polyphyletic, and shows independent origins for derived similarities in widely separated genera, rather than simple plesiomorphy in the basal/Austral restricted branches of a worldwide family. Within the Australian desid taxa, the intertidal species has shifted to ecribellate silk production, supporting the general supposition that Austral taxa exhibit shifts between these cribellate and ecribellate morphologies (Forster and Wilton 1973; Forster and Gray 1979; Griswold *et al*. 1999), while reinforcing the more recent hypothesis that the inclusion of Austral taxa in worldwide families may make them polyphyletic (Spagna and Gillespie 2008).

The expansion of the 'Austral clade' to include members of another worldwide family (Desidae) makes it clear that a major unsolved aspect of spider phylogeny is the true extent of this clade (or clades, as the case may turn out to be under broader sampling) and its relation to other RTA-clade spiders. While some Austral taxa clearly belong in worldwide families (e.g. Australian Lycosidae Sundevall, 1833, Australian Selenopidae Simon, 1897, Australian orb-weavers and derived orb-weavers), this

work casts doubt on the proper phylogenetic placement of a large number of taxa, particularly those restricted to New Zealand, many of which were described as, and remain in, Agelenidae (Forster and Wilton 1973; Platnick 2010).

### *An 'aquatic' clade?*

In the current study, one shift back to the aquatic habitat (the *Argyroneta* + *Paratheuma* + *Saltonia* line) led to freshwater, saltwater and salt-flat species. Relative to taxonomy, the close relationship between these taxa is a novel result. However, given the documented similarities in some characters, including well developed tracheae and an oversized colulus, this result is not altogether surprising. The most intriguing aspect of this result is that, while these environments are fundamentally different (under fresh water, intertidal, dry salt lakes), and present very different physiological challenges, it may be that the spiders share an ancient history, and are an early offshoot of Dictynidae, a family with the oldest known fossils, being dated to the mid-Cretaceous (Penney *et al*. 2003; Penney and Ortuño 2006). By contrast, dictynids of more typical terrestrial environments are most commonly associated with vegetation, making small irregular webs under tree bark, below (or in curled) leaves, or in clusters of sticks or flowers, although some may be found in leaf litter or soil (Chamberlin and Gertsch 1958; Hagley and Allen 1989). Accordingly, these taxa may be relictual members of a previously more widespread, and more physiologically homogeneous, lineage adapted to an aquatic existence. Although intriguing, this result might not be surprising, as aquatic habitats generally appear to foster relictualism (Humphreys 2006). The current occurrence of one of the aquatic taxa in what are now dry, alkaline salt-flats is likely the result of these relictual taxa adapting over time as their aquatic environment dried up, leaving the current disjunct distribution of *Saltonia* (S. C. Crews and R. G. Gillespie, unpubl. data). This highlights the extreme niche conservatism found among spiders, at least with regards to the terrestrial versus aquatic lifestyle.

It should be noted that although freshwater and saltwater physiologies are highly divergent in many organisms due to the opposite osmotic pressures presented by the two environments, this may not be true for spiders, in which fresh and salt-water spiders share many characteristics. The difference is that a spider's method for dealing with both kinds of aquatic environment is similar, employing strategies to minimise interactions with the water directly. The spiders more or less insulate themselves from contact with the water using some behavioural or mechanical trick, such as carrying around an air bubble (*Argyroneta*) (Shirtcliffe *et al*. 2006) or sealing themselves in barnacle shells during high tide (*Paratheuma*) (Roth and Brown 1975*a*). The close relationships among the different aquatic extremophiles shown in this study raise some important questions regarding the underlying behavioural or ecological adaptations that allow spiders to adapt to large fluctuations of moisture, temperature and salinity.

## Conclusions

By providing evidence for the relative rarity of shifts by spiders to aquatic, semi-aquatic, and historically aquatic environments, this study contributes to the general picture of long-term morphological stasis in arachnids, and extreme age (>90MY) for their major adaptations (such as evolution of venoms, use of silks as prey snares, and the shift from hackled-silk webs to glue-coated webs (Eberhard 1990; Coddington and Levi 1991; Opell 1999). Moreover, this work suggests that convergent evolution of some aquatic features hinder a correct understanding of the historical pattern of spider evolution. In particular, it appears that Austral lineages may not provide a basal framework of the overall entelegyne tree (Forster and Wilton 1973; Griswold *et al*. 1999). Rather, there may be an Austral versus worldwide division deep within the RTA clade followed by independent sets of shifts of important characters, such as production of cribellate silk and intertidal habitat affinity. The study provides strong evidence for an important role for conservative evolution of aquatic specialisation, followed by ancient relictualism, allowing the signature of aquatic lifestyle to be maintained long after the geological environment itself has been modified.

Understanding of deep divergences between spider lineages, like those we have explored here, has heretofore been hindered by the relatively small number of genetic markers of proven utility for phylogenetic study (Ayoub *et al*. 2007). We have demonstrated methodological improvements, in particular enhanced partitioned analysis through modelling of secondary structure, that – when applied to the most common 'slow' markers (18S rRNA and 28S rRNA) – can help compensate for this shortage of data sources.

## Authors' contributions

J.C.S. and S.C.C. designed the study, modelled the secondary structures, and drafted the manuscript. S.C.C. amplified and sequenced genetic loci and analysed the data. R.G.G. assisted with drafts of the manuscript, and provided project guidance and material support for the laboratory work. All authors read, edited and approved the final manuscript.

## Acknowledgements

The authors would like to thank Charles Griswold, Joel Ledford and Brent Opell for loan of specimens crucial to this work. All new specimens were collected under appropriate permits. We thank Jim McGuire and Matt Brandley for use of and assistance with computing resources at UCB, and Mark Miller and Lucie Chan for assistance with the CIPRES web portal. We thank David Gilley for comments leading to many improvements in the manuscript. Funding for the project was provided by the Schlinger Foundation, with additional support from the Division of Insect Biology at UC Berkeley and the National Science Foundation (Grant # DGE 0231877).

## References

Angelini, D. R., and Jockusch, E. L. (2008). Relationships among pest flour beetles of the genus *Tribolium* (Tenebrionidae) inferred from multiple molecular markers. *Molecular Phylogenetics and Evolution* **46**, 127–141. doi:10.1016/j.ympev.2007.08.017

Ayoub, N. A., and Riechert, S. E. (2004). Molecular evidence for Pleistocene glacial cycles driving diversification of a North American desert spider, *Agelenopsis aperta*. *Molecular Ecology* **13**, 3453–3465. doi:10.1111/j.1365-294X.2004.02335.x

Ayoub, N. A., Garb, J. E., Hedin, M., and 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). *Molecular Phylogenetics and Evolution* **42**, 394–409. doi:10.1016/j.ympev.2006.07.018

Beatty, J. A., and Berry, J. W. (1988*a*). The spider genus *Paratheuma* Bryant (Araneae, Desidae). *The Journal of Arachnology* **16**, 47–54.

Beatty, J. A., and Berry, J. W. (1988*b*). Four new species of *Paratheuma* (Araneae, Desidae) from the Pacific. *The Journal of Arachnology* **16**, 339–347.

Bennett, R. G. (1991). 'The Systematics of the North American Cybaeid Spiders (Araneae, Dictynoidea, Cybaeidae).' (University of Guelph: Guelph, Ontario.)

Bergsten, J. (2005). A review of long-branch attraction. *Cladistics* **21**, 163–193. doi:10.1111/j.1096-0031.2005.00059.x

Bernhart, S., Hofacker, I., Will, S., Gruber, A., and Stadler, P. (2008). RNAalifold: improved consensus structure prediction for RNA alignments. *BMC Bioinformatics* **9**, 474. doi:10.1186/1471-2105-9-474

Blackledge, T. A., Scharff, N., Coddington, J. A., Szüts, T., Wenzel, J. W., *et al*. (2009). Reconstructing web evolution and spider diversification in the molecular era. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5229–5234. doi:10.1073/pnas.0901377106

Brandley, M. C., Shmitz, A., and Reeder, T. W. (2005). Partitioned Bayesian analyses, partition choice, and the phylogeny of scincid lizards. *Systematic Biology* **54**, 373–390. doi:10.1080/10635150590946808

Chamberlin, R. V., and Gertsch, W. J. (1958). The spider family Dictynidae in America north of Mexico. *Bulletin of the American Museum of Natural History* **116**, 1–152.

Coddington, J. A., and Levi, H. W. (1991). Systematics and evolution of spiders (Araneae). *Annual Review of Ecology and Systematics* **22**, 565–592. doi:10.1146/annurev.es.22.110191.003025

Coddington, J. A., Giribet, G., Harvey, M. S., Prendini, L., and Walter, D. E. (2004). Arachnida. In 'Assembling the Tree of Life'. (Eds J. Cracraft and M. Donoghue.) pp. 296–318. (Oxford University Press: New York.)

Colgan, D., McLauchlan, A., Wilson, G., and Livingston, S. (1998). Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. *Australian Journal of Zoology* **46**, 419–437. doi:10.1071/ZO98048

Crews, S. C., Puente-Rolon, A. R., Rutstein, E., and Gillespie, R. G. (2010). A comparison of populations of island and adjacent mainland species of Caribbean *Selenops* (Araneae: Selenopidae) spiders. *Molecular Phylogenetics and Evolution* **54**, 970–983. doi:10.1016/j.ympev.2009.10.012

Deans, A. R., Gillespie, J. J., and Yoder, M. J. (2006). An evaluation of ensign wasp classification (Hymenoptera: Evaniidae) based on molecular data and insights from ribosomal RNA secondary structure. *Systematic Entomology* **31**, 517–528. doi:10.1111/j.1365-3113.2006.00327.x

Dixon, M. T., and Hillis, D. M. (1993). Ribosomal-RNA secondary structure – compensatory mutations and implications for phylogenetic analysis. *Molecular Biology and Evolution* **10**, 256–267.

Eberhard, W. G. (1990). Function and phylogeny of spider webs. *Annual Review of Ecology and Systematics* **21**, 341–372. doi:10.1146/annurev.es.21.110190.002013

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797. doi:10.1093/nar/gkh340

Erpenbeck, D., Nichols, S. A., Voigt, O., Dohrmann, M., Degnan, B. M., *et al*. (2007). Phylogenetic analyses under secondary structure-specific substitution models outperform traditional approaches: case studies with diploblast LSU. *Journal of Molecular Evolution* **64**, 543–557. doi:10.1007/s00239-006-0146-3

Fares, M. A., Byrne, K. P., and Wolfe, K. H. (2006). Rate asymmetry after genome duplication causes substantial long-branch attraction artifacts in the phylogeny of saccharomyces species. *Molecular Biology and Evolution* **23**, 245–253. doi:10.1093/molbev/msj027

Felsenstein, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* **27**, 401–410. doi:10.2307/2412923

Foelix, R. F. (1996). 'Biology of Spiders.' 2nd edn. (Oxford University Press: New York.)

Forster, R. R. (1987). A review of the spider superfamilies Hypochiloidea and Austrochiloidea (Araneae, Araneomorphae). *Bulletin of the American Museum of Natural History* **185**, 1–116.

Forster, R. R., and Forster, L. (1999). 'Spiders of New Zealand and Their Worldwide Kin.' (University of Otago Press: Dunedin.)

Forster, R. R., and Gray, M. R. (1979). *Progradungula*, a new cribellate genus of the spider family Gradungulidae (Araneae). *Australian Journal of Zoology* **27**, 1051–1071. doi:10.1071/ZO9791051

Forster, R. R., and Wilton, C. L. (1973). The spiders of New Zealand. Part IV. *Otago Museum Bulletin* **4**, 1–309.

Gillespie, J. J., Cannone, J., Gutell, R., and Cognato, A. I. (2004). A secondary structural model of the 28S rRNA expansion segments D2 and D3 from rootworms and related leaf beetles (Coleoptera: Chrysomelidae; Galerucinae). *Insect Molecular Biology* **13**, 495–518. doi:10.1111/j.0962-1075.2004.00509.x

Giribet, G., Carranza, S., Baguña, N., Riutort, M., and Ribera, C. (1996). First molecular evidence for the existence of a Tardigrada + Arthropoda clade. *Molecular Biology and Evolution* **13**, 76–84.

Givnish, T. J., and Systma, K. J. (Eds) (1997). 'Molecular Evolution and Adaptive Radiation.' (Cambridge University Press: Cambridge, UK.)

Glesne, R. (1998). Terrestrial riparian arthropod investigations in the Big Beaver Creek Research Natural Area, North Cascades National Park Service Complex, 1995–1996: Part III, Arachnida: Araneae. Technical Report NPS/NRNOCA/NRTR/98-03. United States Department of Interior – National Park Service – Pacific West Region.

Gowri-Shankar, V., and Jow, H. (2006). PHASE: A Software Package for Phylogenetics and Sequence Evolution; Manual version 2.0: Available at: http://www.cs.manchester.ac.uk/ai/Software/phase/phase-2.0-manual.pdf [Accessed 1 October 2009].

Griswold, C. E., Coddington, J. A., Platnick, N. I., and Forster, R. R. (1999). Towards a phylogeny of entelegyne spiders (Araneae, Araneomorphae, Entelegynae). *The Journal of Arachnology* **27**, 53–63.

Griswold, C. E., Ramirez, M. J., Coddington, J. A., and Platnick, N. I. (2005). Atlas of phylogenetic data for entelegyne spiders. *Proceedings of the California Academy of Sciences* **56**(Suppl. II), 1–324.

Hagley, E. A. C., and Allen, W. R. (1989). Prey of the cribellate spider *Dictyna annulipes* (Araneae: Dictynidae), on apple tree foliage. *The Journal of Arachnology* **17**, 366–377.

Hedin, M. C., and Maddison, W. P. (2001). A combined molecular approach to phylogeny of the jumping spider subfamily Dendryphantinae (Araneae: Salticidae). *Molecular Phylogenetics and Evolution* **18**, 386–403. doi:10.1006/mpev.2000.0883

Hendriks, L., Van Broeckhoven, C., Vandenberghe, A., Van de Peer, Y., and De Wachter, R. (1988). Primary and secondary structure of the 18S ribosomal RNA of the bird spider *Eurypelma californica* and evolutionary relationships among eukaryotic phyla. *European Journal of Biochemistry* **177**, 15–20. doi:10.1111/j.1432-1033.1988.tb14339.x

Hofacker, I. L., Fekete, M., and Stadler, P. F. (2002). Secondary structure prediction for aligned RNA sequences. *Journal of Molecular Biology* **319**, 1059–1066. doi:10.1016/S0022-2836(02)00308-X

Huelsenbeck, J. P., and Rannala, B. (2004). Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. *Systematic Biology* **53**, 904–913. doi:10.1080/10635150490522629

Huelsenbeck, J. P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755. doi:10.1093/bioinformatics/17.8.754

Humphreys, W. F. (2006). Aquifers: the ultimate groundwater-dependent ecosystems. *Australian Journal of Botany* **54**, 115–132. doi:10.1071/BT04151

Jeram, A. J., Selden, P. A., and Edwards, D. (1990). Land Animals in the Silurian: arachnids and myriapods from Shropshire, England. *Science* **250**, 658–661. doi:10.1126/science.250.4981.658

Kelchner, S. A., and Thomas, M. A. (2007). Model use in phylogenetics: nine key questions. *Trends in Ecology & Evolution* **22**, 87–94. doi:10.1016/j.tree.2006.10.004

Kim, S., Kjer, K. M., and Duckett, C. N. (2003). Comparison between molecular and morphological-based phylogenies of galerucine/alticine leaf beetles. *Insect Systematics & Evolution* **34**, 53–64.

Kjer, K. M. (1995). Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from the frogs. *Molecular Phylogenetics and Evolution* **4**, 314–330. doi:10.1006/mpev.1995.1028

Kjer, K. M. (2004). Aligned 18S and insect phylogeny. *Systematic Biology* **53**, 506–514. doi:10.1080/10635150490445922

Kolaczkowski, B., and Thornton, J. W. (2009). Long-branch attraction bias and inconsistency in Bayesian phylogenetics. *PLoS ONE* **4**, e7891. doi:10.1371/journal.pone.0007891

Lee, M. S. Y. (1998). Convergent evolution and character correlation in burrowing reptiles: towards a resolution of squamate relationships. *Biological Journal of the Linnean Society. Linnean Society of London* **65**, 369–453. doi:10.1111/j.1095-8312.1998.tb01148.x

Lehtinen, P. (1967). Classification of the cribellate spiders and some allied families. *Annales Zoologici Fennici* **5**, 199–468.

MacNaughton, R., Cole, J., Dalrymple, R., Braddy, S., Briggs, D., and Lukie, T. (2002). First steps on land: arthropod trackways in Cambrian-Ordovician eolian sandstone, southeastern Ontario, Canada. *Geology* **30**, 391–394. doi:10.1130/0091-7613(2002)030<0391:FSOLAT>2.0.CO;2

Maddison, W. P., and Hedin, M. C. (2003). Jumping spider phylogeny. *Invertebrate Systematics* **17**, 529–549. doi:10.1071/IS02044

Mallatt, J. M., Garey, J. R., and Shultz, J. W. (2004). Ecdysozoan phylogeny and Bayesian inference first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin. *Molecular Phylogenetics and Evolution* **31**, 178–191. doi:10.1016/j.ympev.2003.07.013

Martins, E. P. (2000). Adaptation and the comparative method. *Trends in Ecology & Evolution* **15**, 296–299. doi:10.1016/S0169-5347(00)01880-2

Mathews, D. H. (2010). RNA Structure, version 5.03. Available at http://rna.urmc.rochester.edu/RNAstructure.html

Mathews, D. H., and Turner, D. H. (2002). Dynalign: an algorithm for finding the secondary structure common to two RNA sequences. *Journal of Molecular Biology* **317**, 191–203. doi:10.1006/jmbi.2001.5351 [Verified July 2010].

McQueen, D. J., and McLay, C. L. (1983). How does the intertidal spider *Desis marina* (Hector) remain under water for such a long time? *New Zealand Journal of Zoology* **10**, 383–392.

Miller, M. A., Holder, M. T., Vos, R., Midford, P. E., Liebowitz, T., Chan, L., Hoover, P., and Warnow, T. (2009). 'The CIPRES Portals.' Available at http://www.phylo.org/sub_sections/portal [Accessed 23 December 2009].

Miller, J. A., Carmichael, A., Ramírez, M. J., Spagna, J. C., Haddad, C. R., Rezac, M., Johannesen, J., Kral, J., Wang, X.-P., and Griswold, C. E. (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). *Molecular Phylogenetics and Evolution* **55**, 786–804. doi:10.1016/j.ympev.2010.02.021

Murphy, N. P., Framenau, V. W., Donnellan, S. C., Harvey, M. S., Park, Y.-C., and Austin, A. D. (2006). Phylogenetic reconstruction of the wolf spiders (Araneae: Lycosidae) using sequences from the 12S rRNA, 28S rRNA, and NADH1 genes: implications for classification, biogeography, and the evolution of web building behavior. *Molecular Phylogenetics and Evolution* **38**, 583–602. doi:10.1016/j.ympev.2005.09.004

Nylander, J. A. A. (2004). 'MrModeltest (Version 2.2).' Program distributed by the author. Uppsala University. Available at http://www.abc.se/~nylander/mrmodeltest2/mrmodeltest2.html [Verified July 2010].

Opell, B. D. (1999). Changes in spinning anatomy and thread stickiness associated with the origin of orb-weaving spiders. *Biological Journal of the Linnean Society. Linnean Society of London* **68**, 593–612. doi:10.1111/j.1095-8312.1999.tb01190.x

Penney, D. (2003). Does the fossil record of spiders track that of their principal prey, the insects? *Transactions of the Royal Society of Edinburgh. Earth Sciences* **94**, 275–281. doi:10.1017/S0263593300000675

Penney, D., and Ortuño, V. M. (2006). Oldest true orb-weaving spider (Araneae: Araneidae). *Biology Letters* **2**, 447–450. doi:10.1098/rsbl.2006.0506

Penney, D., Wheater, C. P., and Selden, P. A. (2003). Resistance of spiders to Cretaceous–Tertiary extinction events. *Evolution* **57**, 2599–2607.

Pisani, D., Poling, L. L., Lyons-Weilier, M., and Hedges, S. B. (2004). The colonization of land by animals: molecular phylogeny and divergence times among arthropods. *BMC Biology* **2**, 1–10. doi:10.1186/1741-7007-2-1

Platnick, N. I. (2010). The world spider catalog version 10.5. American Museum of Natural History, Available at http://research.amnh.org/entomology/spiders/catalog/index.html [Accessed April 10 2010].

Poinar, G., Kerp, H., and Hass, H. (2007). *Palaeonema phyticum* gen. n., sp. n. (Nematoda: Palaeonematidae fam. n.), a Devonian nematode associated with early land plants. *Nematology* **10**, 9–14.

Posada, D., and Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* **53**, 793–808. doi:10.1080/10635150490522304

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

Ronquist, F., and Huelsenbeck, J. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574. doi:10.1093/bioinformatics/btg180

Roth, V. D. (1967). A review of the South American spiders of the family Agelenidae (Arachnida, Araneae). *Bulletin of the American Museum of Natural History* **134**, 297–346.

Roth, V. D., and Brown, W. L. (1975*a*). A new genus of Mexican intertidal zone spider (Desidae) with biological and behavioral notes. *American Museum Novitates* **2568**, 1–7.

Roth, V. D., and Brown, W. L. (1975*b*). Comments on the spider *Saltonia incerta* Banks (Agelenidae?). *The Journal of Arachnology* **3**, 53–56.

SantaLucia, J. J. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 1460–1465. doi:10.1073/pnas.95.4.1460

Schöniger, M., and von Haeseler, A. (1994). A stochastic model for the evolution of autocorrelated DNA sequences. *Molecular Phylogenetics and Evolution* **3**, 240–247. doi:10.1006/mpev.1994.1026

Selden, P. A. (1990). Terrestrialization**:** invertebrates. In 'Paleobiology: A Synthesis'. (Eds D. E. G. Briggs and P. R. Crowther.) pp. 64–68. (Blackwell: Oxford, UK.)

Selden, P. A. (2002). Missing links between *Argyroneta* and Cybaeidae revealed by fossil spiders. *The Journal of Arachnology* **30**, 189–200. doi:10.1636/0161-8202(2002)030[0189:MLBAAC]2.0.CO;2

Shirtcliffe, N. J., McHale, G., Newton, M. I., Perry, C. C., and Pyatt, F. B. (2006). Plastron properties of superhydrophobic surface. *Applied Physics Letters* **89**, 104–106. doi:10.1063/1.2347266

Simon, E. (1892). 'Histoire naturelle des araignees (Vol. 1).' (Paris.)

Spagna, J. C., and Gillespie, R. G. (2006). Unusually long *Hyptiotes* (Araneae: Uloboridae) sequence for small subunit (18S) ribosomal RNA supports secondary structure model utility in spiders. *The Journal of Arachnology* **34**, 557–565. doi:10.1636/H05-54.1

Spagna, J. C., and Gillespie, R. G. (2008). More data, fewer shifts: molecular insights into the evolution of the spinning apparatus in non-orb-weaving spiders. *Molecular Phylogenetics and Evolution* **46**, 347–368. doi:10.1016/j.ympev.2007.08.008

Stamatakis, A. (2006*a*). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690. doi:10.1093/bioinformatics/btl446

Stamatakis, A. (2006*b*). Phylogenetic models of rate heterogeneity: a high performance computing perspective. In 'Proceedings of the 20th IEEE/ACM International Parallel and Distributed Processing Symposium (IPDPS2006), Rhodos, Greece'.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A fast bootstrapping algorithm for the RAxML web-servers. *Systematic Biology* **57**, 758–771. doi:10.1080/10635150802429642

Thomas, B. A. (1972). A probable moss from the Lower Carboniferous of the Forest of Dean, Gloucestershire. *Annals of Botany* **36**, 155–161.

Thorp, J. H., and Kovich, A. P. (1991). 'Ecology and Classification of North American Freshwater Invertebrates.' (Academic Press: New York.)

Tsagkogeorga, G., Turon, X., Hopcroft, R. R., Tilak, M. K., Feldstein, T., *et al*. (2009). An updated 18S rRNA phylogeny of tunicates based on mixture and secondary structure models. *BMC Evolutionary Biology* **9**, 187. doi:10.1186/1471-2148-9-187

Ubick, D., Paquin, P., Cushing, P. E., and Roth, V. (2005). 'Spiders of North America: An Identification Manual.' (American Arachnological Society: Gainesville, FL.)

Wake, D. B. (1991). Homoplasy: the result of natural selection, or evidence of design limitations? *American Naturalist* **138**, 543–567. doi:10.1086/285234

Wheeler, W. C., and Hayashi, C. Y. (1998). The phylogeny of the extant chelicerate orders. *Cladistics* **14**, 173–192. doi:10.1111/j.1096-0031.1998.tb00331.x

Wheeler, W. C., Whiting, M. F., Carpenter, J. C., and Wheeler, Q. D. (2001). The phylogeny of the insect orders. *Cladistics* **12**, 1–57. doi:10.1111/j.1096-0031.1996.tb00189.x

Wilgenbusch, J. C., Warren, D. L., and Swofford, D. L. (2004). AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. Available at http://king2.scs.fsu.edu/CEBProjects/awty/awty_start.php [Accessed 1 May 2008].

Wise, D. H. (1993). 'Spiders in Ecological Webs.' (Cambridge University Press: Cambridge, UK.)

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**, 3406–3415. doi:10.1093/nar/gkg595

Manuscript received 8 January 2010, accepted 3 June 2010

**Appendix 1. Specimen and genetic data information for exemplars used**
N.S. indicates no sequence was available for a particular locus

| OTU | Family | Voucher/reference | 28S accession | 18S accessions | Histone H3 accession |
|---|---|---|---|---|---|
| *Agelena labyrinthica* | Agelenidae | NJNU03002 | AY633851 | AY633862 | N.S. |
| *Agelenopsis aperta* | Agelenidae | Spagna and Gillespie (2008) | DQ628659 | DQ628695<br>DQ628732 | DQ628632 |
| *Allagelena gracilens* | Agelenidae | Spagna and Gillespie (2008) | DQ628661 | DQ628697<br>DQ628733 | DQ628634 |
| *Barronopsis barrowsi* | Agelenidae | Spagna and Gillespie (2008) | DQ628664 | DQ628700<br>DQ628737 | DQ628636 |
| *Huangyuannia tibetana* | Agelenidae | NJNU02001 | AY633857 | AY633859 | N.S. |
| *Neoramia janus* | Agelenidae | Spagna and Gillespie (2008) | DQ628680 | DQ628716<br>DQ628753 | DQ628645 |
| *Novalena intermedia* | Agelenidae | Spagna and Gillespie (2008) | DQ628679 | DQ628715<br>DQ628752 | N.S. |
| *Tegenaria domestica* | Agelenidae | NJNU03003 | AY633852 | AY633852 | N.S. |
| *Tegenaria domestica* | Agelenidae | Spagna and Gillespie (2008) | DQ628683 | DQ628719<br>DQ628756 | DQ628648 |
| *Textrix denticulata* | Agelenidae | Spagna and Gillespie (2008) | DQ628682 | DQ628718<br>DQ628755 | DQ628647 |
| *Amaurobius similis* | Amaurobiidae | Spagna and Gillespie (2008) | DQ628663 | DQ628699<br>DQ628736 | N.S. |
| *Callobius gertschi* | Amaurobiidae | Spagna and Gillespie (2008) | DQ628668 | DQ628704<br>DQ628741 | N.S. |
| *Coelotes terrestris* (specimen 1) | Amaurobiidae | Spagna and Gillespie (2008) | DQ628688 | DQ628724<br>DQ628761 | DQ628651 |
| *Coelotes terrestris* (specimen 2) | Amaurobiidae | Spagna and Gillespie (2008) | DQ628689 | DQ628725<br>DQ628762 | DQ628652 |
| *Eurocoelotes inermis* | Amaurobiidae | Spagna and Gillespie (2008) | DQ628690 | DQ628726<br>DQ628763 | DQ628653 |
| *Pimus* sp. 1 | Amaurobiidae | Spagna and Gillespie (2008) | DQ628675 | DQ628711<br>DQ628748 | DQ628642 |
| *Wadotes dixiensis* | Amaurobiidae | Spagna and Gillespie (2008) | DQ628685 | DQ628721<br>DQ628758 | N.S. |
| *Pimus* sp. 2 | Amaurobiidae | Spagna and Gillespie (2008) | DQ628681 | DQ628717<br>DQ628754 | DQ628646 |
| *Metaltella simoni* | Amphinectidae | Spagna and Gillespie (2008) | DQ628677 | DQ628713<br>DQ628750 | N.S. |
| *Anyphaena californica* | Anyphaenidae | Spagna and Gillespie (2008) | DQ628660 | DQ628696<br>DQ628733 | DQ628633 |
| *Zygiella x-notata* | Araneidae | Spagna and Gillespie (2008) | DQ628658 | DQ628694<br>DQ628731 | N.S. |
| New genus #4 | Cybaeidae | JSSC_004 | HM576653 | HM576638 | HM576672 |
| New genus #4 | Cybaeidae | JSSC_012 | HM576652 | HM576637 | HM576671 |
| *Argyroneta aquatica* | Cybaeidae | Spagna and Gillespie (2008) | DQ628687 | DQ628723<br>DQ628760 | DQ628650 |
| *Argyroneta aquatica* | Cybaeidae | ARAQ1 | HM576655 | N.S. | N.S. |
| *Cybaeina* sp 1 | Cybaeidae | JSSC_006 | HM576647 | HM576631 | HM576666 |
| *Cybaeina* sp. 2 | Cybaeidae | JSSC_009 | HM576648 | HM576632 | HM576667 |
| *Cybaeota* sp. 2 | Cybaeidae | JSSC_005 | HM576656 | N.S. | HM576674 |
| *Cybaeota* sp. 3 | Cybaeidae | JSSC_007 | HM576645 | HM576629 | HM576664 |
| *Cybaeus morosus* | Cybaeidae | Spagna and Gillespie (2008) | DQ628671 | DQ628744<br>DQ628707 | DQ628641 |
| *Cybaeus* sp. 1 | Cybaeidae | JSSC_001 | N.S. | HM576635 | HM576669 |
| *Cybaeus* sp. 2 | Cybaeidae | JSSC_003 | HM576654 | HM576639 | HM576673 |
| *Cybaeus* sp. 3 | Cybaeidae | JSSC_013 | HM576651 | HM576636 | HM576670 |
| New genus #1 | Cybaeidae | JSSC_008 | HM576646 | HM576630 | HM576665 |
| *Badumna longinqua* | Desidae | Spagna and Gillespie (2008) | DQ628665 | DQ628701<br>DQ628738 | DQ628637 |
| *Desis marina* | Desidae | DesisNZ | HM576644 | HM576628 | HM576663 |
| *Desis formidabilis* | Desidae | DDf1 | HM576643 | HM576627 | HM576662 |

(*continued next page*)

**Appendix 1.** (*continued*)

| OTU | Family | Voucher/reference | 28S accession | 18S accessions | Histone H3 accession |
|---|---|---|---|---|---|
| *Paratheuma armata* | Desidae | Spagna and Gillespie (2008) | DQ628674 | DQ628710<br>DQ628747 | N.S. |
| *Paratheuma insulana* | Desidae | TCI_045 | HM576642 | HM576626 | HM576661 |
| *Blabomma* sp. | Dictynidae | JSSC_011 | HM576649 | HM576633 | N.S. |
| *Dictyna* sp. | Dictynidae | Spagna and Gillespie (2008) | DQ628673 | DQ628709<br>DQ628746 | N.S. |
| *Lathys alberta* | Dictynidae | Spagna and Gillespie (2008) | DQ628676 | DQ628712<br>DQ628749 | DQ628643 |
| *Saltonia incerta* – CA | Dictynidae | salt_033 | HM576657 | N.S. | HM576675 |
| *Saltonia incerta* – CA | Dictynidae | salt_053 | HM576640 | HM576624 | HM576659 |
| *Saltonia incerta* – NM | Dictynidae | salt_085 | HM576641 | HM576625 | HM576660 |
| *Yorima* sp. | Dictynidae | JSSC_010 | HM576650 | HM576634 | HM576668 |
| *Cicurina* sp. | Dicynidae | Spagna and Gillespie (2008) | DQ628699 | DQ628705<br>DQ628742 | DQ628640 |
| *Stegodyphus* sp. | Eresidae | Spagna and Gillespie (2008) | DQ628691 | DQ628727<br>DQ628764 | DQ628654 |
| *Zelotes* sp. 1 | Gnaphosidae | Spagna and Gillespie (2008) | DQ628686 | DQ628722<br>DQ628759 | N.S. |
| *Calymmaria* sp. 1 | Hahniidae | Spagna and Gillespie (2008) | DQ628666 | DQ628702<br>DQ628739 | DQ628638 |
| *Calymmaria* sp. 2 | Hahniidae | Spagna and Gillespie (2008) | DQ628667 | DQ628703<br>DQ628740 | DQ628639 |
| *Cryphoeca* sp. | Hahniidae | Spagna and Gillespie (2008) | DQ628672 | DQ628708<br>DQ628614 | N.S. |
| *Neoantistea agilis* | Hahniidae | Spagna and Gillespie (2008) | DQ628678 | DQ628714<br>DQ628751 | DQ628644 |
| *Alopecosa kochi* | Lycosidae | Spagna and Gillespie (2008) | DQ628662 | DQ628698<br>DQ628735 | DQ628635 |
| *Cheiracanthium mildei* | Miturgidae | Spagna and Gillespie (2008) | DQ628670 | DQ628706<br>DQ628743 | N.S. |
| *Neolana dalmasi* | Neolanidae | Spagna and Gillespie (2008) | N.S. | DQ628728<br>DQ628765 | DQ628655 |
| *Megadictyna thilenii* | Nicodamidae | Spagna and Gillespie (2008) | DQ628692 | DQ628729<br>DQ628766 | DQ628656 |
| *Stiphidion facetum* | Stiphidiidae | Spagna and Gillespie (2008) | DQ628693 | DQ628730<br>DQ628767 | DQ628657 |
| *Tengella radiata* | Tengellidae | Spagna and Gillespie (2008) | DQ628684 | DQ628720<br>DQ628757 | DQ628649 |

**Appendix 2. Primers used**

| Genetic locus/primer name | Primer sequence 5′→ 3′ | Reference |
|---|---|---|
| 28S/28S O | ACT GCT CAA AGG TAA ACG G | Hedin and Maddison (2001) |
| 28S/28S C | GGT TCG ATT AGT CTT TCG CC | Hedin and Maddison (2001) |
| 18S/18S 1F | TAC CTG GTT GAT CCT GCC AGT AG | Giribet *et al*. (1996) |
| 18S/18S 5R | CTT GGC AAA TGC TTT CGC | Giribet *et al*. (1996) |
| 18S/18S 5F | GCG AAA GCA TTT GCC AAG AA | Giribet *et al*. (1996) |
| 18S/18S 9R | ATG GCT CGT ACC AAG CAGACV GC | Giribet *et al*. (1996) |
| Histone H3/H3aF | GCT CGT ACC AAG CAG ACV GC | Colgan *et al*. (1998) |
| Histone H3/H3aR | ATA TCC TTR GGC ATR ATR GTG AC | Colgan *et al*. (1998) |