A barcode-based phylogenetic scaffold for *Xysticus* and its relatives
(Araneae: Thomisidae: Coriarachnini)

RAINER BREITLING

Faculty of Science and Engineering, University of Manchester, Manchester M1 7DN, United Kingdom. E-mail: rainer.breitling@manchester.ac.uk

Received 12 March 2019 │ Accepted by V. Pešić: 26 March 2019 │ Published online 27 March 2019.

Abstract
The phylogenetic relationships and taxonomy of the crab spider genus *Xysticus* and its closest relatives (i.e., the tribe Coriarachnini, also including, e.g., *Ozyptila, Coriarachne* and *Bassaniana*) have long been controversial, with several alternative classifications being proposed, none of which has gained universal acceptance. As Coriarachnini is largely confined to the Holarctic region, the main target area of recent DNA barcoding projects for spiders, a large amount of genetic data for the group is now publicly available. The results of a phylogenetic analysis of this sequence dataset are largely congruent with earlier morphology-based results regarding the evolutionary structure of the group. In particular, they highlight the fact that *Xysticus* s. lat. is a paraphyletic assembly and that several species groups need to be placed in separate genera to achieve monophyly of *Xysticus* s. str. Similarly, *Coriarachne* and *Bassaniana* appear as independent clades rather than a joined monophyletic *Coriarachne* s. lat. In contrast, further subdivision of *Ozyptila* is not supported by the genetic data. Importantly, the analysis also shows that anapophysate members of *Xysticus* s. lat. form two widely separated groups: a primarily anapophysate division, also including *Coriarachne* and *Bassaniana*, at the base of *Xysticus* s. lat., and a secondarily anapophysate clade deeply nested within *Xysticus* s. str. This might explain some of the earlier difficulties when trying to define generally accepted subgroups within *Xysticus* s. lat. The phylogenetic scaffold based on barcode sequences is sufficiently dense and well resolved to attempt the tentative and provisional placement of the majority of species in *Xysticus* s. lat. in the independent genera *Xysticus* s. str., *Bassaniodes, Psammitis* and *Spiracme* as a starting point for a future more formal revision of the group.

Key words: Araneae, DNA barcoding, cladistics, phylogenetic systematics.

Introduction
The crab spider genus *Xysticus* s.lat. and its closest relatives, the genera *Ozyptila* s. lat. and *Coriarachne* s.lat., represent a well-defined morphologically homogeneous clade, constituting the core of the tribe Coriarachnini sensu Ono (1988). While the group is relatively well delimited (but see Lehtinen 2002), its internal structure has been a matter of considerable debate and confusion. A large number of competing and not always mutually compatible subdivisions of the group have been proposed, either using species groups and subgenera, or raising some parts of the large genera *Xysticus* and *Ozyptila* to genus rank (Dondale & Redner 1975, Gertsch 1939, 1953, Jantscher 2001, 2002, Lehtinen 2002, Marusik et al. 2005, Ono 1978, 1988, Schick 1965, Turnbull et al. 1965, Wunderlich 1987, 1992, 1995). Few other major groups of spiders
have attracted such intense scrutiny using traditional morphological approaches, with so little consensus emerging. Particularly contentious has been the position and taxonomic treatment of the anapophysate members of *Xysticus*, i.e. those species lacking bulbar apophyses on the male pedipalp. Anapophysate *Xysticus* have repeatedly been placed into separate (sub)genera, such as *Psammitis*, *Proxysticus*, *Spiracme* or *Bassaniodes* in various combinations (see Jantscher 2001 and Lehtinen 2002 for detailed recent reviews). Moreover, it has been proposed that the genus *Ozyptila* is possibly not monophyletic and may require subdivision into multiple genera based on present species groups (Lehtinen 2002, Marusik et al. 2005). The possible synonymy of *Coriarachne* and *Bassaniiana* has also been controversial (Gertsch 1953, Bowling & Sauer 1975, Lehtinen 2002). With the exception of the recent removal of the anapophysate genera *Cozyptila* and *Modysticus* from *Ozyptila* s. lat. by Marusik et al. (2005), none of the proposed subdivisions has gained widespread acceptance, and the World Spider Catalog still uses both *Xysticus* and *Ozyptila* in the broad sense, while treating *Coriarachne* and *Bassaniiana* as separate genera.

Coriarachnini is a largely Holarctic clade, being particularly diverse in those areas that have been the major focus of recent DNA barcoding projects for spiders (e.g., Astrin et al. 2016, Blagoev et al. 2013, 2016). It was therefore interesting to explore if the phylogenetic signal contained in the available barcodes for members of this group would provide sufficient complementary evidence to resolve some of these problems (Breitling 2017, Coddington et al. 2016).

It could be argued that such an approach is futile, given the known limitations of mitochondrial barcodes as indicators of phylogeny (Maddison 2018) and the recently announced plans for a Earth BioGenome Project, intending to sequence the full genome of every plant and animal on the planet, thus making single-marker molecular phylogenies hopelessly outdated (Lewin et al. 2018). However, two important reasons support publication of the resulting hypotheses in their present preliminary state. (1) The analysis of a closely related group of genera focuses on the “sweet spot” of barcode-based phylogenetic analysis. Exploration of evolutionary patterns at the species boundary is frequently impeded by incomplete lineage sorting (van Velsen et al. 2017, Maddison & Knowles 2006) and introgression between closely related sibling species and semispecies (Roux et al. 2016), possibly driven by intracellular parasites like *Wolbachia* (Klopfstein et al. 2016). This is illustrated by a failure to distinguish closely related species pairs in many recent spider barcoding projects (e.g., Astrin et al. 2016, Blagoev et al. 2013, 2016, Oxford & Bolzern 2018). On the other hand, spurious signals from nuclear copies of the mitochondrial barcode sequence can result in apparent discrepancies between gene and species phylogenies (e.g., Hawlitschek et al. 2017, Ermakov et al. 2015). Moreover, about 600 bp of rapidly evolving protein-coding DNA sequence can obviously not contain sufficient information to resolve deep phylogenetic branching patterns – for this purpose whole-genome sequencing and related phylogenomic approaches are essential (e.g., Hedin et al. 2018, Fernández et al. 2018). In contrast, when focusing on evolutionary affinities at the genus level in the context of an existing morphology-based system, limited resolution between sibling species due to introgression and incomplete lineage sorting has no adverse consequences, and contamination by spurious nuclear sequences would be identifiable by severe misplacements of individual species, incongruent with morphological evidence. (2) An extensive framework of morphological revisions is available as a starting point for the present analysis. While a barcode-based phylogenetic analysis in isolation would indeed be reckless and unjustifiable, the molecular evidence is here provided as part of an integrative attempt to synthesize numerous contradictory taxonomic proposals based on traditional morphological analysis. A delay in publishing the barcode-based results would unnecessarily prolong the unsatisfactory lack of nomenclatural stability within Coriarachnini. Moreover, it is hoped that making these results available in their present form, rather than waiting for a future whole-genome project, will make these molecular resources accessible to a much larger group of spider taxonomists, who may wish to test (and potentially refute) the suggested phylogenetic hypotheses using alternative character sets.

**Materials and methods**

All public DNA barcode data annotated as derived from “Thomisidae” were downloaded in FASTA format from the BOLD database (www.boldsystems.org; Ratnasingham & Hebert 2007) in April 2018. Sequences for Cytochrome Oxidase Subunit 1 5’ region barcodes (COI-5P) were selected from this dataset, and processed using the DECIPHER package (Wright 2016) and custom-made scripts in R (R Core Team 2018). Majority-rule consensus barcodes for each species were inferred from alignments of all barcodes assigned to
a given species, and species with incomplete barcodes were removed. Sequences were aligned in BioEdit v7.2.5 (Hall 1999), and positions containing gaps (at the ends of the sequences) were removed.

The final dataset included 3585 sequences for 108 species of Thomisidae (some of them identified only to genus level), including 49 species of *Xysticus* C.L. Koch, 1835 (including the type species, *X. audax*, as well as the type species of *Psammitis* Menge, 1868, *X. sabulosus*, and *Spiracme* Menge, 1868, *X. striatipes*, and representatives of the *pretiosus*, *locuples*, *pellax*, *cunctator*, *montanensis*, *lucutosus*, *triangulosus*, *lalandei*, *sabulosus*, and *labradorensis* species groups of Schick 1965), 17 species of *Ozyptila* Simon, 1864 (including the type species, *O. claveata*, as well as representatives of the *raudu, brevipes, praticola*, and *trux* species groups suggested by Lehtinen 2002), two representatives of *Coriarachne* Thorell, 1870 (including the type species, *C. depressa*) and two of *Bassaniida*. No representatives of *Cozyptila* Marusik & Kovblyuk, 2005, *Modysticus* Gertsch 1953, or *Bassaniodes* Pocock, 1903 (=*Proxysticus* Dalmas, 1922, s. str.) were included in the final dataset. The alignment of consensus sequences (https://doi.org/10.13140/RG.2.2.15312.76803) and a table detailing the taxonomic treatment of all represented species by earlier authors as well as the individual reconstructed trees (https://doi.org/10.13140/RG.2.2.17829.35047) are provided as electronic supplementary material.

No attempt was made to manually curate the barcode data, and all data are presented under the species names assigned in the BOLD database. Obvious misidentifications are readily apparent: e.g., a specimen labelled as *Xysticus sicus* matches the barcode of *Oxytate striatipes* and is clearly misidentified, and a single, partial barcode attributed to *X. bliteus* appears highly dubious, as it is assigned the same Barcode Index Number (BIN; Ratnasingham & Hebert 2013) as two *X. nubilus* specimens from Portugal. Nevertheless, in most cases the use of consensus sequences minimizes problems created by individual misidentified specimens, as well as the influence of contamination by nuclear copies of the barcode and mitochondrial introgression by rare hybridization events between distantly related species, as outliers will not contribute to a majority-rule consensus sequence (which on the other hand considerably reduces the computational burden).

Phylogenetic trees were reconstructed using phylogeny.fr (Dereeper et al. 2008) as described in Breitling (2017), using the default workflow and parameters to infer Maximum Likelihood, Maximum Parsimony and Bayesian trees. For the Bayesian analysis using the MrBayes algorithm within phylogeny.fr, the dataset was divided into overlapping subsets of 30 sequences to minimize computational burden. Neighbour joining trees were constructed using custom R scripts and the DECIPHER package, employing either pairwise distances between consensus sequences or the median pairwise distance between all individual sequences in each species pair, in an attempt to explore the potential impact of misidentified sequences. The neighbour-joining trees also include species represented by incomplete barcodes; the model-free neighbour-joining approach is least sensitive to this kind of incomplete data and allows examining the placement of these additional species in the context of the reconstructed phylogeny.

**Results and discussion**

As shown in Figure 1, the barcode sequences indeed contain a robust phylogenetically informative signal. As expected when using such relatively short sequences, not all internal branches are fully and reproducibly resolved, but the four inference methods agree on the major points of interest.

*Coriarachnini* is found to be a monophyletic clade within the barcoded sample of Thomisidae. Within this tribe, *Ozyptila* is the sister group of a large group containing *Xysticus* s.lat. together with *Coriarachne* and *Bassaniida*.

The *Ozyptila* species included in the barcode analysis (all of them apophysate) are recovered consistently as a monophyletic group. The internal structure of the genus is only poorly resolved by the barcoding data, indicating that the group is quite homogeneous, despite extensive diversity of the genitalia. There seems to be very little support for splitting the genus further, especially as *O. arctica*, the single representative of the most distinct species group, the *raudu* group (Hippa et al. 1986, Marusik 2008), turns out to be most closely related to the type species of the genus, *O. claveata*. The *raudu* group is thus merely a particularly derived subgroup at the core of *Ozyptila*. In light of this result, it will be interesting to see barcoding data for representatives of *Cozyptila* and *Modysticus*; however, given their anapophysate condition, they may very well be justifiably placed outside of *Ozyptila* s. str.
Figure 1. Majority-rule consensus tree of Coriarachnini barcode sequences. Barcodes of all other members of Thomisidae are used as outgroup to root the tree, and support by each of the four phylogenetic inference methods (Neighbour Joining, MrBayes, Maximum Likelihood, and Maximum Parsimony) is indicated at each resolved branch (filled circles: supported; grey circle: not resolved by this method; open circle: not supported).

Most importantly, in the context of earlier attempts to subdivide *Xysticus* s. lat. into smaller coherent groups, the anapophysate *Xysticus* species form two well separated groups: one group, the anapophysate division in the strict sense, groups together with the equally anapophysate genera *Coriarachne* and *Bassaniana*. The
second group, a secondarily anapophysate clade as suggested by the topology, is nested as a monophyletic subgroup deeply within the apophysate Xysticus s. str. Members of both of these groups were regularly placed together in earlier attempts at subdividing Xysticus s.lat., e.g., in Schick’s (1965) and Ono’s (1978) Proxysticus, in Gertsch’s (1953) sabulosus group, Turnbull et al.’s (1965) triangulosus group, and in Wunderlich’s (1995) and Lehtinen’s (2002) Psammitis. This might explain some of the difficulties in defining generally accepted subgroups with unambiguous diagnostic criteria.

Within the (primarily) anapophysate division, two clearly distinct clades are apparent, which due to a fortunate coincidence or a stroke of Mengean taxonomic genius appear to contain the type species of Psammitis and Spiracme, respectively (Menge 1868). The Psammitis group also very consistently includes typical members of the labradorensis group of previous authors, while Spiracme includes X. triangulosus (but not the other sequenced member of the triangulosus group sensu Turnbull et al. 1965, X. winnipegensis), as well as representatives of the durus and nigromaculatus groups, albeit with slightly weaker support. Spiracme and Psammitis as delimited here show unresolved affinities to Coriarachne and Bassaniana in the present dataset. All four tree-building methods agree that the latter two genera are not closely related but form independent members of the anapophysate sister group of Xysticus s. str., at the same rank as Spiracme and Psammitis. The taxonomic implication is that either all four groups are joined in a large Coriarachne s. lat. (this option was considered but apparently rejected by Lehtinen 2002), or that they are all treated as independent genera. For the time being, the latter approach seems to be least disruptive and maintains the largest amount of information to facilitate subsequent reclassifications. Importantly, in some of the reconstructed trees the anapophysate division is paraphyletic relative to Xysticus s. str., which also argues against a hasty merging of these groups into a single genus.

This decision also has implications for the remaining anapophysate members of Xysticus s. lat., not represented in the barcode dataset: In his revision of the Coriarachnini, Lehtinen (2002) re-delimited the genera Proxysticus and Psammitis as used by Schick (1965) or Wunderlich (1992, 1995). He correctly observed that Bassaniodes Pocock, 1903, is a senior synonym of Proxysticus Dalmas, 1922, based on the characteristic female genitalia of the type species of the two genera (syn. conf.), and restricted the scope of this genus to Paleaeartic species, mostly from the Canary Islands, the Mediterranean and Western Central Asia, based on a set of clear morphological criteria. While no representative of Bassaniodes sensu Lehtinen (2002) was present in the barcode dataset analysed here, it seems clear that this Palaeaeartic group of species is morphologically the most divergent and distinct anapophysate clade of Xysticus s. lat., with obvious affinities to Özyptila. While the exact position of the group within Coriarachnini remains unresolved for the time being, for consistency it appears necessary to treat Bassaniodes as a separate genus, as already convincingly argued by Lehtinen (2002).

Naturally, this analysis, which is based on only a small subset of the diversity of Coriarachnini, will not be the last word on the taxonomy of this group. Nevertheless, the results presented here might help clearing up some long-standing confusion and provide a scaffold to support a more extensive integrative taxonomic study. They, therefore, will be an important contribution to the hypothesis building required for future more thorough studies. In preparation for such future work, an attempt has been made to reorganize the majority of the species currently classified as members of Xysticus s. lat. into the four genera Xysticus s. str., Psammitis, Spiracme, and Bassaniodes based on published morphological assessments of their affinities to barcoded species, as well as illustrations in the literature. For example, all apophysate species are considered members of Xysticus s. str., based on this putative synapomorphy of the genus; the members of the labradorensis group as defined by Marusik & Logunov (1995) are considered members of Psammitis; species of the durus and nigromaculatus groups of Gertsch (1953), Turnbull et al. (1965) and Fomichev et al. (2014) are assigned to Spiracme; while species listed under Bassaniodes by Lehtinen (2002) and their closest relatives are assigned to the latter genus. A large number of anapophysate species, as well as many species inquirendae, often known only from female or juvenile specimens, typically the type material, remain incertae sedis within Xysticus, pending additional analysis. One could question whether the new combinations proposed here are a potential threat to nomenclatural stability (analogous to the case discussed by Kropf et al. 2019) and whether any major rearrangement would better be published as part of a “final” comprehensive revision of the group. Obviously, the present work does build on and complement several major revisions, which have already proposed a large number of new combinations without any adverse effect on nomenclatural practice. The present work does not neglect the earlier morphological evidence, but complements it with a new set of molecular data and suggests a number of new hypotheses that could explain the previous failure to reach consensus on the systematics of this group. As Norman Platnick states
in a recent very readable defence of Linnaean taxonomy (2013), “Linnaean names are very bold hypotheses, indeed; they allow us to divide the virtually infinite universe of potential three-taxon statements about the world's species into two groups; two-thirds of those potential statements are prohibited by the classification, which makes Linnaean names exceedingly strong and testable.” Consequently, it would be detrimental to the status of taxonomy as a scientific enterprise should the publication of new combinations be restricted for the sake of an illusory “stability”. Ultimately, stability will only be achieved if competing hypotheses are explicitly spelled out in a testable format. For a hypothesis to be testable by new independent evidence and analysis, it first needs to be shared by publication. This seems particularly important in the case of molecular evidence, which tends to be difficult to access and interpret for those traditionally trained taxonomist that can make the most important contribution to future classification attempts.

The following preliminary composition of the four genera of species formerly contained within Xysticus s.lat. is tentatively and provisionally proposed here (bibliographic details are available in the WSC):


**Spiracme**: Spiracme balistianus, S. durus, S. keyserlingi, S. lehtineni, S. nigromaculatus, X. quadratius, X. striatipes (type species), S. triangulosus, X. vachoni. (9 species; all except the type species comb. nov.)
All other species currently listed in *Xysticus* remain *incertae sedis* within *Xysticus* s.lat. Some members of this latter group might, e.g., belong to *Bassariania*, as Marusik & Logunov (1995) place several of these species in the neighbourhood of *Xysticus albomaculatus*, i.e. *Bassariania baueduier* (Breitling et al. 2016). This could affect, e.g., *X. bakanas*, *X. kuzgi*, *X. lapidarius*, *X. marusiki*, *X. turkmenicus*, or *X. urgumchak*. Divergent species groups, such as the *maculatus* group (*X. alevs*, *X. bacurianensis*, *X. maculatus*, *X. nitidus*), or some South African *Xysticus* species (*X. fagei*, *X. havilandi*, *X. lucifugus*, *X. natalensis*, *X. sagittifer*, *X. urbensis*), might well deserve independent genus status of their own. Other species, such as *X. bimaculatus* from Australia, *X. palawanicus* from the Philippines, and the majority of the *Xysticus* species described from India, are obviously misplaced in *Xysticus* and probably not even members of Coriarachne. These examples show that we are still far from a final classification of this tribe. A most urgent desideratum will be a consistent morphological definition of the genera involved, which are here tentatively defined mostly on the basis of general similarity to barcoded specimens; obviously an unsatisfactory situation. Thus, it is hoped that the present preliminary attempt will stimulate further analysis to fully resolve the relationships within this largest and most abundant Holarctic group of crab spiders.

Acknowledgements
I thank Theo Blick and two anonymous referees for their very critical and constructive comments on an earlier version of the manuscript.

References


Ecológica Montenegrina, 20, 2019, 198-206 205


