DNA barcoding of flat bugs (Hemiptera: Aradidae) with phylogenetic implications

VASILY V. GREBENNIKOV 1 & ERNST HEISS 2

1 Canadian Food Inspection Agency, Ottawa Plant Laboratory, KW Neatby Bldg, 960 Carling Ave, Ottawa, Ontario K1A 0C6, Canada; Vasily V. Grebennikov [vasily.grebennikov@inspection.gc.ca] — 2 Tiroler Landesmuseum, Josef-Schraffl-Straße 2a, A-6020 Innsbruck, Austria; Ernst Heiss [aradus@aon.at]

Accepted 17.vi.2014. Published online at www.senckenberg.de/arthropod-systematics on 18.vii.2014.

Abstract

This work introduces an open online library of Aradidae DNA barcodes, specimen images and geographical data with intent of promoting further community-based DNA barcoding of flat bugs. We report the results of an attempt to DNA barcode 191 dry specimens of the flat bugs representing all 8 extant subfamilies of the Aradidae. 145 sequences > 300 nt in length were obtained (76% success rate; all can be seen in a publicly accessible dataset “World Aradidae”, doi: dx.doi.org/10.5883/DS-WAHG), 98 of them representing 55 species and 29 genera were > 500 nt (51%). These 98 sequences were combined with 10 Heteroptera-Trichophora outgroup taxa into an aligned matrix of 658 positions and variously analysed to obtain the first DNA-based phylogeny for the Aradidae, which was our second goal. Aradidae and five of its subfamilies (Aradinae, Aneurinae, Mezirinae, Isoderminae, Calisiinae) were recovered as monophyletic, although the latter two were markedly underrepresented. The subfamily Carventinae was consistently polyphyletic. Isoderminae + Prosympiestinae was recovered as a clade, while Aradinae formed the sister group to the rest of the family in three of four analyses. The genus Mezira appeared polyphyletic while the genus Neuroctenus was paraphyletic with respect to Ctenoneurus.

Key words

Aneurinae, Aradinae, Calisiinae, Carventinae, Chinamyersiinae, Isoderminae, Mezirinae, Prosympiestinae, Termitaphididae, Pentatomomorpha, CO1.

1. Introduction

The nearly cosmopolitan terrestrial heteropteran family Aradidae, colloquially known as “flat bugs”, contains at least 1,798 markedly depressed medium-sized darkish species arranged in at least 211 genera, as per the somewhat outdated catalog by KORMILEV & FROESCHNER (1987). Since then at least 56 new genera and 168 new species have been described, of them 35 genera and 88 species by the second author. Flat bugs are thought to be mycophagous, but very little of their biology is known (LESCHEN & TAYLOR 1987). The majority of Aradidae in temperate zone ecosystems are winged and are normally found under bark, while most tropical species are wingless and inhabit forest leaf litter. For years the latter specimens were thought to be nymphs, until MILLER (1938) described the first exclusively apterous genus Chelonecoris (Fig. 1) from Malaysia. Since the pivotal volume of USINGER & MATSUDA (1959), eight extant subfamilies have been consistently recognized. Among them, Mezirinae is by far the largest, embracing more than half of all Aradidae species; the other being Aneurinae, Aradinae, Calisiinae, Carventinae, Chinamyersiinae, Isoderminae and Prosympiestinae. Only New Zealand and Australia have representatives of all subfamilies (LARIVIÈRE & LAROCHETTE 2006). The ninth subfamily, Archaearadi-
nae, is extinct and contains two monotypic genera from Burmese Amber of the mid- and upper-Cretaceous, respectively (Heiss & Poinar 2012). The mid-Cretaceous Cretotriesma suukyiae Grimaldi & Engel, 2008 from Myanmar possibly represents another subfamily (Cassis & Schuh 2010), while at least 52 other extinct aradid species have been attributed to four of the eight extant subfamilies (references to more than half of them can be found in Kauffuss et al. 2011).

The phylogenetic placement of Aradidae is less controversial than that for many other bug families. The family is either the sister taxon to Termitaphididae (Schuh & Slater 1995; Henry 1997) or more likely paraphyletic (Grimaldi & Engel 2005, 2008; Cassis & Schuh 2010) with respect to that family. The latter includes 13 exclusively termitophilous, blind and wingless small-bodied species (Fig. 1) arranged in two genera and known from all main tropical regions, as well as from Mexican and Dominican Miocene amber (Poinar & Heiss 2011). Owing to the rarity of specimens, Termitaphididae eluded phylogenetic analyses until the morphological analysis in Cassis & Schuh (2010) consistently placed it deeply nested within the strongly supported Aradidae: the single representative used in their analysis, Termitaradus guianae Morrison, 1923, was placed as sister to the Aradinae-Chinamyersinae. Among other characters, the two families share a flattened body and elongate mouthparts coiled inside the head; the latter character is a seemingly complex and unique synapomorphy. The superfamily Aradoidea (or Aradidae sensu lato, i.e. including Termitaphididae) is consistently placed as sister to the rest of Pentatomomorpha, which is often termed “Trichophora”, reflecting the presence of trichobothria on the ventral surface of the abdomen (see Henry 1997; Cassis & Schuh 2010). Pentatomomorpha is one of the seven heteropteran infraorders (others being Enicocephalomorpha, Dipsocoromorpha, Gerromorpha, Nepomorpha, Leptopodomorpha and Cimicomorpha; see Weirauch & Schuh 2011).

Three publications have addressed relationships within the Aradidae using morphological characters. Vasarhelyi (1987) and Grozeva & Kerzinger (1992) provided intuitive topologies with the eight extant subfamilies as terminals. Cassis & Schuh (2010) performed a cladistic analysis based on 34 terminals (including eight extant Aradidae) and 78 characters and included two critically important terminals: a species of Termitaphididae and the fossil Cretotriesma suukyiae. The latter work was, however, not designed to resolve relationships within the Aradidae, but to test the placement of the fossil and that of Termitaphididae. No DNA-based topology has ever been proposed for the Aradidae.

We used the BOLD online platform to develop the first open library of Aradidae DNA barcodes (Hebert et al. 2003) in order to facilitate future research on this group similarly as it was done with many other animal groups (see, for example, Hogg et al. 2009 for New Zealand Trichoptera, De Waard et al. 2011 for British Columbia Geometridae, Speleja et al. 2011 for Bavarian Myriapoda). While most projects have examined specimens from a limited geographic area, we removed this constraint by analyzing specimens from the highly representative collection of Aradidae assembled by Ernst Heiss. Until now the family was poorly represented in the Barcode of Life Database (BOLD) with only 15 CO1 sequences other than ours. Our second goal was to utilize the sequence data, to the extent possible, to propose the first DNA-based Aradidae phylogenetic hypothesis and to compare it with those advanced previously on the basis of morphological characters.

2. Material and methods

Specimen origin, depository and choice of taxa. All specimens barcoded in this project came from, and are deposited in, the research collection of Ernst Heiss in Innsbruck, Austria. They were collected by various people under different circumstances and were stored as dry pinned specimens for 8–12 years before being used for DNA extraction. Specimens and species were chosen by EH to proportionally represent each of the eight extant Aradidae subfamilies and some of the most diversified genera without any geographical bias, although Ecuador (18 records), Malaysia (15), India (13), Madagascar (13) and Austria (13) were the five most strongly represented countries. A single leg of a dry specimen was removed for DNA extraction, as customary for medium-sized insects and other arthropods.

Aradidae genetic barcodes. In June 2014 the project “Barcoding World Aradidae” contained 191 records, each represented by an imaged specimen, 59 with georeferenced data. These specimens belonged to 73 identified species, plus a few species not identified beyond the genus, representing all eight currently recognized subfamilies of the Aradidae, although four (Calisiinae, Prosympiestinae, Isoderminae, Chinamyersinae) were represented by only 2, 2, 1 and 1 sequences, respectively. Chinamyersinae was particularly poorly represented as only one 325 nt sequence was obtained from the Australian Kumaressa scutellata Monteith, 1966, so it was excluded from subsequent analyses.

To maximize the amplification success rate, two amplification attempts were made: the first one with the PCR primers LepF2_t1/LepR1 targeting the entire standard 658 nt “barcoding” region of the CO1 mtDNA gene (Hebert et al. 2003), and the second one with a cocktail of two PCR primer sets LepF2_t1/MHemR and MHemF/LepR1 targeting two shorter and partly overlapping sequences of the same 658 nt region. Among the total of 191 specimens, 46 failed to amplify and contained no associated DNA data, while the remaining 145 were each represented by sequences > 300 nt (GenBank accession numbers KF809495–KF809639). Information on 145
sequenced Aradidae voucher specimens with their digital images and all relevant data such as primers and original chromatograms, was deposited in the Barcode Of Life Database in the publicly accessible dataset “World Aradidae”, doi: dx.doi.org/10.5883/DS-WAHG. Among the 145 sequences, 98 were > 500 nt and represented 55 species and 29 genera of seven subfamilies (see Table 1).

Matrix construction and outgroup. The analysed dataset contained 108 sequences with the minimal/maximal lengths of 518/658 nt having no indels and was unambiguously aligned. It included all 98 barcode-compliant Aradidae records, plus 10 outgroup sequences. The latter were chosen among the publicly available “barcoding” Aradidae records, plus 10 outgroup sequences. The latter were chosen among the publicly available “barcoding” sequences of “Trichophora”, the putative sister-group of Aradoidea, representing eight families and four all superfamilies: (BOLD process ID followed by GenBank accession number, whenever available, are given in brackets): Pentatomomorpha: Pentatomidae: Eusichus tus servus (HCNCS839-09 HQ105678) and Podisus serieventris (HCNCT785-09 HQ106267); Scutelleridae: Vanduzeina balli (HCNCT778-09 HQ106544) and Homaemus aeneifrons (HCNCT197-09 HQ105749). Pyrrhocoroidea: Largidae: Physopelta australis (MAMTF1450-12); Pyrrhocoridae: Probergrothius sp. (GBMHH432-10 GU247509). Lygaeoidea: Lygaeidae Lygaeus kalmii (HCNCT755-09 HQ105855); Piesmatidae: Piesia cineraris (RBI-NA1761-13). Coreoidea: Coreidae Leptoglossus oculatus (HCNCS251-09 HQ105828); Rhopalidae: Rhopalus tigrinus (HCNCS316-09 HQ106310).

DNA substitution model and topology building. The search for a substitution model required for two out of four analyses (BI and ML; see below) was performed using MEGA5 (Tamura et al. 2011). Four different topology-building analyses were implemented and are as follows (in order of the increasing computation complexity): the Neighbor-Joining method (NJ) using uncorrected p-distances with 1000 replicates bootstrapping was performed using MEGA5; the Maximum Parsimony method (MP) with 1000 replicates bootstrapping was performed using MEGA5; the Maximum Likelihood method (ML) using the GTR+G+I model (see Results) with 1000 replicates bootstrapping was performed using MEGA5; the Maximum Likelihood method (ML) using the GTR+G+I model (see Results) with 1000 replicates bootstrapping was performed using MEGA5; the Bayesian Inference method (BI) using the commands “lset nst=6 rates=gamma” and three million generations with the default burn-in using MrBases 3.2.2. (Ronquist et al. 2012). The NJ analysis, being a not phylogenetic approach, is provided here only for comparative purposes since it is the most common and fast-performing algorithm widely used with the DNA barcoding data.

Tree manipulation and visualization. The resulting BI and three majority consensus topologies obtained by bootstrapping the dataset by using NJ, MP and ML methods were visualized in FigTree (Rambaut 2013) for subsequent comparison. All four obtained topologies were first examined unrooted. Since in all four of them all Aradidae and all non-Aradidae taxa were always forming two sister clusters, the root was forced between them and the bootstrap support value for this internode was taken as that indicating monophyly of Aradidae. The majority consensus topology from the ML analysis showing only branches with > 10% support was then exported as a Windows Enhanced Metafile (emf), the outgroup manually deleted, and the tree enhanced in CorelDraw and then Photoshop (Fig. 1). Adult habitus images illustrating terminal taxa are specimens actually genetically barcoded and can be seen at higher resolution in the online BOLD project.

Phylogenetic limitations. Our phylogenetic interpretations presented below should be taken with caution considering their unavoidable limitations. The size of the matrix could have been much greater, either in number of nucleotides or in number of terminals. Two critically important taxa were not represented: the obligatory inquilinous family Termitaphididae and the south-Pacific subfamily Chinamyersiinae (see Montetith 1980; the single available sequence of Kumaressa scutellata was judged too short to be analysed). Three oligotopic subfamilies Prosypriestinae, Isoderminae and Calisinae were represented by a single specimen, species and genus, respectively, thus depriving us of an opportunity to test their monophyly. Perhaps even more acute was the limitation imposed by use of only the “barcoding” region of the CO1 mitochondrial gene. The latter is generally considered too quickly saturating for adequate representation evolutionary events as old as the early radiation of Aradidae, which likely dates back to at least the mid-Cretaceous (120–90 Myr; see Heiss & Poinar 2012), and, therefore, the reported results should be treated with much caution.

3. Results and discussion

Statistics. The generalised time-reversible substitution model (GTR) with gamma distributed rate heterogeneity (G=0.7120) and inferred proportion of invariable sites (I=37.9976%) was of the best overall fit. The Bayesian Inference method after 3,000,000 generations had the standard deviation of split frequencies 0.009363. The BI tree and majority rule bootstrapping consensus trees representing different analytical method (NJ, MP and ML) consistently recovered monophyletic Aradidae with relatively similar internal branching patterns summarized on Table 1.

DNA barcoding of Aradidae. The method performed adequately with our Aradidae samples, even though only half of the submitted samples resulted in full sequence-recovery after trying three sets of generalized insect primers. Only 76% (145 of 191 samples submit-
Phylogeny of Aradidae. The monophyly of Aradidae has never been questioned based on any available evidence, and, therefore, the consistent recovery of the family as a clade with each of the four DNA-based analyses is not surprising (clade A on Fig. 1). It is interesting that such a well-supported group (Cassis & Schuh 2010) was only strongly recovered by BI (99%), but only weakly in NJ, MP and ML (51 – 53%; see Table 1). This fact reflects the well-known limitation of using a short fragment of the fast-evolving CO1 gene to recover early evolutionary events (Maddison 2012).

Besides recovering the Aradidae as a clade, the largest subfamily Mezirinae, represented in the analysis by 72% of all ingroup genera, is the only diversified group with more than three genera that was consistently recovered in the analysis. The monophyly of the Mezirinae (clade J on Fig. 1) is arguably our most important result, since it corroborates continuous recognition of the subfamily embracing over 50% of all flat bug genera and species. Statistical support for Mezirinae was markedly higher than that for Aradidae (see Table 1) and with less discrepancy between BI and other three analytical approaches.

Four of the seven analysed subfamilies we recovered as clades: Aneurinae, Aradinae, Calisiinae and Isoderminae. Among these, Calisiinae (clade I on Fig. 1)
were represented by only two records of two congeneric species, Aradininae (clade B on Fig. 1) by only 10 congeneric species, Aneurinae (clade G on Fig. 1) by only four species from two genera, and Isoderminae (clade F on Fig. 1) by only two records of a single species. Each of these four subfamilies was, therefore, significantly underrepresented in the analysis and their monophyly could not be rigorously tested. Prosympietiniae was represented in our analysis by a single record, thus no evidence at all could be obtained on its monophyly. Carventinae was split in two (BI) or three (NJ, MP, ML, Fig. 1) separate clades, each consistently placed on the tree. Thus the genus Birauna formed a clade with Prosympietiniae and Isoderminae (clade D on Fig. 1), while four other Carventinae genera (Stysaperta, Atractocoris, Malgasyaperta and Pericartaperta) grouped with Calissinae (clade H on Fig. 1).

Two taxonomic groups of lower rank were consistently recovered as non-monophyletic. The genus Mezira, the type genus of the most speciose subfamily, was consistently recovered in two subunits (Fig. 1). The genus Neuroctenus (clade K on Fig. 1) was found parageneric with respect to Cienoneurus (Fig. 1). The inconsistencies discovered for the three latter genera are not surprising, since they all urgently need a revision, particularly with respect to numerous species in the Oriental Region, which, as stated by Kormilev (1971), cannot be assuredly assigned to either Neuroctenus or Mezira.

Three recent studies bear on subfamily arrangement of Aradidae. Vásárhelyi (1987) suggested monophyletic Chinamyrserini to be the sister to the rest consisting of: (Aradininae + Calissinae) + (Isoderminae + (Prosympietinae + (Aneurinae + (Mezirinae + Carventinae)))). Grozeva & Kerzhner (1992) accepted (Aradininae + Calissinae) as sister to Chinamyrserini-Tretocorini, and the entire group as sister to Chinamyrseriniae-Chinamyserini + ((Prosympietinae + Isoderminae) + (Aneurinae + (Mezirinae + Carventinae))). The subfamily-level arrangement of Aradidae was not the primary goal of Cassis & Schuh (2010) and their topologies were not dogmatic in this respect, however three points of their results relevant to our work are: (1) Prosympietinae and Isoderminae do form a clade, (2) this clade is sister to the rest of the family (with or without a representative of non-monophyletic Chinamyrserinae) and (3) Aradininae is never sister to the rest of the family. Most of our results cannot be adequately compared with those above, except that in three among four analyses we recovered Aradininae as sister to the rest of the family (clade C on Fig. 1, except for MP) and Prosympietinae and Isoderminae forming a clade (clade E on Fig. 1, except for BI). At present we attribute all such discrepancies to the lack of knowledge and relative immaturity of phylogenetic Aradidae research and, therefore, have to be content with the results pending further more sizable and focussed efforts.

In conclusion, it is appropriate to speculate on what next might be the most logical step in revealing the phylogenetic history of Aradidae. A morphology-based analysis similar in scope and implementation to those of other comparably diversified family-level clades, like the Otalouline Group of rove beetles (Netwon & Thayer 1995), or Reduviidae assassin bugs (Wierauch 2008), is long overdue. Analysis of larger DNA matrices is another obvious approach. Supplementing future topologies with geographical and, particularly, biological information will markedly add to their informative value and eventually lead to a revised family classification at all levels. Reliable species identification and revived interest to flat bugs will be necessary to do so, and the herein released online DNA barcode and image collection is expected to stimulate future phylogenetic work. Until then the present status quo eight subfamily arrangement of visionary Usinger & Matsuda (1959) seems an adequate approximation.

4. Acknowledgments

Karen McLachlan Hamilton (Ottawa, Canada) critically read an early draft of this paper. Sequence analysis was carried out at the Canadian Centre for DNA Barcoding (CCDB) at the University of Guelph with support provided by the Government of Canada through Genome Canada and the Ontario Genomics Institute for the International Barcode of Life project. We thank Paul Hebert (Guelph, Canada) for overseeing the sequence analyses and for comments on our manuscript.

5. References

HENRY T.J. 1997. Phylogenetic analysis of family groups within the infraorder Pentatomomorpha (Hemiptera: Heteroptera), with
emphasis on the Lygaeoidea. – Annals of the Entomological Society of America 90: 275–301.


Rambaut A. 2013. FigTree, Version 1.4. Available at: http://tree.bio.ed.ac.uk/software/figtree/.


