Phylogenomics and the evolution of hemipteroid insects

Kevin P. Johnson[a,1], Christopher H. Dietrich[a], Frank Friedrich[b], Rolf G. Beutel[c], Benjamin Wipfler[c,d], Ralph S. Peters[d], Julie M. Allen[a,e], Malte Petersen[f], Alexander Donath[f], Kimberly K. O. Walden[g], Alexey M. Kozlov[h], Lars Podsiadlowski[f,i], Christoph Mayer[f], Karen Meusemann[f,j,k], Alexandros Vasilikopoulos[f], Robert M. Waterhouse[l], Stephen L. Cameron[m], Christiane Weirauch[n], Daniel R. Swanson[a], Diana M. Percy[o,p], Nate B. Hardy[q], Irene Terry[r], Shanlin Liu[s], Xin Zhou[t], Bernhard Misof[f], Hugh M. Robertson[g], and Kazunori Yoshizawa[u]

[a]Illinois Natural History Survey, Prairie Research Institute, University of Illinois at Urbana–Champaign, Champaign, IL 61820; [b]Institut für Zoologie, Universität Hamburg, 20146 Hamburg, Germany; [c]Institut für Zoologie und Evolutionsforschung, Friedrich-Schiller-Universität Jena, 07743 Jena, Germany; [d]Center of Taxonomy and Evolutionary Research, Arthropoda Department, Zoological Research Museum Alexander Koenig, 53113 Bonn, Germany; [e]Department of Biology, University of Nevada, Reno, NV 89557; [f]Center for Molecular Biodiversity Research, Zoological Research Museum Alexander Koenig, 53113 Bonn, Germany; [g]Department of Entomology, University of Illinois at Urbana–Champaign, Urbana, IL 61801; [h]Scientific Computing Group, Heidelberg Institute for Theoretical Studies, 69118 Heidelberg, Germany; [i]Institute of Evolutionary Biology and Ecology, University of Bonn, 53121 Bonn, Germany; [j]Evolutionary Biology and Ecology, Institute for Biology I (Zoology), University of Freiburg, 79104 Freiburg, Germany; [k]Australian National Insect Collection, Commonwealth Scientific and Industrial Research Organisation National Research Collections Australia, Acton, ACT 2601 Canberra, Australia; [l]Department of Ecology and Evolution, University of Lausanne and Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland; [m]Department of Entomology, Purdue University, West Lafayette, IN 47907; [n]Department of Entomology, University of California, Riverside, CA 92521; [o]Department of Life Sciences, Natural History Museum, London, SW7 5BD United Kingdom; [p]Department of Botany, University of British Columbia, Vancouver V6T 1Z4, Canada; [q]Department of Entomology and Plant Pathology, Auburn University, Auburn, AL 36849; [r]School of Biological Sciences, University of Utah, Salt Lake City, UT 84112; [s]BGI-Shenzhen, Shenzhen, 518083 Guangdong Province, People's Republic of China; [t]Department of Entomology, China Agricultural University, 100193 Beijing, People's Republic of China; and [u]Systematic Entomology, Hokkaido University, Sapporo, 060-8589 Japan

Hemipteroid insects (Paraneoptera), with over 10% of all known insect diversity, are a major component of terrestrial and aquatic ecosystems. Previous phylogenetic analyses have not consistently resolved the relationships among major hemipteroid lineages. We provide maximum likelihood-based phylogenomic analyses of a taxonomically comprehensive dataset comprising sequences of 2,395 single-copy, protein-coding genes for 193 samples of hemipteroid insects and outgroups. These analyses yield a well-supported phylogeny for hemipteroid insects. Monophyly of each of the three hemipteroid orders (Psocodea, Thysanoptera, and Hemiptera) is strongly supported, as are most relationships among suborders and families. Thysanoptera (thrips) is strongly supported as sister to Hemiptera. However, as in a recent large-scale analysis sampling all insect orders, trees from our data matrices support Psocodea (bark lice and parasitic lice) as the sister group to the holometabolous insects (those with complete metamorphosis). In contrast, four-cluster likelihood mapping of these data does not support this result. A molecular dating analysis using 23 fossil calibration points suggests hemipteroid insects began diversifying before the Carboniferous, over 365 million years ago. We also explore implications for understanding the timing of diversification, the evolution of morphological traits, and the evolution of mitochondrial genome organization. These results provide a phylogenetic framework for future studies of the group.

phylogeny | systematics | transcriptomes | Hemiptera | Psocodea

The hemipteroid insect orders, Psocodea (bark lice and parasitic lice), Thysanoptera (thrips), and Hemiptera (true bugs and allies; i.e., hemipterans), with over 120,000 described species, comprise well over 10% of known insect diversity. However, the evolutionary relationships among the major lineages of these insects are not yet resolved. Recent phylogenomic analyses questioned the monophyly of this group (1) demanding a reconsideration of the evolution of hemipteroid and holometabolous insects. We assess these prior results, which placed Psocodea as the sister taxon to Holometabola (insects with complete metamorphosis; e.g., wasps, flies, beetles, butterflies), and uncover relationships within and among hemipteroid insect orders by analyzing a large phylogenomic dataset covering all major lineages of hemipteroid insects.

Knowledge of the phylogeny of these insects is important for several reasons. First, major transitions between the mandibulate (chewing) mouthpart insect groundplan and "piercing–sucking" mouthparts occurred in this group. In particular, thrips and hemipterans, and some ectoparasite lice in Psocodea, have highly modified mouthparts adapted for feeding on fluids and, hence, differ markedly from their mandibulate ancestors. Through a series of remarkable modifications, hemipteroids acquired a piercing–sucking mode of feeding in both immature and adult stages that enabled them to feed not only on plant vascular fluids, but also on blood and other liquid diets. Resolution of the evolutionary tree of hemipteroid insects is needed to provide a framework for

Significance

Hemipteroid insects constitute a major fraction of insect diversity, comprising three orders and over 120,000 described species. We used a comprehensive sample of the diversity of this group involving 193 genome-scale datasets and sequences from 2,395 genes to uncover the evolutionary tree for these insects and provide a timescale for their diversification. Our results indicated that thrips (Thysanoptera) are the closest living relatives of true bugs and allies (Hemiptera) and that these insects started diversifying before the Carboniferous period, over 365 million years ago. The evolutionary tree from this research provides a backbone framework for future studies of this important group of insects.

EVOLUTION

understanding morphological transitions that occurred in this group, as well as to provide a timeframe over which these changes occurred.

In addition, several lineages of hemipteroid insects (particularly thrips and Psocodea) underwent major reorganizations of their mitochondrial genomes, including the emergence of minicircles (2). Understanding how these changes in mitochondrial genome organization occurred requires knowledge of evolutionary relationships to document in which lineages these changes first arose. Finally, hemipteroids are among the most abundant insects (3) and are therefore key components of terrestrial and aquatic food webs (4). Thus, a robust backbone phylogenetic framework is needed to place ecological studies in their evolutionary context and for use in comparative genomic and macroevolutionary analyses.

Despite their importance, relatively few studies have addressed the relationships among the major groups of hemipteroid insects [Paraneoptera, sensu stricto (excluding Zoraptera), also termed Acercaria]. While a recent large transcriptome-based phylogenomic analysis of insects (1) provided a well-resolved and strongly supported phylogenetic framework for the insect orders in general, it did not sample intensively within individual orders and recovered some unexpected relationships. Among the most puzzling was the nonmonophyly of the hemipteroid insects, with Psocodea as the sister taxon of holometabolous insects rather than as sister to thrips plus hemipterans (Condylognatha). Although this result was congruent with one earlier analysis based on three nuclear protein-coding genes (5), it had not been proposed in other molecular phylogenetic or morphological studies. Previous morphological studies indicated monophyly of hemipteroid insects with Psocodea sister to thrips plus hemipterans (6–9), or sometimes a group comprising thrips plus Psocodea (10, 11).

Another unexpected relationship recovered by Misof et al. (1) was the placement of moss bugs (Coleorrhyncha) as sister to a group comprising leafhoppers, cicadas, and relatives (Auchenorrhyncha) instead of sister to true bugs (Heteroptera). A recent morphological study also found some support for moss bugs sister to Auchenorrhyncha (12). In contrast, prior analyses based on morphology (e.g., ref. 9) and DNA sequence data (e.g., ref. 13) consistently placed moss bugs as sister to true bugs. An analysis of a reduced gene set from transcriptome data (14) also recovered moss bugs as sister to true bugs, while the full gene set placed moss bugs as sister to Auchenorrhyncha. Analysis of mitochondrial genomes (15) produced an even more unconventional result, with moss bugs placed as the sister taxon of planthoppers (Fulgoroidea), making Auchenorrhyncha paraphyletic. Thus, it is important to investigate the placement of moss bugs in more detail with both expanded taxon and gene sampling.

We evaluated these possible conflicts among analyses by analyzing a more comprehensive dataset comprising an increased number of clusters of orthologous sequence groups (2,395 protein-coding, single-copy genes) as well as an increased taxon sample within hemipteroid insects: 160 samples vs. 22 sampled by Misof et al. (1). We included representatives of all major hemipteroid lineages (sub- and infraorders). Outgroups comprised 33 species of holometabolous and nonholometabolous insect orders. This dataset enabled us to test the hypothesis of nonmonophyly of hemipteroid insects and also provides a more detailed backbone framework for the hemipteroid phylogeny. We evaluate the implications of this phylogeny for understanding the evolution of feeding strategy, morphology, and mitochondrial genome organization of this major group of insects.

Results

Phylogeny of Hemipteroid Insect Orders. Separate amino acid sequence alignments of the 2,395 single-copy genes across 193 terminal taxa (*SI Appendix*, Tables S1–S4) yielded a concatenated supermatrix of 859,518 aligned amino acid positions, which was used in subsequent phylogenetic analyses. A concatenated nucleotide sequence supermatrix of only first and second codon positions resulted in ~1.72 million aligned nucleotide sequence sites. Tree reconstructions based on the nucleotide sequence data supported a phylogenetic tree (Fig. 1 and *SI Appendix*, Figs. S1 and S2) with 172/

190 (~90%) of all nodes supported in 100% of bootstrap replicates. The tree based on amino acid sequence data (*SI Appendix*, Fig. S3) was highly concordant with that based on nucleotide data. Analysis of an optimized amino acid dataset (*SI Appendix, Supplemental Materials and Methods*) produced a tree (*SI Appendix*, Fig. S4) that was identical to that based on all amino acids with respect to relationships among orders, suborders, infraorders, and superfamilies, but had some minor rearrangements within these groups.

Considering relationships within and among orders in more detail, the thrips (Thysanoptera) were recovered with 100% bootstrap support as the sister taxon of Hemiptera (i.e., monophyletic Condylognatha), although only 68% of quartets supported this result in four-cluster likelihood mapping (FcLM) (*SI Appendix*, Tables S5 and S6). As in the study of Misof et al. (1), Psocodea was placed as the sister taxon of Holometabola in 100% of bootstrap replicates, rendering hemipteroid insects paraphyletic. However, only 25% of quartets supported Psocodea as sister to Holometabola, compared with 67% of the quartets supporting hemipteroid insect monophyly. Results from the FcLM imply that the placement of Psocodea as sister to Holometabola is unstable and may be due to confounding phylogenetic signal (e.g., from heterogeneous composition of amino acid sequences, nonstationarity of substitution processes, or nonrandom distribution of missing data) and is also dependent on the taxon sample. However, permutation tests of these results suggested the impact of these potential confounding signals on the topology was minor (*SI Appendix*, Table S6). To evaluate whether the parasitic lice in particular (Phthiraptera), which have elevated substitution rates compared with other hemipteroids (16), were a possible source of conflicting signal, we compared quartets with and without these ectoparasitic insects as the representative of Psocodea. However, the support from FcLM for monophyly of hemipteroid insects was highly similar whether parasitic lice were included (66%) or not (67%).

Morphological character mapping over three possible alternative topologies (*SI Appendix*, Fig. S5) revealed no apomorphies supporting Psocodea + Holometabola. In contrast, there are 14 potential apomorphies for the monophyly of Paraneoptera. These results indicate that there is more agreement between morphology and the FcLM results, compared with the supermatrix analyses with all taxa. For Coleorrhyncha (moss bugs), three characters are apomorphies for a sister relationship to Auchenorrhyncha (leafhoppers and relatives) but two other characters appear to support a sister relationship to Heteroptera (true bugs).

In general, the phylogenetic results from transcriptomes are congruent with the generally accepted classification schemes within these insect orders. Bark lice and parasitic lice (Psocodea) together are monophyletic. As has been suggested based on both morphological (17) and molecular (16, 18) analyses, the parasitic lice are embedded within free-living bark lice, being the sister taxon of book lice (Liposcelididae), which makes the bark lice ("Psocoptera") paraphyletic. In contrast to results based on 18S rDNA sequences (18), parasitic lice (Phthiraptera) were supported as a monophyletic group in our analyses, which included representatives of all four suborders of parasitic lice.

The thrips (Thysanoptera) were found to be monophyletic. The thrips family Phlaeothripidae was recovered as the sister taxon to the remaining thrips (Aeolothripidae + Thripidae), congruent with previous molecular analyses and the current classification of Thysanoptera into the suborders Tubulifera (i.e., Phlaeothripidae) and Terebrantia (all other thrips) (19).

The order Hemiptera was also monophyletic. Within Hemiptera, Sternorrhyncha (whiteflies, psyllids, scales, and aphids) was recovered as the sister taxon of the remaining hemipterans. Recent classification schemes (20) and prior molecular studies (13, 21) have placed the enigmatic moss bugs as the sister taxon of true bugs. However, our results recovered moss bugs as the sister taxon of Auchenorrhyncha (leafhoppers, planthoppers, and relatives), which was also found by Misof et al. (1). In FcLM analyses, 96% of quartets placed moss bugs with Auchenorrhyncha, suggesting little underlying conflict in the data for this result (*SI Appendix*, Table S6).
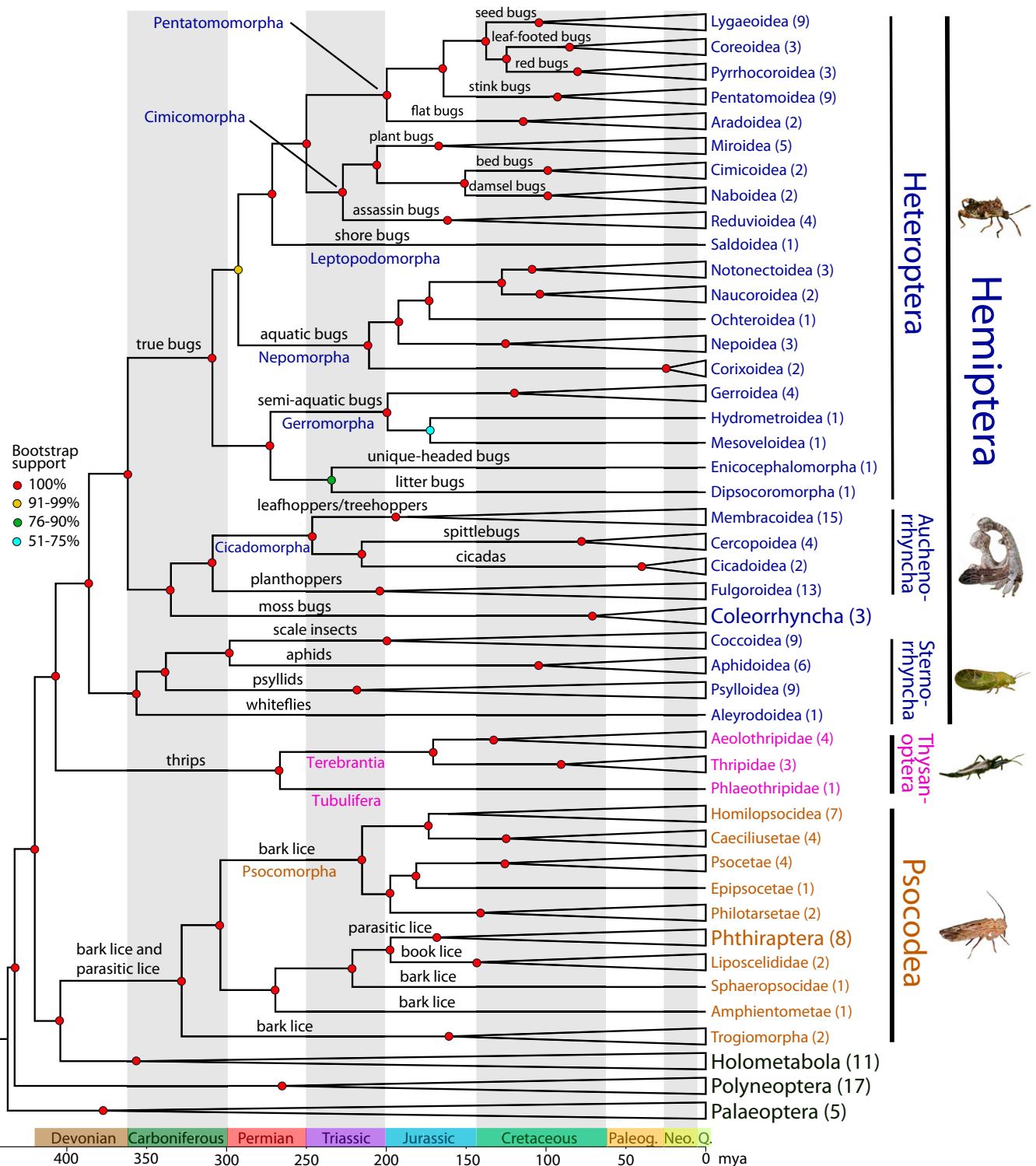
Fig. 1. Dated phylogeny of hemipteroid insects (Hemiptera, Thysanoptera, and Psocodea) based on maximum likelihood analysis of a supermatrix of first and second codon position nucleotides corresponding to 859,518 aligned amino acid positions from transcriptome or genome sequences of 193 samples. Colored circles indicate bootstrap support. Timescale in millions of years (*Bottom*) estimated from MCMCTree Bayesian divergence time analyses using 23 fossil calibration points and a reduced dataset. Number of species sampled from each group indicated in parentheses. Higher taxa are indicated as taxon labels and below branches; most convenient generalized common names are above branches. Images represent five major groups: Heteroptera, Auchenorrhyncha, Sternorrhyncha, Thysanoptera, and Psocodea.

Within Sternorrhyncha, whiteflies (Aleyrodoidea) were sister to the remainder of the suborder, and psyllids (Psylloidea) were sister to a clade composed of aphids (Aphidoidea) + scale insects (Coccoidea), also supported by 91% of quartets in FcLM analyses. Previous phylogenetic analyses of Sternorrhyncha have tended to focus within particular superfamilies or families (e.g., refs. 22–24)

rather than addressing relationships among major lineages (superfamilies).

The earliest molecular phylogenetic analyses of Hemiptera (e.g., refs. 25 and 26) failed to recover Auchenorrhyncha as a monophyletic group, as has a more recent analysis of mitochondrial genomes (15). However, our analyses provided strong support for monophyly of this group, corroborating results of other studies based on multiple loci (13, 14). Within Auchenorrhyncha, our results strongly support the taxonomic status of the two recognized infraorders Fulgoromorpha (i.e., Fulgoroidea, planthoppers) and Cicadomorpha (leafhoppers/treehoppers, spittlebugs, and cicadas) as monophyletic, as found previously (13). However, relationships among the three superfamilies of Cicadomorpha were inconsistently resolved. Cicadas (Cicadoidea) plus spittlebugs (Cercopoidea) were sister to leafhoppers/treehoppers (Membracoidea) in the analysis of nucleotide sequences (Fig. 1, FcLM 52% of quartets), but cicadas were sister to spittlebugs plus leafhoppers/treehoppers in the analysis of amino acid sequence data (*SI Appendix*, Fig. S1), which was also found in 48% of quartets of nucleotide data in FcLM analyses.

Relationships among the earlier diverging lineages of true bugs (Heteroptera) have not been resolved consistently across previous analyses (14, 27–29), in which the deepest divergences received low statistical branch support and recovered different relationships among infraorders. In our analysis, which included representatives of all seven currently recognized infraorders, the four infraorders for which more than one species was included were found to be monophyletic. Like two recent studies based on combined molecular and morphological data (29) and transcriptome data (14), we found 100% bootstrap support for (*i*) a clade comprising litter bugs (Dipsocoromorpha), unique-headed bugs (Enicocephalomorpha), and semiaquatic bugs (Gerromorpha) (also found in 100% of quartets in FcLM analyses) and (*ii*) shore bugs (Leptopodomorpha) as the sister to Cimicomorpha + Pentatomomorpha (also found in 100% of quartets in FcLM analyses).

Divergence Time Analysis. The estimate of the root age for our tree, the split between Paleoptera (dragonflies, damselflies, and mayflies) and Neoptera (all other insects) at 437 million years ago (mya) (95% CI 401–486) was only slightly older than that estimated for this node by Misof et al. (1), at 406 mya. Divergence dates for more interior nodes tended to be older than those estimated by Misof et al. (1) and more similar to those of Tong et al. (30), possibly due either to much denser sampling of minimum age fossil calibration points throughout this part of the insect tree or to different methodology (e.g., MCMCtree versus BEAST or different prior distributions of expected ages for Bayesian analyses). Analyses of divergence times postulated a common ancestor of thrips and hemipterans as early as the Devonian (~407 mya, 95% CI 373–451). Radiation within Hemiptera is also inferred to have begun in this period (~386 mya, 95% CI 354–427), with radiations within Sternorrhyncha, Auchenorrhyncha, and Heteroptera having commenced by the late Carboniferous (all before 300 mya). Radiation within modern Psocodea dates to the Carboniferous (328 mya, 95% CI 292–376), with divergence of this lineage from other insects as early as 404 mya (95% CI 367–451).

Discussion

Analysis of 2,395 protein-coding, single-copy genes derived from transcriptomes of hemipteroid insects and outgroups provided strong support for a backbone tree of hemipteroid insects largely congruent with previous analyses and classification schemes. In particular, we recovered with strong support monophyly of the three orders of hemipteroid insects: Psocodea, Thysanoptera, and Hemiptera. We also recovered monophyly of most currently recognized suborders, infraorders, and superfamilies within these groups as well as resolving relationships among these major groups. Although the unconventional result of a sister relationship between Psocodea and Holometabola of Misof et al. (1) appeared to be robust to our substantially increased taxon sampling based on maximum likelihood bootstrapping, it was not

supported by four-cluster likelihood mapping analyses. FcLM, which can detect potentially confounding signal, suggests extensive underlying conflict for this result, with the majority of quartets placing Psocodea with thrips and hemipterans, which would imply monophyly of Paraneoptera in rooted trees. However, permutations appear to rule out several possible types of confounding signal (e.g., among-lineage heterogeneity or nonrandom distribution of missing data) in our dataset. Recent work has suggested that bootstrap support from very large datasets may provide an overestimate of confidence for phylogenetic results (31–33). Thus, the position of Psocodea in the insect tree is still an open question. Monophyly of hemipteroid insects is supported by several morphological autapomorphies (34); therefore, nonmonophyly of the group would imply homoplasy in these traits. In addition, there is no known morphological apomorphy supporting Psocodea + Holometabola (*SI Appendix*, Fig. S5). In contrast, the other less conventional relationship, a clade comprising Coleorrhyncha and Auchenorrhyncha uncovered by Misof et al. (1), was recovered by our trees with increased taxon sampling and is supported by 96% of quartets in the FcLM analyses and three morphological apomorphies, suggesting that this result is robust.

Divergence time estimates using a dense sampling of 23 fossil calibration points suggest that the radiation of the hemipteroid insect orders is relatively ancient, beginning before the early Carboniferous, considerably older than initial expectations based on available fossils. However, the insect fossil record of this period is extremely fragmentary, and relatively old fossils of modern lineages that are used as calibration points imply that branches uniting these lineages must be older still, given that fossil ages represent minimum ages.

Implications for Evolution of Feeding Strategy. Our phylogenetic results generally agree with evidence from the fossil record that the earliest hemipteroids fed on detritus, pollen, fungi, or spores (as in most modern bark lice and thrips). Plant-fluid feeding probably coincided with the origin of Hemiptera and was independently derived in thrips. Today, Hemiptera is the fifth largest insect order, surpassed only by the four major holometabolous orders (Hymenoptera, Coleoptera, Lepidoptera, and Diptera). It remains one of the most abundant and diverse groups of plant-feeding insects. Within Hemiptera, the origin of true bugs apparently coincided with a shift from herbivory to predation, with subsequent shifts back to herbivory (29, 35) in the more derived lineages (Pentatomomorpha and Cimicomorpha). The two other large suborders of Hemiptera (Auchenorrhyncha and Sternorrhyncha) feed almost exclusively on vascular plant fluids.

Our results also suggest that the earliest hemipterans fed preferentially on phloem. Phloem feeding remains predominant in extant plant-feeding hemipterans, including nearly all Sternorrhyncha and most Auchenorrhyncha (36), while modern moss bugs feed on phloem-like tissues in mosses (37). A shift to xylem feeding appears to have coincided with the origin of Cicadomorpha (at least the crown group of this lineage), in which all cicadas and spittlebugs retain this preference. This is also supported by the fossil record in which the earliest leafhoppers had inflated faces (38), indicating a preference for xylem feeding, despite the predominance of phloem feeding among modern leafhoppers and treehoppers (Membracoidea). A shift to phloem feeding apparently occurred early in the evolution of Membracoidea but at least one reversal to xylem feeding [in Cicadellinae (sharpshooters)] has been inferred previously (39), consistent with our results.

Implications for Morphological Evolution. Based on the conflicting statistical support between the supermatrix analysis and four-cluster likelihood mapping, the position of lice (Psocodea) appears to be unstable. Morphological evidence, in contrast, supports the monophyly of hemipteroid insects (Paraneoptera). Our parsimony mapping of 142 morphological characters (*SI Appendix*, Fig. S5) found no apomorphies supporting Psocodea + Holometabola but 14 apomorphies supporting hemipteroid insect monophyly.

Some of these are reductions or losses, including the reduced number of tarsomeres (three in modern hemipteroids), reduced number of Malpighian tubules (four), and presence of only one abdominal ganglionic complex. Nevertheless, these characters, together with characters of the forewing base, still appear to support the sister group relationship between Psocodea and thrips plus hemipterans (11, 34, 40). Thus, the phylogenetic position of Psocodea requires further study of morphological and molecular data.

In contrast to the equivocal support for Paraneoptera, Condylognatha is strongly supported not only in the phylogenomic analyses, but also with six morphological apomorphies. The origin of this group apparently coincided with a distinct shift in mouthpart morphology and feeding habits toward piercing and sucking. These changes include anterior shifting of tentorial pits, elongated and slender mandibles, stylet-like laciniae, and a narrowed labium (*SI Appendix*, Fig. S5). Subsequent evolutionary transformations led to the very distinct and unique piercing–sucking mouthparts of hemipterans that facilitate ingestion of liquid from plant or animal tissues.

The sister-group relationship that we found between moss bugs (Coleorrhyncha) and Auchenorrhyncha has not, to our knowledge, been proposed previously in any explicit phylogenetic analysis other than in recent phylogenomic analyses of transcriptomes (1, 14). Traditionally, moss bugs were treated as one of three suborders of "Homoptera" (along with Sternorrhyncha and Auchenorrhyncha), largely based on the structure of the head. The mouthparts of moss bugs arise posteroventrally (41), as in leafhoppers and relatives, rather than anteriorly as in true bugs (42). Nevertheless, morphological evidence from fossil and living moss bugs, primarily from wing structure and musculature, suggested a closer relationship to true bugs (9, 41, 43). However, a recent comparative morphological study (12) revealed that moss bugs share a unique derived feature of the wing base with Auchenorrhyncha; a membranous proximal median plate. The same study also showed that some previously suggested morphological synapomorphies of moss bugs and true bugs (*SI Appendix*, Fig. S5C) are either ambiguous or have been misinterpreted (12). Prior molecular evidence supporting moss bugs plus true bugs was also somewhat equivocal [ref. 13: maximum likelihood (ML) bootstrap 83% and maximum parsimony (MP) bootstrap 63%]. Our results support those of other transcriptome studies (1, 14) in placing Coleorrhyncha sister to Auchenorrhyncha.

Implications for Evolution of Mitochondrial Genome Organization.
Several groups of hemipteroid insects have been shown to have highly rearranged mitochondrial genomes (2). The sister relationship between thrips and hemipterans indicates that the heightened rates of mitochondrial (mt) genome rearrangements observed in the lice (44) and thrips (45) are the result of convergence between these two clades. Even if Psocodea is sister to thrips plus hemipterans, and not to holometabolous insects, recent analyses indicating that the ancestor of all Psocodea had a generally standard insect mitochondrial gene order still result in an interpretation involving convergence (46). This phylogenetic evidence is also consistent with the absence of any shared, derived gene arrangements between Psocodea and thrips, as both have independently diverged from the inferred ancestral insect mt genome arrangement (2, 45).

An interpretation involving convergence is also consistent with the varying degrees of rearrangement observed within each order. Within Psocodea, mt genomes vary wildly across different taxonomic scales, from a single derived arrangement found in all Psocomorpha (46), to wide variation within a single genus (*Liposcelis*, ref. 47), and between closely related species of parasitic lice. In contrast, for the thrips, mitochondrial genome arrangements are relatively consistent at the family level (with only tRNA rearrangements observed), albeit still highly rearranged relative to the ancestral insect mt genome (48). Very few rearrangements of any type are observed in the Hemiptera, with the vast majority of families possessing the inferred ancestral arrangement (2).

In summary, although the exact phylogenetic position of Psocodea remains to be resolved convincingly, our results based on transcriptomes for hemipteroid insects provide a strong phylogenetic framework for future studies of genomic, morphological, ecological, and behavioral characteristics of this important group of insects.

Materials and Methods

Our general approach closely followed methods described previously by Misof et al. (1) and Peters et al. (49) for phylogenomic analyses of insect transcriptomes (*SI Appendix*, Dryad repository, 10.5061/dryad.t4f4g85). Transcriptomes of 140 samples of Paraneoptera were newly sequenced with 100 bp paired-end reads for this study using Illumina HiSeq2000 or HiSeq2500 machines to achieve at least 2.5 Gbp per taxon. The final taxon sample of 193 includes representatives of 97 hemipteroid families with several larger families represented by multiple subfamilies.

All paired-end reads were assembled with SOAPdenovo-Trans (version 1.01; ref. 50) and the assembled transcripts were filtered for possible contaminants (*SI Appendix, Table S2*) as described in Peters et al. (49). The raw reads and filtered assemblies were submitted to the NCBI SRA and TSA archives (*SI Appendix, Table S1*). We searched the assemblies for transcripts of 2,395 protein-coding genes that the OrthoDB v7 database (51) suggested to be single copy across the genomes of six species (*SI Appendix, Table S3*) using the software Orthograph (version beta4, ref. 52; for results of the orthology search see *SI Appendix, Table S4*). Orthologous transcripts were aligned with MAFFT (version 7.123; ref. 53) at the translational (amino acid) level. Corresponding nucleotide multiple sequence alignments were generated with a modified version of the software Pal2Nal (54) (version 14).

Alignment sections that could not be discriminated from randomly aligned regions at the amino acid level of each gene were identified with Aliscore version 1.2 (55, 56). To maximize the fit of our substitution models, we identified for each gene the protein domains (clans, families) and unannotated regions using the Pfam database (refs. 1 and 57 and *SI Appendix, Supplemental Materials and Methods*). The phylogenetic information content of each data block was assessed with MARE (version 0.1.2-rc) (58), and all uninformative data blocks (IC = 0) were removed. We subsequently used PartitionFinder (developer version 2.0.0-pre14, ref. 59) to simultaneously infer the best partitioning scheme and amino acid or nucleotide (removing third positions because of heterogeneity, *SI Appendix*, Fig. S6) substitution models, using the rclusterf algorithm.

Phylogenetic trees were inferred using a maximum likelihood approach with ExaML version 3.0.17 (60) for both the nucleotide and amino acid datasets. We performed 50 nonparametric bootstrap replicates mapping the support on the best ML tree after checking for bootstrap convergence with the default bootstopping criteria (61). An optimized dataset, which requires the presence of at least one species from a given taxonomic group (*SI Appendix, Table S5*) in each data block of the supermatrix (62), was used for testing the possible impact of missing data at the partition level. Four-cluster likelihood mapping (63) was used for assessing the phylogenetic signal for alternative phylogenetic relationships (*SI Appendix, Tables S5 and S6*). Permutation tests in these analyses assessed the impact of heterogeneous amino acid sequence composition among lineages, nonstationarity of substitution processes, and non-random distribution of missing data on the inferred phylogenetic tree (1).

To understand the morphological transformations underlying the evolution of the hemipteroid groups and to identify potential shared derived characters (synapomorphies), we used the morphological data matrix of Friedemann et al. (9) with 118 characters of the entire body (with modifications from ref. 12) and additionally 25 characters associated with the wing base (8). By tracing characters over the tree using maximum parsimony using Winclada (64), we evaluated three possible phylogenetic alternatives: (i) paraphyletic Paraneoptera and Coleorrhyncha sister to Auchenorrhyncha (result from ML analysis of transcriptomes); (ii) monophyletic Paraneoptera (as suggested by FcLM analyses); and (iii) paraphyletic Paraneoptera, but with Coleorrhyncha sister to Heteroptera (as suggested in previous literature).

To estimate divergence dates, we used the topology resulting from ML analysis of first and second position nucleotides as the input tree and assigned 23 ingroup fossil calibration points (65) throughout the tree (*SI Appendix, Table S7*). These calibrations were used as minimum ages in soft bound uniform priors with a root age of 406 mya (1) as a soft bound maximum. These priors were used in a Bayesian MCMCTree (66) molecular dating analysis of a first and second position nucleotide dataset for which sites were present in at least 95% of taxa.

1. Misof B, et al. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346:763–767.
2. Cameron SL (2014) Insect mitochondrial genomics: Implications for evolution and phylogeny. *Annu Rev Entomol* 59:95–117.
3. Adis J, Lubin YD, Montgomery GG (1984) Arthropods from the canopy of inundated and terra firme forests near Manaus, Brazil, with critical considerations on the pyrethrum-fogging technique. *Stud Neotrop Fauna Environ* 19:223–236.
4. Schaefer CW, Panizzi AR (2000) *Heteroptera of Economic Importance* (CRC Press, Boca Raton, FL).
5. Ishiwata K, Sasaki G, Ogawa J, Miyata T, Su Z-H (2011) Phylogenetic relationships among insect orders based on three nuclear protein-coding gene sequences. *Mol Phylogenet Evol* 58:169–180.
6. Beutel RG, Gorb SN (2001) Ultrastructure of attachment specializations of hexapods (Arthropoda): Evolutionary patterns inferred from a revised ordinal phylogeny. *J Zool Syst Evol Res* 39:177–207.
7. Wheeler WC, Whiting M, Wheeler QD, Carpenter JM (2001) The phylogeny of the extant hexapod orders. *Cladistics* 17:113–169.
8. Yoshizawa K, Saigusa T (2001) Phylogenetic analysis of paraneopteran orders (Insecta: Neoptera) based on forewing base structure, with comments on monophyly of Auchenorrhyncha (Hemiptera). *Syst Ent* 26:1–13.
9. Friedemann K, Spangenberg R, Yoshizawa K, Beutel RG (2014) Evolution of attachment structure in the highly diverse Acercaria (Hexapoda). *Cladistics* 30:170–201.
10. Whiting MF, Carpenter JC, Wheeler QD, Wheeler WC (1997) The Strepsiptera problem: Phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Syst Biol* 46:1–68.
11. Kristensen NP (1991) Phylogeny of extant hexapods. *The Insects of Australia*, ed CSIRO (Melbourne Univ Press, Melbourne), pp 125–140.
12. Yoshizawa K, Ogawa N, Dietrich CH (2017) Wing base structure supports Coleorrhyncha + Auchenorrhyncha (Insecta: Hemiptera). *J Zool Syst Evol Res* 55:199–207.
13. Cryan JR, Urban JM (2012) Higher-level phylogeny of the insect order Hemiptera: Is Auchenorrhyncha really paraphyletic? *Syst Ent* 37:7–21.
14. Wang Y-H, et al. (2017) When did the ancestor of true bugs become stinky? Disentangling the phylogenomics of Hemiptera-Heteroptera. *Cladistics*, in press.
15. Li H, et al. (2017) Mitochondrial phylogenomics of Hemiptera reveals adaptive innovations driving the diversification of true bugs. *Proc Biol Sci* 284:20171223.
16. Yoshizawa K, Johnson KP (2013) Changes in base composition bias of nuclear and mitochondrial genes in lice (Insecta: Psocodea). *Genetica* 141:491–499.
17. Lyal CHC (1985) Phylogeny and classification of the Psocodea, with particular reference to the lice (Psocodea: Phthiraptera). *Syst Ent* 10:145–165.
18. Johnson KP, Yoshizawa K, Smith VS (2004) Multiple origins of parasitism in lice. *Proc Biol Sci* 271:1771–1776.
19. Buckman RS, Mound LA, Whiting MF (2013) Phylogeny of thrips (Insecta: Thysanoptera) based on five molecular loci. *Syst Ent* 38:123–133.
20. Bourgoin T, Campbell BC (2002) Inferring a phylogeny for Hemiptera: Falling into the 'autapomorphic trap'. *Denisia* 4:67–82.
21. Ouvrard D, Campbell BC, Bourgoin T, Chan KL (2000) 18S rRNA secondary structure and phylogenetic position of Peloridiidae (Insecta, Hemiptera). *Mol Phylogenet Evol* 16:403–417.
22. Von Dohlen CD, Moran NA (2000) Molecular data support a rapid radiation of aphids in the Cretaceous and multiple origins of host alternation. *Biol J Linn Soc Lond* 71:689–717.
23. Gullan PJ, Cook LG (2007) Phylogeny and higher classification of the scale insects (Hemiptera: Sternorrhyncha: Coccoidea). *Zootaxa* 1668:413–425.
24. Percy DM, et al. (2018) Resolving the psyllid tree of life: Phylogenomic analyses of the superfamily Psylloidea (Hemiptera). *Syst Ent* 43:762–776.
25. Campbell BC, Steffen-Campbell JD, Sorensen HT, Gill RJ (1995) Paraphyly of Homoptera and Auchenorrhyncha inferred from 18S rDNA nucleotide sequences. *Syst Ent* 20:175–194.
26. von Dohlen CD, Moran NA (1995) Molecular phylogeny of the Homoptera: A paraphyletic taxon. *J Mol Evol* 41:211–223.
27. Li H, et al. (2012) The complete mitochondrial genome and novel gene arrangement of the unique-headed bug *Stenopirates* sp. (Hemiptera: Enicocephalidae). *PLoS One* 7:e29419.
28. Weirauch C, Štys P (2014) Litter bugs exposed: Phylogenetic relationships of Dipsocoromorpha (Hemiptera: Heteroptera) based on molecular data. *Insect Syst Evol* 45:351–370.
29. Weirauch C, Schuh RT, Cassis G, Wheeler WC (2018) Revisiting habitat and lifestyle transitions in Heteroptera (Insecta: Hemiptera): Insights from combined morphological and molecular phylogeny. *Cladistics*, in press.
30. Tong KJ, Duchêne S, Ho SYW, Lo N (2015) INSECT PHYLOGENOMICS. Comment on "Phylogenomics resolves the timing and pattern of insect evolution". *Science* 349:487.
31. Salichos L, Rokas A (2013) Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature* 497:327–331.
32. Kobert K, Salichos L, Rokas A, Stamatakis A (2016) Computing the internode certainty and related measures from partial gene trees. *Mol Biol Evol* 33:1606–1617.
33. Shen X-X, Hittinger CT, Rokas A (2017) Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nature Ecol Evol* 1:126.
34. Yoshizawa K, Lienhard C (2016) Bridging the gap between chewing and sucking in the hemipteroid insects: New insights from Cretaceous amber. *Zootaxa* 4079:229–245.
35. Cobben RH (1979) On the original feeding habits of the Hemiptera (Insecta): A reply to Merrill sweet. *Ann Entomol Soc Am* 72:711–715.
36. Backus EA (1988) Sensory systems and behaviors which mediate hemipteran plant-feeding: A taxonomic overview. *J Insect Physiol* 34:151–165.
37. Cronk QCB, Forest F (2017) The evolution of Angiosperm trees: From Palaeobotany to genomics. *Comparative and Evolutionary Genomics of Angiosperm Trees*, eds Groover A, Cronk Q (Springer, New York), pp 1–17.
38. Shcherbakov D (1996) Origin and evolution of the Auchenorrhyncha as shown by the fossil record. *Studies on Hemipteran Phylogeny*, ed Schaeffer CW (Entomol Soc Am, Lanham, MD).
39. Dietrich CH, et al. (2017) Anchored hybrid enrichment-based phylogenomics of leafhoppers and treehoppers (Hemiptera: Cicadomorpha: Membracoidea). *Insect Syst Diver* 1:57–72.
40. Beutel RG, Friedrich F, Ge S-Q, Yang X-K (2014) *Insect Morphology and Phylogeny: A Textbook for Students of Entomology* (Walter de Gruyter, Berlin).
41. Spangenberg R, et al. (2013) The cephalic morphology of the Gondwanan key taxon *Hackeriella* (Coleorrhyncha, Hemiptera). *Arthropod Struct Dev* 42:315–337.
42. Spangenberg R, Friedemann K, Weirauch C, Beutel RG (2013) The head morphology of the potentially basal heteropteran lineages Enicocephalomorpha and Dipsocoromorpha (Insecta: Hemiptera: Heteroptera). *Arthropod Syst Phyl* 71:103–136.
43. Shcherbakov D, Popov YA (2002) Superorder Cimicidea Laicharting, 1781, Order Hemiptera Linné, 1758. The bugs, cicadas, plantlice, scale insects, etc. *History of Insects*, eds Rasnitsyn AP, Quicke DLJ (Kluwer, Dordrecht, The Netherlands), pp 143–157.
44. Cameron SL, Yoshizawa K, Mizukoshi A, Whiting MF, Johnson KP (2011) Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). *BMC Genomics* 12:394.
45. Dickey AM, et al. (2015) A novel mitochondrial genome architecture in thrips (Insecta: Thysanoptera): Extreme size asymmetry among chromosomes and possible recent control region duplication. *BMC Genomics* 16:439.
46. Yoshizawa K, et al. (2018) Mitochondrial phylogenomics and genome rearrangements in the barklice (Insecta: Psocodea). *Mol Phylogenet Evol* 119:118–127.
47. Shi Y, et al. (2016) The mitochondrial genome of booklouse, *Liposcelis sculptilis* (Psocoptera: Liposcelididae) and the evolutionary timescale of *Liposcelis*. *Sci Rep* 6:30660.
48. Yan D, et al. (2014) The mitochondrial genome of *Frankliniella intonsa*: Insights into the evolution of mitochondrial genomes at lower taxonomic levels in Thysanoptera. *Genomics* 104:306–312.
49. Peters RS, et al. (2017) Evolutionary history of the Hymenoptera. *Curr Biol* 27:1013–1018.
50. Xie Y, et al. (2014) SOAPdenovo-Trans: De novo transcriptome assembly with short RNA-seq reads. *Bioinformatics* 30:1660–1666.
51. Kriventseva EV, Rahman N, Espinosa O, Zdobnov EM (2008) OrthoDB: The hierarchical catalog of eukaryotic orthologs. *Nucleic Acids Res* 36:D271–D275.
52. Petersen M, et al. (2017) Orthograph: A versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. *BMC Bioinformatics* 18:111.
53. Katoh K, Standley DM (2016) A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* 32:1933–1942.
54. Suyama M, Torrents D, Bork P (2006) PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* 34:W609–W612.
55. Misof B, Misof K (2009) A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: A more objective means of data exclusion. *Syst Biol* 58:21–34.
56. Kück P, et al. (2010) Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Front Zool* 7:10.
57. Finn RD, et al. (2014) Pfam: The protein families database. *Nucleic Acids Res* 42:D222–D230.
58. Misof B, et al. (2013) Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. *BMC Bioinformatics* 14:348.
59. Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A (2014) Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol Biol* 14:82.
60. Kozlov AM, Aberer AJ, Stamatakis A (2015) ExaML version 3: A tool for phylogenomic analyses on supercomputers. *Bioinformatics* 31:2577–2579.
61. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A (2010) How many bootstrap replicates are necessary? *J Comput Biol* 17:337–354.
62. Dell'Ampio E, et al. (2014) Decisive data sets in phylogenomics: Lessons from studies on the phylogenetic relationships of primarily wingless insects. *Mol Biol Evol* 31:239–249.
63. Strimmer K, von Haeseler A (1997) Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. *Proc Natl Acad Sci USA* 94:6815–6819.
64. Nixon K (2002) Winclada (Ithaca, NY), version 1.00. 08. Available at http://www.diversityoflife.org/winclada/. Accessed January 24, 2018.
65. Parham JF, et al. (2012) Best practices for justifying fossil calibrations. *Syst Biol* 61:346–359.
66. Yang Z (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591.

Supplementary Information for

Phylogenomics and the evolution of hemipteroid insects

Kevin P. Johnson, Christopher H. Dietrich, Frank Friedrich, Rolf G. Beutel, Benjamin Wipfler, Ralph S. Peters, Julie M. Allen, Malte Petersen, Alexander Donath, Kimberly K. O. Walden, Alexey M. Kozlov, Lars Podsiadlowski, Christoph Mayer, Karen Meusemann, Alexandros Vasilikopoulos, Robert M. Waterhouse, Stephen L. Cameron, Christiane Weirauch, Daniel R. Swanson, Diana M. Percy, Nate B. Hardy, Irene Terry, Shanlin Liu, Xin Zhou, Bernhard Misof, Hugh M. Robertson, Kazunori Yoshizawa

Kevin P. Johnson

Email: kpjohnso@illinois.edu

This PDF file includes:

       Supplemental text

       Figs. S1 to S6

       Tables S1 to S7

       References for SI reference citations

Supplemental Materials and Methods

Sampling of transcriptomes. We sampled RNA of 139 species of Paraneoptera for transcriptome sequencing. Collected samples were ground and preserved in RNAlater (Qiagen, Hilden, Germany) and stored at +4 °C or -80 °C or dry at -80 °C until further processing. For all species, we sampled the RNA from the entire body of adult specimens (in a few cases nymphs were also sampled or information on life-stage is unknown). Detailed information on the samples (e.g., collecting information and date) are provided in NCBI along with deposited data (Table S1). For one species (*Largus californicus*), we sequenced two samples because they had not been identified before sequencing. We also included in our study new assemblies of raw data from 48 whole body transcriptomes (18 additional Paraneoptera, 5 Palaeoptera, 16 Polyneoptera, and 9 Holometabola; marked with asterisks in Table S1) published in a preceding investigation (1). Our sampling of transcriptomes thus comprised 158 samples of Paraneoptera, comprising 157 different species, and 30 outgroup species.

Transcriptome sequencing. RNA extraction, next generation sequencing (NGS) library preparation, and sequencing of the prepared libraries on Illumina HiSeq 2000 or 2500 sequencers (Illumina, San Diego, CA, USA) followed the protocol given by Misof et al. (1). In one species (i.e., *Diaphorina citri*), from which a total RNA yield < 3 µg was obtained, we constructed the NGS library using the TruSeq mRNA Library Prep Kit (Illumina). We sheared the purified mRNA into fragments of 160–170 bp in length using divalent cations at 98 °C. Fragment sizes and concentrations were determined with the aid of an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) and a StepOnePlus Real-Time PCR thermocycler (Applied Biosystems, Waltham, MA, USA). All NGS libraries were paired-end sequenced on a HiSeq 2000 or 2500 (Illumina) with 150 bp (all libraries except those prepared using the TruSeq kit) and 90 bp (libraries prepared with the TruSeq kit) read length. Per library, we collected at least 2.5 Gbp of raw data.

De novo assembly of transcriptomes. Transcript raw reads were assembled using the assembler SOAPdenovo-Trans 31 or 49 kmer (version 1.01) (2). All raw reads were quality checked and trimmed, and all remaining reads were used for *de novo* assembly. Quality check/trimming and *de novo* assembly were done as described by Peters et al. (3).

Identification and removal of contaminating sequences. Local installations of VecScreen (http://www.ncbi.nlm.nih.gov/tools/vecscreen/) and the UniVec database build 7.1 (or 9.0; Table S2) (http://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/) were used to remove terminal and internal contamination with vector or linker/adapter sequences from transcriptome assemblies. The removal of internal contamination resulted in a split of contigs/scaffolds. We next searched the assembled transcriptomes for cross-library contamination, as it often occurs when index-tagged NGS libraries are pooled on the same Illumina NGS sequencer lane. The search strategy is outlined in Mayer et al. (4) and Peters et al. (3). Specifically, we compared each transcriptome assembly with all other assemblies sequenced in context of the 1KITE project using BLASTN of the BLAST+ (version 2.2.29) program suite (5). We explicitly refrained from restricting the contamination search to only those transcriptomes sequenced on the same lane, in order to be able to detect also contamination that may have occurred in pre-sequencing steps (e.g. library construction). In cases where BLASTN identified transcripts that shared a sequence similarity of at least 98% over a length of at least 180 bp, we used the coverage depth to decide which transcript is likely the original sequence (by being more than twice as abundant) and which one likely

represents the contamination (by being less abundant). We used as coverage depth of a given transcript the average k-mer coverage statistic provided by the assembly software SOAPdenovo-Trans-31 or 49 kmer (2). Having identified transcripts sharing a high sequence identity, we applied the following procedure: (*i*) If two transcripts differed more than 2-fold in their relative coverage, we removed the transcript with the lower relative coverage from the corresponding assembly. (*ii*) If the coverage of the two transcripts in question differed 2-fold or less, we conservatively removed both of them from the two corresponding assemblies. This procedure allowed us to remove putative foreign contamination (e. g., from third party libraries sequenced on the same lane, but not present in our analyses). In case of multiple highly similar sequences, we retained only the single transcript with the highest relative coverage, given that its coverage was more than 2-fold higher than the coverage of the second-best matching transcript. We additionally removed transcripts from the assemblies that NCBI identified as possibly foreign contamination when submitting the assemblies to the NCBI Transcriptome Shotgun Assembly (TSA) database (see below). Information on how many transcripts were removed from each transcript library is summarized in Table S2. Due to the different assembly strategies some of the assemblies had significantly shorter and less abundant contigs filtered out due to contamination issues. All assemblies and raw data (reads) are available at NCBI via the Umbrella BioProjects accession number NCBI: PRJNA183205 (''The 1KITE project: evolution of insects'') or NCBI: PRJNA296358 ("Hemipteroid Insect Assembling the Tree of Life Project").

Identification of orthologous transcripts of single-copy protein-coding genes. We used the program Orthograph (version beta4) (6) to map transcripts to a set of target genes. For this search, Orthograph requires a set of reference gene groups in which all sequences within a group of genes are suggested to be orthologous among the reference taxa ("orthologous groups"). This set is also referred to as the "ortholog reference set". A customized profile query in OrthoDB7 (7) identified clusters of orthologous sequence groups (COGs) of single-copy orthologs within Pancrustacea. Specifically, we selected the Pancrustacea node and required the presence of exactly one gene copy (=1) in the genomes of six reference species: *Daphnia pulex* (8), *Zootermopsis nevadensis* (9), *Pediculus humanus* (10), *Rhodnius prolixus* (11), *Tribolium castaneum* (12), and *Acromyrmex echinatior* (13). We also required the presence of zero or a single gene copy (<=1) in the genome of *Acyrthosiphon pisum* (14). We did not restrict the number of gene copies in the remaining of the Pancrustacean genomes. Our specific profile query returned 2,453 clusters of COGs. We did not include the sequences of *Acyrthosiphon pisum* in our ortholog clusters, so every cluster contained exactly six sequences, one from each of the six reference species. We next downloaded the official gene sets (OGSs) for each reference species, from the respective genome databases (see Table S3). OGSs were downloaded both on the amino acid and the corresponding nucleotide level (transcript sequences and cds). The OGS of *Zootermopsis nevadensis* was kindly provided by J. Liebig and R. Waterhouse prior to their official release. We modified the OGSs by the use of custom Perl scripts as following: 1) sequence headers were shortened so that they correspond with the public gene identifiers in the OrthoDB-tab-delimited and OrthoDB-fasta files, 2) sequences with ambiguous/duplicate headers were removed in the OGSs files, 3) amino acid sequences with no corresponding nucleotide sequence were removed and vice versa, 4) sequence headers were made corresponding on the amino acid and the nucleotide level, 5) alternative isoforms of the same gene were removed from the OGSs only if the gene belonged to our reference set of COGs, 6) terminal stop codons were removed from the protein OGSs , 7) If the protein sequence used by OrthoDB for inferring *de novo* orthology (fetched from the OrthoDB-fasta file) did not exactly match the sequence with the same identifier in the protein OGS, we replaced the sequence in the OGS with the corresponding

OrthoDB sequence. The screening for alternative isoforms of the same gene was only possible in those cases where such information was available from the headers of the sequences. Next, we checked whether all gene identifiers in the OrthoDB-tab-delimited file were present in the respective OGS files. We found 58 COGs for which *Daphnia pulex* identifiers had no equivalent in the respective protein and nucleotide OGS files. We deleted those COGs from the ortholog reference set, which reduced the total number of orthologous groups to 2,395. The OrthoDB7 IDs of COGs which were deleted from the initial set are the following: EOG7JTJQ7, EOG7HFCC3, EOG7FNQ2F, EOG73RP58, EOG7HXQJZ, EOG79H519, EOG75XTFF, EOG7D8HQK, EOG7H7MX4, EOG7813F7, EOG7VXMQN, EOG7RZHHV, EOG7CKF38, EOG76B9C8, EOG73NSX7, EOG7WDZWG, EOG728MKZ, EOG7WXB40, EOG7M9M92, EOG7DZMFK, EOG7RVNBJ, EOG7RZHKM, EOG7WB3Q1, EOG7F873J, EOG7QZT5N, EOG7WDZWD, EOG74V7SB, EOG74NH9C, EOG75N6RK, EOG7GJJ72, EOG75BM0H, EOG7KDS50, EOG7RRS2S, EOG7JMT9D, EOG7XMDS3, EOG7HXQJW, EOG7B3CFV, EOG75R2XS, EOG72K79M, EOG7XT583, EOG7M9M97, EOG7XHJJH, EOG7H4RP4, EOG7D5NGJ, EOG7BGVGZ, EOG70GZ94, EOG7161Z2, EOG77DW2F, EOG7NH272, EOG7CGJST, EOG72VV23, EOG7QGFXZ, EOG7KMHN3, EOG74Z422, EOG7DZMDN, EOG7V7J1V, EOG7SFVQF, EOG73JXPV. The transcript libraries were searched for these 2,395 single-copy protein-coding target genes with Orthograph (6). Orthograph uses a graph-based approach with the best-reciprocal hit (BRH) criterion to extend clusters of known orthologous sequence groups in transcript sequences. The algorithm maps sequences from the transcript library to the globally best matching COG.

Orthograph depends on the following bioinformatics software packages, of which we used the hereafter specified versions: HMMER (version 3.1b2) (15), BLAST+ (version 2.2.26) (5), Exonerate (version 2.2) (16), and MAFFT (version 7.123) (17). We applied a non-strict reciprocal search. Thus, the BRH criterion was fulfilled if the first reciprocal hit was a protein sequence that was also part of the pHMM, irrespective of the species. We allowed frame shift-corrected transcriptional extension of each transcript beyond the region of the transcript for which the BRH criterion was established (option: extend-orf = 1), as long as this region was not longer than the region for which the BRH was established (see above). We used the accompanying helper script to summarize the Orthograph results from all taxa. When summarizing the results, we removed terminal and masked internal stop codons with X and NNN, respectively, in all amino acid and corresponding nucleotide sequences (option: -s). All other Orthograph parameters were left at the default values.

Searching for transcripts being orthologous to one of the 2,395 target genes in all 188 analyzed transcript libraries returned transcripts of, on average, 2,068 target genes (median: 2,113.5; minimum: 1,451; maximum 2,339; Table S4). All 2,395 target genes were present in at least one of the 188 transcriptomes.

Inference of multiple sequence alignments. We constructed multiple sequence alignments (MSA) using the amino acid sequences with MAFFT (version 7.123we) (17), using the L-INS-i algorithm. We subsequently checked the resulting MSAs and, if necessary, refined the MSAs as outlined by Misof et al. (1). Assessing the quality of the inferred Multiple Sequence Alignments (MSAs) revealed putatively misaligned amino acid transcripts (hereafter referred to as outliers) in 651 of the 2,395 investigated single copy genes. Alignment refinement of the 1,642 outlier sequences succeeded for 449 sequences. The remaining 1,193 sequences, referring to a total of 532 single copy genes, remained outliers and were consequently removed from the MSAs. We also removed the corresponding nucleotide sequences from the respective nucleotide file. In addition, we removed all sequences of *Daphnia pulex* from both the amino acid and nucleotide data as well as sites with gaps only.

We inferred the MSA of each OG also on the transcriptional (nucleotide) level using a modified version of the software Pal2Nal (18) (version 14; see Misof et al. [1] for details on the software modification). All MSAs have been deposited in and are available from Dryad Data Repository (doi:10.5061/dryad.t4f4g85).

Protein domain identification. To infer a protein domain-based sequence data partitioning scheme, we applied the procedure described by Misof et al. (1). Protein domains and protein domain clans (i. e., evolutionary related protein domains) (19) were identified on the amino acid level in each MSA by using pHMMs from the Pfam database (release 29) and the software pfam_scan.pl (version 1.5, release date: May 24, 2016) (20) in conjunction with HMMER (version 3.1b2) (15). For additional details, see Misof et al. (1) and Peters et al. (3).

Search for protein domains in the refined amino acid sequence alignments of the 2,395 single-copy genes assigned 33.7% of the alignment sites to domains of Pfam-A. A total of 66.3% of the alignment sites remained unannotated. Based on the domain identification results, we split the 2,395 MSAs and rearranged their sites into 3,717 data blocks. More specifically, by pooling Pfam A domain regions according to their clan annotation (19), we obtained 295 clan-based data blocks. Pooling Pfam-A domains without a clan annotation according to the domain name, we obtained 1,027 data blocks. Finally, pooling unannotated regions (voids) according to their gene origin we obtained 2,395 data blocks corresponding to the 2,395 analyzed genes.

Multiple sequence alignment masking. All amino acid alignments were analyzed with Aliscore version 1.2 (21, 22) to identify blocks of putative ambiguously aligned sequence regions or regions indistinguishable from randomized data. We used the '-e' option, which is able to cope with transcript sequence alignments containing many gaps (23). Aliscore was run with the default sliding window size and the option –r in order to compare all sequence pairs in each sliding window. Alignment segments identified by Aliscore as being ambiguously aligned or randomized in the MSAs were subsequently removed from domain and void data blocks at the amino acid and corresponding nucleotide level. Gap symbols (-) at the beginning and at the end of the resulting MSAs were replaced with 'X' (translational level) and 'N' (transcriptional level), respectively. Finally, we concatenated protein domain- and gene-based data blocks (the latter comprising void regions) to supermatrices using custom Perl scripts.

Removal of phylogenetically non-informative data blocks. We assessed the potential phylogenetic signal of each data block with the software MARE (version 0.1.2-rc) (24). Data blocks whose information content (IC) was zero at the translational level were removed from the supermatrix on the translational and on the transcriptional level using custom Perl scripts.

After (*i*) removing ambiguous alignment sites in each data block resulting from steps outlined above, (*ii*) deleting data blocks that contained an IC = 0 (792 in total), and (*iii*) concatenating all remaining data blocks, the resulting supermatrices consisted of 859,518 amino acid and 1,719,036 nucleotide sites (1st and 2nd codon position only), respectively. The amino acid supermatrix comprised 2,902 data blocks (279 clan data blocks from pooled Pfam-A protein domains with clan association, 855 data blocks from pooled Pfam-A protein domains without clan association, 1,768 data blocks from void regions). Both supermatrices are deposited at Dryad Data Repository (doi:10.5061/dryad.t4f4g85).

Partitioning and substitution model selection. We used PartitionFinder (developer version 2.0.0-pre14) (25, 26) to identify combinations of data blocks that can be modeled with the same substitution model and parameters and which therefore should be combined into

partitions. The corrected Akaike information criterion (AICc) (27) was used to assess whether or not to combine data blocks to partitions. PartitionFinder was run with the following settings: branchlengths [linked], models [LG+G, LG+G+F, WAG+G, WAG+G+F, BLOSUM62+G, BLOSUM62+G+F, DCMUT+G, DCMUT+G+F, JTT+G, JTT+G+F, LG4X], model_selection [AICc], search [rclusterf] (Lanfear et al. 2014), weights [rate = 1.0, base = 1.0, model = 0.0, alpha = 1.0], rcluster-percent [100.0], rcluster-max [10,000]. PartitionFinder was started with the following additional command line arguments: '--raxml --ml-tree --all-states --min-subset-size 50'. The restriction to the selected substitution matrices was done because of computational limitations.

PartitionFinder suggested merging the 2,902 data blocks of the amino acid supermatrix into 1,857 partitions. Furthermore, PartitionFinder determined the best amino acid substitution model for each partition. This partitioning scheme and the determined models were used for subsequent analysis.

We obtained a partitioning scheme on the nucleotide level by transferring the partitioning scheme that we used in the amino acid supermatix to the nucleotide supermatrix. An important extension of the partitioning scheme on the nucleotide level is that we treated all 1st and 2nd codon positions of each partition as separate partitions. Third codon positions were discarded and not considered in our analyses, since the nucleotide data set with third codon positions exhibits a much higher level of among lineage compositional heterogeneity. As a result, the nucleotide partitioning scheme had twice as many partitions (3,714) as the amino acid partitioning scheme. In the nucleotide phylogenetic analysis, we used the GTR+G substitution model for all partitions, which is the only model that can be analyzed with ExaML. The final data sets and files containing the inferred partition schemes and model selection results are deposited at Dryad Data Repository (doi:10.5061/dryad.t4f4g85). To distinguish the supermatrices mentioned here from the optimized dataset (see below), we refer to the here-mentioned matrices as the "primary matrices".

Testing for compositional heterogeneity. We evaluated whether or not our datasets have evolved under globally stationary, reversible, and homogeneous (SRH) conditions (28). To test for possible compositional heterogeneity of amino acid and nucleotide datasets, we used SymTest version 2.0.47 (https://github.com/ottmi/symtest) (29), which uses matched-pairs tests of homogeneity (for details see Misof et al. [1]). Tests were applied on i) the amino acid supermatrix, ii) the nucleotide supermatrix with all codon positions and iii) the nt supermatrix including 1st and 2nd codon positions only. We generated heatmaps based on the p-values obtained from Bowker's matched-pairs test of symmetry in order to determine which sequence pairs matched SRH conditions. All tested datasets exhibit among-lineage compositional heterogeneity, albeit to very different extents (Fig. S5). For our phylogenetic tree inference, we used the amino acid dataset and the nucleotide dataset with the 1st and 2nd codon positions included and the 3rd codon positions excluded because 1st and 2nd codon positions showed less among-lineage heterogeneity than a dataset including all codon positions. We assessed the possible impact of compositional heterogeneity among lineages on the inferred relationships using quartet permutation tests (see below).

We additionally evaluated the coverage of the (amino acid) dataset with respect to pairwise sequence coverage of unambiguous data using AliStat v.1.6 (https://github.com/thomaskf/AliStat) (30, see also 1). The overall completeness score was 60%.

Phylogenetic tree inference and bootstrap analysis. We applied the maximum likelihood optimality criterion (ML) as implemented in ExaML (version 3.0.17) (31) for the phylogenetic inferences. We selected the tree with the best log-likelihood score found in 50

independent tree searches (25 randomized stepwise addition parsimony starting trees, 25 completely random starting trees) per dataset. All starting trees were inferred with RAxML (version 8.0.26) (32). We applied the GTR substitution model when analyzing the nucleotide sequence data and applied the partition-specific substitution models suggested by PartitionFinder when analyzing the amino acid sequences in ExaML. Rate heterogeneity was approximated with gamma distributed site rates, using four discrete rate categories and the median for each category. Node support was estimated with non-parametric bootstrap analyses (33). Bootstrap replicate alignments and random starting trees were generated with RAxML and a custom shell script. Subsequently, ExaML was used to infer one ML tree per bootstrap replicate, applying the original partitioning scheme suggested by PartitionFinder for the respective supermatrix. In order to determine the minimum number of replicates needed for a reliable estimation of bootstrap support values, we used the "autoMRE" bootstrap convergence criterion (34), as implemented in RAxML, with the default threshold setting of 0.03. Convergence was evaluated after analyzing each batch of 50 bootstrap replicates. All inferred trees were rooted with the palaeopteran outgroups.

The majority rule extended (MRE) bootstrap convergence criterion (34) indicated that 50 bootstrap replicates are sufficient to accurately assess node support irrespective of whether we analyzed the primary supermatrix containing amino acid sequences or the primary supermatrix containing the corresponding nucleotide sequences. The node support values in Fig. 1 and Figs. S2 and S3 are consequently based on 50 non-parametric bootstrap replicates each. The results from the phylogenetic tree inference and bootstrap analysis have been deposited at Dryad Data Repository (doi:10.5061/dryad.t4f4g85).

Testing for impact of missing data via analysis of a decisive dataset. Following Dell'Ampio et al. (35) and Misof et al. (1), we generated an optimized dataset, which contained at least one species from a given taxonomic group in each partition of the analyzed supermatrix (also known as a "decisive dataset" [35]). We defined 23 taxonomic groups (see Table S5 for which species are part of a given group). Given that the generation of an optimized dataset resulted in some data blocks being excluded from subsequent analyses, we inferred new partitioning schemes and corresponding partition-specific substitution models with PartitionFinder and RAxML seperately for the decisive dataset (with the same versions and settings as outlined above for the primary supermatrix). Note that we considered only amino acid sequences for the phylogenetic analyses of the optimized dataset. The decisive dataset comprises 550,395 amino acid positions, analyzed in 852 partitions. We conducted 50 tree searches with ExaML (version 3.0.17) (25 with randomized stepwise addition parsimony starting trees, 25 with completely random starting trees), identical software and settings to what was used for analysis of the primary supermatrix. The results from analyzing the decisive dataset have been deposited at Dryad Data Repository (doi:10.5061/dryad.t4f4g85).

Four-cluster Likelihood Quartet Mapping (FcLM). We applied FcLM to assess the phylogenetic support for conflicting hypotheses (1, 3, 36). We focused on nine phylogenetic hypotheses: (*i*) Psocodea sister to Holometabola. This hypothesis was tested with three different analyses, including or excluding subgroups of Condylognatha to assess the impact of these subgroups. First, it was tested with all Condylognatha included (*ia*), second with Thysanoptera excluded and Hemiptera included (*ib*), and third with Hemiptera excluded and Thysanoptera included (*ic*) (*ii*) Thysanoptera sister to Hemiptera; (*iii*) Coleorrhyncha sister to Auchenorrhyncha; (*iv*) Myerslopiidae sister to remainder of Membracoidea; (*v*) Aphidoidea sister to Coccoidea; (*vi*) Nepomorpha sister to the clade including Leptopodomorpha, Cimicomorpha, and Pentatomomorpha; (*vii*) Leptopodomorpha sister to a clade including Cimicomorpha and Pentatomomorpha; (*viii*) Gerromorpha sister to a clade including

7

Dipsocoromorpha and Enicocephalomorpha; (*ix*) Cercopoidea sister to Cicadoidea. For hypothesis *ia*, we also performed four variations excluding quartets including representatives of certain subgroups: (*iaI*) excluding all Phthiraptera, (*iaII*) excluding all Phthiraptera plus Liposcelidae, (*iaIII*) excluding all "Psocoptera", and (*iaIV*) excluding all Phthiraptera and all Thysanoptera.

For each hypothesis tested, we defined four taxonomic groups. Specific information about which species were part of which group is given in Table S5. For testing each hypothesis, we generated optimized datasets, including only those partitions of the primary supermatrix which contained sequences of at least one species from the four taxa specified for testing the respective hypothesis (see Table S5).

We used the PTHREADS implementation of RAxML version 8.2.8 (32) to generate parsimony start trees and used RAxML version 8.2.8 PTHREADS and Examl version 3.0.17 (31) to infer the support for each quartet using the partition scheme and substitution models inferred for the complete supermatrix at the translational (amino acid) level. Results were finally mapped into 2D simplex graphs with a custom-made Perl script. Results are provided in Table S6. The optimized analyzed amino acid data sets used for FcLM have been deposited along with partition information files at Dryad Data Repository (doi:10.5061/dryad.t4f4g85).

Results from FcLM testing the putative sister group relationship of Psocodea and Holometabola did not support the ML tree topology, i.e., Psocodea and Holometabola as sister groups, irrespective of whether all condylognathan taxa were included or certain taxa (Thysanoptera, Hemiptera) were excluded. The majority of quartets in all analyses support monophyletic Acercaria, with about half to two thirds of the quartets. We examined these results in further detail: based on the results including all Condylognathan taxa we only mapped selected quartets excluding certain psocodean or condylognathan taxa *a posteriori* (see Tables S5; S6) to see if the support for monophyletic Acercaria could be associated with the inclusion of certain taxonomic groups. Specifically, we excluded quartets including representatives of the following taxa from the results of (*ia*): (*iaI*) all Phthiraptera, (*iaII*) all Phthiraptera plus Liposcelididae, (*iaIII*) all "Psocoptera", (*iaIV*) all Phthiraptera plus all Thysanoptera (see Tables S5; S6). In all tests, the results only marginally differed from the analysis including all taxa, the percentage of quartets supporting monophyletic Acercaria varied between 66% and 68%. This implies that the incongruence between the ML tree and the FcLM results is not caused by signal of certain taxonomic groups.

FcLM results of testing whether Nepomorpha are sister group of the clade including Leptopodomorpha, Cimicomorpha, and Pentatomomorpha (hypothesis (*vi*)) are not in line with the results of the ML tree reconstructions. All three possible unambiguous placements of Nepomorpha are supported by roughly one third of the quartets. However, the placement of Nepomorpha as sister group of the clade including Leptopodomorpha, Cimicomorpha, and Pentatomomorpha is also not fully supported in the ML tree. The placement of Nepomorpha should be addressed again in future studies, using a more specifically designed dataset with a broader taxon sampling in the respective groups.

Testing for non-phylogenetic signal biases via permutation tests. To assess the possible impact of confounding signal, we adopted the permutation approach suggested by Misof et al. (1): the permutation tests were performed using FcLM (see above). Confounding signal could arise from a variety of sources: 1) heterogeneous composition of amino acid sequences (i.e., among-lineage heterogeneity), 2) non-stationary substitution processes, and 3) non-random distribution of missing data on our phylogenetic inferences. All three phenomena are present in our datasets. For more information on the applied permutation schemes and their explanatory power, see Misof et al. (1). FcLM was conducted on the translational level and

using the same partition scheme as before, but using the LG substitution matrix across all partitions for permutation I, II and III. Permutation tests were done for the hypotheses and with the identical taxonomic groups as described above (section 14), using RAxML (versions 8.2.6 and 8.2.8) (32) and ExaML (version 3.0.17) (31). Results of the permutation tests are provided in Table S6.

In none of the permutation tests did any of the three possible unambiguous topologies receive more than 20% support except when testing the sister group relationship of Aphidoidea and Coccoidea (hypothesis *v*). Here, support for the three unambiguous topologies was rather uniformly distributed in the permutation tests, with peaks of 33% (permutation II) and 37% (permutation III). However, this artificial signal supported only topologies incompatible with a sister group relationship of Aphidioidea and Coccoidea (which is supported by both the ML tree and the FcLM analysis on the original data). In summary, we interpret the impact of confounding signal on the inferred phylogeny shown in Figure 1 as minor.

Therefore, as also discussed in the main text, reasons for the observed incongruence between ML tree and FcLM results concerning the placement of Psocodea (hypothesis *i*) are unclear and the underlying phylogenetic question still remains unanswered.

Divergence time estimation. We assigned 23 calibration points throughout the tree (Fig. S1, Table S7). These included four calibration points used by Misof et al. (1) as well as 19 additional fossil calibration points based on best-practice recommendations (37) using only fossils that had an apomorphy for the taxon they represented. Minimum age soft bound (0.025 probability density of age below minimum) uniform priors were applied to these fossil calibration points. For computational efficiency, using the dataset comprising first and second codon positions, we discarded all sites that were not present in at least 95% of the taxa using trimAl (38). The maximum root age (split between Palaeoptera and remaining pterygote insects) was set to be 406 mya (1) for baseml analysis of substitution rates as well as root calibration. Because this maximum age was a point estimate from a prior study, it is not appropriate to use it has a hard bound maximum, so we used a soft bound (0.025 probability density of maximum age exceeding this value) uniform maximum prior.

We used an approximate likelihood calibration by first estimating the branch lengths with maximum likelihood calculation in *baseml* and then estimating divergence times with a GTR + gamma model with Markov Chain Monte Carlo (MCMC). Priors for the lineage specific rate variation *rgene gamma* and *sigma2 gamma* were set to (1, 2.708) and (1, 4.5), respectively, under an independent rates model (39). MCMC approximations of branching times were derived from 10,000 samples taken every 50 cycles after burnin of 50,000 cycles. The parameters were checked for stationarity in Tracer (40). Two independent runs were performed which produced nearly identical estimates.
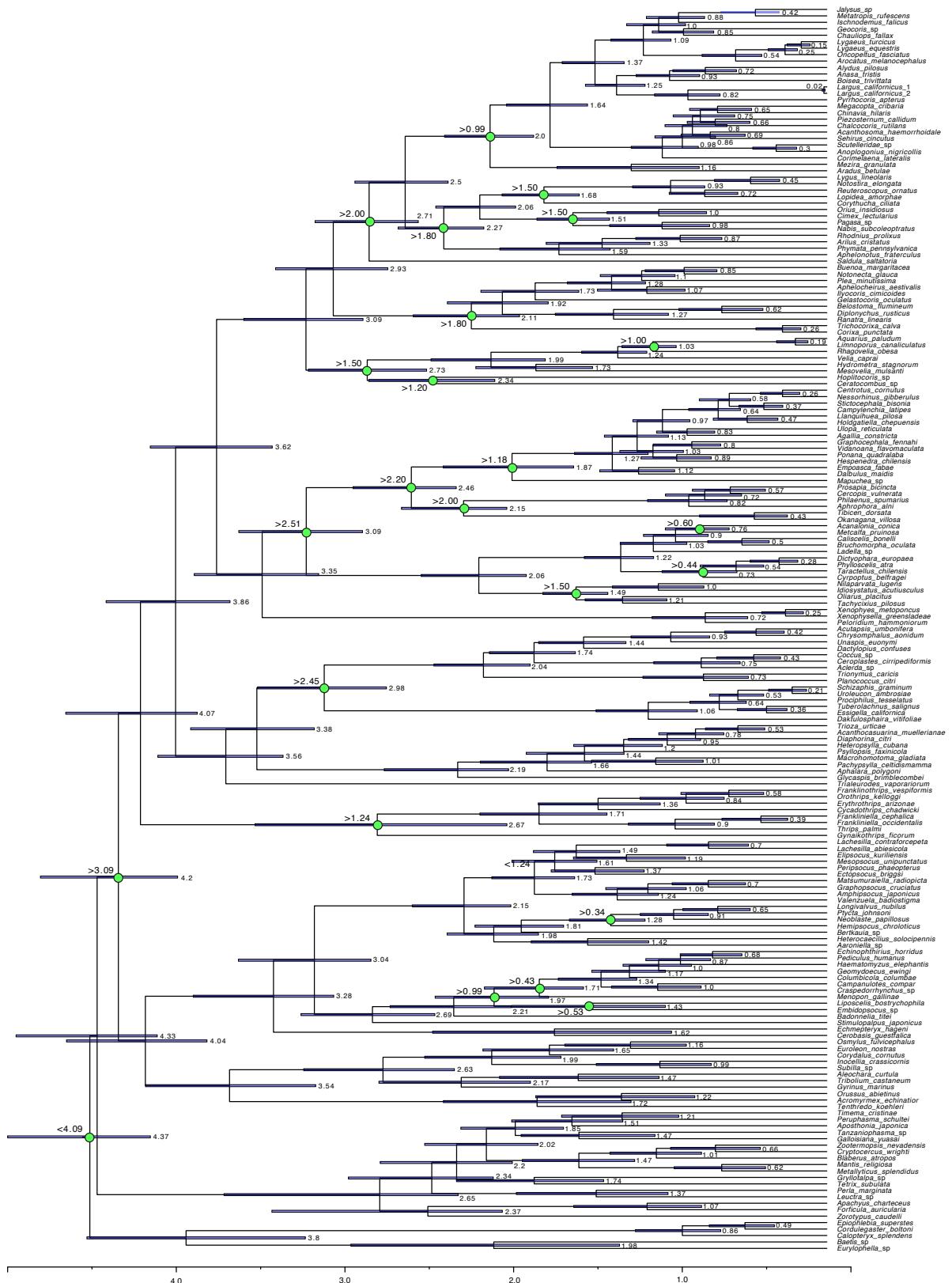
Fig. S1

Dated tree from MCMCTree Bayesian search of first and second position nucleotides filtered to include sites for which 95% of taxa are present (input topology as Fig. 1). Numbers to right of nodes indicate estimate of node age, with 95% confidence intervals noted

10

by light blue bar. Green circles indicate nodes used as calibration points in the analysis, with age of calibration point given to left of node (< is soft maximum, and > is soft minimum). *Indicate calibrations from Misof et al. (1), all other calibration points detailed in Table S7. Scale is in units of 100 million years.
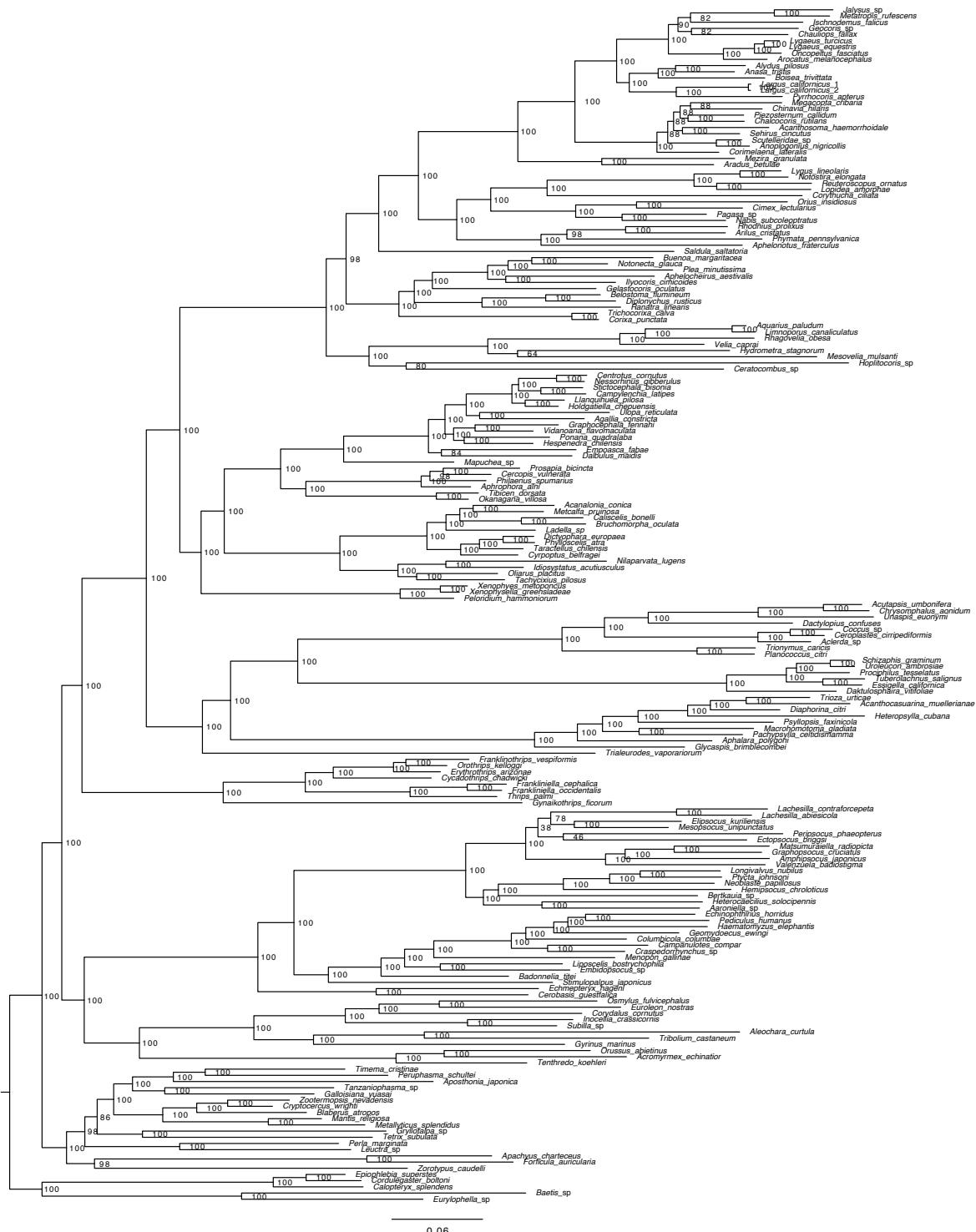
Fig. S2: Tree from ExaML analysis of first and second position nucleotides (see Materials and Methods for details). Numbers indicate support from bootstrap replicates.
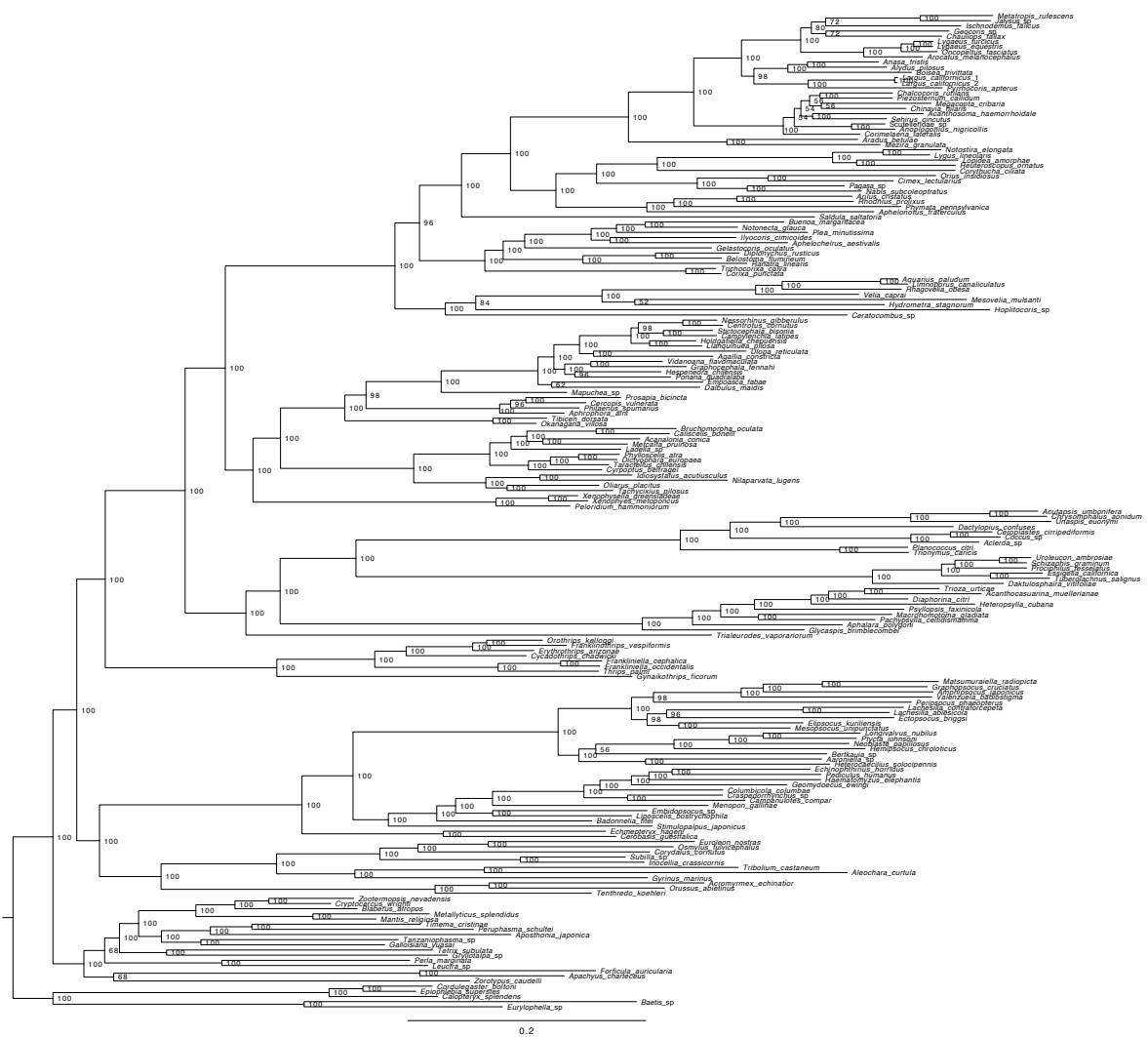
Fig. S3: Tree from ExaML analysis of amino acid sequences (see Materials and Methods for details). Numbers indicate support from bootstrap replicates.
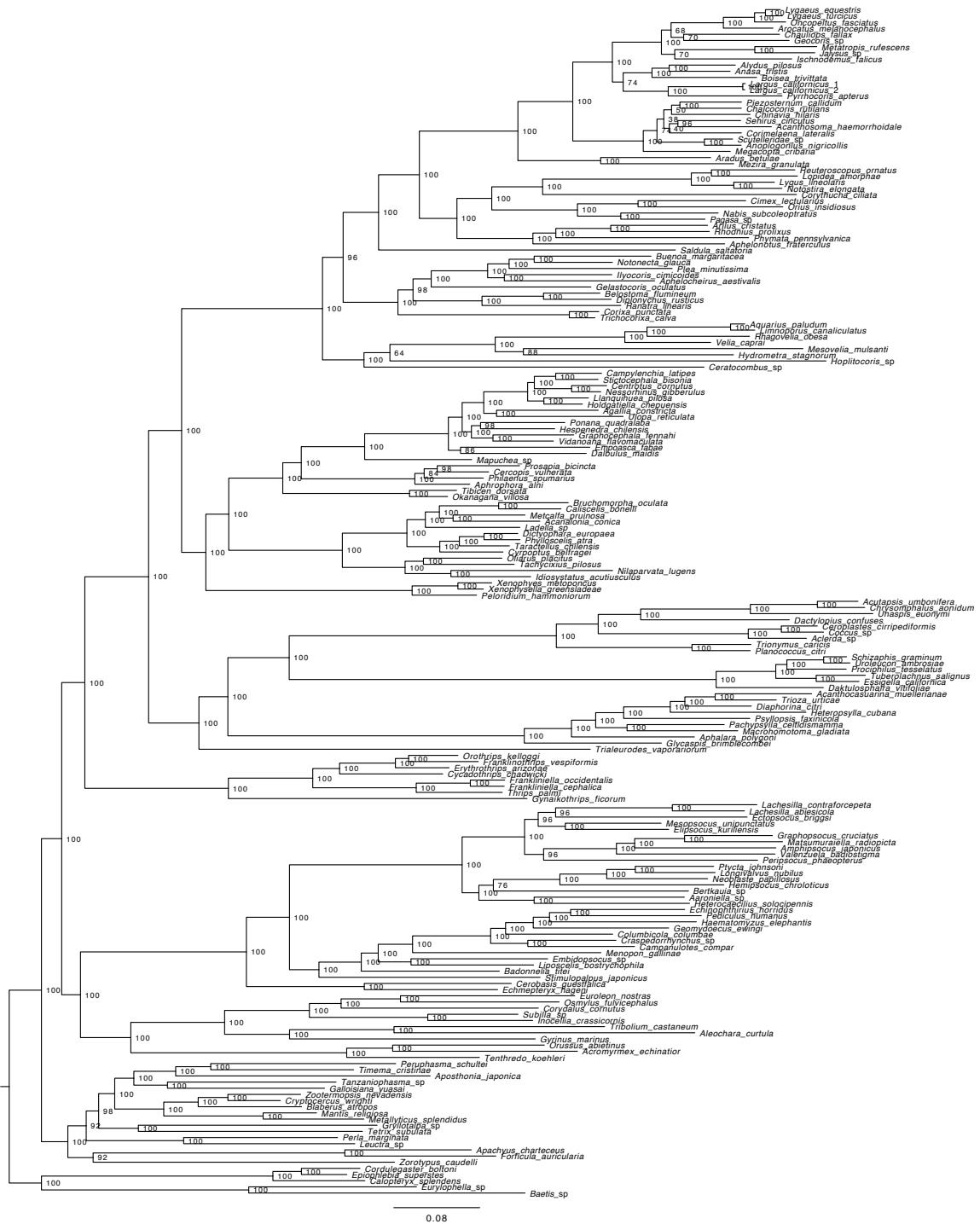
Fig. S4: Phylogenetic tree from ExaML analysis of a "decisive dataset" (see Materials and Methods) for amino acid sequences.

A paraphyletic Paraneoptera

B monophyletic Paraneoptera

C Coleorrhyncha as sistergroup to Heteroptera
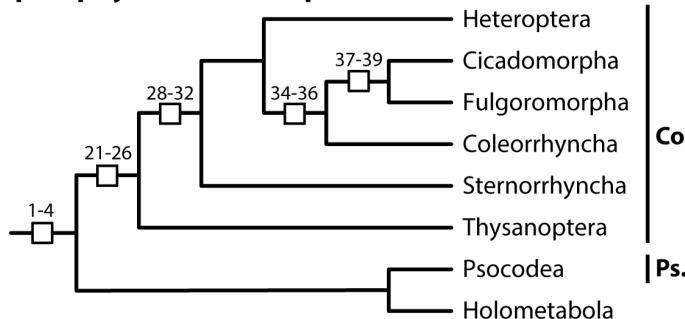
Characters:
1. jugal bar present
2. ocelli of immature stages absent
3. cerci of immature stages absent
4. paired ocelli in nymphs or larvae absent
5. polytrophic ovarioles
6. insertion of lacinia detached from stipes
7. lacinia elongate and stylet-like
8. abdominal ganglia one single mass
9. anterior region of 2. axillary sclerite inflated
10. abdominal sternite 1 absent
11. four or less Malpighian tubules
12. two tarsal segments
13. humeral plate and basisubcostale united with each other
14. anterior region of 2. axillary sclerite swollen
15. posterior median plate located posterodistally to 2. axillary sclerite
16. posterior median plate deeply concave
17. distal margin of posterior median plate sclerotized more strongly than its other regions
18. dorsal median plate divided into two sclerites
19. distal arm of 3. axillary sclerite and dorsal median plate do not articulate with each other
20. basalare and posterior median plate fused with each other
21. mandibles elongated
22. dorsal median plate placed next to 2. axillary sclerite, articulating along a convex hinge
23. proboscis not shifted posteriorly between bases of procoxae
24. anterior tentorial pits shifted dorsally
25. labrum narrowed
26. basisubcostale fused with anteroproximal part of 2. axillary sclerite
27. maxillary palps absent or reduced number of segments
28. cardo strongly reduced or absent
29. labial rostrum present
30. labial palps absent or strongly reduced
31. second axillary sclerite divided into two sclerites
32. labial palp absent
33. tubular labium with three segments
34. posterior median plate reduced, often completely membranous
35. Evan's organ present
36. arolium present
37. complex tymbal acoustic system present
38. aristate antennal flagellum present
39. three tarsal segments
40. tubular labium with four segments
41. four or less antennal flagellomeres
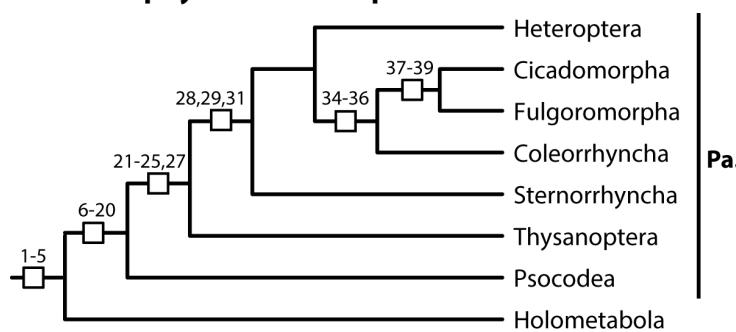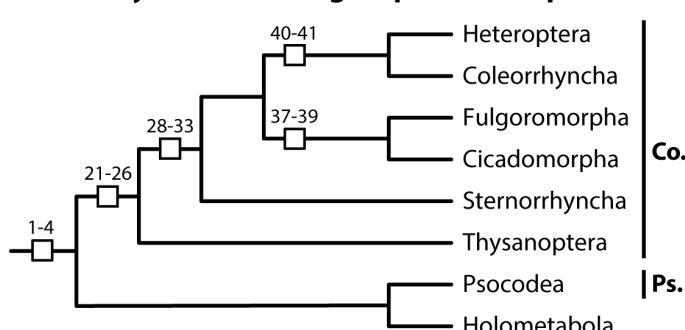
Fig. S5: Maximum parsimony character mapping of 41 key morphological characters across three phylogenetic scenarios for major lineages of Paraneoptera (A-C). Only characters that are synapomorphic for non-terminal clades are shown. Some identical clades have different synapomorphies shown on different trees because some of the character changes were ambiguous (i.e., may have occurred earlier or later) under that particular phylogenetic scenario under parsimony optimization. Note that, in contrast to previous studies, there are three apomorphies supporting Coleorrhyncha + Auchenorrhyncha but only two supporting Coleorrhyncha + Heteroptera. The full morphological data matrix and character list are available in Dryad (doi:10.5061/dryad.t4f4g85).
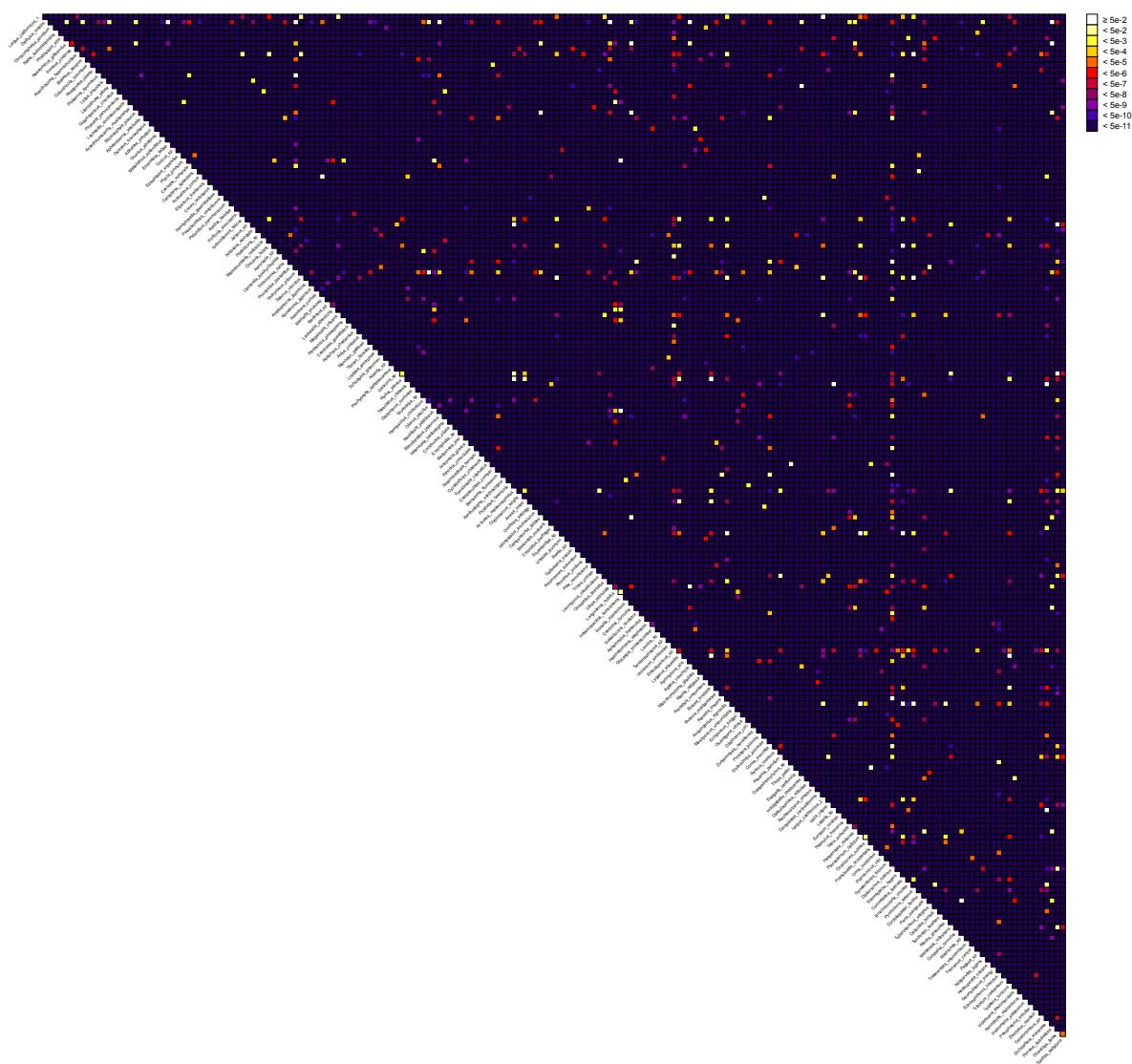
Fig. S6A
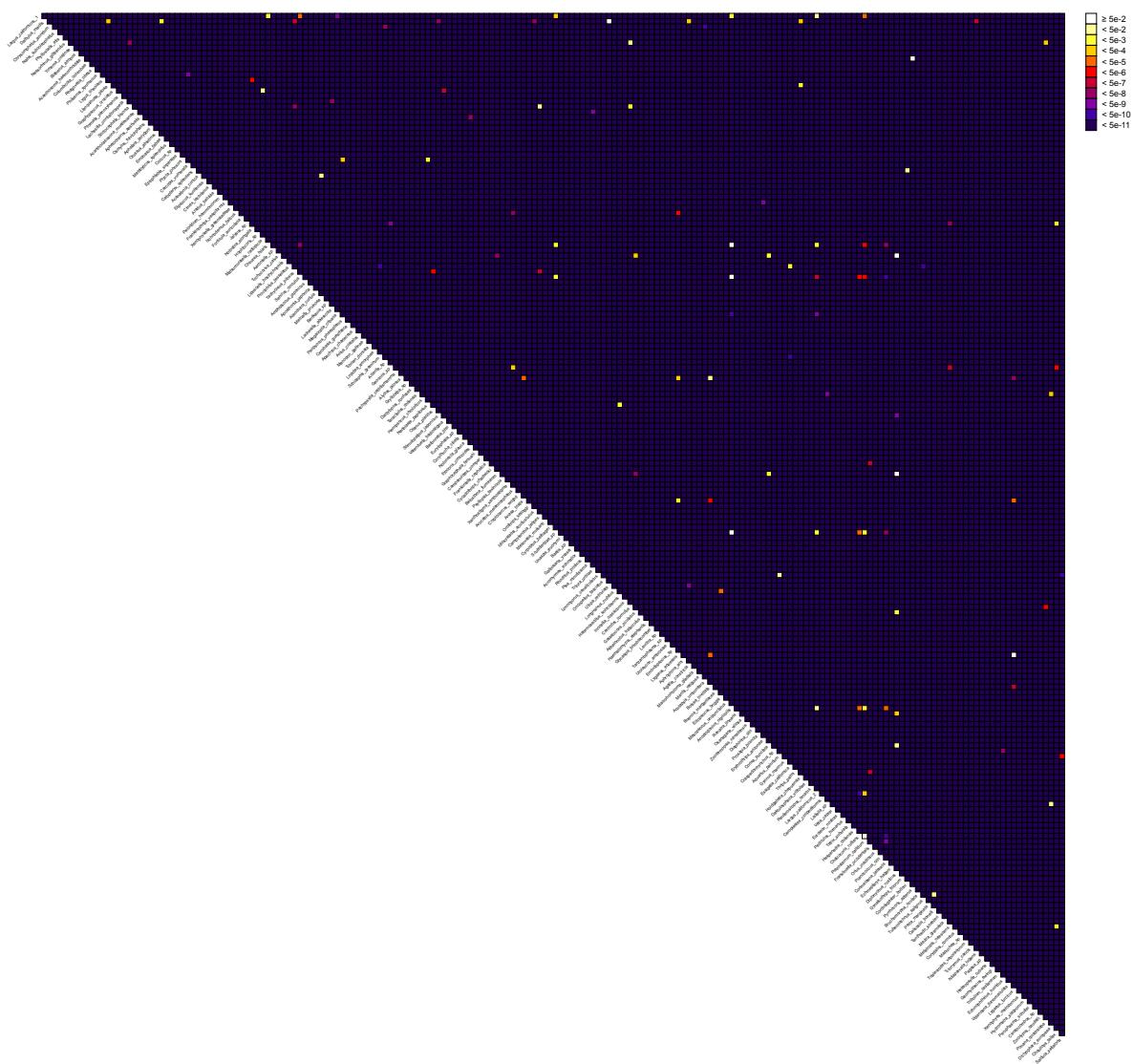
Fig. S6B
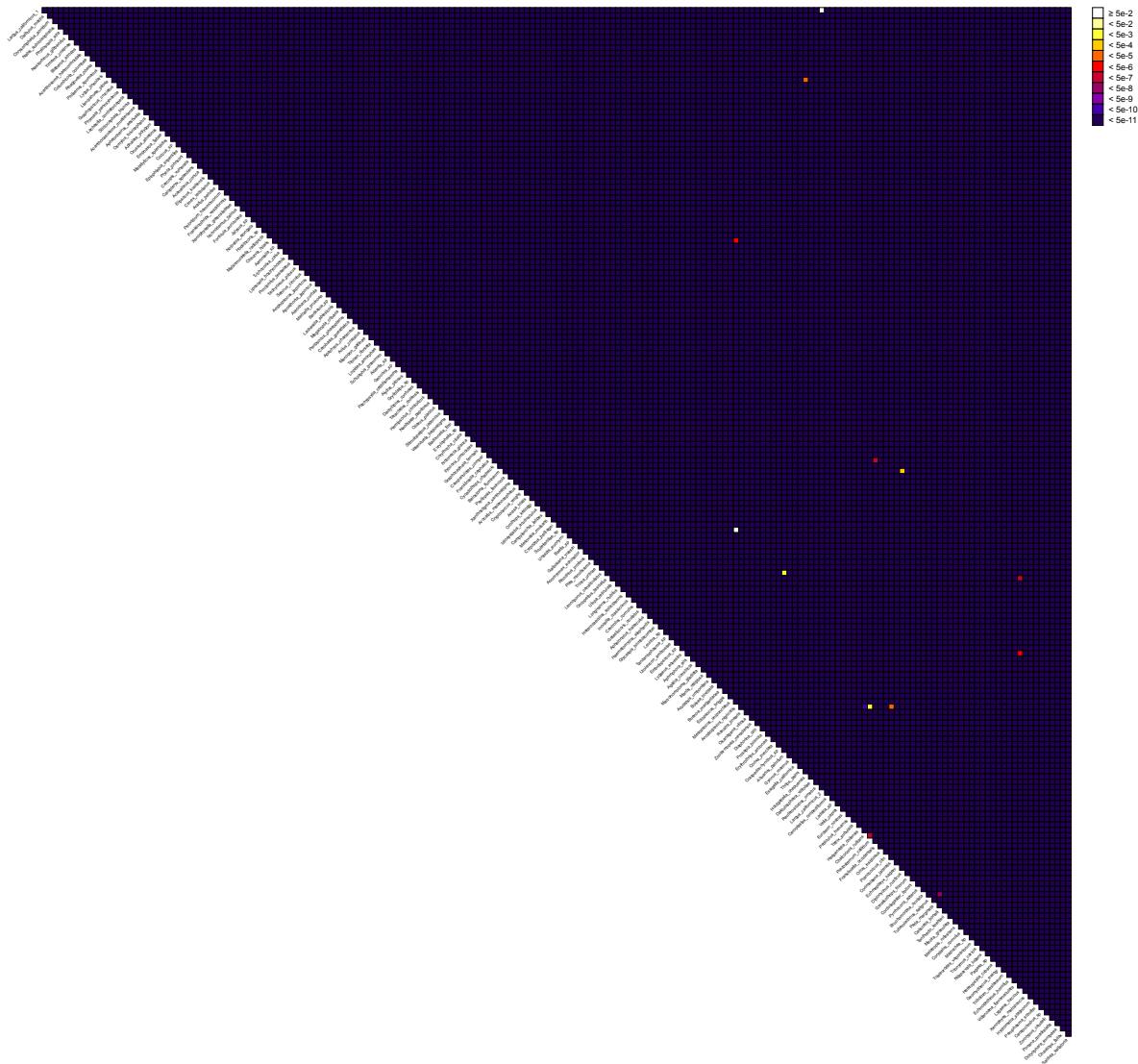
Fig. S6C

Fig. S6 Test for among lineage compositional heterogeneity. Heat map showing pairwise Bowker's tests of the full supermatrix of aligned amino acid sequences (A) and of nucleotide sequences of 1.+2. codon position (B) and of 1.,2.+3. codon position (C). P-values > 0.05 are colored in white and indicate sequence pairs that fully match SRH conditions.

Table S1. NCBI accession numbers of the sequenced and assembled transcriptomes. (*raw data from Misof et al. [1])

Library ID	Species	NCBI Tax ID	BioProject No.	BioSample No.	Experiment No.	Run No.	TSA Project No.	TSA Version
INSobdTACRAAPEI-89	*Aaroniella* sp.	1585309	PRJNA272159	SAMN03331946	SRX1048894	SRR2051465	GDEY00000000	GDEY01000000
INSeqtTCCRAAPEI-13	*Amphipsocus japonicus*	297952	PRJNA272160	SAMN03331947	SRX1048895	SRR2051466	GDDX00000000	GDDX01000000
INSlupTBMRAAPEI-34	*Aphelocheirus aestivalis*	280209	PRJNA272161	SAMN03331948	SRX1048896	SRR2051467	GDEQ00000000	GDEQ01000000
INSlupTAFRAAPEI-44	*Aphrophora alni*	295201	PRJNA272162	SAMN03331949	SRX1048897	SRR2051468	GDEN00000000	GDEN01000000
INSytvTBARAAPEI-94	*Aquarius paludum*	95691	PRJNA272163	SAMN03331950	SRX1048898	SRR2051469	GDFO00000000	GDFO01000000
INSlupTALRAAPEI-62	*Aradus betulae*	1452765	PRJNA272164	SAMN03331951	SRX1048899	SRR2051470	GDFD00000000	GDFD01000000
INSqiqTBURAAPEI-171	*Arocatus melanocephalus*	1561047	PRJNA272165	SAMN03331952	SRX1048900	SRR2051471	GDEZ00000000	GDEZ01000000
INSofmTBCRAAPEI-37	*Badonnelia titei*	297922	PRJNA272166	SAMN03331953	SRX1048901	SRR2051472	GDFG00000000	GDFG01000000
INSobdTABRAAPEI-88	*Bertkauia* sp.	1585311	PRJNA272167	SAMN03331954	SRX1048902	SRR2051473	GDET00000000	GDET01000000
INSqiqTBMRAAPEI-84	*Caliscelis bonellii*	1585281	PRJNA272168	SAMN03331955	SRX1048903	SRR2051474	GDFE00000000	GDFE01000000
INSytvTAQRAAPEI-43	*Centrotus cornutus*	1585357	PRJNA272169	SAMN03331956	SRX1048904	SRR2051475	GDFF00000000	GDFF01000000
INSeqtTCGRAAPEI-17	*Cerobasis guestfalica*	209976	PRJNA272170	SAMN03331957	SRX1048905	SRR2051476	GDEA00000000	GDEA01000000
INSjdsTAFRAAPEI-31	*Cimex lectularius*	79782	PRJNA272171	SAMN03331958	SRX1048906	SRR2051477	GDEK00000000	GDEK01000000
INSlupTAORAAPEI-79	*Corixa punctata*	1545103	PRJNA272172	SAMN03331959	SRX1048907	SRR2051478	GDDR00000000	GDDR01000000
INSbusTBLRAAPEI-77	*Corythucha ciliata*	369451	PRJNA272173	SAMN03331960	SRX1048908	SRR2051479	GDDW00000000	GDDW01000000
INSeqtTDYRAAPEI-20	*Daktulosphaira vitifoliae*	58002	PRJNA272174	SAMN03331961	SRX1048909	SRR2051480	GDEB00000000	GDEB01000000
INSpmbTABRAAPEI-227	*Diaphorina citri*	121845	PRJNA272175	SAMN03331962	SRX1048910	SRR2051481	GDFN00000000	GDFN01000000
INSlupTASRAAPEI-89	*Dictyophara europaea*	1240369	PRJNA272176	SAMN03331963	SRX1048911	SRR2051482	GDEP00000000	GDEP01000000
INSerlTAHRAAPEI-72	*Diplonychus rusticus*	575839	PRJNA272177	SAMN03331964	SRX1048912	SRR2051483	GDED00000000	GDED01000000
INSinlTASRAAPEI-36	*Echinophthirius horridus*	138816	PRJNA272178	SAMN03331965	SRX1048913	SRR2051484	GDEJ00000000	GDEJ01000000
RINSinlTAERACPEI-57	*Elipsocus kuriliensis*	1585294	PRJNA272179	SAMN03331966	SRX1048914	SRR2051485	GDFL00000000	GDFL01000000
INSeqtTCBRAAPEI-12	*Embidopsocus* sp.	1585315	PRJNA272180	SAMN03331967	SRX1048915	SRR2051486	GDFK00000000	GDFK01000000
INSqiqTBZRAAPEI-21	*Franklinothrips vespiformis*	297892	PRJNA272181	SAMN03331968	SRX1048916	SRR2051487	GDFM00000000	GDFM01000000
INSytvTAERAAPEI-14	*Glycaspis brimblecombei*	121847	PRJNA272182	SAMN03331969	SRX1048917	SRR2051488	GDFP00000000	GDFP01000000
INShauTAORAAPEI-8	*Graphocephala fennahi*	1464911	PRJNA272183	SAMN03331970	SRX1048918	SRR2051489	GDEF00000000	GDEF01000000
INSeqtTCERAAPEI-15	*Graphopsocus cruciatus*	239255	PRJNA272184	SAMN03331971	SRX1048919	SRR2051490	GDDY00000000	GDDY01000000
INSinlTAJRAAPEI-75	*Haematomyzus elephantis*	160133	PRJNA272185	SAMN03331972	SRX1048920	SRR2051491	GDEI00000000	GDEI01000000
INSeqtTCFRAAPEI-16	*Hemipsocus chroloticus*	209955	PRJNA272186	SAMN03331973	SRX1048921	SRR2051492	GDDZ00000000	GDDZ01000000
INSobdTADRAAPEI-90	*Heterocaecilius solocipennis*	297992	PRJNA272187	SAMN03331974	SRX1048922	SRR2051493	GDEU00000000	GDEU01000000
INShkeTCCRAAPEI-37	*Hydrometra stagnorum*	446445	PRJNA272188	SAMN03331975	SRX1048923	SRR2051494	GDEG00000000	GDEG01000000
INSlupTBERAAPEI-18	*Ilyocoris cimicoides*	280203	PRJNA272189	SAMN03331976	SRX1048924	SRR2051495	GDEW00000000	GDEW01000000
INSkzdTABRAAPEI-136	*Lachesilla abiesicola*	1585296	PRJNA272191	SAMN03331978	SRX1048926	SRR2051497	GDEL00000000	GDEL01000000
INSobdTAGRAAPEI-109	*Longivalvus nubilus*	469694	PRJNA272192	SAMN03331979	SRX1048927	SRR2051498	GDEV00000000	GDEV01000000
INSqiqTBYRAAPEI-19	*Lygaeus equestris*	696229	PRJNA272193	SAMN03331980	SRX1048928	SRR2051499	GDFJ00000000	GDFJ01000000
INSeqtTCDRAAPEI-14	*Matsumuraiella radiopicta*	297957	PRJNA272194	SAMN03331981	SRX1048929	SRR2051500	GDDQ00000000	GDDQ01000000
INSobdTAERAAPEI-93	*Mesopsocus unipunctatus*	297979	PRJNA272196	SAMN03331983	SRX1048931	SRR2051502	GDDT00000000	GDDT01000000
INSytvTAVRAAPEI-87	*Metatropis rufescens*	1545162	PRJNA272197	SAMN03331984	SRX1048932	SRR2051503	GDFA00000000	GDFA01000000
INSqiqTBERAAPEI-57	*Metcalfa pruinosa*	1185500	PRJNA272198	SAMN03331985	SRX1048933	SRR2051504	GDFH00000000	GDFH01000000
INSobdTAHRAAPEI-118	*Neoblaste papillosa*	1585345	PRJNA272199	SAMN03331986	SRX1048934	SRR2051505	GDFC00000000	GDFC01000000
INShkeTCJRAAPEI-41	*Notonecta glauca*	260537	PRJNA272200	SAMN03331987	SRX1048935	SRR2051506	GDEH00000000	GDEH01000000
INSobdTAFRAAPEI-95	*Peripsocus phaeopterus*	1585301	PRJNA272201	SAMN03331988	SRX1048936	SRR2051507	GDDU00000000	GDDU01000000
INSnfrTABRAAPEI-14	*Plea minutissima*	280219	PRJNA272202	SAMN03331989	SRX1048937	SRR2051508	GDES00000000	GDES01000000
INSbusTBJRAAPEI-85	*Pyrrhocoris apterus*	37000	PRJNA272203	SAMN03331990	SRX1048938	SRR2051509	GDFI00000000	GDFI01000000
INSlupTBPRAAPEI-37	*Saldula saltatoria*	280090	PRJNA272204	SAMN03331991	SRX1048939	SRR2051510	GDER00000000	GDER01000000
INSobdTAARAAPEI-87	*Stimulopalpus japonicus*	209965	PRJNA272205	SAMN03331992	SRX1048940	SRR2051511	GDDS00000000	GDDS01000000
INSlupTBKRAAPEI-31	*Tachycixius pilosus*	166225	PRJNA272206	SAMN03331993	SRX1048941	SRR2051512	GDEX00000000	GDEX01000000
INSlupTAKRAAPEI-57	*Ulopa reticulata*	139497	PRJNA272207	SAMN03331994	SRX1048942	SRR2051513	GDEO00000000	GDEO01000000
INSytvTCGRAAPEI-39	*Valenzuela badiostigma*	1407767	PRJNA272208	SAMN03331995	SRX1048943	SRR2051514	GDFB00000000	GDFB01000000
INSjdsTBFRAAPEI-57	*Xenophyes metoponcus*	1585308	PRJNA272209	SAMN03331996	SRX1048944	SRR2051515	GDEM00000000	GDEM01000000
Accon	*Acanalonia conica*	1585284	PRJNA272210	SAMN03341923	SRX893485	SRR1821891	GCXC00000000	GCXC01000000
AcspC	*Aclerda* sp.	1585310	PRJNA272211	SAMN03341924	SRX893486	SRR1821892	GCWU00000000	GCWU01000000
Acumb	*Acutaspis umbonifera*	1585280	PRJNA272212	SAMN03341925	SRX893487	SRR1821893	GCXJ00000000	GCXJ01000000
Agcon	*Agallia constricta*	1585285	PRJNA272213	SAMN03341926	SRX893488	SRR1821894	GCWT00000000	GCWT01000000
Alpil	*Alydus pilosus*	1585286	PRJNA272214	SAMN03341927	SRX893489	SRR1821895	GCVZ00000000	GCVZ01000000
AntriUL	*Anasa tristis*	236421	PRJNA272215	SAMN03341928	SRX893490	SRR1821896	GCWG00000000	GCWG01000000
ScspS2	*Anoplogonius nigricollis*	1585353	PRJNA272216	SAMN03341929	SRX893562	SRR1821968	GCYN00000000	GCYN01000000
Appol	*Aphalara polygoni*	1585287	PRJNA272217	SAMN03341930	SRX893491	SRR1821897	GCWO00000000	GCWO01000000
ApspP	*Aphelonotus fraterculus*	1585355	PRJNA272218	SAMN03341931	SRX893492	SRR1821898	GCWA00000000	GCWA01000000
Arcri	*Arilus cristatus*	72517	PRJNA272219	SAMN03341932	SRX893493	SRR1821899	GCWE00000000	GCWE01000000
Beflu	*Belostoma flumineum*	236416	PRJNA272220	SAMN03341933	SRX893494	SRR1821900	GCWK00000000	GCWK01000000
Botri	*Boisea trivittata*	1255142	PRJNA272221	SAMN03341934	SRX893495	SRR1821901	GCWC00000000	GCWC01000000
Brocu	*Bruchomorpha oculata*	130607	PRJNA272222	SAMN03341935	SRX893496	SRR1821902	GCWF00000000	GCWF01000000
BuspN	*Buenoa margaritacea*	1585289	PRJNA272223	SAMN03341936	SRX893497	SRR1821903	GCXE00000000	GCXE01000000
Cacom	*Campanulotes compar*	135595	PRJNA272225	SAMN03341937	SRX893577	SRR1821983	GCWD00000000	GCWD01000000
Calat	*Campylenchia latipes*	54654	PRJNA272226	SAMN03341938	SRX893498	SRR1821904	GCWI00000000	GCWI01000000
CespC	*Ceratocombus* sp.	1585312	PRJNA272227	SAMN03341939	SRX893500	SRR1821906	GCWS00000000	GCWS01000000
Cecir	*Ceroplastes cirripediformis*	1585290	PRJNA272228	SAMN03341940	SRX893499	SRR1821905	GCWZ00000000	GCWZ01000000
TespT	*Chalcocoris rutilans*	1592895	PRJNA272224	SAMN03341941	SRX893565	SRR1821971	GCYS00000000	GCYS01000000
McspM	*Chauliops fallax*	1244200	PRJNA272229	SAMN03341942	SRX893532	SRR1821938	GCXK00000000	GCXK01000000
Chhil	*Chinavia hilaris*	244443	PRJNA272230	SAMN03341943	SRX893503	SRR1821909	GCWR00000000	GCWR01000000
DispD1	*Chondrodire chilensis*	1585349	PRJNA272294	SAMN03341944	SRX893509	SRR1821915	GCYG00000000	GCYG01000000
Chaon	*Chrysomphalus aonidum*	340536	PRJNA272231	SAMN03341945	SRX893501	SRR1821907	GCVU00000000	GCVU01000000
CospC	*Coccus* sp.	1585313	PRJNA272232	SAMN03341946	SRX893505	SRR1821911	GCWW00000000	GCWW01000000
Cocol	*Columbicola columbae*	128991	PRJNA272233	SAMN03341947	SRX893578	SRR1821984	GCWB00000000	GCWB01000000
Colat	*Corimelaena lateralis*	1585293	PRJNA272234	SAMN03341948	SRX893504	SRR1821910	GCXA00000000	GCXA01000000
Crspl	*Craspedorrhynchus* sp.	1585369	PRJNA272235	SAMN03341949	SRX893506	SRR1821912	GCWN00000000	GCWN01000000
Chcha2	*Cycadothrips chadwicki*	1291257	PRJNA272236	SAMN03341950	SRX893502	SRR1821908	GCZM00000000	GCZM01000000

Cybel	*Cyrpoptus belfragei*	1464866	PRJNA272237	SAMN03341951	SRX893507	SRR1821913	GCWQ00000000	GCWQ01000000
Dacon	*Dactylopius confusus*	38117	PRJNA272238	SAMN03341952	SRX893508	SRR1821914	GCWV00000000	GCWV01000000
Damai	*Dalbulus maidis*	74065	PRJNA272239	SAMN03341953	SRX893575	SRR1821981	GCWP00000000	GCWP01000000
Echag	*Echmepteryx hageni*	209971	PRJNA272240	SAMN03341954	SRX893576	SRR1821982	GCWY00000000	GCWY01000000
Emfab	*Empoasca fabae*	139595	PRJNA272241	SAMN03341955	SRX893511	SRR1821917	GCVV00000000	GCVV01000000
Thsp	*Frankliniella occidentalis*	133901	PRJNA272242	SAMN03341956	SRX893566	SRR1821972	GCYR00000000	GCYR01000000
Geocu	*Gelastocoris oculatus*	38104	PRJNA272243	SAMN03341957	SRX893514	SRR1821920	GCXI00000000	GCXI01000000
GespG	*Geocoris* sp.	1585316	PRJNA272244	SAMN03341958	SRX893515	SRR1821921	GCWM00000000	GCWM01000000
Geewi	*Geomydoecus ewingi*	35641	PRJNA272245	SAMN03341959	SRX893513	SRR1821919	GCXD00000000	GCXD01000000
Hechi	*Hespenedra chilensis*	1585359	PRJNA272247	SAMN03341961	SRX893516	SRR1821922	GCXG00000000	GCXG01000000
Hecub	*Heteropsylla cubana*	121849	PRJNA272248	SAMN03341962	SRX893517	SRR1821923	GCXB00000000	GCXB01000000
Hoche	*Holdgatiella chepuensis*	1585361	PRJNA272249	SAMN03341963	SRX893518	SRR1821924	GCXW00000000	GCXW01000000
EnspE	*Hoplitocoris* sp.	1585371	PRJNA272250	SAMN03341964	SRX893512	SRR1821918	GCWH00000000	GCWH01000000
PespD	*Idiosystatus acutiusculus*	1585363	PRJNA272251	SAMN03341965	SRX893544	SRR1821950	GCXZ00000000	GCXZ01000000
Isfal	*Ischnodemus falicus*	881479	PRJNA272252	SAMN03341966	SRX893519	SRR1821925	GCXH00000000	GCXH01000000
JaspB	*Jalysus* sp.	1585318	PRJNA272253	SAMN03341967	SRX893520	SRR1821926	GCXL00000000	GCXL01000000
Lacon	*Lachesilla contraforcepeta*	1585297	PRJNA272254	SAMN03341968	SRX893521	SRR1821927	GCWJ00000000	GCWJ01000000
Lasp	*Ladella* sp.	1585373	PRJNA272255	SAMN03341969	SRX893522	SRR1821928	GCXO00000000	GCXO01000000
LrspL1	*Largus californicus*	1585298	PRJNA272256	SAMN03341970	SRX893526	SRR1821932	GCXX00000000	GCXX01000000
LrspL2	*Largus californicus*	1585298	PRJNA272803	SAMN03341971	SRX893527	SRR1821933	GCXR00000000	GCXR01000000
Lican	*Limnoporus canaliculatus*	56088	PRJNA272257	SAMN03341972	SRX893523	SRR1821929	GCYA00000000	GCYA01000000
Llpil	*Llanquihuea pilosa*	1585299	PRJNA272258	SAMN03341973	SRX893524	SRR1821930	GCWX00000000	GCWX01000000
LospM	*Lopidea amorphae*	1585300	PRJNA272259	SAMN03341974	SRX893525	SRR1821931	GCXF00000000	GCXF01000000
Lytur	*Lygaeus turcicus*	1291711	PRJNA272260	SAMN03341975	SRX893529	SRR1821935	GCYB00000000	GCYB01000000
Lylin	*Lygus lineolaris*	50650	PRJNA272261	SAMN03341976	SRX893528	SRR1821934	GCXM00000000	GCXM01000000
Magla	*Macrohomotoma gladiata*	1585365	PRJNA272262	SAMN03341977	SRX893530	SRR1821936	GCXQ00000000	GCXQ01000000
MaspM8	*Mapuchea* sp.	1585375	PRJNA272263	SAMN03341978	SRX893531	SRR1821937	GCXN00000000	GCXN01000000
PlspP	*Megacopta cribraria*	299257	PRJNA272264	SAMN03341979	SRX893550	SRR1821956	GCXV00000000	GCXV01000000
Olpla	*Melanoliarus placitus*	491322	PRJNA272269	SAMN03341980	SRX893537	SRR1821943	GCZE00000000	GCZE01000000
Memul	*Mesovelia mulsanti*	236398	PRJNA272265	SAMN03341981	SRX893534	SRR1821940	GCXS00000000	GCXS01000000
Megra	*Mezira granulata*	236326	PRJNA272266	SAMN03341982	SRX893533	SRR1821939	GCYM00000000	GCYM01000000
Nasub	*Nabis subcoleoptratus*	1594575	PRJNA272267	SAMN03341983	SRX893535	SRR1821941	GCVW00000000	GCVW01000000
Tidor	*Neotibicen dorsatus*	1464820	PRJNA272295	SAMN03342009	SRX893567	SRR1821973	GCYV00000000	GCYV01000000
Negib	*Nessorhinus gibberulus*	1594574	PRJNA272268	SAMN03341984	SRX893536	SRR1821942	GCYF00000000	GCYF01000000
Onfas	*Oncopeltus fasciatus*	7536	PRJNA272270	SAMN03341985	SRX893538	SRR1821944	GCXY00000000	GCXY01000000
Orins	*Orius insidiosus*	83647	PRJNA272271	SAMN03341986	SRX893539	SRR1821945	GCYL00000000	GCYL01000000
Orkel	*Orothrips kelloggi*	1291281	PRJNA272273	SAMN03341988	SRX893540	SRR1821946	GCXT00000000	GCXT01000000
Pacel	*Pachypsylla celtidismamma*	121851	PRJNA272274	SAMN03341989	SRX893542	SRR1821948	GCYC00000000	GCYC01000000
PespN	*Pagasa* sp.	1592893	PRJNA272275	SAMN03341990	SRX893545	SRR1821951	GCXU00000000	GCXU01000000
Peham	*Peloridium pomponorum*	1172124	PRJNA272276	SAMN03341991	SRX893543	SRR1821949	GCZG00000000	GCZG01000000
Phspu	*Philaenus spumarius*	36667	PRJNA272277	SAMN03341992	SRX893549	SRR1821955	GCZA00000000	GCZA01000000
PhspP	*Phlaeothripidae* sp.	1585319	PRJNA272278	SAMN03341993	SRX893548	SRR1821954	GCYQ00000000	GCYQ01000000
Phatr	*Phylloscelis atra*	1465156	PRJNA272279	SAMN03341994	SRX893546	SRR1821952	GCYH00000000	GCYH01000000
Phpen	*Phymata pennsylvanica*	237174	PRJNA272280	SAMN03341995	SRX893547	SRR1821953	GCYI00000000	GCYI01000000
DispD2	*Piezosternum calidum*	1585282	PRJNA272281	SAMN03341996	SRX893510	SRR1821916	GCWL00000000	GCWL01000000
Poqua	*Ponana quadralaba*	1465165	PRJNA272282	SAMN03341997	SRX893551	SRR1821957	GCZF00000000	GCZF01000000
Prtes	*Prociphilus tesselatus*	1585303	PRJNA272283	SAMN03341998	SRX893553	SRR1821959	GCZH00000000	GCZH01000000
Prbic	*Prosapia bicincta*	139605	PRJNA272284	SAMN03341999	SRX893552	SRR1821958	GCYJ00000000	GCYJ01000000
Psfax	*Psyllopsis fraxinicola*	1585347	PRJNA272285	SAMN03342000	SRX893554	SRR1821960	GCYW00000000	GCYW01000000
Ptjoh	*Ptycta johnsoni*	469707	PRJNA272286	SAMN03342001	SRX893556	SRR1821962	GCYD00000000	GCYD01000000
RespM	*Reuteroscopus ornatus*	1010780	PRJNA272288	SAMN03342003	SRX893558	SRR1821964	GCYK00000000	GCYK01000000
Rhobe	*Rhagovelia obesa*	1190598	PRJNA272289	SAMN03342004	SRX893559	SRR1821965	GCYP00000000	GCYP01000000
Scgra	*Schizaphis graminum*	13262	PRJNA272290	SAMN03342005	SRX893560	SRR1821966	GCZI00000000	GCZI01000000
ScspS1	*Scutelleridae* sp.	1585322	PRJNA272291	SAMN03342006	SRX893561	SRR1821967	GCYE00000000	GCYE01000000
Secin	*Sehirus cinctus*	1585283	PRJNA272292	SAMN03342007	SRX893563	SRR1821969	GCZB00000000	GCZB01000000
Stbis	*Stictocephala bisonia*	1585304	PRJNA272293	SAMN03342008	SRX893564	SRR1821970	GCZK00000000	GCZK01000000
TrspC	*Trichocorixa calva*	1585305	PRJNA272296	SAMN03342010	SRX893566	SRR1821975	GCYZ00000000	GCYZ01000000
Trcar	*Trionymus caricis*	1585306	PRJNA272297	SAMN03342011	SRX893568	SRR1821974	GCYT00000000	GCYT01000000
Trurt	*Trioza urticae*	121826	PRJNA272298	SAMN03342012	SRX893570	SRR1821976	GCYX00000000	GCYX01000000
Tusal	*Tuberolachnus salignus*	96551	PRJNA272299	SAMN03342013	SRX893571	SRR1821977	GCYY00000000	GCYY01000000
Uneuo	*Unaspis euonymi*	340633	PRJNA272300	SAMN03342014	SRX893572	SRR1821978	GCZJ00000000	GCZJ01000000
Uramb	*Uroleucon ambrosiae*	87311	PRJNA272301	SAMN03342015	SRX893573	SRR1821979	GCZC00000000	GCZC01000000
Vifla	*Vidanoana flavomaculata*	1585307	PRJNA272302	SAMN03342016	SRX893574	SRR1821980	GCZD00000000	GCZD01000000
INSytvTALRAAPEI-35*	*Acanthocasuarina muelleriana*	1323547	PRJNA219519	SAMN02047100	SRX314816	SRR921560	GAYY00000000	GAYY02000000
INSnfrTAPRAAPEI-36*	*Acanthosoma haemorrhoidale*	483950	PRJNA219520	SAMN02047154	SRX314817	SRR921561	GAUV00000000	GAUV02000000
INShauTBERAAPEI-33*	*Aleochara curtula*	135881	PRJNA219522	SAMN02047128	SRX314819	SRR921563	GATW00000000	GATW02000000
INSfrgTALRAAPEI-22*	*Apachyus charteceus*	1323532	PRJNA219524	SAMN02047175	SRX314821	SRR921565	GAUW00000000	GAUW02000000
INSfrgTAZRAAPEI-46*	*Aposthonia japonica*	911381	PRJNA219525	SAMN02047170	SRX314822	SRR921566	GAWU00000000	GAWU02000000
INShauTAKRAAPEI-90*	*Baetis* sp.	1476881	PRJNA219528	SAMN02047149	SRX314825	SRR921569	GATU00000000	GATU02000000
INSfrgTAVRAAPEI-41*	*Blaberus atropos*	36942	PRJNA219531	SAMN02047121	SRX314828	SRR921572	GAYD00000000	GAYD02000000
INStmbTAARAAPEI-84*	*Calopteryx splendens*	52612	PRJNA219534	SAMN02047184	SRX314831	SRR921575	GAYM00000000	GAYM02000000
INSnfrTANRAAPEI-34*	*Cercopis vulnerata*	279269	PRJNA219537	SAMN02047155	SRX314834	SRR921578	GAUN00000000	GAUN02000000
INSjdsTBHRAAPEI-74*	*Cordulegaster boltoni*	126173	PRJNA219542	SAMN02047156	SRX314839	SRR921583	GAYO00000000	GAYO02000000
INSbttTKRAAPEI-18*	*Corydalus cornutus*	559164	PRJNA219543	SAMN02047201	SRX314840	SRR921584	GATG00000000	GATG02000000
INSytvTCDRAAPEI-35*	*Cryptocercus wrighti*	89837	PRJNA219546	SAMN02047199	SRX314843	SRR921587	GAZN00000000	GAZN02000000
INSytvTCFRAAPEI-37*	*Ectopsocus briggsi*	322492	PRJNA219550	SAMN01801569	SRX215455	SRR645929	GAPT00000000	GAPT02000000
INSfrgTAHRAAPEI-18*	*Epiophlebia superstes*	126247	PRJNA219553	SAMN02047171	SRX314846	SRR921592	GAVW00000000	GAVW02000000
INSytvTBHRAAPEI-14*	*Essigella californica*	759921	PRJNA219554	SAMN02047099	SRX314848	SRR921594	GAZF00000000	GAZF02000000

INSjdsTATRAAPEI-57*	*Euroleon nostras*	516507	PRJNA219555	SAMN02047165	SRX314849	SRR921595	GAXW00000000	GAXW02000000
INSytvTCERAAPEI-36*	*Eurylophella* sp.	1323562	PRJNA219556	SAMN02047200	SRX314850	SRR921596	GAZG00000000	GAZG02000000
INSjdsTBNRAAPEI-89*	*Forficula auricularia*	13068	PRJNA219558	SAMN02047143	SRX314852	SRR921598	GAYQ00000000	GAYQ02000000
INSjdsTABRAAPEI-20*	*Frankliniella cephalica*	407008	PRJNA219559	SAMN02047110	SRX314853	SRR921599	GAYE00000000	GAYE02000000
INSfrgTAKRAAPEI-21*	*Galloisiana yuasai*	378494	PRJNA219560	SAMN02047172	SRX314854	SRR921600	GAWN00000000	GAWN02000000
INSfrgTAXRABPEI-44*	*Gryllotalpa* sp.	1323563	PRJNA219562	SAMN02047167	SRX314856	SRR921602	GAWZ00000000	GAWZ02000000
INSjdsTADRAAPEI-22*	*Gynaikothrips ficorum*	59752	PRJNA219563	SAMN02047111	SRX314857	SRR921603	GAXG00000000	GAXG02000000
INSnfrTBERAAPEI-19*	*Gyrinus marinus*	1323538	PRJNA219564	SAMN02047132	SRX314858	SRR921604	GAUY00000000	GAUY02000000
INSytvTCFRAAPEI-43*	*Inocellia crassicornis*	279416	PRJNA219567	SAMN02047103	SRX314861	SRR921607	GAZH00000000	GAZH02000000
INSnfrTALRAAPEI-31*	*Leuctra* sp.	1323566	PRJNA219571	SAMN02047153	SRX314865	SRR921611	GAUF00000000	GAUF02000000
INStmbTBGRAAPEI-33*	*Liposcelis bostrychophila*	185214	PRJNA219573	SAMN02047187	SRX314867	SRR921613	GAYV00000000	GAYV02000000
INShauTAARAAPEI-90*	*Mantis religiosa*	7507	PRJNA219575	SAMN02047157	SRX314869	SRR921615	GASW00000000	GASW02000000
INSfrgTAFRAAPEI-31*	*Menopon gallinae*	328185	PRJNA219579	SAMN02047141	SRX314873	SRR921619	GAWR00000000	GAWR02000000
INShauTAMRAAPEI-94*	*Metallyticus splendidus*	627751	PRJNA219580	SAMN02047174	SRX314874	SRR921620	GATB00000000	GATB02000000
INSfrgTBCRAAPEI-57*	*Nilaparvata lugens*	108931	PRJNA219582	SAMN02047185	SRX314876	SRR921622	GAYF00000000	GAYF02000000
INShauTAPRAAPEI-9*	*Notostira elongata*	537464	PRJNA219583	SAMN02047151	SRX314877	SRR921623	GASV00000000	GASV02000000
INSnfrTBLRAAPEI-11*	*Okanagana villosa*	1323542	PRJNA219585	SAMN02047193	SRX314879	SRR921625	GAWQ00000000	GAWQ02000000
INSnfrTAARAAPEI-13*	*Orussus abietinus*	222816	PRJNA219586	SAMN02047118	SRX314880	SRR921626	GAUJ00000000	GAUJ02000000
INSjdsTBJRAAPEI-79*	*Osmylus fulvicephalus*	446453	PRJNA219587	SAMN02047166	SRX314881	SRR921627	GAYC00000000	GAYC02000000
INShauTALRAAPEI-93*	*Perla marginata*	227363	PRJNA219591	SAMN02047115	SRX314885	SRR921631	GATV00000000	GATV02000000
INSfrgTAORAAPEI-33*	*Peruphasma schultei*	614134	PRJNA219592	SAMN02047114	SRX314886	SRR921632	GAWJ00000000	GAWJ02000000
INSjdsTAIRAAPEI-35*	*Planococcus citri*	170843	PRJNA219593	SAMN02047127	SRX314887	SRR921633	GAXF00000000	GAXF02000000
INSytvTANRAAPEI-37*	*Ranatra linearis*	280158	PRJNA219599	SAMN02047138	SRX314893	SRR921639	GAYZ00000000	GAYZ02000000
INSnfrTAGRAAPEI-19*	*Subilla* sp.	2054904	PRJNA219617	SAMN02047106	SRX314911	SRR921657	GAUI00000000	GAUI02000000
INSfrgTBBRAAPEI-56*	*Tanzaniophasma* sp.	1323568	PRJNA219604	SAMN02047176	SRX314898	SRR921644	GAXB00000000	GAXB02000000
INSfrgTATRAAPEI-37*	*Tenthredo koehleri*	1323543	PRJNA219605	SAMN02047159	SRX314899	SRR921645	GAWW00000000	GAWW02000000
INShauTANRAAPEI-95*	*Tetrix subulata*	288127	PRJNA219606	SAMN02047150	SRX314900	SRR921646	GASQ00000000	GASQ02000000
INSjdsTACRAAPEI-21*	*Thrips palmi*	161013	PRJNA219609	SAMN02047112	SRX314903	SRR921649	GAXC00000000	GAXC02000000
INSnfrTBPRAAPEI-15*	*Timema cristinae*	61476	PRJNA219610	SAMN02047191	SRX314904	SRR921650	GAVX00000000	GAVX02000000
INSfrgTAPRAAPEI-33*	*Trialeurodes vaporariorum*	88556	PRJNA219611	SAMN02047126	SRX314905	SRR921651	GAWX00000000	GAWX02000000
INSnfrTAORAAPEI-35*	*Velia caprai*	301293	PRJNA219616	SAMN02047131	SRX314910	SRR921656	GAUO00000000	GAUO02000000
INSjdsTARRAAPEI-47*	*Xenophysella greensladeae*	1254495	PRJNA219618	SAMN02047181	SRX314912	SRR921658	GAYI00000000	GAYI02000000
INSjdsTAQRAAPEI-46*	*Zorotypus caudelli*	1323545	PRJNA219620	SAMN02047173	SRX314914	SRR921660	GAYA00000000	GAYA02000000

Table S2. Sequences removed during various contamination filtering steps.

Library-ID	Species	No. seq. after assembly	No. seq. omitted after cross-contam. screen	No. seq. filtered by NCBI	No. seq. published	Local UniVec version
INSobdTACRAAPEI-89	*Aaroniella* sp.	27,402	239	16	27,147	7.1
INSeqtTCCRAAPEI-13	*Amphipsocus japonicus*	29,615	412	1	29,202	7.1
INSlupTBMRAAPEI-34	*Aphelocheirus aestivalis*	40,934	316	5	40,613	7.1
INSlupTAFRAAPEI-44	*Aphrophora alni*	40,522	544	16	39,962	7.1
INSytvTBARAAPEI-94	*Aquarius paludum*	71,669	163	3	71,503	7.1
INSlupTALRAAPEI-62	*Aradus betulae*	55,623	62	2	55,559	7.1
INSqiqTBURAAPEI-171	*Arocatus melanocephalus*	45,590	387	47	45,156	7.1
INSofmTBCRAAPEI-37	*Badonnelia titei*	33,779	186	10	33,583	7.1
INSobdTABRAAPEI-88	*Bertkauia* sp.	27,453	177	1	27,275	7.1
INSqiqTBMRAAPEI-84	*Caliscelis bonellii*	32,529	135	7	32,387	7.1
INSytvTAQRAAPEI-43	*Centrotus cornutus*	40,730	201	20	40,509	7.1
INSeqtTCGRAAPEI-17	*Cerobasis guestfalica*	51,021	167	4	50,850	7.1
INSjdsTAFRAAPEI-31	*Cimex lectularius*	33,465	74	1	33,390	7.1
INSlupTAORAAPEI-79	*Corixa punctata*	22,917	77	0	22,840	7.1
INSbusTBLRAAPEI-77	*Corythucha ciliata*	19,467	1,477	5	17,985	7.1
INSeqtTDYRAAPEI-20	*Daktulosphaira vitifoliae*	26,964	450	9	26,505	7.1
INSpmbTABRAAPEI-227	*Diaphorina citri*	29,098	220	10	28,868	7.1
INSlupTASRAAPEI-89	*Dictyophara europaea*	50,399	243	4	50,152	7.1
INSerlTAHRAAPEI-72	*Diplonychus rusticus*	49,277	512	55	48,710	7.1
INSinlTASRAAPEI-36	*Echinophthirius horridus*	22,274	142	25	22,107	7.1
RINSinlTAERACPEI-57	*Elipsocus kuriliensis*	47,053	120	43	46,890	7.1
INSeqtTCBRAAPEI-12	*Embidopsocus* sp.	59,243	195	1,041	58,007	7.1
INSqiqTBZRAAPEI-21	*Franklinothrips vespiformis*	32,690	116	3	32,571	7.1
INSytvTAERAAPEI-14	*Glycaspis brimblecombei*	81,857	80	5	81,772	7.1
INShauTAORAAPEI-8	*Graphocephala fennahi*	31,250	72	3	31,175	7.1
INSeqtTCERAAPEI-15	*Graphopsocus cruciatus*	22,995	163	7	22,825	7.1
INSinlTAJRAAPEI-75	*Haematomyzus elephantis*	22,032	120	114	21,798	7.1
INSeqtTCFRAAPEI-16	*Hemipsocus chroloticus*	28,341	185	6	28,150	7.1
INSobdTADRAAPEI-90	*Heterocaecilius solocipennis*	31,242	1,018	1	30,223	7.1
INShkeTCCRAAPEI-37	*Hydrometra stagnorum*	33,023	738	1,276	31,009	7.1
INSlupTBERAAPEI-18	*Ilyocoris cimicoides*	30,092	153	1	29,938	7.1
INSkzdTABRAAPEI-136	*Lachesilla abiesicola*	18,822	770	125	17,927	7.1
INSobdTAGRAAPEI-109	*Longivalvus nubilus*	25,942	402	1	25,539	7.1
INSqiqTBYRAAPEI-19	*Lygaeus equestris*	56,420	61	1	56,358	7.1
INSeqtTCDRAAPEI-14	*Matsumuraiella radiopicta*	25,357	136	0	25,221	7.1
INSobdTAERAAPEI-93	*Mesopsocus unipunctatus*	24,933	107	0	24,826	7.1
INSytvTAVRAAPEI-87	*Metatropis rufescens*	24,505	649	8	23,848	7.1
INSqiqTBERAAPEI-57	*Metcalfa pruinosa*	59,425	793	5	58,627	7.1
INSobdTAHRAAPEI-118	*Neoblaste papillosa*	30,050	254	12	29,784	7.1
INShkeTCJRAAPEI-41	*Notonecta glauca*	22,739	94	9	22,636	7.1
INSobdTAFRAAPEI-95	*Peripsocus phaeopterus*	25,252	1,578	0	23,674	7.1
INSnfrTABRAAPEI-14	*Plea minutissima*	42,153	954	23	41,176	7.1
INSbusTBJRAAPEI-85	*Pyrrhocoris apterus*	57,571	2,813	7	54,751	7.1
INSlupTBPRAAPEI-37	*Saldula saltatoria*	29,656	396	13	29,247	7.1
INSobdTAARAAPEI-87	*Stimulopalpus japonicus*	22,728	163	0	22,565	7.1
INSlupTBKRAAPEI-31	*Tachycixius pilosus*	33,958	289	1	33,668	7.1
INSlupTAKRAAPEI-57	*Ulopa reticulata*	47,519	207	5	47,307	7.1
INSytvTCGRAAPEI-39	*Valenzuela badiostigma*	19,269	69	12	19,188	7.1
INSjdsTBFRAAPEI-57	*Xenophyes metoponcus*	76,720	616	28	76,076	7.1
Accon	*Acanalonia conica*	186,341	710	74,923	110,708	9.0
AcspC	*Aclerda* sp.	82,501	42	34,766	47,693	9.0
Acumb	*Acutaspis umbonifera*	80,407	541	29,101	50,765	9.0
Agcon	*Agallia constricta*	133,937	141	75,416	58,380	9.0
Alpil	*Alydus pilosus*	77,309	169	31,305	45,835	9.0
AntriUL	*Anasa tristis*	96,206	76	37,903	58,227	9.0
ScspS2	*Anoplogonius nigricollis*	27,094	55	7,660	19,379	9.0
Appol	*Aphalara polygoni*	136,201	101	73,995	62,105	9.0
ApspP	*Aphelonotus fraterculus*	52,518	48	16,998	35,472	9.0
Arcri	*Arilus cristatus*	50,450	146	17,710	32,594	9.0
Beflu	*Belostoma flumineum*	54,409	109	19,878	34,422	9.0
Botri	*Boisea trivittata*	86,679	192	33,080	53,407	9.0
Brocu	*Bruchomorpha oculata*	137,320	172	65,978	71,170	9.0

22

BuspN	*Buenoa margaritacea*	99,056	91	50,075	48,890	9.0
Cacom	*Campanulotes compar*	86,458	496	25,271	60,691	9.0
Calat	*Campylenchia latipes*	140,739	547	66,396	73,796	9.0
CespC	*Ceratocombus* sp.	65,213	25	24,322	40,866	9.0
Cecir	*Ceroplastes cirripediformis*	54,822	141	17,407	37,274	9.0
TespT	*Chalcocoris rutilans*	52,061	61	14,027	37,973	9.0
McspM	*Chauliops fallax*	44,369	133	15,922	28,314	9.0
Chhil	*Chinavia hilaris*	83,065	693	31,674	50,698	9.0
DispD1	*Chondrodire chilensis*	73,874	95	22,229	51,550	9.0
Chaon	*Chrysomphalus aonidum*	78,237	229	27,923	50,085	9.0
CospC	*Coccus* sp.	72,792	174	28,209	44,409	9.0
Cocol	*Columbicola columbae*	64,966	1,659	24,868	38,439	9.0
Colat	*Corimelaena lateralis*	82,865	1,020	38,196	43,649	9.0
Crspl	*Craspedorrhynchus* sp.	61,147	71	21,971	39,105	9.0
Chcha2	*Cycadothrips chadwicki*	27,167	29	8,942	18,196	9.0
Cybel	*Cyrpoptus belfragei*	67,604	45	22,692	44,867	9.0
Dacon	*Dactylopius confusus*	66,214	44	21,570	44,600	9.0
Damai	*Dalbulus maidis*	26,176	0	514	25,662	9.0
Echag	*Echmepteryx hageni*	22,137	0	499	21,638	9.0
Emfab	*Empoasca fabae*	112,926	60	59,323	53,543	9.0
Thsp	*Frankliniella occidentalis*	54,713	241	24,912	29,560	9.0
Geocu	*Gelastocoris oculatus*	73,286	261	29,504	43,521	9.0
GespG	*Geocoris* sp.	59,350	37	22,441	36,872	9.0
Geewi	*Geomydoecus ewingi*	45,032	103	16,500	28,429	9.0
Hechi	*Hespenedra chilensis*	62,911	170	19,947	42,794	9.0
Hecub	*Heteropsylla cubana*	74,761	96	31,476	43,189	9.0
Hoche	*Holdgatiella chepuensis*	67,417	492	23,730	43,195	9.0
EnspE	*Hoplitocoris* sp.	48,290	49	18,177	30,064	9.0
PespD	*Idiosystatus acutiusculus*	118,120	194	56,938	60,988	9.0
Isfal	*Ischnodemus falicus*	109,469	100	53,238	56,131	9.0
JaspB	*Jalysus* sp.	162,007	113	94,412	67,482	9.0
Lacon	*Lachesilla contraforcepeta*	72,841	238	43,105	29,498	9.0
Lasp	*Ladella* sp.	203,456	525	107,149	95,782	9.0
LrspL1	*Largus californicus*	120,350	95	52,101	68,154	9.0
LrspL2	*Largus californicus*	118,794	81	48,568	70,145	9.0
Lican	*Limnoporus canaliculatus*	209,846	287	111,792	97,767	9.0
Llpil	*Llanquihuea pilosa*	71,727	325	24,842	46,560	9.0
LospM	*Lopidea amorphae*	60,119	160	24,819	35,140	9.0
Lytur	*Lygaeus turcicus*	117,789	358	44,815	72,616	9.0
Lylin	*Lygus lineolaris*	121,382	114	75,718	45,550	9.0
Magla	*Macrohomotoma gladiata*	83,386	92	37,223	46,071	9.0
MaspM8	*Mapuchea* sp.	123,033	181	52,373	70,479	9.0
PlspP	*Megacopta cribraria*	35,716	34	9,547	26,135	9.0
Olpla	*Melanoliarus placitus*	143,486	273	59,160	84,053	9.0
Memul	*Mesovelia mulsanti*	183,199	506	102,766	79,927	9.0
Megra	*Mezira granulata*	111,446	65	51,021	60,360	9.0
Nasub	*Nabis subcoleoptratus*	55,264	92	22,481	32,691	9.0
Tidor	*Neotibicen dorsatus*	113,041	159	49,640	63,242	9.0
Negib	*Nessorhinus gibberulus*	67,149	462	26,966	39,721	9.0
Onfas	*Oncopeltus fasciatus*	84,660	174	30,974	53,512	9.0
Orins	*Orius insidiosus*	60,138	160	34,065	25,913	9.0
Orkel	*Orothrips kelloggi*	54,987	168	22,111	32,708	9.0
Pacel	*Pachypsylla celtidismamma*	246,968	50	162,680	84,238	9.0
PespN	*Pagasa* sp.	51,885	62	20,506	31,317	9.0
Peham	*Peloridium pomponorum*	112,698	154	49,839	62,705	9.0
Phspu	*Philaenus spumarius*	204,741	313	115,426	89,002	9.0
PhspP	*Phlaeothripidae* sp.	88,571	1,653	28,867	58,051	9.0
Phatr	*Phylloscelis atra*	92,285	66	39,666	52,553	9.0
Phpen	*Phymata pennsylvanica*	117,819	106	62,328	55,385	9.0
DispD2	*Piezosternum calidum*	38,887	53	11,388	27,446	9.0
Poqua	*Ponana quadralaba*	150,261	403	69,097	80,761	9.0
Prtes	*Prociphilus tesselatus*	76,345	35	24,111	52,199	9.0
Prbic	*Prosapia bicincta*	75,730	66	27,430	48,234	9.0

Psfax	*Psyllopsis fraxinicola*	142,181	114	78,524	63,543	9.0
Ptjoh	*Ptycta johnsoni*	56,831	284	27,768	28,779	9.0
RespM	*Reuteroscopus ornatus*	50,102	45	16,610	33,447	9.0
Rhobe	*Rhagovelia obesa*	78,443	52	28,093	50,298	9.0
Scgra	*Schizaphis graminum*	48,786	103	17,148	31,535	9.0
ScspS1	*Scutelleridae* sp.	45,583	25	13,708	31,850	9.0
Secin	*Sehirus cinctus*	88,121	91	32,383	55,647	9.0
Stbis	*Stictocephala bisonia*	103,720	210	41,014	62,496	9.0
TrspC	*Trichocorixa calva*	104,655	31	57,173	47,451	9.0
Trcar	*Trionymus caricis*	42,615	29	14,341	28,245	9.0
Trurt	*Trioza urticae*	162,494	112	105,037	57,345	9.0
Tusal	*Tuberolachnus salignus*	58,768	56	18,953	39,759	9.0
Uneuo	*Unaspis euonymi*	97,153	451	37,495	59,207	9.0
Uramb	*Uroleucon ambrosiae*	52,904	25	17,345	35,534	9.0
Vifla	*Vidanoana flavomaculata*	92,727	944	35,285	56,498	9.0
INSytvTALRAAPEI-35	*Acanthocasuarina muellerianae*	43,114	212	23	42,879	7.1
INSnfrTAPRAAPEI-36	*Acanthosoma haemorrhoidale*	28,356	138	0	28,218	7.1
INSnfrTANRAAPEI-34	*Cercopis vulnerata*	34,017	133	12	33,872	7.1
INSytvTCFRAAPEI-37	*Ectopsocus briggsi*	35,614	116	4	35,494	7.1
INSytvTBHRAAPEI-14	*Essigella californica*	58,003	1,347	0	56,656	7.1
INSjdsTABRAAPEI-20	*Frankliniella cephalica*	28,220	2,201	3	26,016	7.1
INSjdsTADRAAPEI-22	*Gynaikothrips ficorum*	111,931	543	138	111,250	7.1
INStmbTBGRAAPEI-33	*Liposcelis bostrychophila*	33,956	211	234	33,511	7.1
INSfrgTAFRAAPEI-31	*Menopon gallinae*	25,281	136	0	25,145	7.1
INSfrgTBCRAAPEI-57	*Nilaparvata lugens*	40,064	198	0	39,866	7.1
INShauTAPRAAPEI-9	*Notostira elongata*	27,842	415	14	27,413	7.1
INSnfrTBLRAAPEI-11	*Okanagana villosa*	51,383	69	0	51,314	7.1
INSjdsTAIRAAPEI-35	*Planococcus citri*	37,278	120	9	37,149	7.1
INSytvTANRAAPEI-37	*Ranatra linearis*	32,052	84	11	31,957	7.1
INSjdsTACRAAPEI-21	*Thrips palmi*	24,671	305	0	24,366	7.1
INSfrgTAPRAAPEI-33	*Trialeurodes vaporariorum*	50,050	155	0	49,895	7.1
INSnfrTAORAAPEI-35	*Velia caprai*	25,734	70	0	25,664	7.1
INSjdsTARRAAPEI-47	*Xenophysella greensladeae*	75,081	460	37	74,584	7.1
INShauTBERAAPEI-33	*Aleochara curtula*	21,787	113	0	21,674	7.1
INSjdsTATRAAPEI-57	*Euroleon nostras*	40,636	352	63	40,221	7.1
INSytvTCFRAAPEI-43	*Inocellia crassicornis*	36,526	133	15	36,378	7.1
INSjdsTBJRAAPEI-79	*Osmylus fulvicephalus*	23,176	144	0	23,032	7.1
INSbttTKRAAPEI-18	*Corydalus cornutus*	18,806	142	0	18,664	7.1
INSnfrTBERAAPEI-19	*Gyrinus marinus*	23,637	127	19	23,491	7.1
INSnfrTAARAAPEI-13	*Orussus abietinus*	26,905	253	3	26,649	7.1
INSnfrTAGRAAPEI-19	*Subilla* sp.	59,469	1,513	155	57,801	7.1
INSfrgTATRAAPEI-37	*Tenthredo koehleri*	25,456	499	0	24,957	7.1
INShauTAKRAAPEI-90	*Baetis* sp.	23,766	67	2	23,697	7.1
INStmbTAARAAPEI-84	*Calopteryx splendens*	21,389	511	0	20,878	7.1
INSjdsTBHRAAPEI-74	*Cordulegaster boltoni*	30,660	132	1	30,527	7.1
INSfrgTAHRAAPEI-18	*Epiophlebia superstes*	30,449	61	0	30,388	7.1
INSytvTCERAAPEI-36	*Eurylophella* sp.	36,421	43	6	36,372	7.1
INSfrgTALRAAPEI-22	*Apachyus charteceus*	33,943	175	13	33,755	7.1
INSfrgTAZRAAPEI-46	*Aposthonia japonica*	54,664	204	2	54,458	7.1
INSfrgTAVRAAPEI-41	*Blaberus atropos*	63,927	258	17	63,652	7.1
INSytvTCDRAAPEI-35	*Cryptocercus wrighti*	56,337	63	0	56,274	7.1
INSjdsTBNRAAPEI-89	*Forficula auricularia*	48,947	1,106	5	47,836	7.1
INSfrgTAKRAAPEI-21	*Galloisiana yuasai*	49,615	93	0	49,522	7.1
INSfrgTAXRABPEI-44	*Gryllotalpa* sp.	45,804	73	6	45,725	7.1
INSnfrTALRAAPEI-31	*Leuctra* sp.	27,748	108	1	27,639	7.1
INShauTAARAAPEI-90	*Mantis religiosa*	45,453	59	0	45,394	7.1
INShauTAMRAAPEI-94	*Metallyticus splendidus*	62,765	341	0	62,424	7.1
INShauTALRAAPEI-93	*Perla marginata*	34,553	269	0	34,284	7.1
INSfrgTAORAAPEI-33	*Peruphasma schultei*	48,023	54	0	47,969	7.1
INSfrgTBBRAAPEI-56	*Tanzaniophasma* sp.	36,828	221	0	36,607	7.1
INShauTANRAAPEI-95	*Tetrix subulata*	30,056	248	0	29,808	7.1
INSnfrTBPRAAPEI-15	*Timema cristinae*	41,806	845	10	40,951	7.1
INSjdsTAQRAAPEI-46	*Zorotypus caudelli*	54,883	420	129	54,334	7.1

Table S3. Official gene sets used to search for single-copy genes in the assembled transcriptomes.

Species name	No. aa seq (adj.)	No. nt seq. (adj.)	Version	Source	URL	Date of download
Daphnia pulex	30940	30940	jgi060905	wFleabase	http://wfleabase.org/	2015-12-09
Zootermopsis nevadensis	17737	17737	2.1	J. Liebig and R. Waterhouse		2012-03-06
Rhodnius prolixus	15437	15437	1.1	Vectorbase	https://www.vectorbase.org/	2015-12-09
Pediculus humanus	10773	10773	1.2	Vectorbase	https://www.vectorbase.org/	2015-12-09
Acromyrmex echinatior	17277	17277	3.8	Hymenoptera genome database	http://hymenopteragenome.org/acromyrmex/	2015-12-09
Tribolium castaneum	16643	16643	3.0 (GLEAN set)	Beetlebase	http://beetlebase.org/	2012-09-24

Table S4. Assembly statistics of and number of identified target genes in the analyzed transcriptomes.

species name	library ID	orthologs	assembly size	no. of contigs	longest contig	GC content
Aaroniella_ sp	INSobdTACRAAPEI-89	1927	23766864	27404	16212	38.0
Acanalonia_conica	Accon	2238	78405065	110708	26636	33.5
Acanthocasuarina_muellerianae	INSytvTALRAAPEI-35	1948	32423115	43114	26973	40.7
Acanthosoma_haemorrhoidale	INSnfrTAPRAAPEI-36	2038	21839689	28363	21664	37.2
Aclerda_sp	AcspC	2194	52568191	47693	28558	36.3
Acutapsis_umbonifera	Acumb	2204	56800197	50765	34303	37.9
Agallia_constricta	Agcon	2224	43799808	58380	16244	38.5
Aleochara_curtula	INShauTBERAAPEI-33	2205	26220864	21790	32053	42.7
Alydus_pilosus	Alpil	2152	34803354	45835	15679	36.2
Amphipsocus_japonicus	INSeqtTCCRAAPEI-13	2205	29689878	29617	15982	41.3
Anasa_tristis	AntriUL	2232	41737148	58227	17616	36.2
Anoplogonius_nigricollis	ScspS2	1801	13678498	19379	6237	35.1
Apachyus_charteceus	INSfrgTALRAAPEI-22	2137	22858423	33954	18263	34.4
Aphalara_polygoni	Appol	2018	40370361	62105	26200	36.7
Aphelocheirus_aestivalis	INSlupTBMRAAPEI-34	1740	24403060	40938	31110	37.1
Aphelonotus_fraterculus	ApspP	1876	28334676	35472	11008	35.8
Aphrophora_alni	INSlupTAFRAAPEI-44	2029	24934951	40530	14315	34.6
Aposthonia_japonica	INSfrgTAZRAAPEI-46	1977	38245803	54667	28491	35.0
Aquarius_paludum	INSytvTBARAAPEI-94	1768	40444876	71674	19731	36.4
Aradus_betulae	INSlupTALRAAPEI-62	2059	36374496	55635	20022	40.1
Arilus_cristatus	Arcri	2099	26079491	32594	12288	33.0
Arocatus_melanocephalus	INSqiqTBURAAPEI-171	1624	23080808	45600	23878	36.4
Badonnelia_titei	INSofmTBCRAAPEI-37	2213	31991002	33784	12278	40.4
Baetis_ sp	INShauTAKRAAPEI-90	2187	24843866	23767	15601	47.7
Belostoma_flumineum	Beflu	2043	31958003	34422	25271	38.7
Bertkauia_ sp	INSobdTABRAAPEI-88	1967	24712390	27453	15801	39.9
Blaberus_atropos	INSfrgTAVRAAPEI-41	1630	42412602	63930	43709	35.5
Boisea_trivittata	Botri	2239	44811122	53407	40947	34.7
Bruchomorpha_oculata	Brocu	2245	44927908	71170	25096	34.3
Buenoa_margaritacea	BuspN	2209	42732228	48890	26142	39.7
Caliscelis_bonelli	INSqiqTBMRAAPEI-84	2098	24859751	32533	22272	34.7
Calopteryx_splendens	INStmbTAARAAPEI-84	1681	16132415	21395	12927	40.6
Campanulotes_compar	Cacom	2279	37003730	60691	10662	42.3
Campylenchia_latipes	Calat	2248	53986922	73796	26892	36.2
Centrotus_cornutus	INSytvTAQRAAPEI-43	2041	29417059	40732	10054	38.3
Ceratocombus_ sp	CespC	2119	32858075	40866	26482	35.3
Cercopis_vulnerata	INSnfrTANRAAPEI-34	2144	29859012	34019	19358	36.5
Cerobasis_guestfalica	INSeqtTCGRAAPEI-17	2219	35966251	51022	8101	40.2
Ceroplastes_cirripediformis	Cecir	2107	39278582	37274	16944	37.4
Chalcocoris_rutilans	TespT	2176	32002254	37973	10167	34.6
Chauliops_fallax	McspM	1523	18071929	28314	10202	38.4
Chinavia_hilaris	Chhil	2200	40045896	50698	15495	34.1
Chrysomphalus_aonidum	Chaon	2218	59630753	50085	26873	36.8
Cimex_lectularius	INSjdsTAFRAAPEI-31	2282	32073011	33466	23766	39.5
Coccus_ sp	CospC	2205	54712771	44409	23342	37.8
Columbicola_columbae	Cocol	2263	23469018	38439	12188	39.2
Cordulegaster_boltoni	INSjdsTBHRAAPEI-74	2123	29371935	30664	33264	40.9
Corimelaena_lateralis	Colat	2193	34742355	43649	20654	35.3
Corixa_punctata	INSlupTAORAAPEI-79	1812	14994749	22921	26887	40.9
Corydalus_cornutus	INSbttTKRAAPEI-18	2184	20260587	18806	35896	37.3
Corythucha_ciliata	INSbusTBLRAAPEI-77	2261	23062128	19469	13226	44.2
Craspedorrhynchus_ sp	CrspI	2329	43044996	39105	47085	36.8
Cryptocercus_wrighti	INSytvTCDRAAPEI-35	2170	48853407	56338	54303	39.4
Cycadothrips_chadwicki	Chcha2	1451	13337924	18196	8752	40.2
Cyrpoptus_belfragei	Cybel	2121	33202685	44867	27489	32.6
Dactylopius_confuses	Dacon	2210	48584386	44600	31937	33.8
Daktulosphaira_vitifoliae	INSeqtTDYRAAPEI-20	2169	23156504	26967	11232	34.0
Dalbulus_maidis	Damai	1708	28305469	25662	9283	36.9
Diaphorina_citri	INSpmbTABRAAPEI-227	1853	25035188	29097	10258	38.7
Dictyophara_europaea	INSlupTASRAAPEI-89	1715	27538128	50408	45016	32.0
Diplonychus_rusticus	INSerlTAHRAAPEI-72	1549	23091850	49286	14091	39.7
Echinophthirius_horridus	INSinlTASRAAPEI-36	1657	11524277	22275	6618	38.0
Echmepteryx_hageni	Echag	1890	26911408	21638	11802	36.3
Ectopsocus_briggsi	INSytvTCFRAAPEI-37	2153	25227588	35615	9531	43.4

Elipsocus_kuriliensis	RINSinITAERACPEI-57	2191	40016327	47063	8877	44.1
Embidopsocus_ sp	INSeqtTCBRAAPEI-12	1758	21905982	59250	10702	40.4
Empoasca_fabae	Emfab	2088	38163234	53543	15054	39.2
Epiophlebia_superstes	INSfrgTAHRAAPEI-18	2151	31652091	30452	23676	39.9
Erythrothrips_arizonae	PhspP	2267	59986664	58051	23741	42.1
Essigella_californica	INSytvTBHRAAPEI-14	2176	42976089	58013	14678	35.4
Euroleon_nostras	INSjdsTATRAAPEI-57	2149	25501862	40647	12436	34.1
Eurylophella_ sp	INSytvTCERAAPEI-36	2255	33582280	36423	24697	38.5
Forficula_auricularia	INSjdsTBNRAAPEI-89	2034	32695435	48954	13835	36.4
Frankliniella_cephalica	INSjdsTABRAAPEI-20	2175	27030171	28223	11155	45.2
Frankliniella_occidentalis	Thsp	2210	32367596	29560	15554	43.9
Franklinothrips_vespiformis	INSqiqTBZRAAPEI-21	2084	31768835	32692	14897	40.6
Galloisiana_yuasai	INSfrgTAKRAAPEI-21	2061	41946837	49619	40128	39.3
Gelastocoris_oculatus	Geocu	2161	32973781	43521	36273	37.9
Geocoris_ sp	GespG	1850	28981020	36872	10545	37.3
Geomydoecus_ewingi	Geewi	2339	30368452	28429	27623	35.1
Glycaspis_brimblecombei	INSytvTAERAAPEI-14	1924	45376001	81874	24880	33.7
Graphocephala_fennahi	INShauTAORAAPEI-8	1953	26595304	31251	14969	43.1
Graphopsocus_cruciatus	INSeqtTCERAAPEI-15	2127	24293525	22995	11545	44.3
Gryllotalpa_ sp	INSfrgTAXRABPEI-44	2093	35496701	45808	12684	43.4
Gynaikothrips_ficorum	INSjdsTADRAAPEI-22	2064	66024375	111941	16079	40.9
Gyrinus_marinus	INSnfrTBERAAPEI-19	2039	20417344	23637	13197	41.7
Haematomyzus_elephantis	INSinITAJRAAPEI-75	1926	12833632	22048	10375	34.2
Hemipsocus_chroloticus	INSeqtTCFRAAPEI-16	2059	20416847	28340	10115	38.5
Hespenedra_chilensis	Hechi	1973	34845676	42794	16141	37.3
Heterocaecilius_solocipennis	INSobdTADRAAPEI-90	2156	28503963	31247	27624	41.0
Heteropsylla_cubana	Hecub	2074	31314294	43189	27561	31.6
Holdgatiella_chepuensis	Hoche	2089	37743759	43195	12186	39.3
Hoplitocoris_ sp	EnspE	1731	23360899	30064	8092	31.6
Hydrometra_stagnorum	INShkeTCCRAAPEI-37	1835	19418800	33031	7700	41.2
Idiosystatus_acutiusculus	PespD	2186	58492799	60988	27335	37.4
Ilyocoris_cimicoides	INSlupTBERAAPEI-18	2076	23705280	30096	23234	39.5
Inocellia_crassicornis	INSytvTCFRAAPEI-43	2045	29780369	36528	26505	37.2
Ischnodemus_falicus	Isfal	2227	46332497	56131	19540	35.8
Jalysus_ sp	JaspB	2145	48312152	67482	26786	37.5
Lachesilla_abiesicola	INSkzdTABRAAPEI-136	2144	15723521	18834	75872	43.6
Lachesilla_contraforcepeta	Lacon	2305	30598256	29498	23120	43.6
Ladella_sp	Lasp	2029	56403614	95782	22438	32.0
Largus_californicus_1	LrspL1	2223	52822694	68154	27408	34.3
Largus_californicus_2	LrspL2	2284	57081533	70145	30587	34.4
Leuctra_sp	INSnfrTALRAAPEI-31	2262	34586468	27749	16698	44.4
Limnoporus_canaliculatus	Lican	2116	57721725	97767	23521	34.8
Liposcelis_bostrychophila	INStmbTBGRAAPEI-33	2234	30313310	33962	11842	37.0
Llanquihuea_pilosa	Llpil	1913	37598185	46560	14020	38.9
Longivalvus_nubilus	INSobdTAGRAAPEI-109	1748	20513523	25943	27936	40.8
Lopidea_amorphae	LospM	2236	42042747	35140	45764	44.1
Lygaeus_equestris	INSqiqTBYRAAPEI-19	2082	33073221	56428	23814	37.7
Lygaeus_turcicus	Lytur	2209	57465981	72616	20964	33.4
Lygus_lineolaris	Lylin	2197	41384616	45550	16081	41.4
Macrohomotoma_gladiata	Magla	1663	28762605	46071	14458	36.8
Mantis_religiosa	INShauTAARAAPEI-90	1941	32398303	45459	30306	38.3
Mapuchea_sp	MaspM8	2196	58929596	70479	23498	37.0
Matsumuraiella_radiopicta	INSeqtTCDRAAPEI-14	2236	29010145	25357	12093	44.2
Megacopta_cribaria	PlspP	2107	23186814	26135	14075	33.9
Menopon_gallinae	INSfrgTAFRAAPEI-31	2289	30203641	25285	22431	44.1
Mesopsocus_unipunctatus	INSobdTAERAAPEI-93	2015	21829923	24932	20770	44.4
Mesovelia_mulsanti	Memul	2081	51536645	79927	26120	36.9
Metallyticus_splendidus	INShauTAMRAAPEI-94	1879	41919841	62765	28076	38.5
Metatropis_rufescens	INSytvTAVRAAPEI-87	1545	17936255	24507	12345	38.2
Metcalfa_pruinosa	INSqiqTBERAAPEI-57	2057	39141482	59431	46763	37.3
Mezira_granulata	Megra	2270	52025883	60360	23101	34.2
Nabis_subcoleoptratus	Nasub	2231	31787638	32691	31678	35.9
Neoblaste_papillosus	INSobdTAHRAAPEI-118	2080	26179780	30050	21400	39.8
Nessorhinus_gibberulus	Negib	2051	26821667	39721	14307	38.8
Nilaparvata_lugens	INSfrgTBCRAAPEI-57	2177	34901049	40068	14006	41.0

Notonecta_glauca	INShkeTCJRAAPEI-41	1861	18412270	22742	17892	39.5
Notostira_elongata	INShauTAPRAAPEI-9	2010	27035672	27842	20340	42.6
Okanagana_villosa	INSnfrTBLRAAPEI-11	2034	41652891	51387	27820	36.6
Oliarus_placitus	Olpla	2079	65922569	84053	23378	37.2
Oncopeltus_fasciatus	Onfas	2229	43620981	53512	26876	34.6
Orius_insidiosus	Orins	2218	26926668	25913	25544	38.9
Orothrips_kelloggi	Orkel	2145	35528455	32708	14480	42.0
Orussus_abietinus	INSnfrTAARAAPEI-13	2250	41103769	26905	30157	45.0
Osmylus_fulvicephalus	INSjdsTBJRAAPEI-79	2146	26355321	23179	19514	41.5
Pachypsylla_celtidismamma	Pacel	2102	61715315	84238	24914	34.6
Pagasa_sp	PespN	2280	30390488	31317	23975	37.6
Peloridium_hammoniorum	Peham	2111	57726145	62705	13029	37.1
Peripsocus_phaeopterus	INSobdTAFRAAPEI-95	2066	21477391	25252	9485	41.4
Perla_marginata	INShauTALRAAPEI-93	1648	25914117	34552	49442	40.1
Peruphasma_schultei	INSfrgTAORAAPEI-33	1801	39144177	48026	29663	41.5
Philaenus_spumarius	Phspu	2248	65543966	89002	17088	34.4
Phylloscelis_atra	Phatr	2199	36849431	52553	27734	32.2
Phymata_pennsylvanica	Phpen	2176	38379194	55385	17434	35.0
Piezosternum_callidum	DispD2	1985	22251963	27446	13970	34.0
Planococcus_citri	INSjdsTAIRAAPEI-35	2073	33725955	37278	19452	37.4
Plea_minutissima	INSnfrTABRAAPEI-14	1985	29142902	42160	17548	39.6
Ponana_quadralaba	Poqua	2267	64045803	80761	37426	36.3
Prociphilus_tesselatus	Prtes	2192	47009400	52199	20954	32.8
Prosapia_bicincta	Prbic	2156	39526530	48234	26849	34.3
Psyllopsis_faxinicola	Psfax	2012	46898347	63543	26501	39.6
Ptycta_johnsoni	Ptjoh	2039	25929616	28779	8858	37.8
Pyrrhocoris_apterus	INSbusTBJRAAPEI-85	2211	50904712	57571	17493	45.8
Ranatra_linearis	INSytvTANRAAPEI-37	1879	25565273	32053	20494	39.2
Reuteroscopus_ornatus	RespM	2229	38334317	33447	28771	45.4
Rhagovelia_obesa	Rhobe	2053	37228276	50298	27798	34.3
Saldula_saltatoria	INSlupTBPRAAPEI-37	1923	17750697	29659	8499	41.4
Schizaphis_graminum	Scgra	2200	36200869	31535	28132	34.3
Scutelleridae_sp	ScspS1	2189	29353446	31850	27028	34.6
Sehirus_cincutus	Secin	2296	45570966	55647	26998	34.3
Stictocephala_bisonia	Stbis	2230	53718825	62496	27522	36.7
Stimulopalpus_japonicus	INSobdTAARAAPEI-87	2097	22669149	22727	10168	39.9
Tachycixius_pilosus	INSlupTBKRAAPEI-31	1648	20914042	33969	20868	38.4
Tanzaniophasma_ sp	INSfrgTBBRAAPEI-56	1908	34502039	36833	56132	42.1
Taractellus_chilensis	DispD1	1993	37581699	51550	18195	31.6
Tenthredo_koehleri	INSfrgTATRAAPEI-37	2259	34190889	25456	28472	39.9
Tetrix_subulata	INShauTANRAAPEI-95	2168	27272730	30059	26747	41.2
Thrips_palmi	INSjdsTACRAAPEI-21	2156	26952606	24676	9288	46.5
Tibicen_dorsata	Tidor	2062	46190617	63242	50110	34.3
Timema_cristinae	INSnfrTBPRAAPEI-15	1667	28302605	41810	18112	37.4
Trialeurodes_vaporariorum	INSfrgTAPRAAPEI-33	2093	41324811	50053	26890	38.4
Trichocorixa_calva	TrspC	2202	37866798	47451	29164	38.5
Trionymus_caricis	Trcar	2169	34841673	28245	16367	37.2
Trioza_urticae	Trurt	2009	41034766	57345	14025	39.7
Tuberolachnus_salignus	Tusal	2177	42294348	39759	24256	34.9
Ulopa_reticulata	INSlupTAKRAAPEI-57	1730	27100790	47527	7949	39.3
Unaspis_euonymi	Uneuo	2241	54887968	59207	24867	40.4
Uroleucon_ambrosiae	Uramb	2197	42211539	35534	27583	34.4
Valenzuela_badiostigma	INSytvTCGRAAPEI-39	2204	21728266	19271	19268	46.4
Velia_caprai	INSnfrTAORAAPEI-35	1796	19253870	25735	16929	39.1
Vidanoana_flavomaculata	Vifla	2109	45681665	56498	26349	39.0
Subilla_ sp	INSnfrTAGRAAPEI-19	2162	42967462	59482	40221	36.0
Xenophyes_metoponcus	INSjdsTBFRAAPEI-57	2001	54285874	76720	12828	38.5
Xenophysella_greensladeae	INSjdsTARRAAPEI-47	1951	48507682	75086	9682	39.6
Zorotypus_caudelli	INSjdsTAQRAAPEI-46	2003	35582821	54890	15225	35.6

Table S5. Group composition for Four-Cluster Likelihood Mapping (FcLM).

Hypothesis (*i*): Psocodea are the sister group of Holometabola

(*ia*) All Condylognatha included

Group 1 (G1): Psocodea (31species)

Echmepteryx_hageni, Cerobasis_guestfalica, Bertkauia_sp, Amphipsocus_japonicus, Valenzuela_badiostigma, Matsumuraiella_radiopicta, Graphopsocus_cruciatus, Aaroniella_sp, Heterocaecilius_solocipennis, Lachesilla_abiesicola, Mesopsocus_unipunctatus, Peripsocus_phaeopterus, Hemipsocus_chroloticus, Longivalvus_nubilus, Neoblaste_papillosus, Elipsocus_kuriliensis, Lachesilla_contraforcepeta, Ectopsocus_briggsi, Ptycta_johnsoni, Stimulopalpus_japonicus, Embidopsocus_sp, Badonnelia_titei, Liposcelis_bostrychophila, Pediculus_humanus, Haematomyzus_elephantis, Echinophthirius_horridus, Columbicola_columbae, Campanulotes_compar, Craspedorrhynchus_sp, Geomydoecus_ewingi, Menopon_gallinae

Group 2 (G2): Condylognatha (129 species)

Franklinothrips_vespiformis, Thrips_palmi, Gynaikothrips_ficorum, Frankliniella_cephalica, Erythrothrips_arizonae, Frankliniella_occidentalis, Orothrips_kelloggi, Cycadothrips_chadwicki, Diaphorina_citri, Trionymus_caricis, Macrohomotoma_gladiata, Psyllopsis_faxinicola, Heteropsylla_cubana, Pachypsylla_celtidismamma, Aphalara_polygoni, Trioza_urticae, Glycaspis_brimblecombei, Acanthocasuarina_muellerianae, Trialeurodes_vaporariorum, Daktulosphaira_vitifoliae, Planococcus_citri, Aclerda_sp, Ceroplastes_cirripediformis, Coccus_sp, Dactylopius_confuses, Chrysomphalus_aonidum, Acutapsis_umbonifera, Unaspis_euonymi, Essigella_californica, Prociphilus_tesselatus, Schizaphis_graminum, Uroleucon_ambrosiae, Tuberolachnus_salignus, Cercopis_vulnerata, Aphrophora_alni, Philaenus_spumarius, Prosapia_bicincta, Tibicen_dorsata, Okanagana_villosa, Dalbulus_maidis, Ponana_quadralaba, Agallia_constricta, Hespenedra_chilensis,Empoasca_fabae, Vidanoana_flavomaculata, Llanquihuea_pilosa, Stictocephala_bisonia, Campylenchia_latipes, Nessorhinus_gibberulus, Holdgatiella_chepuensis, Centrotus_cornutus, Graphocephala_fennahi, Ulopa_reticulata, Mapuchea_sp, Dictyophara_europaea, Tachycixius_pilosus, Nilaparvata_lugens, Metcalfa_pruinosa, Caliscelis_bonelli, Acanalonia_conica, Bruchomorpha_oculata, Oliarus_placitus, Idiosystatus_acutiusculus, Phylloscelis_atra, Taractellus_chilensis, Cyrpoptus_belfragei, Ladella_sp, Xenophyes_metoponcus, Xenophysella_greensladeae, Peloridium_hammoniorum, Hoplitocoris_sp, Ceratocombus_sp, Mesovelia_mulsanti, Rhagovelia_obesa, Limnoporus_canaliculatus, Hydrometra_stagnorum, Aquarius_paludum, Velia_caprai, Belostoma_flumineum, Gelastocoris_oculatus, Trichocorixa_calva, Buenoa_margaritacea, Ranatra_linearis, Aphelocheirus_aestivalis, Diplonychus_rusticus, Corixa_punctata, Ilyocoris_cimicoides, Notonecta_glauca, Plea_minutissima, Saldula_saltatoria, Cimex_lectularius, Arilus_cristatus, Phymata_pennsylvanica, Rhodnius_prolixus, Corythucha_ciliata, Aphelonotus_fraterculus, Lygus_lineolaris, Reuteroscopus_ornatus, Lopidea_amorphae, Notostira_elongata, Nabis_subcoleoptratus, Pagasa_sp, Orius_insidiosus, Arocatus_melanocephalus, Lygaeus_equestris, Metatropis_rufescens, Pyrrhocoris_apterus, Aradus_betulae, Mezira_granulata, Sehirus_cinctus, Corimelaena_lateralis, Piezosternum_callidum, Chinavia_hilaris, Megacopta_cribaria, Scutelleridae_sp, Anoplogonius_nigricollis, Chalcocoris_rutilans, Jalysus_sp, Lygaeus_turcicus, Oncopeltus_fasciatus, Ischnodemus_falicus, Geocoris_sp, Chauliops_fallax, Largus_californicus_1, Largus_californicus_2, Alydus_pilosus, Anasa_tristis, Boisea_trivittata, Acanthosoma_haemorrhoidale

Group 3 (G3): Holometabola (11 species)

Corydalus_cornutus, Osmylus_fulvicephalus, Euroleon_nostras, Inocellia_crassicornis, Subilla_sp, Aleochara_curtula, Tribolium_castaneum, Gyrinus_marinus, Orussus_abietinus, Tenthredo_koehleri, Acromyrmex_echinatior

Group 4 (G4): outgroups (17 species)

Blaberus_atropos, Cryptocercus_wrighti, Apachyus_charteceus, Forficula_auricularia, Aposthonia_japonica, Galloisiana_yuasai, Mantis_religiosa, Metallyticus_splendidus, Tanzaniophasma_sp, Tetrix_subulata, Gryllotalpa_sp, Peruphasma_schultei, Timema_cristinae, Leuctra_sp, Perla_marginata, Zorotypus_caudelli, Zootermopsis_nevadensis

(*ib*) With Thysanoptera excluded

Group 1 (G1): Psocodea (31 species)

Echmepteryx_hageni, Cerobasis_guestfalica, Bertkauia_sp, Amphipsocus_japonicus, Valenzuela_badiostigma, Matsumuraiella_radiopicta, Graphopsocus_cruciatus, Aaroniella_sp, Heterocaecilius_solocipennis, Lachesilla_abiesicola, Mesopsocus_unipunctatus, Peripsocus_phaeopterus, Hemipsocus_chroloticus, Longivalvus_nubilus, Neoblaste_papillosus, Elipsocus_kuriliensis, Lachesilla_contraforcepeta, Ectopsocus_briggsi, Ptycta_johnsoni, Stimulopalpus_japonicus, Embidopsocus_sp, Badonnelia_titei, Liposcelis_bostrychophila, Pediculus_humanus, Haematomyzus_elephantis, Echinophthirius_horridus, Columbicola_columbae, Campanulotes_compar, Craspedorrhynchus_sp, Geomydoecus_ewingi, Menopon_gallinae

Group 2 (G2): Hemiptera (121 species)

Diaphorina_citri, Trionymus_caricis, Macrohomotoma_gladiata, Psyllopsis_faxinicola, Heteropsylla_cubana, Pachypsylla_celtidismamma, Aphalara_polygoni, Trioza_urticae, Glycaspis_brimblecombei, Acanthocasuarina_muellerianae, Trialeurodes_vaporariorum, Daktulosphaira_vitifoliae, Planococcus_citri, Aclerda_sp, Ceroplastes_cirripediformis, Coccus_sp, Dactylopius_confuses, Chrysomphalus_aonidum, Acutapsis_umbonifera, Unaspis_euonymi, Essigella_californica, Prociphilus_tesselatus, Schizaphis_graminum, Uroleucon_ambrosiae, Tuberolachnus_salignus, Cercopis_vulnerata, Aphrophora_alni, Philaenus_spumarius, Prosapia_bicincta, Tibicen_dorsata, Okanagana_villosa, Dalbulus_maidis, Ponana_quadralaba, Agallia_constricta, Hespenedra_chilensis,Empoasca_fabae, Vidanoana_flavomaculata, Llanquihuea_pilosa, Stictocephala_bisonia, Campylenchia_latipes, Nessorhinus_gibberulus, Holdgatiella_chepuensis, Centrotus_cornutus, Graphocephala_fennahi, Ulopa_reticulata, Mapuchea_sp, Dictyophara_europaea, Tachycixius_pilosus, Nilaparvata_lugens, Metcalfa_pruinosa, Caliscelis_bonelli, Acanalonia_conica, Bruchomorpha_oculata, Oliarus_placitus, Idiosystatus_acutiusculus, Phylloscelis_atra, Taractellus_chilensis, Cyrpoptus_belfragei, Ladella_sp,

Xenophyes_metoponcus, Xenophysella_greensladeae, Peloridium_hammoniorum, Hoplitocoris_sp, Ceratocombus_sp, Mesovelia_mulsanti, Rhagovelia_obesa, Limnoporus_canaliculatus, Hydrometra_stagnorum, Aquarius_paludum, Velia_caprai, Belostoma_flumineum, Gelastocoris_oculatus, Trichocorixa_calva, Buenoa_margaritacea, Ranatra_linearis, Aphelocheirus_aestivalis, Diplonychus_rusticus, Corixa_punctata, Ilyocoris_cimicoides, Notonecta_glauca, Plea_minutissima, Saldula_saltatoria, Cimex_lectularius, Arilus_cristatus, Phymata_pennsylvanica, Rhodnius_prolixus, Corythucha_ciliata, Aphelonotus_fraterculus, Lygus_lineolaris, Reuteroscopus_ornatus, Lopidea_amorphae, Notostira_elongata, Nabis_subcoleoptratus, Pagasa_sp, Orius_insidiosus, Arocatus_melanocephalus, Lygaeus_equestris, Metatropis_rufescens, Pyrrhocoris_apterus, Aradus_betulae, Mezira_granulata, Sehirus_cinctus, Corimelaena_lateralis, Piezosternum_callidum, Chinavia_hilaris, Megacopta_cribaria, Scutelleridae_sp, Anoplogonius_nigricollis, Chalcocoris_rutilans, Jalysus_sp, Lygaeus_turcicus, Oncopeltus_fasciatus, Ischnodemus_falicus, Geocoris_sp, Chauliops_fallax, Largus_californicus_1, Largus_californicus_2, Alydus_pilosus, Anasa_tristis, Boisea_trivittata, Acanthosoma_haemorrhoidale

Group 3 (G3): Holometabola (11 species)

Corydalus_cornutus, Osmylus_fulvicephalus, Euroleon_nostras, Inocellia_crassicornis, Subilla_sp, Aleochara_curtula, Tribolium_castaneum, Gyrinus_marinus, Orussus_abietinus, Tenthredo_koehleri, Acromyrmex_echinatior

Group 4 (G4): outgroups (17 species)

Blaberus_atropos, Cryptocercus_wrighti, Apachyus_charteceus, Forficula_auricularia, Aposthonia_japonica, Galloisiana_yuasai, Mantis_religiosa, Metallyticus_splendidus, Tanzaniophasma_sp, Tetrix_subulata, Gryllotalpa_sp, Peruphasma_schultei, Timema_cristinae, Leuctra_sp, Perla_marginata, Zorotypus_caudelli, Zootermopsis_nevadensis

(ic) With Hemiptera excluded

Group 1 (G1): Psocodea (31 species)

Echmepteryx_hageni, Cerobasis_guestfalica, Bertkauia_sp, Amphipsocus_japonicus, Valenzuela_badiostigma, Matsumuraiella_radiopicta, Graphopsocus_cruciatus, Aaroniella_sp, Heterocaecilius_solocipennis, Lachesilla_abiesicola, Mesopsocus_unipunctatus, Peripsocus_phaeopterus, Hemipsocus_chroloticus, Longivalvus_nubilus, Neoblaste_papillosus, Elipsocus_kuriliensis, Lachesilla_contraforcepeta, Ectopsocus_briggsi, Ptycta_johnsoni, Stimulopalpus_japonicus, Embidopsocus_sp, Badonnelia_titei, Liposcelis_bostrychophila, Pediculus_humanus, Haematomyzus_elephantis, Echinophthirius_horridus, Columbicola_columbae, Campanulotes_compar, Craspedorrhynchus_sp, Geomydoecus_ewingi, Menopon_gallinae

Group 2 (G2): Thysanoptera (8 species)

Franklinothrips_vespiformis, Thrips_palmi, Gynaikothrips_ficorum, Frankliniella_cephalica, Erythrothrips_arizonae, Frankliniella_occidentalis, Orothrips_kelloggi, Cycadothrips_chadwicki

Group 3 (G3): Holometabola (11 species)

Corydalus_cornutus, Osmylus_fulvicephalus, Euroleon_nostras, Inocellia_crassicornis, Subilla_sp, Aleochara_curtula, Tribolium_castaneum, Gyrinus_marinus, Orussus_abietinus, Tenthredo_koehleri, Acromyrmex_echinatior

Group 4 (G4): outgroups (17 species)

Blaberus_atropos, Cryptocercus_wrighti, Apachyus_charteceus, Forficula_auricularia, Aposthonia_japonica, Galloisiana_yuasai, Mantis_religiosa, Metallyticus_splendidus, Tanzaniophasma_sp, Tetrix_subulata, Gryllotalpa_sp, Peruphasma_schultei, Timema_cristinae, Leuctra_sp, Perla_marginata, Zorotypus_caudelli, Zootermopsis_nevadensis

Hypothesis (ii): Thysanoptera are the sister group of Hemiptera

Group 1 (G1): Psocodea (31 species)

Echmepteryx_hageni, Cerobasis_guestfalica, Bertkauia_sp, Amphipsocus_japonicus, Valenzuela_badiostigma, Matsumuraiella_radiopicta, Graphopsocus_cruciatus, Aaroniella_sp, Heterocaecilius_solocipennis, Lachesilla_abiesicola, Mesopsocus_unipunctatus, Peripsocus_phaeopterus, Hemipsocus_chroloticus, Longivalvus_nubilus, Neoblaste_papillosus, Elipsocus_kuriliensis, Lachesilla_contraforcepeta, Ectopsocus_briggsi, Ptycta_johnsoni, Stimulopalpus_japonicus, Embidopsocus_sp, Badonnelia_titei, Liposcelis_bostrychophila, Pediculus_humanus, Haematomyzus_elephantis, Echinophthirius_horridus, Columbicola_columbae, Campanulotes_compar, Craspedorrhynchus_sp, Geomydoecus_ewingi, Menopon_gallinae

Group 2 (G2): Thysanoptera (8 species)

Franklinothrips_vespiformis, Thrips_palmi, Gynaikothrips_ficorum, Frankliniella_cephalica, Erythrothrips_arizonae, Frankliniella_occidentalis, Orothrips_kelloggi, Cycadothrips_chadwicki

Group 3 (G3): Hemiptera (121 species)

Diaphorina_citri, Trionymus_caricis, Macrohomotoma_gladiata, Psyllopsis_faxinicola, Heteropsylla_cubana, Pachypsylla_celtidismamma, Aphalara_polygoni, Trioza_urticae, Glycaspis_brimblecombei, Acanthocasuarina_muellerianae, Trialeurodes_vaporariorum, Daktulosphaira_vitifoliae, Planococcus_citri, Aclerda_sp, Ceroplastes_cirripediformis, Coccus_sp, Dactylopius_confuses, Chrysomphalus_aonidum, Acutapsis_umbonifera, Unaspis_euonymi, Essigella_californica, Prociphilus_tesselatus, Schizaphis_graminum, Uroleucon_ambrosiae, Tuberolachnus_salignus, Cercopis_vulnerata, Aphrophora_alni, Philaenus_spumarius, Prosapia_bicincta, Tibicen_dorsata, Okanagana_villosa, Dalbulus_maidis, Ponana_quadralaba, Agallia_constricta, Hespenedra_chilensis,Empoasca_fabae, Vidanoana_flavomaculata, Llanquihuea_pilosa, Stictocephala_bisonia, Campylenchia_latipes, Nessorhinus_gibberulus, Holdgatiella_chepuensis, Centrotus_cornutus, Graphocephala_fennahi, Ulopa_reticulata, Mapuchea_sp, Dictyophara_europaea, Tachycixius_pilosus, Nilaparvata_lugens, Metcalfa_pruinosa, Caliscelis_bonelli, Acanalonia_conica, Bruchomorpha_oculata,

Oliarus_placitus, Idiosystatus_acutiusculus, Phylloscelis_atra, Taractellus_chilensis, Cyrpoptus_belfragei, Ladella_sp,
Xenophyes_metoponcus, Xenophysella_greensladeae, Peloridium_hammoniorum, Hoplitocoris_sp, Ceratocombus_sp, Mesovelia_mulsanti,
Rhagovelia_obesa, Limnoporus_canaliculatus, Hydrometra_stagnorum, Aquarius_paludum, Velia_caprai, Belostoma_flumineum,
Gelastocoris_oculatus, Trichocorixa_calva, Buenoa_margaritacea, Ranatra_linearis, Aphelocheirus_aestivalis, Diplonychus_rusticus,
Corixa_punctata, Ilyocoris_cimicoides, Notonecta_glauca, Plea_minutissima, Saldula_saltatoria, Cimex_lectularius, Arilus_cristatus,
Phymata_pennsylvanica, Rhodnius_prolixus, Corythucha_ciliata, Aphelonotus_fraterculus, Lygus_lineolaris, Reuteroscopus_ornatus,
Lopidea_amorphae, Notostira_elongata, Nabis_subcoleoptratus, Pagasa_sp, Orius_insidiosus, Arocatus_melanocephalus,
Lygaeus_equestris, Metatropis_rufescens, Pyrrhocoris_apterus, Aradus_betulae, Mezira_granulata, Sehirus_cincutus,
Corimelaena_lateralis, Piezosternum_callidum, Chinavia_hilaris, Megacopta_cribaria, Scutelleridae_sp, Anoplogonius_nigricollis,
Chalcocoris_rutilans, Jalysus_sp, Lygaeus_turcicus, Oncopeltus_fasciatus, Ischnodemus_falicus, Geocoris_sp, Chauliops_fallax,
Largus_californicus_1, Largus_californicus_2, Alydus_pilosus, Anasa_tristis, Boisea_trivittata, Acanthosoma_haemorrhoidale

Group 4 (G4): Holometabola (11 species)

Corydalus_cornutus, Osmylus_fulvicephalus, Euroleon_nostras, Inocellia_crassicornis, Subilla_sp, Aleochara_curtula,
Tribolium_castaneum, Gyrinus_marinus, Orussus_abietinus, Tenthredo_koehleri, Acromyrmex_echinatior

Hypothesis (*iii*): Coleorrhyncha are the sister group of Auchenorrhyncha

Group 1 (G1): Coleorrhyncha (3 species)

Xenophyes_metoponcus, Xenophysella_greensladeae, Peloridium_hammoniorum

Group 2 (G2): Auchenorrhyncha (34 species)

Cercopis_vulnerata, Aphrophora_alni, Philaenus_spumarius, Prosapia_bicincta, Tibicen_dorsata, Okanagana_villosa, Dalbulus_maidis,
Ponana_quadralaba, Agallia_constricta, Hespenedra_chilensis,Empoasca_fabae, Vidanoana_flavomaculata, Llanquihuea_pilosa,
Stictocephala_bisonia, Campylenchia_latipes, Nessorhinus_gibberulus, Holdgatiella_chepuensis, Centrotus_cornutus,
Graphocephala_fennahi, Ulopa_reticulata, Mapuchea_sp, Dictyophara_europaea, Tachycixius_pilosus, Nilaparvata_lugens,
Metcalfa_pruinosa, Caliscelis_bonelli, Acanalonia_conica, Bruchomorpha_oculata, Oliarus_placitus, Idiosystatus_acutiusculus,
Phylloscelis_atra, Taractellus_chilensis, Cyrpoptus_belfragei, Ladella_sp

Group 3 (G3): Heteroptera (59 species)

Hoplitocoris_sp, Ceratocombus_sp, Mesovelia_mulsanti, Rhagovelia_obesa, Limnoporus_canaliculatus, Hydrometra_stagnorum,
Aquarius_paludum, Velia_caprai, Belostoma_flumineum, Gelastocoris_oculatus, Trichocorixa_calva, Buenoa_margaritacea,
Ranatra_linearis, Aphelocheirus_aestivalis, Diplonychus_rusticus, Corixa_punctata, Ilyocoris_cimicoides, Notonecta_glauca,
Plea_minutissima, Saldula_saltatoria, Cimex_lectularius, Arilus_cristatus, Phymata_pennsylvanica, Rhodnius_prolixus,
Corythucha_ciliata, Aphelonotus_fraterculus, Lygus_lineolaris, Reuteroscopus_ornatus, Lopidea_amorphae, Notostira_elongata,
Nabis_subcoleoptratus, Pagasa_sp, Orius_insidiosus, Arocatus_melanocephalus, Lygaeus_equestris, Metatropis_rufescens,
Pyrrhocoris_apterus, Aradus_betulae, Mezira_granulata, Sehirus_cincutus, Corimelaena_lateralis, Piezosternum_callidum,
Chinavia_hilaris, Megacopta_cribaria, Scutelleridae_sp, Anoplogonius_nigricollis, Chalcocoris_rutilans, Jalysus_sp, Lygaeus_turcicus,
Oncopeltus_fasciatus, Ischnodemus_falicus, Geocoris_sp, Chauliops_fallax, Largus_californicus_1, Largus_californicus_2,
Alydus_pilosus, Anasa_tristis, Boisea_trivittata, Acanthosoma_haemorrhoidale

Group 4 (G4): Sternorrhyncha (25 species)

Diaphorina_citri, Trionymus_caricis, Macrohomotoma_gladiata, Psyllopsis_faxinicola, Heteropsylla_cubana, Pachypsylla_celtidismamma,
Aphalara_polygoni, Trioza_urticae, Glycaspis_brimblecombei, Acanthocasuarina_muellerianae, Trialeurodes_vaporariorum,
Daktulosphaira_vitifoliae, Planococcus_citri, Aclerda_sp, Ceroplastes_cirripediformis, Coccus_sp, Dactylopius_confuses,
Chrysomphalus_aonidum, Acutapsis_umbonifera, Unaspis_euonymi, Essigella_californica, Prociphilus_tesselatus, Schizaphis_graminum,
Uroleucon_ambrosiae, Tuberolachnus_salignus

Hypothesis (*iv*): Myerslopiidae are the sister group of the remainder of Membracoidea

Group 1 (G1): Cercopoidea (4 species)

Cercopis_vulnerata, Aphrophora_alni, Philaenus_spumarius, Prosapia_bicincta

Group 2 (G2): Cicadoidea (2 species)

Tibicen_dorsata, Okanagana_villosa

Group 3 (G3): Membracoidea minus Myerslopiidae (14 species)

Dalbulus_maidis, Ponana_quadralaba, Agallia_constricta, Hespenedra_chilensis, Empoasca_fabae, Vidanoana_flavomaculata,
Llanquihuea_pilosa, Stictocephala_bisonia, Campylenchia_latipes, Nessorhinus_gibberulus, Holdgatiella_chepuensis, Centrotus_cornutus,
Graphocephala_fennahi, Ulopa_reticulata

Group 4 (G4): Myerslopiidae (1 species)

Mapuchea_sp

Hypothesis (*v*): Aphidoidea are the sister group of Coccoidea

Group 1 (G1): Psylloidea (10 species)

Diaphorina_citri, Trionymus_caricis, Macrohomotoma_gladiata, Psyllopsis_faxinicola, Heteropsylla_cubana, Pachypsylla_celtidismamma, Aphalara_polygoni, Trioza_urticae, Glycaspis_brimblecombei, Acanthocasuarina_muellerianae

Group 2 (G2): Aleyrodoidea (1 species)

Trialeurodes_vaporariorum

Group 3 (G3): Coccoidea (9 species)

Daktulosphaira_vitifoliae, Planococcus_citri, Aclerda_sp, Ceroplastes_cirripediformis, Coccus_sp, Dactylopius_confuses, Chrysomphalus_aonidum, Acutapsis_umbonifera, Unaspis_euonymi

Group 4 (G4): Aphidoidea (5 species)

Essigella_californica, Prociphilus_tesselatus, Schizaphis_graminum, Uroleucon_ambrosiae, Tuberolachnus_salignus

Hypothesis (*vi*): Nepomorpha are the sister group of a clade including Leptopodomorpha, Cimicomorpha, and Pentatomomorpha

Group 1 (G1): Auchenorrhyncha+Coleorrhyncha (37 species)

Cercopis_vulnerata, Aphrophora_alni, Philaenus_spumarius, Prosapia_bicincta, Tibicen_dorsata, Okanagana_villosa, Dalbulus_maidis, Ponana_quadralaba, Agallia_constricta, Hespenedra_chilensis,Empoasca_fabae, Vidanoana_flavomaculata, Llanquihuea_pilosa, Stictocephala_bisonia, Campylenchia_latipes, Nessorhinus_gibberulus, Holdgatiella_chepuensis, Centrotus_cornutus, Graphocephala_fennahi, Ulopa_reticulata, Mapuchea_sp, Dictyophara_europaea, Tachycixius_pilosus, Nilaparvata_lugens, Metcalfa_pruinosa, Caliscelis_bonelli, Acanalonia_conica, Bruchomorpha_oculata, Oliarus_placitus, Idiosystatus_acutiusculus, Phylloscelis_atra, Taractellus_chilensis, Cyrpoptus_belfragei, Ladella_sp, Xenophyes_metoponcus, Xenophysella_greensladeae, Peloridium_hammoniorum

Group 2 (G2): Enicocephalomorpha+Dipsocoromorpha+Gerromorpha (8 species)

Hoplitocoris_sp, Ceratocombus_sp, Mesovelia_mulsanti, Rhagovelia_obesa, Limnoporus_canaliculatus, Hydrometra_stagnorum, Aquarius_paludum, Velia_caprai

Group 3 (G3): Nepomorpha (11 species)

Belostoma_flumineum, Gelastocoris_oculatus, Trichocorixa_calva, Buenoa_margaritacea, Ranatra_linearis, Aphelocheirus_aestivalis, Diplonychus_rusticus, Corixa_punctata, Ilyocoris_cimicoides, Notonecta_glauca, Plea_minutissima

Group 4 (G4): Leptopodomorpha+Cimicomorpha+Pentatomomorpha (40 species)

Saldula_saltatoria, Cimex_lectularius, Arilus_cristatus, Phymata_pennsylvanica, Rhodnius_prolixus, Corythucha_ciliata, Aphelonotus_fraterculus, Lygus_lineolaris, Reuteroscopus_ornatus, Lopidea_amorphae, Notostira_elongata, Nabis_subcoleoptratus, Pagasa_sp, Orius_insidiosus, Arocatus_melanocephalus, Lygaeus_equestris, Metatropis_rufescens, Pyrrhocoris_apterus, Aradus_betulae, Mezira_granulata, Sehirus_cincutus, Corimelaena_lateralis, Piezosternum_callidum, Chinavia_hilaris, Megacopta_cribaria, Scutelleridae_sp, Anoplogonius_nigricollis, Chalcocoris_rutilans, Jalysis_sp, Lygaeus_turcicus, Oncopeltus_fasciatus, Ischnodemus_falicus, Geocoris_sp, Chauliops_fallax, Largus_californicus_1, Largus_californicus_2, Alydus_pilosus, Anasa_tristis, Boisea_trivittata, Acanthosoma_haemorrhoidale

Hypothesis (*vii*): Leptopodomorpha sister to a clade comprising Cimicomorpha and Pentatomomorpha

Group 1 (G1): Cimicomorpha+Pentatomomorpha (39 species)

Arocatus_melanocephalus, Lygaeus_equestris, Metatropis_rufescens, Pyrrhocoris_apterus, Aradus_betulae, Mezira_granulata, Sehirus_cincutus, Corimelaena_lateralis, Piezosternum_callidum, Chinavia_hilaris, Megacopta_cribaria, Scutelleridae_sp, Anoplogonius_nigricollis, Chalcocoris_rutilans, Jalysus_sp, Lygaeus_turcicus, Oncopeltus_fasciatus, Ischnodemus_falicus, Geocoris_sp, Chauliops_fallax, Largus_californicus_1, Largus_californicus_2, Alydus_pilosus, Anasa_tristis, Boisea_trivittata, Acanthosoma_haemorrhoidale, Cimex_lectularius, Arilus_cristatus, Phymata_pennsylvanica, Rhodnius_prolixus, Corythucha_ciliata, Aphelonotus_fraterculus, Lygus_lineolaris, Reuteroscopus_ornatus, Lopidea_amorphae, Notostira_elongata, Nabis_subcoleoptratus, Pagasa_sp, Orius_insidiosus

Group 2 (G2): Leptopodomorpha (1 species)

Saldula_saltatoria

Group 3 (G3): Nepomorpha (11 species)

Belostoma_flumineum, Gelastocoris_oculatus, Trichocorixa_calva, Buenoa_margaritacea, Ranatra_linearis, Aphelocheirus_aestivalis, Diplonychus_rusticus, Corixa_punctata, Ilyocoris_cimicoides, Notonecta_glauca, Plea_minutissima

Group 4 (G4): Enicocephalomorpha+Dipsocoromorpha+Gerromorpha (8 species)

Hoplitocoris_sp, Ceratocombus_sp, Mesovelia_mulsanti, Rhagovelia_obesa, Limnoporus_canaliculatus, Hydrometra_stagnorum, Aquarius_paludum, Velia_caprai

Hypothesis (*viii*): Gerromorpha sister to a clade including Dipsocoromorpha and Enicocephalomorpha

Hoplitocoris_sp, Ceratocombus_sp

Group 2 (G2): Gerromorpha (6 species)

Mesovelia_mulsanti, Rhagovelia_obesa, Limnoporus_canaliculatus, Hydrometra_stagnorum, Aquarius_paludum, Velia_caprai

Group 3 (G3): Nepomorpha+Letopodomorpha+Cimicomorpha+Pentatomomorpha (51 species)

Belostoma_flumineum, Gelastocoris_oculatus, Trichocorixa_calva, Buenoa_margaritacea, Ranatra_linearis, Aphelocheirus_aestivalis, Diplonychus_rusticus, Corixa_punctata, Ilyocoris_cimicoides, Notonecta_glauca, Plea_minutissima, Saldula_saltatoria, Arocatus_melanocephalus, Lygaeus_equestris, Metatropis_rufescens, Pyrrhocoris_apterus, Aradus_betulae, Mezira_granulata, Sehirus_cincutus, Corimelaena_lateralis, Piezosternum_callidum, Chinavia_hilaris, Megacopta_cribaria, Scutelleridae_sp, Anoplogonius_nigricollis, Chalcocoris_rutilans, Jalysus_sp, Lygaeus_turcicus, Oncopeltus_fasciatus, Ischnodemus_falicus, Geocoris_sp, Chauliops_fallax, Largus_californicus_1, Largus_californicus_2, Alydus_pilosus, Anasa_tristis, Boisea_trivittata, Acanthosoma_haemorrhoidale, Cimex_lectularius, Arilus_cristatus, Phymata_pennsylvanica, Rhodnius_prolixus, Corythucha_ciliata, Aphelonotus_fraterculus, Lygus_lineolaris, Reuteroscopus_ornatus, Lopidea_amorphae, Notostira_elongata, Nabis_subcoleoptratus, Pagasa_sp, Orius_insidiosus

Group 4 (G4): Auchenorrhyncha+Coleorrhyncha (37 species)

Cercopis_vulnerata, Aphrophora_alni, Philaenus_spumarius, Prosapia_bicincta, Tibicen_dorsata, Okanagana_villosa, Dalbulus_maidis, Ponana_quadralaba, Agallia_constricta, Hespenedra_chilensis,Empoasca_fabae, Vidanoana_flavomaculata, Llanquihuea_pilosa, Stictocephala_bisonia, Campylenchia_latipes, Nessorhinus_gibberulus, Holdgatiella_chepuensis, Centrotus_cornutus, Graphocephala_fennahi, Ulopa_reticulata, Mapuchea_sp, Dictyophara_europaea, Tachycixius_pilosus, Nilaparvata_lugens, Metcalfa_pruinosa, Caliscelis_bonelli, Acanalonia_conica, Bruchomorpha_oculata, Oliarus_placitus, Idiosystatus_acutiusculus, Phylloscelis_atra, Taractellus_chilensis, Cyrpoptus_belfragei, Ladella_sp, Xenophyes_metoponcus, Xenophysella_greensladeae, Peloridium_hammoniorum

Hypothesis (*ix*): Cercopoidea sister to Cicadoidea

Group 1 (G1): Membracodea (15 species)

Dalbulus_maidis, Ponana_quadralaba, Agallia_constricta, Hespenedra_chilensis, Empoasca_fabae, Vidanoana_flavomaculata, Llanquihuea_pilosa, Stictocephala_bisonia, Campylenchia_latipes, Nessorhinus_gibberulus, Holdgatiella_chepuensis, Centrotus_cornutus, Graphocephala_fennahi, Ulopa_reticulata, Mapuchea_sp

Group 2 (G2): Cercopoidea (4 species)

Cercopis_vulnerata, Aphrophora_alni, Philaenus_spumarius, Prosapia_bicincta

Group 3 (G3): Cicadoidea (2 species)

Tibicen_dorsata, Okanagana_villosa

Group 4 (G4): Fulgoroidea (13 species)

Dictyophara_europaea, Tachycixius_pilosus, Nilaparvata_lugens, Metcalfa_pruinosa, Caliscelis_bonelli, Acanalonia_conica, Bruchomorpha_oculata, Oliarus_placitus, Idiosystatus_acutiusculus, Phylloscelis_atra, Taractellus_chilensis, Cyrpoptus_belfragei, Ladella_sp

Table S6. Results from Four-Cluster Likelihood Mapping (FcLM), including permutation tests.

Hypothesis (*i*): Psocodea are the sister group of Holometabola

(*ia*) All Condylognatha included

quartets: 747,813

aa sites: 859,005

partitions: 1,856

Groups: G1: Psocodea, G2: Condylognatha, G3: Holometabola, G4: outgroups

T1: Psocodea,Condylognatha | Holometabola,Outgroups

T2: Psocodea,Holometabola | Condylognatha,Outgroups

T3: Psocodea,Outgroups | Condylognatha,Holometabola

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	67%	25%	8%	0	0	0	0
Perm I	15%	7%	5%	8%	2%	1%	61%
Perm II	6%	4%	4%	5%	6%	6%	68%
Perm III	0%	0%	0%	4%	4%	4%	86%

Results from modifications excluding quartets including representatives of the following taxa from the results of (*ia*) (original data only):

(*iaI*) Excluding all Phthiraptera

quartets: 554,829

T1: 67%, T2: 26%, T3: 7%, T12, T13, T23, T*: 0%

(*iaII*) Excluding all Phthiraptera plus Liposcelidae

quartets: 506,583

T1: 67%, T2: 27%, T3: 6%, T12, T13, T23, T*: 0%

(*iaIII*) Excluding all "Psocoptera"

quartets: 192,984

T1: 66%, T2: 22%, T3: 12%, T12, T13, T23, T*: 0%

(*iaIV*) Excluding all Phthiraptera from G1 and all Thysanoptera from G2

quartets: 520,421

T1: 68%, T2: 25%, T3: 7%, T12, T13, T23, T*: 0%

(*ib*) With Thysanoptera excluded

quartets: 701,437

aa sites: 858,762

partitions: 1,856

Groups: G1: Psocodea, G2: Hemiptera, G3: Holometabola, G4: outgroups

T1: Psocodea,Hemiptera | Holometabola,Outgroups

T2: Psocodea,Holometabola | Hemiptera,Outgroups

T3: Psocodea,Outgroups | Hemiptera,Holometabola

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	68%	24%	8%	0%	0%	0%	0%
Perm I	14%	8%	6%	8%	2%	1%	60%
Perm II	5%	5%	5%	5%	6%	6%	67%
Perm III	0%	0%	0%	4%	4%	4%	86%

(*ic*) With Hemiptera excluded

quartets: 46,376

aa sites: 846,928

partitions: 1,831

Groups: G1: Psocodea, G2: Thysanoptera, G3: Holometabola, G4: outgroups

T1: Psocodea, Thysanoptera | Holometabola,Outgroups

T2: Psocodea,Holometabola | Thysanoptera,Outgroups

T3: Psocodea,Outgroups | Thysanoptera,Holometabola

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	51%	46%	3%	0%	0%	0%	0%
Perm I	20%	5%	4%	5%	1%	1%	64%
Perm II	11%	17%	16%	11%	9%	12%	24%
Perm III	0%	0%	0%	6%	5%	5%	84%

Hypothesis (*ii*): Thysanoptera are the sister group of Hemiptera

quartets: 330,088

aa sites: 846,234

partitions: 1,806

Groups: G1: Psocodea, G2: Thysanoptera, G3: Hemiptera, G4: Holometabola

T1: Psocodea,Thysanoptera | Hemiptera,Holometabola

T2: Psocodea,Hemiptera | Thysanoptera,Holometabola

T3: Psocodea,Holometabola | Thysanoptera,Hemiptera

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	7%	25%	68%	0%	0%	0%	0%
Perm I	19%	6%	10%	2%	3%	10%	51%
Perm II	5%	4%	4%	5%	5%	6%	70%
Perm III	0%	0%	0%	5%	5%	4%	85%

Hypothesis (*iii*): Coleorrhyncha are the sister group of Auchenorrhyncha

quartets: 150,450

aa sites: 833,738

partitions: 1,757

Groups: G1: Coleorrhyncha, G2: Auchenorrhyncha, G3: Heteroptera, G4: Sternorrhyncha

T1: Coleorrhyncha,Auchenorrhyncha | Heteroptera,Sternorrhyncha

T2: Coleorrhyncha,Heteroptera | Auchenorrhyncha,Sternorrhyncha

T3: Coleorrhyncha,Sternorrhyncha | Auchenorrhyncha,Heteroptera

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	96%	4%	0%	0%	0%	0%	0%
Perm I	13%	8%	7%	5%	2%	10%	56%
Perm II	6%	7%	8%	8%	8%	8%	54%
Perm III	0%	0%	0%	4%	4%	4%	86%

Hypothesis (*iv*): Myerslopiidae are the sister group of the remainder of Membracoidea

quartets: 112

aa sites: 779,558

partitions: 1,659

Groups: G1: Cercopoidea, G2: Cicadoidea, G3: Membracoidea minus Myerslopiidae, G4: Myerslopiidae

T1: Cercopoidea,Cicadoidea | Membracoidea minus Myerslopiidae,Myerslopiidae

T2: Cercopoidea,Membracoidea minus Myerslopiidae | Cicadoidea,Myerslopiidae

T3: Cercopoidea,Myerslopiidae | Cicadoidea,Membracoidea minus Myerslopiidae

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	100%	0%	0%	0%	0%	0%	0%
Perm I	7%	12%	2%	2%	0%	10%	68%
Perm II	21%	18%	46%	4%	6%	1%	5%
Perm III	31%	37%	11%	6%	3%	11%	2%

Hypothesis (*v*): Aphidoidea are the sister group of Coccoidea

quartets: 450

aa sites: 780,773

partitions: 1,610

Groups: G1: Psylloidea, G2: Aleyrodoidea, G3: Coccoidea, G4: Aphidoidea

T1: Psylloidea,Aleyrodoidea | Coccoidea,Aphidoidea

T2: Psylloidea, Coccoidea | Aleyrodoidea,Aphidoidea

T3: Psylloidea,Aphidoidea | Aleyrodoidea, Coccoidea

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	91%	9%	0%	0%	0%	0%	0%
Perm I	5%	5%	5%	1%	0%	1%	82%
PermII	23%	19%	37%	5%	5%	8%	4%
PermIII	23%	33%	21%	8%	7%	8%	2%

Hypothesis (*vi*): Nepomorpha are the sister group of a clade including Leptopodomorpha, Cimicomorpha, and Pentatomomorpha

quartets: 130,240

aa sites: 841,766

partitions: 1,817

Groups: G1: Auchenorrhyncha+Coleorrhyncha,

G2: Enicocephalomorpha+Dipsocoromorpha+Gerromorpha, G3: Nepomorpha,

G4: Leptopodomorpha+Cimicomorpha+Pentatomomorpha

T1: Nepomorpha, Enicocephalomorpha+Dipsocoromorpha+Gerromorpha | Auchenorrhyncha+Coleorrhyncha, Leptopodomorpha+Cimicomorpha+Pentatomomorpha

T2: Nepomorpha, Auchenorrhyncha+Coleorrhyncha | Enicocephalomorpha+Dipsocoromorpha+Gerromorpha, Leptopodomorpha+Cimicomorpha+Pentatomomorpha

T3: Nepomorpha, Leptopodomorpha+Cimicomorpha+Pentatomomorpha | Enicocephalomorpha+Dipsocoromorpha+Gerromorpha, Auchenorrhyncha+Coleorrhyncha

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	31%	37%	32%	0%	0%	0%	0%
Perm I	13%	6%	20%	3%	6%	15%	38%
Perm II	7%	7%	10%	8%	9%	10%	49%
Perm III	0%	0%	0%	4%	4%	4%	87%

Hypothesis (*vii*): Leptopodomorpha sister to a clade comprising Cimicomorpha and Pentatomomorpha

quartets: 3432

aa sites: 770,296

partitions: 1,502

Groups: G1: Cimicomorpha+Pentatomomorpha, G2: Leptopodomorpha, G3: Nepomorpha, G4: Gerromorpha+Dipsocoromorpha+Enicocephalomorpha

T1: Cimicomorpha+Pentatomomorpha, Leptopodomorpha | Nepomorpha, Gerromorpha+Dipsocoromorpha+Enicocephalomorpha

T2: Cimicomorpha+Pentatomomorpha, Nepomorpha | Leptopodomorpha, Gerromorpha+Dipsocoromorpha+Enicocephalomorpha

T3: Cimicomorpha+Pentatomomorpha, Gerromorpha+Dipsocoromorpha+Enicocephalomorpha | Leptopodomorpha, Nepomorpha

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	100%	0%	0%	0%	0%	0%	0%
Perm I	6%	5%	12%	20%	3%	3%	61%
Perm II	21%	8%	15%	10%	15%	11%	21%
Perm III	9%	15%	24%	9%	9%	10%	23%

Hypothesis (*viii*): Gerromorpha sister to a clade including Dipsocoromorpha and Enicocephalomorpha

quartets: 22,032

aa sites: 828,041

partitions: 1,725

Groups: G1: Enicocephalomorpha+Dipsocoromorpha, G2: Gerromorpha, G3: Nepomorpha+Leptopodomorpha+Cimicomorpha+Pentatomomorpha G4: Auchenorrhyncha+Coleorrhyncha

T1: Enicocephalomorpha+Dipsocoromorpha, Gerromorpha | Nepomorpha+Leptopodomorpha+Cimicomorpha+Pentatomomorpha, Auchenorrhyncha+Coleorrhyncha

T2: Enicocephalomorpha+Dipsocoromorpha, Nepomorpha+Leptopodomorpha+Cimicomorpha+Pentatomomorpha | Gerromorpha, Auchenorrhyncha+Coleorrhyncha

T3: Enicocephalomorpha+Dipsocoromorpha, Auchenorrhyncha+Coleorrhyncha | Gerromorpha, Nepomorpha+Leptopodomorpha+Cimicomorpha+Pentatomomorpha

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	100%	0%	0%	0%	0%	0%	0%
Perm I	11%	5%	10%	12%	5%	7%	51%
Perm II	8%	10%	11%	9%	10%	14%	39%
Perm III	0%	0%	0%	4%	4%	4%	87%

Hypothesis (*ix*): Cercopoidea sister to Cicadoidea

quartets: 1,456

aa sites: 829,781

partitions: 1,718

Groups: G1: Membracoidea, G2: Cercopoidea, G3: Cicadoidea G4: Fulgoroidea

T1: Membracoidea, Cercopoidea | Cicadoidea, Fulgoroidea

T2: Membracoidea, Cicadoidea | Cercopoidea, Fulgoroidea

T3: Membracoidea, Fulgoroidea | Cercopoidea, Cicadoidea

data	T1: G1,G2|G3,G4	T2: G1,G3|G2,G4	T3: G1,G4|G2,G3	T12	T13	T23	T*
original	48%	0%	52%	0%	0%	0%	0%
Perm I	16%	6%	6%	2%	3%	0%	68%
Perm II	14%	18%	36%	6%	12%	8%	6%
Perm III	14%	18%	29%	7%	9%	10%	13%

Table S7. Information on fossils used to calibrate the inferred phylogenetic tree and to estimate divergence times between major lineages.

Fossil Taxon	Current Placement	Age Estimate	Citation
Liposcelis sp.	Troctomorpha: Liposcelididae	53 MYA	41
Psocidus multiplex	Psocomorpha: Psocidae	33.9 MYA	42
Liposcelis sp.	Troctomorpha: Liposcelididae	53 MYA	41
Psocidus multiplex	Psocomorpha: Psocidae	33.9 MYA	42
Leaphis prima	Hemiptera: Sternorrhyncha: Aphidomorpha	245 MYA	43
Surijokocixius tomiensis	Hemiptera: Fulgoromorpha: Surijokocixiidae	251 MYA	44
Paraprosbole rotruda	Cicadoidea: Tettigarctidae	200 MYA	45
Undescribed	Hemiptera: Fulgoromorpha: Cixiidae	150 MYA	46
Priscoflata subvexa	Hemiptera: Fulgoromorpha: Flatidae	60 MYA	47
Wedelphus dichopteroides	Hemiptera: Fulgoromorpha: Fulgoridae	44 MYA	48
Triassojassus proavitus	Archijassidae: Archijassinae	220 MYA	49
Proerrhomus rugosus	Cicadellidae: Proerrhominae	118 MYA	50
Enicocephalinus acragrimaldii	Hemiptera: Enicocephalomorpha: Enicocephalidae	120 MYA	51
Gallomesovelia grioti	Hemiptera: Gerromorpha: Mesoveliidae	150MYA	52
Cretogerris albianus	Hemiptera: Gerromorpha: Gerridae	100 MYA	53
Shurabella lepyroniopsis	Hemiptera: Nepmorpha: Corixoidea: Shurabellidae	180 MYA	54
Mesolygaeus laiyangensis	Hemiptera: Leptopodomorpha: Saldoidea: Archegocimicidae	200 MYA	55
Ceresopsis costalis	Hemiptera: Cimicomorpha: Reduvioidea: Ceresopsidae	180 MYA	45
Vetanthocoris decorus	Hemiptera: Cimicomorpha: Cimoicoidea: Vetanthocoridae	150 MYA	56
Miridoides mesozoicus	Hemiptera: Cimicomorpha: Miridae	150 MYA	57
Cretopiesma suukyae	Hemiptera: Pentatomomorpha: Aradidae	99 MYA	58

References

1. Misof B, et al. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346:763-767.
2. Xie Y, et al. (2014). SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Bioinf* 30:1660–1666.
3. Peters RS, et al. (2017) Evolutionary history of Hymenoptera. *Current Biol* 27:1-6.
4. Mayer C, et al. (2016) BaitFisher: a software package for multispecies target DNA enrichment probe design. *Mol Biol Evol* 33:1875–1886.
5. Camacho C., et al. (2009). BLAST+: architecture and applications. *BMC Bioinf* 10:421.
6. Petersen M, et al. (2017) Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. *BMC Bioinf* 18:111.
7. Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva EV (2013) OrthoDB: A hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Res* 41(D1):1–8.
8. Colbourne JK, et al. (2011) The ecoresponsive genome of *Daphnia pulex*. *Science* 331:555–561.
9. Terrapon N, et al. (2014) Molecular traces of alternative social organization in a termite genome. *Nat Commun* 5(May). doi:10.1038/ncomms4636.
10. Kirkness EF, et al. (2010) Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc Natl Acad Sci* 107:12168–12173.
11. Mesquita RD, et al. (2015) Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc Natl Acad Sci* 112:14936–14941.
12. Richards S, et al. (2008) The genome of the model beetle and pest *Tribolium castaneum*. *Nature* 452:949–955.
13. Nygaard S, et al. (2011) The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming. *Genome Res* 21:1339–48.
14. Richards S, et al. (2010) Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol* 8(2). doi:10.1371/journal.pbio.1000313.
15. Eddy SR (2011) Accelerated profile HMM searches. *PLOS Comput Biol* 7:e1002195.
16. Slater GS, Birney E (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinf* 6:31.
17. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780.
18. Suyama M, Torrents D, Bork P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nuc Acids Res* 34:W609-W612.
19. Finn RD, et al. (2006) Pfam: clans, web tools and services. *Nuc Acids Res* 34:D247–D251.
20. Finn RD, et al. (2014) Pfam: the protein families database. *Nuc Acids Res*. 42:D222–D230.
21. Misof B, Misof K (2009) A Monte Carlo approach successfully identifies randomness in multiple sequence alignments : A more objective means of data exclusion. *Syst Biol* 58:21-34.
22. Kück P., et al. (2010) Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Front Zool* 7:10.
23. Meusemann K., et al. (2010) A phylogenomic approach to resolve the arthropod tree of life. *Mol Biol Evol* 27:2451–2464.
24. Misof B, et al. (2013). Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. *BMC Bioinf* 14:348.

25. Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A (2014) Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol Biol* 14:82.
26. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B (2017) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol* 34:772-773.
27. Hurvich CM, Tsai C-L (1989) Regression and time series model selection in small samples. *Biometrika* 76:297–307.
28. Jermiin L, Ho SY, Ababneh F, Robinson J, Larkum AW (2004) The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. *Syst Biol* 53:638–643.
29. Jermiin L, Ott M (2017) SymTest version 2.0.47 (https://github.com/ottmi/symtest).
30. Wong T, Jermiin L (2017) AliStat v.1.6 (https://github.com/thomaskf/AliStat).
31. Kozlov AM, Aberer AJ, Stamatakis A (2015) ExaML version 3: a tool for phylogenomic analyses on supercomputers. *Bioinf* 31:2577–2579.
32. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies, *Bioinf* 30:1312–1313.
33. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evol* 39:783–791.
34. Pattengale ND, Alipour M, Bininda-Emonds OR, Moret BM, Stamatakis A (2010) How many bootstrap replicates are necessary? *J Comput Biol* 17:337–354.
35. Dell'Ampio E, et al. (2014) Decisive data sets in phylogenomics: Lessons from studies on the phylogenetic relationships of primarily wingless insects. *Mol Biol Evol* 31:239–249.
36. Strimmer K, von Haeseler A (1997) Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *PNAS* 94:6815–6819.
37. Parham JF, et al. (2012) Best practices for justifying fossil calibrations. *Syst Biol* 61:346-359.
38. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972-1973.
39. Zhong et al. 2009 https://doi.org/10.1371/journal.pone.0005297
40. Rambaut A, Drummond AJ (2007) Tracer v1.4. beast.bio.ed.ac.uk/Tracer
41. Yoshizawa K, Lienhard C (2010) In search of the sister group of the true lice: a systematic review of booklice and their relatives, with an updated checklist of Liposcelididae (Insecta: Psocodea). *Arthrop Syst Phylog* 68:181-195.
42. Engel MS, Perkovsky EE (2006) Psocoptera (Insecta) in Eocene Rovno amber (Ukraine). *Vestnik Zoologii* 40:175-179.
43. Szwedo J, Nel A (2011) The oldest aphid insect from the Middle Triassic of the Vosges, France. *Acta Palaeontologica Polonica* 56:757-776.
44. Shcherbakov DE (2000) Permian faunas of Homoptera (Hemiptera) in relation to phytogeography and the Permo-Triassic crisis. *Paleontol J* 34:S251-267.
45. Shcherbakov DE (2008) Review of the fossil and extant genera of the cicada family Tettigarctidae (Hemiptera: Cicadoidea). *Russian Ent J* 17:343-348.
46. Shcherbakov DE (2004) On Permian and Mesozoic Fulgoroidea. Pp. 68-70 In Kerzhner IM (ed). *Third European Hemiptera Congress Abstracts*. Russian Academy of Sciences, St. Petersburg.
47. Szwedo J, Stroinsky A, Lin Q-B (2013) Discovery of a Flatidae planthopper (Hemiptera: Fulgoromorpha) in the Paleocene of Northern Tibet and its taxonomic and biogeographic significance. *Geodiversitas* 35:767-776.

48. Szwedo J, Wappler T (2006) New planthoppers (Insecta: Hemiptera: Fulgoromorpha) from the Middle Eocene Messel Maar. *Annales Zoologici* 56:555-566.

49. Shcherbakov DE (2012) More on Mesozoic Membracoidea (Homoptera). *Russian Entomol J* 21:15-22.

50. Hamilton KGA (1990) Homoptera. In: Grimaldi DA (ed). Insects from the Santana Formation. Lower Cretaceous of Brazil. *Bulletin of the American Museum of Natural History* 195:82–122.

51. Azar D, Fleck G, Nel A, Solignac M (1999) A new enicocephalid bug, *Enicocephalinus acragrimaldii*, gen. n., sp. n., from the lower Cretaceous amber of Lebanon (Insecta, Heteroptera, Enicocephalidae). *Estudios del Museo de Ciencias Naturales de Alava* 14 (Número especial 2):217-230.

52. Nel A, Nel P, Krieg-Jacquier R, Pouillon J-M, Garrouste R (2014) Exceptionally preserved insect fossils in the Late Jurassic lagoon of Orbagnoux (Rhone Valley, France). *PeerJ* 2:e510.

53. Perrichot V, Nel A, Neraudeau D (2005) Gerromorphan bugs in Early Cretaceous French amber (Insecta: Heteroptera): first representatives of Gerridae and their phylogenetic and palaeoecological implications. *Cretaceous Res* 26:793-800.

54. Popov YA (1971) *Historical development of true bugs of the infraorder Nepomorpha (Heteroptera)*. Paleontological Institute, USSR Academy of Sciences, Moscow.

55. Zhang W, Engel MS, Yao Y, Ren D, Shih C (2014) The Mesozoic family Archegocimicidae and phylogeny of the infraorder Leptopodomorpha (Hemiptera). J Syst Paleontol 12:93-111.

56. Yao Y, Cai W, Ren D (2006) Fossil flower bugs (Heteroptera: Cimicomorpha: Cimicoidea) from the Late Jurassic of Northeast China, including a new family, Vetanthocoridae. *Zootaxa* 1360:1-40.

57. Becker-Migdisova YE, Popov YA (1963) New Heteroptera from Jurassic deposits of Karatau region. *Paleontologicheskiy Zhurnal* 2:74-82.

58. Cassis G, Schuh RT (2010) Systematic methods, fossils, and relationships within Heteroptera (Insecta). *Cladistics* 26:262-280.