BRILL

Insect Systematics & Evolution XX (2024) 1–14

brill.com/ise

Holarctic Lineages Cannot Inform Diversity and Evolution in the Neotropics – the barklice family Psocidae as a case study

Valentina Sarria-Rodríguez[1], Ranulfo Gonzalez-Obando[1], Nelson Rivera-Franco[1,2], Heiber Cardenas-Henao[1], and Cristian Román-Palacios[3*]

[1]Departamento de Biología, Facultad de Ciencias Exactas y Naturales, Universidad del Valle, A. A. 25360, Cali, Valle del Cauca, Colombia
[2]TAO-Lab, Centre for Bioinformatics and Photonics-CIBioFi, Universidad del Valle, Calle 13 # 100-00, Edif. E20, No. 1069, 760032, Cali, Colombia
[3]School of Information, University of Arizona, Tucson, Arizona 85721, USA
*Corresponding author: cromanpa94@arizona.edu

Abstract
Despite tropical species comprising nearly 60% of Psocidae species, previous studies examining the Psocidae phylogeny have undersampled tropical diversity (<40% species in trees). Here we discuss the systematics of the Psocidae based on the most comprehensive species-level sampling of the Psocidae. We infer the phylogenetic position of 43 previously unsampled Neotropical species in the Psocidae phylogeny. We find that Neotropical psocids are generally not closely related to morphologically similar taxa in the Holarctic region. Consequently, the monophyletic status for the major groups within Psocidae (subfamilies and tribes) is recovered only when Holarctic groups are sampled (7–10 of 11 higher-level groups are monophyletic) but violated when Neotropical species are also sampled in the tree (1 of 11 higher-level groups are monophyletic). Our study pinpoints at the downfalls of simply extending taxonomic knowledge from lineages of a certain area (i.e. Holarctic) to inform diversity and evolution of lineages in other regions (i.e. Neotropics).

Keywords
Insecta; Phylogenetics; Neotropical; Barkflies

Introduction

With more than 1,000 species classified in 80 different genera, Psocidae is the largest extant family of free-living lice (Psocodea: 'Psocoptera'; Mockford, 1993; Johnson et al. 2020; Lienhard & Smithers 2002). Although many temperate species (>35° latitude) are currently described (Lienhard & Smithers 2002; Johnson et al. 2020), more than 60% of the family diversity is restricted the tropics (<35° latitude; Supplementary Text S1; Supplementary Table S1; Aldrete & Román-P, 2015; Román-Palacios et al. 2016;

Published with license by Koninklijke Brill NV | DOI: 10.1163/1876312X-BJA10055
 | ISSN: 1399-560X (print) 1876-312X (online)

Oliveira et al. 2017). Contrasting with this observation is the fact that current taxonomic proposals within the family are, however, strongly founded on Holarctic lineages. The implications of extrapolating systematic patterns detected in Holarctic diversity to understanding systematics and evolution at the family level, and specially for tropical lineages, remains unknown and poorly examined in this group.

Species within the Psocidae are currently classified under three subfamilies and ten tribes (Johnson et al. 2020; Lienhard & Smithers 2002; Yoshizawa & Johnson 2008). First, Kaindipsocinae accounts for 36 species (Yoshizawa 1998; Yoshizawa et al. 2011; Johnson et al. 2020). Second, Amphigerontiinae includes 235 species that are classified into three different tribes (Amphigerontini, Blastini, and Stylatopsocini; Yoshizawa 2010). Third, Psocinae includes nearly ~1,000 species that are currently classified under seven tribes ('Ptyctini', Psocini, Atrichadenotecnini, Sigmatoneurini, Metylophorini, Thyrsophorini, and Cycetini; Yoshizawa & Johnson 2008). Although many groups within the Psocidae were erected based on morphology (Yoshizawa 2002, 2005), the recent use of molecular data to study the Psocidae systematics has provided new insights on the natural groups that exist within the family (e.g. Yoshizawa & Johnson 2008).

To date, only a handful of molecular studies have examined the phylogenetic relationships between higher-level groups (subfamilies and tribes) within the Psocidae. For instance, Johnson & Mockford (2003) recovered the family-level monophyly and concluded the paraphyletic status of the Psocinae based on four gene regions (18S, 12S, 16S, and COI) sequenced from four Psocidae species (three Psocinae and a single Amphigerontiinae). More recently, Yoshizawa & Johnson (2008) presented the most comprehensive species-level phylogeny for the Psocidae published to date. This study was sampled on six gene regions (18S, 16S, 12S, COI, H3, and ND5) and 45 Psocidae species. Relative to the morphology-based classical taxonomy (Lienhard & Smithers 2002), Yoshizawa & Johnson (2008) erected a new tribe (Kaindipsocini), synonymized the Oriental Cerastipsocini (*Sigmatoneura* and *Podopterocus*) within Sigmatoneurini, and transferred the remaining Neotropical Cerastipsocini species into Thyrsophorini. Yoshizawa & Johnson (2008) also recovered the monophyly of Psocidae and the paraphyly of both Amphigerontiinae (due to the position of Kaindipsocini; but see below) and 'Ptyctini'. In a follow-up study by Yoshizawa et al. (2011), the taxonomic sampling for Kaindipsocini in Yoshizawa & Johnson (2008) was expanded to six new species. Yoshizawa et al. (2011) also re-defined the taxonomic limits within Amphigerontiinae by suggesting this subfamily to having only two tribes (Amphigerontini and Blastini) and erecting a new subfamily (Kaindipsocinae) from the previously known Kaindipsocini.

The systematics and evolution of Tropical psocids has been historically understood from studies primarily sampling Holarctic species. For instance, tropical lineages comprise 25% of the species sampled in Johnson & Mockford (2003; 1 of 4 taxa), ~17% in Yoshizawa & Johnson (2008; eight of 45), and ~16% in Yoshizawa et al. (2011; eight of 51). This bias in molecular studies towards sampling Holarctic groups questions the practical utility of previous phylogenetic hypotheses to inform the evolution and diversity of neotropical lineages.

In this study, we sequenced three gene regions for 43 Neotropical taxa that have not been sampled in previous molecular phylogenies. We inferred the phylogenetic relationships among psocid species using two molecular datasets that either included (i) only the species that were published sampled in studies of the Psocidae phylogeny, and (ii) or combined published and newly generated sequences for the Neotropical species analyzed in this study. We expected species in the same genera, tribes, and subfamilies (originally classified based on morphology) to be closely related in the Psocidae phylogeny regardless of whether they are neotropical or holarctic. Nevertheless, because taxonomy has been largely based on the morphology of Holarctic groups, and morphological convergence has shown to be widespread in the Psocidae, we suspect that current phylogenetic hypotheses cannot predict the phylogenetic position and diversity of neotropical lineages.

Material and methods

Field work, DNA extraction, amplification, and sequencing

We obtained molecular data for 43 taxa collected from five localities in Colombia. We have conducted extensive collections of psocid taxa in these localities over the last decade: (1) Dagua: El Queremal, Vereda La Elsa (03º33'55.8"N; 76º45'30.0"W; (2) Cali: Los Yes, Quebrada Honda (3°26'01.8''N; 76°38'40.3''W), (3) Cali: La Buitrera (3°32'14.1''N; 76°45'19.0''W; (4) Dagua: Km 23, Via a Buenaventura, El Canasto (3°33'13.5"N; 76°36'34.6"W), y (5) Dagua: Km 18, Via a Zingara (3°32'0.1"N; 76°36'35.1''W). All the collected individuals were dry-stored in vials at –4°C. Morphological identification was conducted using published taxonomic keys (e.g. Smithers, 1990) and recently published diagnoses (e.g. García-Aldrete & Román-P. 2015; Román-P. et al. 2014; Yoshizawa, 1998). All voucher specimens used in this study are deposited in the Psocopteran collection of the Universidad del Valle, Colombia (Grupo de Investigaciones Entomológicas). Photos of the specimens used in the phylogenetic analyses are provided in Supplementary Appendix E1.

We followed Birungi and Munstermann (2002) for the DNA extraction protocol. We used an incubation period of one hour in potassium acetate (Rosero et al. 2010). We followed Ruíz et al. (2010) for reagent concentrations used in PCR. We amplified three gene regions corresponding to one mitochondrial and two nuclear genes (Supplementary Table S2). PCR thermal cycle protocols used to amplify each gene region are summarized in Supplementary Table S3. Sequencing was conducted in Macrogen Inc. Finally, Geneious 7.1.3 (Kearse et al. 2012) was used to assemble the raw sequences.

Retrieval of published sequences

We retrieved molecular data for COI, 18S, and H3 gene regions from GenBank (Benson et al. 2012) and BOLD Systems (Ratnasingham & Hebert 2007). We used public databases to expand the molecular sampling in our study by including 12S,

16S, and Wingless genes. These last three gene regions were sampled in previous studies of the Psocidae phylogeny (e.g. Yoshizawa, 2001, 2004; Bess & Yoshizawa 2007; Yoshizawa & Johnson 2008; Bess et al. 2014). Additionally, we used these databases to sample gene sequences for two outgroup free living lice species in the Hemipsocidae (*Hemipsocus chloroticus*) and Psilopsocidae (*Psilopsocus malayanus*).

Assembly and curation of molecular datasets

We constructed two molecular datasets for the Psocidae by assembling DNA alignments from the (i) sequences obtained through public databases, and (ii) the combination of both newly generated and published sequences. The assembly and curation of each of these two datasets was conducted using SuperCRUNCH version 1.0 (Portik & Wiens 2020). We first combined all the dataset-specific sequences in a fasta file with sequence names according to SuperCRUNCH. We removed duplicated sequences (script Remove_Duplicate_Accessions) and subspecies or ambiguously identified taxa (e.g. sp., aff.; Fasta_Get_Taxa script). Next, loci-specific fasta files were generated (Parse_Loci script) based on the following alternative versions of each gene: COI (COI, COX, and cytochrome), H3 (H3 and Histone 3), wingless (wingless and Wnt), 18S, 12S, and 16S. For each locus, we selected the longest sequence per species (Filter_Seqs_and_Species). We then used CD-HIT version 4.6.8 within the EST package (Li and Godzik, 2006) and BLAST (megablast; Madden, 2013) to test for the sequence orthology within each of the species-level fasta files. For each locus, we kept the largest cluster of orthologous sequences (Cluster_Blast_Extract.py script) and adjusted the direction of all sequences before performing sequence alignment under MAFFT v. 7 (Adjust_Direction script in SuperCRUNCH; Katoh and Standley, 2013). Our phylogenetic analyses are based on the concatenated orthologous clusters that were generated from each of the two datasets assembled in this study.

Sequence alignment

We used SuperCRUNCH to obtain six orthologous gene clusters from each dataset (published sequences and combined sequences). Each of these gene clusters was then aligned using the profile alignment routine in MAFFT v. 7 (Katoh & Standley, 2013). For each sequence alignment in MAFFT, we (i) allowed sequence direction to be adjusted, (ii) aligned length to remain the same as in the existing alignment (–add parameter), and (iii) conducted a local alignment under the L-INS-1 strategy. The remaining parameters were set to default. We selected the following set of published alignments to guide the alignment of our sequences. For COI and 12S genes, we used the alignments in Chesters (2017). We used the H3 sequence alignment from Gamboa et al. (2019). For Wingless, we followed the alignment from Phillips et al. (2017). Finally, we aligned both 16S and 18S genes by following the secondary structure indicated in Viale et al. (2015) and Kjer (2004), respectively. The sequence alignment in Kjer (2004) for 18S was transformed from RNA to DNA using Seqotron (Fourment & Holmes 2016). Finally, we manually removed sequences that did not overlap with the regions sampled in the reference alignments (Kjer 2004; Viale et al.

2015; Chesters 2017; Phillips et al. 2017; Gamboa et al. 2019). We obtained a single concatenated alignment for each dataset (Supplementary File S1 , published sequences; Supplementary File S2, combined sequences). These concatenated alignments, based on profile alignments of individual locus, were then used in the phylogenetic inference steps outlined below.

Partitioning strategies of the concatenated alignment

We obtained one concatenated dataset for published sequences and another for the combined sequences. Given that the analyzed partitioning strategy of the dataset can affect the resulting phylogenetic relationships among species within each dataset, we conducted independent analyses based on alternative partitioning strategies. A partition strategy corresponds to the sequence blocks in an alignment that are selected prior to a statistical analysis of the optimal partitioning (e.g. using PartitionFinder; Lanfear et al. 2017). We used two partitioning strategies for each dataset: (i) gene-based partitioning, and (ii) codon/gene-based partitioning to examine optimal partitioning schemes. A partition scheme results from statistically evaluating partition strategies (results of PartitionFinder). We run PartitionFinder twice in each dataset using two partitioning strategies that resulted in the same number of partitioning schemes per dataset. First, we used gene-based partitions within each concatenated alignment. Alternatively, we used a combination of gene-based (for the non-protein-coding genes 12S, 16S, and 18S) and codon-based (for protein-coding genes COI, H3, and wingless) partitioning for each dataset. PartitionFinder output files are provided in Supplementary File S3.

Phylogenetic analyses

We obtained two different partitioning schemes for each of the two molecular datasets. We followed Baca et al. (2017) to compare the fit of these partitioning schemes. Phylogenetic inference was performed under Maximum Likelihood in RAxML-HPC BlackBox 8.2.10 (Stamatakis, 2014) and Bayesian Inference in MrBayes (Ronquist et al. 2012). We run all phylogenetic analyses in CIPRES Science Gateway V. 3.3 (Miller et al. 2010). Under RAxML, we set a total of 1,000 bootstrap replicates, used a GTRGAMMA model for each partition, and set the remaining parameters to default. Under MrBayes, we performed two simultaneous runs for each combination of dataset and partitioning scheme consisting of eight MCMC chains (one cold and seven heated) chains running for 30 million generations. Trees were sampled every 1,000 generations. We assessed convergence of parameters by investigating the Effective Sample Size (ESS) of all parameters in Tracer 1.7 (Rambaut et al. 2018). We used a value of ESS > 200 as indicative of convergence. We discarded 10% of posterior trees as burn-in and inferred the 50% majority rule consensus tree based on the remaining samples. Finally, we compared the performance of partitioning schemes based on likelihood estimates from RAxML runs. The best partitioning scheme for each dataset was selected based on the highest likelihood score for analyses conducted under maximum likelihood in RAxML. Note that details on taxonomic identifications for the new samples analyzed in the resulting trees are presented in Supplementary Appendix E1.

Extended dataset from Yoshizawa & Johnson (2008)

We also focused on sequences that were exclusively analyzed in Yoshizawa & Johnson (2008). We identified gene partitions in the concatenated alignment provided in the supplement of the relevant study. Next, we used the -merge option in MAFFT version v. 7 (Katoh & Standley, 2013) to add our newly generated sequences to each existing gene-based alignment from Yoshizawa & Johnson (2008). This approach ensured that the original alignment maintained its characteristics while finding an optimal fit for the new sequences. We conducted phylogenetic analyses as indicated in Yoshizawa & Johnson (2008).

Results

We sequenced three gene regions from 43 Neotropical psocopteran species in the Psocinae and Amphigerontiinae (Supplementary Table S4). Four of these samples were not morphologically similar to any of the currently described tribes and subfamilies in the Psocidae. To our knowledge, all the species that were sequenced in this study are exclusively restricted to the Neotropics.

Our phylogenetic analyses were based on 38 of the total 43 Neotropical taxa sequenced in this study – five species were excluded in different stages of the dataset construction under SuperCRUNCH. Since our main interest was on testing if the phylogenetic position of Neotropical species could be predicted from a strongly Holarctic-biased phylogeny, we generated two datasets. First, we retrieved from public databases all available sequences for the Psocidae. Second, we combined our newly generated sequences with published sequences. In total, our published-sequences dataset included 109 Psocidae species from 25 genera, eight tribes, and three subfamilies. The combined dataset included 147 Psocidae species from 30 genera, eight tribes, and three subfamilies. The phylogenetic relationships among the species in each of these datasets was analyzed under two different partition schemes. Our main results for both Maximum Likelihood analyses and Bayesian Inference trees are based on the partitioning scheme resulting in the higher likelihood (under Maximum Likelihood). Specifically, a codon-based partitioning strategy was selected as the best-fitting approach (i.e. the model with the highest likelihood under Maximum Likelihood) for both the combined (Codon = –48.987.549, Genes = –50.082.229) and public datasets (Codon = –44.194.173, Genes = –45.121.622). Nevertheless, all trees recovered congruent phylogenetic relationships among lineages.

Phylogenetic analyses based on the published-only dataset inferred the family-level monophyly (bootstrap = 92%; Figure 1). At the subfamily level, our analyses recovered the monophyly for Amphigerontiinae (bootstrap = 100%) and indicated paraphyly for Kaindipsocinae and Psocinae. Specifically, *Kimunpsocus takumai* (Kaindipsocinae) and multiple Psocinae (*Oreopsocus buholzeri, Loensia conspersa, Camelopsocus monticolus, Loensia variegata*, and *Loensia moesta*) were found to cluster outside the core clades of each these two groups. The species causing the paraphyly of Psocinae and Kaindipsocinae were consistently recovered as being closely related to

Figure 1. Phylogenetic relationships among the Psocidae based on the published-only dataset and reconstructed using Maximum Likelihood using a codon-based partitioning scheme. We summarize the higher-level taxonomy for the species in the tree. The full species-level phylogeny is presented in Supplementary File S4, but additional results under alternative partitioning schemes of the alignment are included in the Supplementary File S5. Results based on Bayesian analyses for the same dataset are similarly included in the Supplementary Files S8–S9.

Amphigerontiinae. At the level of tribes, our analyses recovered (but sometimes weakly supported) the monophyly of Blastini (bootstrap = 64%), Metylophorini (bootstrap = 50%), and Sigmatoneurini (bootstrap = 100%). Our analyses did not test the Amphigerontiini monophyly (we only sampled *Amphigerontia jezoensis*). Pyctini was recovered as a paraphyletic group, with several Pyctini being found closely related to species in almost every other tribe in the Psocinae. We note that although our analyses recover the monophyly of *Trichadenotecnum* and *Ptycta* + *Copostigma* (bootstrap = 82% and 100%, respectively), our results do not support a clade including these three genera: *Trichadenotecnum*, *Copostigma*, and *Ptycta* (bootstrap = 1%). *Atrichadenotecnum* was found to cluster with *Trichadenotecnum*, *Copostigma*, and *Ptycta,* but this clade was not supported. We recovered the paraphyly of Psocini, with several species within this group being found closely related to species in a clade formed by Sigmatoneurini + Thyrsophorini + Metylophoriny (bootstrap = 47%). Finally, we inferred the paraphyly of Thyrsophorini caused by *Longivalvus nubilus* closely related to Sigmatoneurini. We recovered a core clade of Thyrsophorini comprising all *Cerastipsocus* and *Psococerastis* species in our dataset (bootstrap = 63%).

We then examined the phylogenetic relationships within Psocidae based on a second dataset expanding the species-level sampling of the published-only dataset by including 38 Neotropical taxa. Based on the combined dataset, we did not infer the monophyly for any of the three subfamilies (Figure 2). We did not recover the monophyly of Kaindipsocinae due to the position of *Kimunopsocus*. Within Amphigerontini, *Elaphopsocoides* was found nested within a Kaindipsocinae group. However, our analyses do recover a core Amphigerontiinae including most species in *Blaste*, *Blastopsocus*, and *Amphigerontia* in our dataset. Finally, species in Psocinae clustered with species from the other two subfamilies. At the tribal level, we only recovered the monophyly for Sigmatoneurini (bootstrap = 100%). Within Blastini, Neotropical *Blaste* and *Blastopsocus* were closely related to *Amphigerontia* (bootstrap = 97%) and *Chaetoblaste* to *Metylophorus* (bootstrap = 60%). The monophyly of Amphigerontiini was also rejected due to the position of *Elaphopsocoides*. Within Pyctini, all the Holarctic *Trichadenotecnum* were still recovered forming a monophyletic group (bootstrap = 35%). However, we recovered two Neotropical *Trichadenotecnum* in a second clade (including *Atrichadenotecnum* and *Indiopsocus*) that was sister to the remaining *Trichadenotecnum* (bootstrap = 16%). Although Holarctic *Ptycta* and *Copostigma* formed a well-supported clade (bootstrap = 95%), not all Neotropical *Pycta* were clustered within this group. Within Metylophorini, one of the two species sampled in our dataset clustered with Neotropical taxa from *Ptycta*, *Psococerastis*, and *Chaetoblaste* (bootstrap = 93%). Our analyses inferred the monophyly of all non-Neotropical species of Psocini (bootstrap = 33%) but placed a Neotropical species of Psocini as closely related to Neotropical *Trichadenotecnum* (bootstrap = 16%). Most Thyrsophorini formed a single clade that also included two Neotropical *Melophorus* and a neotropical *Ptycta* (bootstrap = 23%). *Longivalvus nubilus* was found to be sister to Sigmatoneurini (bootstrap = 55%) and a single *Psococerastis* closely related to a Neotropical *Ptycta* (bootstrap = 100%). Finally, we note that our results highlighting the lack of correspondence between morphological taxonomy and phylogenetic position for neotropical taxa still holds even when

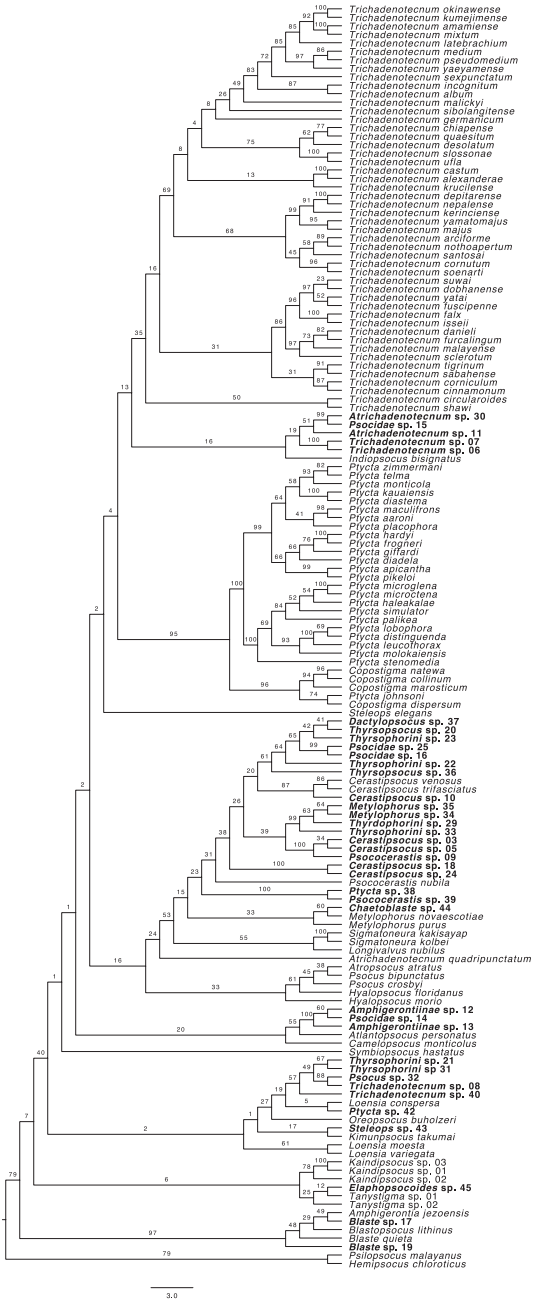

Figure 2. Phylogenetic relationships among the Psocidae based on the combined dataset and reconstructed using Maximum Likelihood using a codon-based partitioning scheme. The full species-level phylogeny is presented in Supplementary File S6, but additional results under alternative partitioning schemes of the alignment in the Supplementary File S7. Results based on Bayesian analyses for the same dataset are shown in the Supplementary Files S10–S11. Newly sampled species are boldfaced in the tree. Taxonomic information on each of the newly sampled species is included in Supplementary Table S4, with corresponding images of relevant morphological features presented in Supplementary Appendix E1.

analyses are based on the alignment from Yoshizawa and Johnson (2008; Supplementary Figure S1; see Supplementary File S12).

Discussion

By leveraging the most comprehensive species-level molecular dataset for the Psocidae including 147 extant species, we inferred the phylogenetic relationships among all extant subfamilies, 80% tribes (8 of 10), and ~38% of genera in the family (30 of ~80). Relative to recent studies on the Psocidae phylogeny, our study increases the sampling of Neotropical taxa in the Psocidae phylogeny by a factor of ~5 (from eight species in the most recent Psocidae phylogeny [Yoshizawa & Johnson 2008] to 38 species in our study). Nevertheless, we acknowledge that conclusions on the systematics within the family derived from our study need further examination given the size of our phylogeny in relation to the total family diversity (15% of ~1000 species). Furthermore, the weak support for deep branches in our tree suggest that further studies using genomic approaches are needed to unravel the actual relationships between clades in the Psocidae. Nevertheless, our study represents an interesting case study for lineages in which (i) most morphological and molecular studies have been based on Holarctic taxa, and (ii) where the systematics of Tropical lineages is understood from morphological resemblance to taxa in other regions. Below, we discuss the implications of our findings on the Psocidae Tree of Life in the context of previous phylogenetic hypotheses.

Can heavily Holarctic sampled phylogenies predict the phylogenetic position of Neotropical taxa?

We inferred similar phylogenetic relationships within and between taxa in the Kaindipsocinae, Ptyctini, Psocini, Thyrsophorini, Sigmatoneurini, Metylophorini, Amphigerontiini, and Blastini relative to published phylogenies (Figs. 1). The inclusion of Neotropical taxa had major implications in the inferred relationships within and between major groups within the Psocidae (Figs. 2, S1). These results are largely independent on whether the newly generated sequences for Neotropical species are included in a newly generated dataset (Figure 2) or based on a previously published sequence alignment (e.g. Yoshizawa & Johnson 2008; Figure 1). We found that several Neotropical species in *Elaphopsocoides*, *Psocus* (code 032; codes follow those in Supplementary Table S4 and Supplementary Appendix E1), *Blaste* (013), *Ptycta* (038), *Psococerastis* (039), *Chaetoblaste* (044), *Metylophorus* (029, 035, 034), *Trichadenotecnum* (006, 007), and *Atrichadenotecnum* (011, 030) did not cluster within their corresponding morphological groups. This incongruence in the systematics of the Psocidae is likely caused by the historical undersampling of Neotropical taxa in previous phylogenetic studies (e.g. Mockford 1993; Yoshizawa & Johnson 2008; Liu et al. 2013). The fact that the evolutionary history in the Psocidae is currently mostly understood from Holarctic lineages, neglects potential alternative drivers of diversity in the tropics while also hindering the potential uniqueness of lineages from the same region.

Morphological convergence: Morphological vs. molecular phylogenetics in the Psocidae

Our analyses indicate that morphological taxonomy largely disagrees with molecular systematics in the Psocidae. Out of the three subfamilies (Kaindipsocinae, Amphigerontiinae, Psocinae) and seven tribes (Amphigerontiini, Blastini, Psocini, Atrichadenoctenini, Sigmatoneurini, Metylophorini, and Thyrsophorini) recovered as monophyletic in previous studies mostly based on Holarctic taxa (e.g. Yoshizawa & Johnson 2008; Yoshizawa et al. 2014), only one tribe (Sigmatoneurini) was inferred as monophyletic after the inclusion of Neotropical lineages. Because our phylogenetic analyses based (i) on published data used in previous studies (Figure 1) and (ii) the expanded dataset including more Neotropical taxa (Figs. 2 and S1), we suggest that the inclusion of Neotropical taxa was responsible for the non-monophyletic status for nine higher-level groups within Psocidae.

We found that morphological classification does not accurately reflect evolutionary closeness in the Psocidae. For instance, our analyses suggest that not all Neotropical and Holarctic *Trichadenoctenum*, a clade that has been historically highly supported by molecular and morphological data, cluster in a single clade. Similarly, *Elaphopsocoides*, an exclusively Neotropical genus (Román-P. et al. 2014), was not recovered within the remaining Amphigerontiini, a tribe that has also been inferred as monophyletic in previous studies (Yoshizawa & Johnson 2008; Yoshizawa et al. 2011). In a more striking example, Neotropical species of Methylophorini were recovered as being closely related to Thyrsophorini. However, the only Holarctic species in this tribe, *Metylophorus novaescotiae*, was found closely related to the Neotropical *Chaetoblaste* (within Amphigerontiinae: Blastini; Aldrete & Román-P. 2015). In short, morphological resemblance between Neotropical and Holarctic taxa is, in many cases, not indicative of recent common ancestry within the Psocidae.

Finally, we note that morphological classification, which is currently largely based on Holarctic taxa, may have hindered a large fraction of diversity and evolutionary uniqueness of Neotropical lineages. While many Neotropical lineages correspond with morphological descriptions of Holarctic taxa, many of these Neotropical groups have an independent evolutionary origin. Multiple debates about the high frequency of morphological convergence in the Psocidae, along with other studies on problematic synapomorphies within groups, further support our conclusions. Our analyses recover many Neotropical lineages to be distantly related to their morphologically closest lineages. This pattern suggests that the diversity and evolutionary differentiation across different taxonomic levels (e.g. genera, tribes, and subfamilies) in the Tropics is potentially higher than what is currently known based on Holarctic groups.

Conclusions

We show that molecular phylogenetics and morphological taxonomy strongly based on Holarctic groups cannot inform the phylogenetic position of Neotropical taxa. In addition to highlighting the need of a new taxonomic classification for the Psocidae,

our analyses suggest that multiple Neotropical Psocidae potentially represent independent lineages to the ones known in the Holarctic region. Although the role geography in affecting taxonomic boundaries within clades remains largely unexplored in this family, our results suggest that, for certain groups within the Psocidae, morphological and phylogenetic classification based on lineages found in certain areas (e.g. Holarctic) do not reflect the evolutionary history of morphologically similar taxa in other regions (Neotropics). Future studies on the Psocidae Tree of Life should rely on a better sampling of non-Holarctic lineages to derive a comprehensive hypothesis of the systematics within the family.

Acknowledgements

We thank the Laboratorio de Estudios Ecogenéticos y Biología Molecular of the Universidad del Valle for allowing us to use their facilities and equipment to conduct molecular protocols. We are also grateful to the Department of Biology, Facultad de Ciencias Naturales y Exactas, Vicerrectoría de Investigaciones, Universidad del Valle, for their constant support. We are especially thankful for the feedback provided by Heidi E. Steiner, and Jose Sergio Hleap on different stages of this project. This study was partially funded by Fundación Banco de la República.

Supplementary material

Supplementary material is available online at:
https://doi.org/10.6084/m9.figshare.25222046

References

Aldrete, A. N. G. & Román-P, C. (2015). A new Neotropical genus of Blastini (Psocodea: 'Psocoptera': Psocidae: Amphigerontiinae). *Zootaxa* 4000 (3), 392–396.

Baca, S. M., Toussaint, E. F., Miller, K. B., Short, A. E. (2017) Molecular phylogeny of the aquatic beetle family Noteridae (Coleoptera: Adephaga) with an emphasis on data partitioning strategies. *Molecular phylogenetics and Evolution* 107, 282–292.

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Sayers, E. W. (2012) GenBank. *Nucleic acids research* 41(D1), D36–D42.

Bess, E. C., Yoshizawa, K. (2007) Redefinition of *Ptycta* Enderlein (Psocodea: 'Psocoptera': Psocidae) and a taxonomic revision of the Japanese species. *Entomological Sciences* 10, 279–290.

Bess, E., Catanach, T. A., Johnson, K. P. (2014) The importance of molecular dating analyses for inferring Hawaiian biogeographical history: a case study with bark lice Psocidae: Ptycta. *Journal of Biogeography* 41(1), 158–167.

Birungi, J., Munstermann, L. E. (2002) Genetic structure of *Aedes albopictus* (Diptera: Culicidae) populations based on mitochondrial ND5 sequences: evidence for an independent invasion into Brazil and United States. *Annals of the Entomological Society of America* 95(1), 125–132.

Chesters, D. (2017) Construction of a species-level tree of life for the insects and utility in taxonomic profiling. *Systematic biology* 66(3), 426–439.

Fourment, M., & Holmes, E. C. (2016). Seqotron: a user-friendly sequence editor for Mac OS X. *BMC research notes*, *9*(1), 1–4.

Gamboa, M., Muranyi, D., Kanmori, S., Watanabe, K. (2019). Molecular phylogeny and diversification timing of the Nemouridae family (Insecta, Plecoptera) in the Japanese Archipelago. *PloS one* 14(1), e0210269.

Johnson, K. P., Mockford, E. L. (2003) Molecular systematics of Psocomorpha (Psocoptera). *Systematic Entomology* 28(3), 409–416.

Johnson, K. P., Smith, V. S., Hopkins, H. H. (2020) *Psocodea Species File Online*. Version 5.0/5.0. http://Psocodea.SpeciesFile.org (accessed 29 September, 2020).

Katoh, K., Standley, D. M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* 30(4), 772–780.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12), 1647–1649.

Kjer, K. M., 2004. Aligned 18S and insect phylogeny. Systematic biology. 53(3), 506–514.

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular biology and evolution* 34(3), 772–773.

Li, W., Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13), 1658–1659.

Lienhard, C., Smithers, C. N. (2002) Psocoptera (Insecta): World catalogue and bibliography. Muséum d'Histoire Naturelle de Genève.

Liu, L., Yoshizawa, K., Li, F., Liu, Z. (2013) *Atrichadenotecnum multispinosus* sp. n. (Psocoptera: Psocidae) from southwestern China, with new synonyms and new combinations from Psocomesites and Clematostigma. *Zootaxa* 3701(4), 460–466.

Madden, T. (2013) The BLAST sequence analysis tool. In The NCBI Handbook [Internet]. 2nd edition. National Center for Biotechnology Information (US).

Miller, M. A., Pfeiffer, W., Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE). 1–8.

Mockford, E. L. (1993) North American Psocoptera (Insecta). Flora and Fauna Handbook No. 10. Sandhill Crane Press, Inc. Gainesville, Florida. Leiden.

Oliveira, J. A., Neto, A. M. S., de Mello Mendes, D. M., Aldrete, A. N. G. (2017) New records of *Dictyopsocus pennicornis* (Burmeister) (Psocodea'Psocoptera': Psocidae: Psocinae). *EntomoBrasilis* 10(2), 127–130.

Phillips, R. D., Brown, G. R., Dixon, K. W., Hayes, C., Linde, C. C., Peakall, R., 2017. Evolutionary relationships among pollinators and repeated pollinator sharing in sexually deceptive orchids. *Journal of evolutionary biology* 30(9), 1674–1691.

Portik, D. M., Wiens, J. J. (2020) SuperCRUNCH: A bioinformatics toolkit for creating and manipulating supermatrices and other large phylogenetic datasets. *Methods in Ecology and Evolution* 11, 763–772.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., Suchard, M. A. (2018) Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67(5), 901–904.

Ratnasingham, S., Hebert, P. D. (2007) BOLD: The Barcode of Life Data System (http://www.barcoding life.org). *Molecular ecology notes* 7(3), 355–364.

Román-P, C., García Aldrete, A. N., González Obando, R. (2014) *Elaphopsocoides*, a new genus of Psocidae (Psocodea: Psocoptera) from Valle del Cauca, Colombia. *Zootaxa* 3873(1), 93–100.

Román-Palacios, C., Obando, R. G., Aldrete, A. N. G., (2016) Uncovering the diversity of the neotropical genus *Elaphopsocus* ('Psocoptera': Psocidae: Amphigerontiinae): from one to ten species. *Zootaxa* 4162(2), 373–380.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., Huelsenbeck, J. P. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic biology* 61(3), 539–542.

Rosero, D. A., Gutiérrez, L. A., Cienfuegos, A. V., Jaramillo, L. M., Correa, M. M. (2010) Optimización de un procedimiento de extracción de ADN para mosquitos anofelinos, *Revista Colombiana de Entomologia* 36(2), pp. 260–263.

Smithers, C. N. (1990) Keys to the families and genera of Psocoptera (Arthropoda: Insecta). *Technical reports of the Australian Museum* 2, 1–82.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9), 1312–1313.

Viale, E., Martinez-Sañudo, I., Brown, J. M., Simonato, M., Girolami, V., Squartini, A., Bressan, A., Faccoli, M., Mazzon, L. (2015) Pattern of association between endemic Hawaiian fruit flies (Diptera, Tephritidae) and their symbiotic bacteria: Evidence of cospeciation events and proposal of "*Candidatus Stammerula trupaneae*". *Molecular phylogenetics and evolution* 90, 67–79.

Yoshizawa, K. (1998) A new genus, *Atrichadenotecnum*, of the tribe Psocini (Psocoptera: Psocidae) and its systematic position. *Entomol. Scand*. 29, 199–209.

Yoshizawa, K. (2001) A systematic revision of Japanese *Trichadenotecnum* Enderlein (Psocodea: 'Psocoptera': Psocidae: Ptyctini), with redefinition and subdivision of the genus. *Inverteb. Taxon*. 15, 159–204.

Yoshizawa, K. (2002) Phylogeny and higher classification of suborder Psocomorpha (Insecta: Psocodea: 'Psocoptera'). *Zool. J. Linn. Soc*. 136, 371–400.

Yoshizawa, K. (2004) Molecular phylogeny of major lineages of Trichadenotecnum and a review of diagnostic morphological characters (Psocoptera: Psocidae). *Syst. Entomol*. 29, 383–394.

Yoshizawa, K. (2005) Morphology of Psocomorpha (Psocodea: 'Psocoptera'). Insecta *Matsumurana*, n. s. 62, 1–44.

Yoshizawa, K. (2010) Systematic revision of the Japanese species of the subfamily Amphigerontiinae (Psocodea:'Psocoptera': Psocidae). *Insecta matsumurana. New series* 66, 11–36.

Yoshizawa, K., Bess, E., Johnson, K. P. (2011) Kaindipsocinae is a sister taxon to the rest of Psocidae (Insecta: Psocodea: 'Psocoptera'). *Invertebrate Systematics* 25(2), 81–90.