Molecular Phylogenetics and Evolution 128 (2018) 112–122

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Topological support and data quality can only be assessed through multiple tests in reviewing Blattodea phylogeny

Dominic Evangelista[a,*], France Thouzé[a], Manpreet Kaur Kohli[b], Philippe Lopez[a], Frédéric Legendre[a]

[a] Institut de Systématique, Evolution, Biodiversité ISYEB – UMR 7205 – MNHN CNRS UPMC EPHE, Muséum national d'Histoire naturelle, Sorbonne Universités, CP50, 57 rue Cuvier, 75005 Paris, France
[b] Department of Biological Sciences, Rutgers, The State University of New Jersey, 195 University Ave., Newark, NJ 07102, United States

ARTICLE INFO

Keywords:
Phylogenetic signal
mtDNA
Termite
Dictyoptera
SAMS
Rogue taxa
Long branch attraction
Signal analysis

ABSTRACT

Assessing support for molecular phylogenies is difficult because the data is heterogeneous in quality and overwhelming in quantity. Traditionally, node support values (bootstrap frequency, Bayesian posterior probability) are used to assess confidence in tree topologies. Other analyses to assess the quality of phylogenetic data (e.g. Lento plots, saturation plots, trait consistency) and the resulting phylogenetic trees (e.g. internode certainty, parameter permutation tests, topological tests) exist but are rarely applied. Here we argue that a single qualitative analysis is insufficient to assess support of a phylogenetic hypothesis and relate data quality to tree quality. We use six molecular markers to infer the phylogeny of Blattodea and apply various tests to assess relationship support, locus quality, and the relationship between the two. We use internode-certainty calculations in conjunction with bootstrap scores, alignment permutations, and an approximately unbiased (AU) test to assess if the molecular data unambiguously support the phylogenetic relationships found. Our results show higher support for the position of Lamproblattidae, high support for the termite phylogeny, and low support for the position of Anaplectidae, Corydioidea and phylogeny of Blaberoidea. We use Lento plots in conjunction with mutation-saturation plots, calculations of locus homoplasy to assess locus quality, identify long branch attraction, and decide if the tree's relationships are the result of data biases. We conclude that multiple tests and metrics need to be taken into account to assess tree support and data robustness.

1. Introduction

Phylogenetic trees are increasingly used across biological disciplines and their interpretation can have deep implications (e.g. Beaulieu et al., 2012; Kozak and Wiens, 2012; Maganga et al., 2014). Yet phylogenetic data are not straightforward and trees inferred from them can be conflicting. Traditional phylogenetic tests such as bootstrapping or tree comparisons are used to assess such conflicts (e.g. Blaimer et al., 2015; Garrison et al., 2016; Johnson et al., 2013; Kjer et al., 2016; Trautwein et al., 2012). These tests, however, have their own limitations and biases and can sometimes give artefactual results (Dell'Ampio et al., 2014). Other tests and metrics have been used to complement traditional ones and provide a more complete picture of data signal and topological support (e.g. spectra analysis, saturation plots, likelihood mapping; Borowiec, 2017; Dell'Ampio et al., 2014; Kobert et al., 2016; Wägele and Mayer, 2007). In addition, they provide better guidance on crafting future datasets through locus choice, which is essential in the age of genomics because next-generation sequencing allows systematists to choose from a variety of loci whose quality varies (as in: Borowiec, 2017; Chen et al., 2015).

1.1. Pre-omics molecular phylogenetic approaches

Pre-omics molecular phylogenetic datasets typically comprise 2–7 markers (e.g. Bradler et al., 2014; Chaboo et al., 2014; Kambhampati et al., 1996; Legendre et al., 2015; Maekawa et al., 2003; Marvaldi et al., 2009; Murienne, 2009; Song et al., 2015). These markers usually come from both mitochondrial DNA (mtDNA) (e.g. Bradler et al., 2014; Kambhampati, 1996; Kambhampati et al., 1996; Legendre et al., 2008; Mandal et al., 2014; Muraji and Tachikawa, 2000; Vogler et al., 2005; Wiens et al., 2010) and nuclear DNA (nucDNA) (e.g. Bradler et al., 2014; Legendre et al., 2015; Song et al., 2015) for animals, and

* Corresponding author.
E-mail addresses: DominicEv@gmail.com (D. Evangelista), francethouze@gmail.com (F. Thouzé), mkk24@njit.edu (M.K. Kohli), philippe.lopez@upmc.fr (P. Lopez), frederic.legendre@mnhn.fr (F. Legendre).

https://doi.org/10.1016/j.ympev.2018.05.007
Received 16 January 2018; Received in revised form 7 May 2018; Accepted 8 May 2018
Available online 30 June 2018
1055-7903/

chloroplast DNA for plants (e.g. Bremer et al., 2004; Schulte et al., 2009). Investigators often choose regions coding for ribosomal RNA (rDNA) because they contain a variety of fast and slow evolving sites and should be able to inform a phylogeny across multiple time-scales (Hillis and Dixon, 1991; Mandal et al., 2014). Indeed, all these types of data have allowed for large progresses in reconstructing the tree of life.

The small number of generic "toolkit" loci typical of pre-omics datasets necessitates consideration of their limitations. For instance, mtDNA markers are known to evolve quickly (Simon et al., 1994) and might be most appropriate to address recent evolutionary questions as opposed to deep ones (Mandal et al., 2014). Also, mtDNA can have very different evolutionary histories from nuclear genes (Fisher-Reid and Wiens, 2011; Kodandaramaiah et al., 2013; Wiens et al., 2010) and phylogenies from nucDNA loci can be different from each other as well (Lanier and Knowles, 2015). Finally, both 18s + 28s and all mtDNA have limited biological independence. All mtDNA is transcribed at once (Cameron, 2013) and similarly the nuclear rDNA are encoded by the same transcription unit (Lodish et al., 2000).

Since there is no ideal molecular marker, biases resulting from their analysis should be accounted for through phylogenetic tests and metrics. One common example in parsimony and maximum likelihood analyses is to calculate clade bootstrap support (Pattengale et al., 2009; Soltis and Soltis, 2003). This test assesses internal alignment conflicts (Soltis and Soltis, 2003), which are dependent on signal conflicts among or within markers. Though, given the complicated nature of phylogenetic inference, a single test is far from universally informative. For instance, bootstrap values are highly dependent upon the number of taxa included (Soltis and Soltis, 2003) and can be biased by artefactual signal (Dell'Ampio et al., 2014). Complementary assays have been proposed to more thoroughly assess data quality and tree support. For example, loci information content and bias can be calculated prior to phylogenetic inference by examining: nucleotide compositional bias (Song et al., 2010), mutation saturation (Wenzel and Siddall, 1999), or information content (Wägele and Mayer, 2007). After tree reconstruction, molecular marker quality can be assessed through calculation of consistency indices (Farris, 1989; Klassen et al., 1991), and locus saturation by branch length (e.g. Borowiec et al., 2015). Support for a given set of relationships can be estimated through statistical tests of alternative relationships (Shimodaira, 2002) by comparing the effect of analysis permutations on tree reconstruction (e.g. Djernæs et al., 2015; Tang et al., 2014), or measures of certainty and stability of trees and clades (Kobert et al., 2016; Legendre et al., 2010). Ideally, synthesis of all tests would allow one to conclude which relationships are strongly supported, which markers are mainly responsible for dark nodes (i.e. unresolved or unsupported relationships), and to suggest further improvements.

1.2. Molecular phylogenetics of Blattodea and dark areas of the tree

Using molecular data to decipher the relationships in Blattodea was revolutionary, as it was for other organisms (Kjer et al., 2016). It helped shed light on the position of termites as sister to *Cryptocercus* and nested within the cockroach clade Blattoidea (Inward et al., 2007; Lo et al., 2003), which had been controversial for decades (e.g. Grandcolas, 1996, 1999; Klass, 2001; McKittrick, 1965; Thorne and Carpenter, 1992). It elucidated the co-evolutionary patterns of cockroaches and one of their most important endosymbionts (Lo et al., 2003). It supported many other relationships and the monophyly of many groups (e.g. Djernæs et al., 2015; Legendre et al., 2015, 2017).

Yet, current molecular datasets are unable to recover many blattodean relationships, despite strong taxon sampling (Fig. 1). For instance, one study highlights the difficulty in reconstructing the relationships within this clade using a set of six markers and 128 taxa (Legendre et al., 2017). In particular, some clades reconstructed with their full dataset were found only part of the time, or even rarely among individual gene trees (Legendre et al., 2017). Similarly, Djernæs et al. (2015) discussed that while the monophyly of Blaberoidea had strong morphological support, the molecular support was weak or lacking for one mtDNA marker (COII). Recent studies have demonstrated low support for evolutionary relationships among Blattoidea prior to the split between sub-social cockroaches (Cryptocercidae) and eusocial termites. This includes the position of Tryonicidae, Lamproblattidae, Blattidae (Djernæs et al., 2015; Legendre et al., 2015) and Anaplectidae (Djernæs et al., 2015).

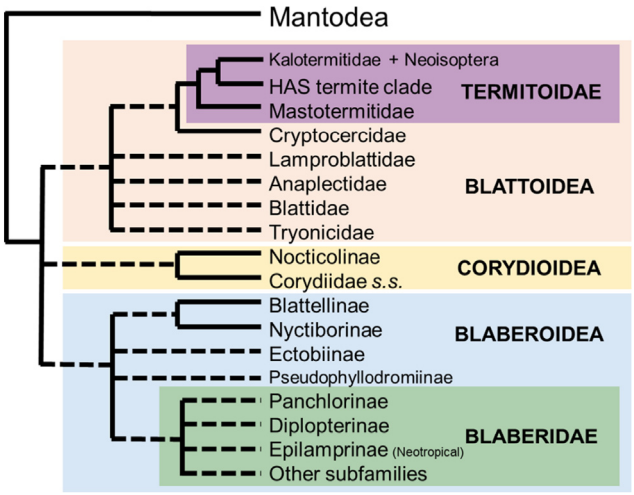

Fig. 1. Phylogenetic context for the present study of Blattodea. Solid edges represent relationships considered well-supported among all previous molecular and morphological studies on Blattodea and Dictyoptera phylogenetics. Dashed edges in polytomies represent clades whose position is uncertain. HAS termite clade consists of Hodotermitidae, Archotermopsidae and Stolotermitidae. Note that we use "Anaplectidae" in a sense equal to that of Anaplectinae Roth (1996) and consider it a family-level lineage as did Wang et al. (2017). Conversely, we use "Nocticolinae" whereas other authors consider this a family, despite its position within Corydiidae *s.l.* (Djernæs et al., 2012, 2015; Legendre et al., 2015; Wang et al., 2017).

There is not just lack of support within studies, but lack of congruence among studies. For example, the competing molecular hypotheses about the position of Tryonicidae (Legendre et al., 2015; Murienne, 2009) confound arguments for (Grandcolas, 1999; 1997) or against (Klass and Meier, 2006) morphological similarity with Blattidae. Also, the existence of some xylophagous species of Tryonicidae (Grandcolas, 1997) possibly suggests shared ancestry with Cryptocercidae and termites, which has sometimes been hypothesized with molecular data (Djernæs et al., 2015; Murienne, 2009). Similarly, putative morphological (Klass and Meier, 2006) and behavioral (McKittrick, 1965) synapomorphies in Lamproblattidae are left ambiguous because of the unclear position of the family with respect to Cryptocercidae (e.g. Djernæs et al., 2015; Legendre et al., 2015). Sister group relationships at the first split in Blattodea are also controversial (e.g. Legendre et al., 2015; Wang et al, 2017). In Blaberoidea, nearly every combination of subfamilial relationships has been obtained in a molecular or combined analysis, the only constant being the monophyly of Blaberidae (Djernæs et al., 2012; Djernæs et al., 2015; Inward et al., 2007; Legendre et al., 2015; Wang et al, 2017). To summarize all these incongruences, Evangelista et al. (2017) reviewed eight phylogenetic studies of Blattodea and showed that the majority of the internal relationships were supported by less than half of the studies.

Using seven tests and metrics, we attempt to assess support of relationships across the Blattodea tree, determine strengths and weaknesses of the dataset contributing to various levels of support, and recommend future improvements based on these findings. We focus on several clades (full list in Appendix A) chosen for biological or methodological reasons. Thus, to better understand: (i) the evolutionary precedents to eusociality, we investigate the position of Tryonicidae

and Lamproblattidae (Djernæs et al., 2015; Legendre et al., 2015; Murienne, 2009); (ii) the early morphological evolution of Blattodea, we focus on Blattoidea, Corydioidea, and Blaberoidea (Djernæs et al., 2012; Klass and Meier, 2006). To assess potential artefactual reconstructions, we focus on taxa supported by: (iii) long branches (Nocticolinae and Pseudophyllodromiinae; Djernæs et al., 2015; Legendre et al., 2015; Wang et al., 2017); (iv) short branches (Blaberidae subfamilies; Legendre et al., 2017); (v) rogue taxa (Diplopterinae; Legendre et al., 2017); or (vi) taxa with incongruent position among studies (Blaberoidea and Blaberidae; Evangelista et al., 2017). Finally, for the purpose of comparison, we focus on (vii) taxa with highly supported and congruent positions among studies (Mastotermitidae, Kalotermitidae, dampwood termites; Bourguignon et al., 2015; Djernæs et al., 2015; Legendre et al., 2015).

2. Methods

2.1. Phylogenetic methodology, tree reconstruction, and support

The alignment was composed of six genetic loci from 575 taxa. Sampling primarily included cockroaches and termites, but with a wide variety of outgroup taxa (Mantodea, Polyneoptera, and Palaeoptera). The six loci were: four mtDNA (12S rDNA, 16S rDNA, COI, COII), and two nucDNA (18S rDNA, 28S rDNA). We obtained the dataset from Legendre et al. (2015) with 65 additional sequences from Genbank (Appendix B), and with many Mantodea removed such that we could focus on relationships in Blattodea.

Starting with the alignment from Legendre et al. (2015) rDNA sequences were manually re-aligned to a structural model from Ware et al. (2008) following Kjer et al. (2009) (Appendix C). This eliminated unalignable hypervariable loop regions but retained hypervariable loop regions for some taxa (mantises and termites separately) that were indeed alignable. Forcing this pattern of missing data could enforce monophyly of the three groups. Yet, the relationships among the three "orders" of Dictyoptera have been settled for at least 10 years (Deitz et al., 2003; Inward et al., 2007; Klass, 2001; Lo et al., 2003) and deviation from these relationships would be a clear sign of error in the analysis. The missing data patterns could still be inflating support for Mantodea or termites, which should be considered when interpreting the results. The retained hypervariable regions were aligned with gaps for all other taxa. The final alignment was 7676 nucleotides long (Appendix D; Table 1), whereas it was about 10 kb in Legendre et al. (2015).

Data blocks were defined as the boundaries of loci, hypervariable regions and codon positions (for the protein coding genes: COI and COII). In PartitionFinder (Lanfear et al., 2012), using these block definitions, output models for RAxML, and the BIC optimality criterion, the following partitioning scheme was determined: 18S + 28S, 12S + 16S (including hypervariable regions), 28S "termite" hypervariable region, 28S "mantis" hypervariable region, COI + COII codon 1, COI + COII codon 2, and COI + COII codon 3. The model GTR + G was chosen for every partition.

A tree search was then implemented in RAxML (Stamatakis, 2014) on CIPRES (Miller et al., 2010), running independent tree searches on 4 threads using the "-f o" option. Preliminary runs with 100, 20 and 10 random starting trees gave the same topology with only changes in branch lengths. Thus, we conducted the final tree search using 10 starting trees to minimize computation time (hereafter called "focal tree"; Appendix D).

2.2. Tests and metrics

Test 1. Testing the effect of removing taxa with low data completeness. Here, five trees were generated with different taxon reduction schemes (one full tree + four reduced trees). For each reduction scheme, taxa were eliminated based on gene sampling completeness. Starting with the full alignment (1geneMin) we removed taxa missing four or more genes (3geneMin), three or more genes (4geneMin), two or more genes (5geneMin), and one or more genes (6geneMin). The last and most complete of these alignments included 92 taxa, and still represented most major lineages of Blattodea.

Maximum likelihood trees were inferred as in Section 2.1, using each of the reduced alignments. To compare trees, we examine both the targeted clades (Appendix A; Section 1.2) and their support, which we calculated as a weighted average of the bootstrap clade support scores of the three clades comprising the relationship (i.e. the clade including the common ancestor, and the two daughter clades individually). The average weights the lowest supported clade by a factor of two. This is based on the following rationale. Given the tree (((A, B), (C, D)), E) the relationship of clade AB sister to clade CD relies on the support of three nodes: (A, B), (C, D) and ((A, B), (C, D)). If any of these nodes are weakly supported one could argue for collapsing them into a tritomy with the superior node [e.g. ((A), (B), (C, D)) or ((A, B), (C, D), (E))]. Collapsing any of these three nodes removes the sister relationship between AB and CD from the tree's topology. Therefore, the support for the relationship depends most heavily on which of the nodes has the lowest support (and is therefore the strongest rationale for collapsing the relationship).

Test 2. Testing the effect of reducing alignment completeness with constant taxon sampling. Six gene trees were generated starting from an alignment of full character sampling of 92 taxa (described in Section 2.2) and 92 alignment blocks were removed at random. Each alignment had subsequently more alignment blocks removed (i.e. a single random gene for a single random taxon; -0Blocks -92Blocks, -184Blocks, -276Blocks, -368Blocks, -460Blocks). Trees were inferred as described in Section 2.1 applied to each alignment. Results were recorded as for Test 1.

Test 3. Assessing node support through bootstrapping and internode certainty. Exhaustive tree searches for bootstrap pseudoreplicate trees using the autoMRE stopping criteria (Pattengale et al., 2009) were done for all the trees generated in the above tests. Additionally, relative tree certainty (RTC) and tree certainty (TC) scores were calculated for all

Table 1
Molecular markers used in this analysis and descriptive information about their contribution to the total alignment. Information was taken from RAXML or manually calculated. Nucleotide composition is illustrated by proportion of each nucleotide, and relative bias is illustrated by base proportion standard deviation.

Alignment		Taxa	Length	Patterns	Patterns /Length	% Gaps	A	T	C	G	Base std. dev.
18S		289	1871	1147	0.61	15.62	0.24	0.24	0.28	0.24	0.019
28S		515	2503	1804	0.72	68.06	0.23	0.26	0.33	0.19	0.059
12S		507	358	321	0.90	15.56	0.26	0.14	0.21	0.40	0.109
16S		294	687	480	0.70	51.68	0.29	0.14	0.23	0.34	0.088
COI	c1	253	1551	387	0.78	35.9	0.29	0.18	0.28	0.25	0.052
	c2			311			0.17	0.25	0.17	0.42	0.117
	c3			517			0.48	0.16	0.04	0.32	0.192
COII	c1	444	677	205	0.89	2.41	0.36	0.21	0.23	0.21	0.073
	c2			173			0.27	0.22	0.14	0.38	0.103
	c3			225			0.50	0.18	0.04	0.28	0.193

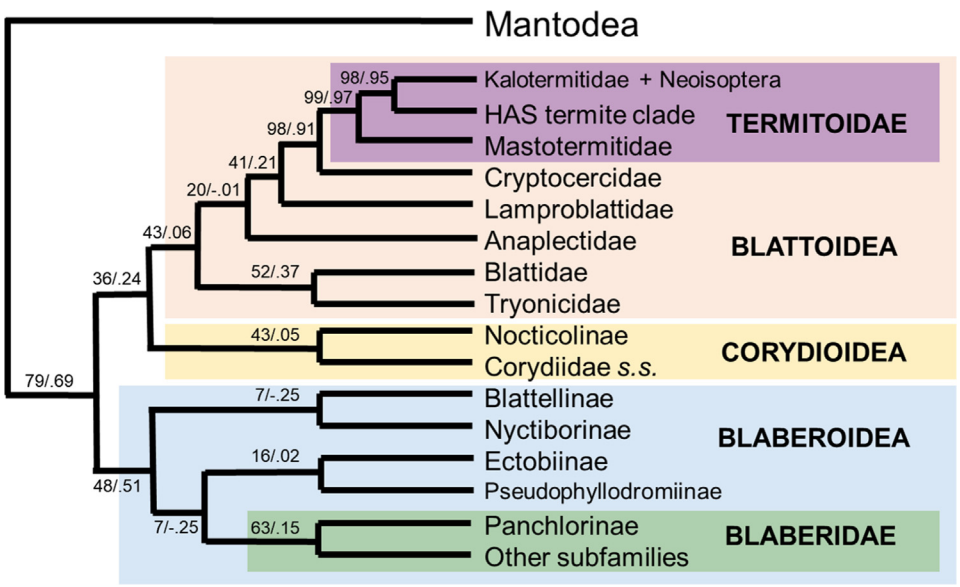

Fig. 2. Focal tree topology (1geneMin tree) for major lineages in Blattodea and support. Tree was inferred in RAXML from molecular data for six loci obtained for 575 taxa. Support values are bootstrap frequencies and internode certainty values. We see strong support for the monophyly of Blattodea, Blaberoidea, termites + Cryptocercidae, and termites; moderate support for Lamproblattidae + Cryptocercidae + termites, Tryonicidae + Blattidae, and Blaberidae; low support for Anaplectidae + sister group, Blattoidea, Corydioidea, and most relationships in Blaberoidea. Full tree with support values is available in Appendix D. See Fig. 1 caption for family-name usages.

bootstrap sets. TC scores are raw sums of all node IC's and RTC are those values normalized in relation to the number of branches whose removal induces a non-trivial bipartition (Kobert et al., 2016). Certainty values are a measure of diversity of alternative topologies among a set of trees (Kobert et al., 2016). Thus certainty values differ from bootstrap frequencies, which only indicate abundance of the reported topology. Using certainty scores, one can differentiate between scenarios where the focal tree topology is the most common topology, or less common than another tree topology within the pseudoreplicate set. For example, a clade with 40% bootstrap support could be the most commonly recovered clade (IC $>$ 0) or less commonly recovered than another clade (IC $<$ 0). Certainty scores were computed using the RAxML "-L MR" options on the bootstrapped pseudoreplicate sets (Kobert et al., 2016). IC-all (ICA) scores, which take into account all recovered clades instead of just the two most abundant, were calculated for the fully-sampled tree (i.e. the "focal tree" obtained as in Section 2.1) using the "-f i" options.

Test 4. Assessing statistical support for alternative topologies. The approximately unbiased (AU) test was used to determine whether the full dataset statistically rules out alternative topologies (Shimodaira, 2002). The AU test is a tree choice test that corrects for the selection and comparison biases of similar tests (Shimodaira, 2002). Yet, all the tests are based on the same underlying principle of assessing variance in the likelihoods of alternative trees inferred from bootstrapped tree sets and assigning p-values to topologies based on those variances (Kishino & Hasegawa, 1989). The "focal tree" described in Section 2.1 was used as a starting tree. The tree was manually edited in Mesquite (Maddison and Maddison, 2017) to create alternative input trees and to remove branch lengths. AU tests were done in the software IQ-Tree (Nguyen et al., 2015) using 10,000 resampled estimated log-likelihood (RELL) bootstraps and the partitioning scheme define in Section 2.1.

Test 5. Visualizing split signal and conflict. The purpose of Test 5 was to evaluate signal and conflict within the data (as opposed to conflict in inferences from the data as in Test 3). Optimized phylogenetic trees inferred from datasets show relationships heavily favored by emergent signal in the dataset but do not show alternative relationships partly supported by the data, as phylogenetic networks can. However, phylogenetic networks cannot be entirely represented in a 2-dimensional space (Wägele and Mayer, 2007) making them difficult to assess qualitatively. Lento plots, also known as spectral plots/analysis, are a simple way of representing the abundance of nucleotide characters supporting specific splits in a phylogenetic network, splits that may conflict with one another. Bars, whose height represent the number of characters supporting a given split, are plotted for any given number of splits in descending order of support. Using the software SAMS (Wägele and Mayer, 2007) split noise and conflict are also indicated. Visually inspecting these plots in parallel with a tree can give evidence for long-branch attraction phenomena, signal/noise ratio and information content of the data (Wägele and Mayer, 2007). Since alignments with high proportions of missing data can be problematic in SAMS, the alignments were trimmed both vertically and horizontally using a custom script in Mathematica (Appendix E) before generating Lento plots. SAMS was used with default parameters (all max noise = 0.25, consensus threshold = 0.5, pairwise comparison, analyze 150 occurring splits) except gaps were treated as missing data and total (tree) split compatibility was visualized.

Test 6. Assessing locus saturation. Quantifying saturation in molecular markers informs on their phylogenetic utility and the time-scale of their optimal utility (Salichos and Rokas, 2013). Mutation saturation was assessed in two ways: plotted raw pairwise genetic distances against GTR model corrected pairwise genetic distances, and plotted raw pairwise genetic distances against pairwise tree branch-length distances using the "focal tree" with all taxa. Genetic distances were calculated in PAUP* (Swofford, 2002) and plotted in Mathematica (Wolfram Research, 2012). Tree distances were calculated in R using the "cophenetic" function (Sneath and Sokal, 1973). Input data can be found in Appendix F.

Test 7. Assessing homoplasy in molecular markers. Here we distinguished markers evolving mostly parsimoniously with tree structure from more homoplastic markers using consistency index (CI) and retention index (RI) scores. Though homoplastic characters are not equivalent to noisy characters, extremely homoplastic characters contribute more to tree error than other characters (Breinholt and Kawahara, 2013; Wenzel and Siddall, 1999; but see Vogler et al., 2005). Scores were calculated for alignment regions in R using the package "phangorn" (Schliep, 2011) and the "focal tree".

3. Results

3.1. Alternative topologies and support (Tests 1–4)

Taxon reduced tree inferences (Test 1; Appendices G & H) were mostly congruent with the "focal tree" (Fig. 2; Appendix D). Among the relationships examined, the position of Anaplectinae and the relationships within Blaberidae differed the most. The results of the gene reduction schemes (Test 2; Appendix H) showed the positions of termite

Table 2
Alignments used in various analyses. 10 alignments used in our main analysis (1geneMin) and subsequent tests of taxon (Test 1) and gene (Test 2) reduction. First 5 alignments are the full alignment with subsequent reductions of taxa by removal of taxa with increasing amounts of missing data (e.g. 3geneMin equals 1geneMin but with all taxa missing more than 3 markers removed). Last 5 alignments are the 6geneMin (alignment with all 92 taxa having all 6 markers present) with the specified number of alignment blocks removed (e.g. -92Blocks alignment has 92 random partitions from random taxa removed). Various alignment statistics are taken from RAXML or calculated in Mathematica 9.

Alignment/Tree name	Num. of taxa	Num. nucleotides	Num. of patterns	Proportion of gaps	Tree likelihood	Relative tree certainty	Num of bootreps	Mean bootscore	Median bootscore	Std. dev. Bootscores
1geneMin (focal tree)	575	7676	5660	62.6%	−290610	0.383	360	59.9	62.5	32.8
3geneMin	321	7651	5210	54.1%	−229581	0.421	360	62.8	64.2	33.0
4geneMin	231	7630	5116	46.2%	−201935	0.458	408	66.6	71.8	31.6
5geneMin	136	7566	4650	32.7%	−151287	0.485	360	68.3	77.6	32.2
6geneMin (-0Blocks)	92	7515	3989	25.9%	−116123	0.551	252	75.0	90.5	29.0
-92Blocks	92	7490	3751	38.5%	−100299	0.513	252	72.1	85.9	30.4
-184Blocks	92	7472	3492	53.1%	−82993	0.318	504	56.2	53.0	30.8
-276Blocks	92	7415	3269	62.5%	−68310	0.278	852	49.8	49.5	33.2
-368Blocks	92	7368	2943	73.1%	−52775	0.174	600	34.2	22.8	32.8
-460Blocks	92	7139	2169	88.2%	−30880	0.302	600	47.5	50.7	36.4

lineages and the sister relationship between Blaberidae and Pseudophyllodromiinae + Ectobiinae are largely robust to perturbation.

The relative tree certainty scores (Table 2; Appendices G & H) correlated with mean bootstrap values (Test 3). The gene reduction tree with 184 data block removals (-184Blocks) had comparatively less certainty relative to its bootstrap scores. In the "focal tree", the most uncertain relationships were the sister group to Blaberidae, and the position of Blattellinae + Nyctiborinae (Appendix H; Fig. 2). Both of these relationships had a negative weighted-average internode certainty, indicating that alternative topologies were more common than the relationships recovered in the "focal tree". The relationships with the highest internode certainty relative to bootstrap support were the positions of *Lamproblatta* and *Tryonicus*.

Test 4 (Appendices H & I) showed that even relationships recovered consistently in Test 1 (and with relatively high bootstrap and internode certainty support values) were not unambiguously supported by the data. Multiple alternative topologies, some of which were never recovered in any of the alignment permutation analyses (Tests 1 and 2), could not be rejected. The alternative topology with Ectobiinae sister to remaining Blaberoidea was never recovered in any of our analyses but actually produced higher log-likelihood (logL) score in the AU test (Appendix I).

3.2. Signal and conflicts (Tests 5–7)

Lento plots (Test 5; Fig. 3; Appendix E) showed that the majority of nucleotide support for splits (signal) are in 18S and 28S, while COI + COII had moderate signal, and 12S + 16S had low signal. Intra-marker conflict was highest in the protein coding genes, and lowest among the nuclear rDNAs. The least conflict occurred in the combined alignment, with only 46% of the first 50 splits conflicting and none of the top 10 splits conflicting. In general, the more markers considered the less internal conflict in the alignment. However, the two nuclear markers together contain more signal and less internal conflict than all four mitochondrial markers (Appendix E).

The two data quality analyses (Tests 6 & 7; Figs. 4–6; Appendix F) both showed that COI, COII and 12S seem to have a more limited phylogenetic utility in Blattodea than 16S, 18S and 28S. 12S and COII in particular had both low consistency (Fig. 6) and high saturation at deep branches (Fig. 5). Consideration by codon position showed that first codon positions in COII and third codon positions in both COI and COII are highly saturated (Fig. 4).

4. Discussion

4.1. Support for the phylogeny of Blattodea

The phylogenetic hypothesis we recovered from the full dataset (1GeneMin tree; Fig. 2; Appendix D) mostly agreed with the topology of Legendre et al. (2015). The tree in Legendre et al. (2015) has higher bootstrap support values comparatively (Appendix H). One notable difference though is in the support for *Panchlora* as sister to the remaining Blaberidae, which is strongly supported in Legendre et al. (2015) but given nearly negligible support here. Most other differences between the two trees are relationships within Blaberidae. Also, since Legendre et al. (2015) did not include most of the Anaplectidae included here, it should be noted that this clade was recovered as polyphyletic, with the majority of the family as highly volatile within Blattoidea. One individual ("*Anaplecta* sp.") fell in Pseudophyllodromiinae and it must be investigated further whether this placement results from a specimen misidentification or a poorly defined taxon that needs revision.

Alignment permutations tests (Appendix H) helped identify some unstable relationships in the tree. Other studies removing taxa and affecting alignment completeness have shown similar variations in tree topology (Chen et al., 2015). The position of Anaplectidae (except "*Anaplecta* sp.") and the lineages of Blaberidae were the most affected by alignment permutations. Most of the other relationships were more stable, which contradicts their volatility when comparing among (Evangelista et al., 2017) or within (Djernæs et al., 2015) past studies. Similarly, the phylogenetic relationships reconstructed here are more stable overall than those Ware et al. (2008) found when experimenting with taxon removal/addition in Dictyoptera phylogenetics. The higher stability here [and in Legendre et al. (2015) because their trees were mostly congruent] is likely due to the larger overall sampling and use of more diverse outgroups than in Ware et al. (2008). This is in agreement with theoretical and empirical studies underlining the importance of large taxon sampling (Heath et al., 2008; Zwickl and Hillis, 2002).

The approximately unbiased (AU) test results mostly corroborated topological support (both strong and weak) from previous tests, but gave some unique insight (Tests 4; Appendix I). The AU tests did not reject any of the topologies that were volatile in the alternative analyses (see Section 2.2, Test 1; Appendix H). Similarly, AU tests did not reject alternative positions for Corydioidea or Nyctiborinae + Blattellinae, yet their placement was not volatile according to Tests 1 and 2. This illustrates that even if the phylogenetic analysis of a dataset results in a best unique tree, this dataset is not necessarily statistically incompatible with alternative topologies (Buddenhagen et al., 2016). Importantly, in

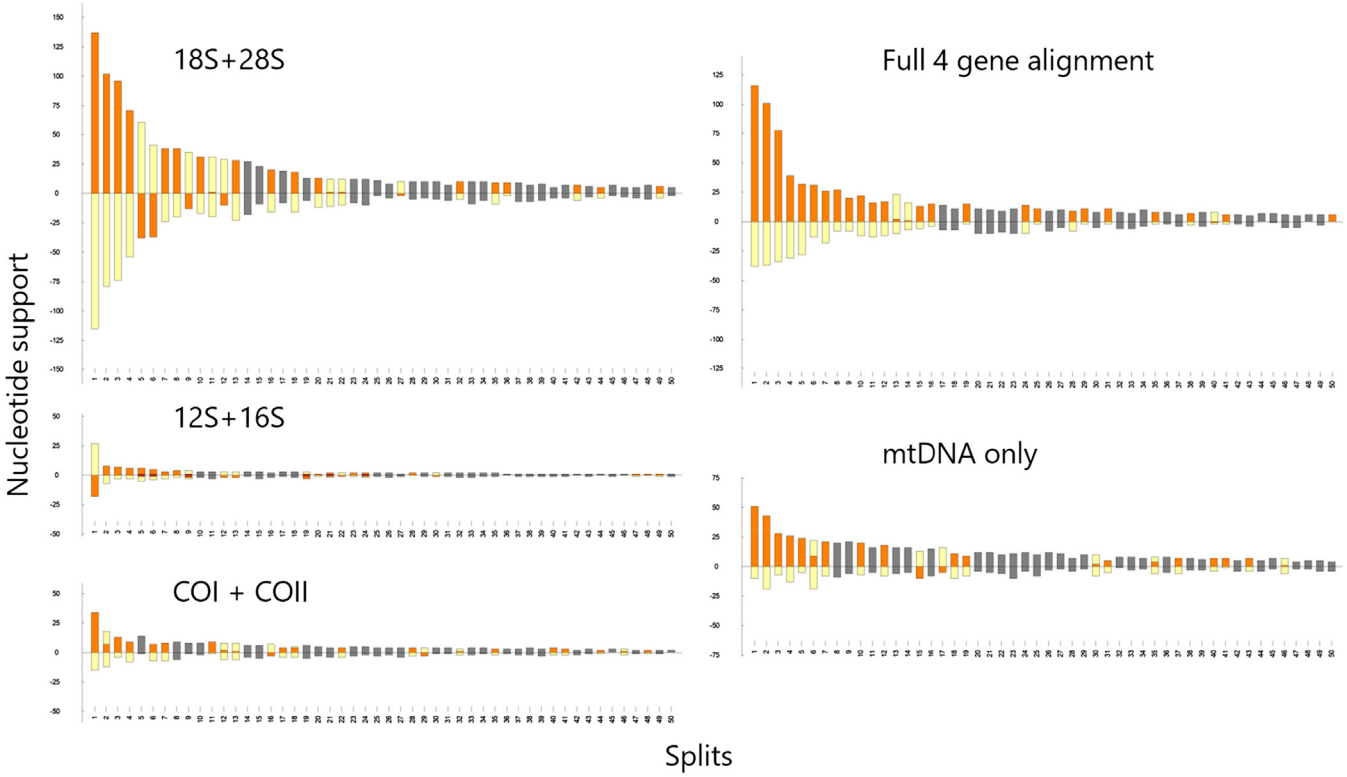

Fig. 3. Splits support spectrograms for alignments. We focus on alignments for 18S and 28S together, 12S and 16S together, COI and COII together, all mtDNA together and a full-length alignment (4geneMin alignment). Horizontal axes are tree splits (bipartitions) supported by the data. Vertical axis shows how many nucleotides support each split. Orange and tan bars indicate noisy data supporting the splits. Grey bars indicate a split that is supported by the alignment but is conflicting with another split that has higher nucleotide support. Only support for the first 50 splits are shown.

a dataset with heterogeneous partition quality and missing data patterns (Table 1) the AU test may lack the statistical power to favor one topology over another, which likely explain why this test was very conservative in comparison to the others. As such, we should consider topologies not uniquely supported by the AU test as given tentatively low support, and other test results should take precedence.

Some relationships we focused on (Appendix H) had negative ICA values. In these cases, the relationships in the 1GeneMin tree were less frequent than other relationships recovered among the pseudo-replicate trees inferred from the same alignment. This reveals conflicts, incongruent phylogenetic signal, among or within the phylogenetic markers used, which is common in datasets with a small number of markers

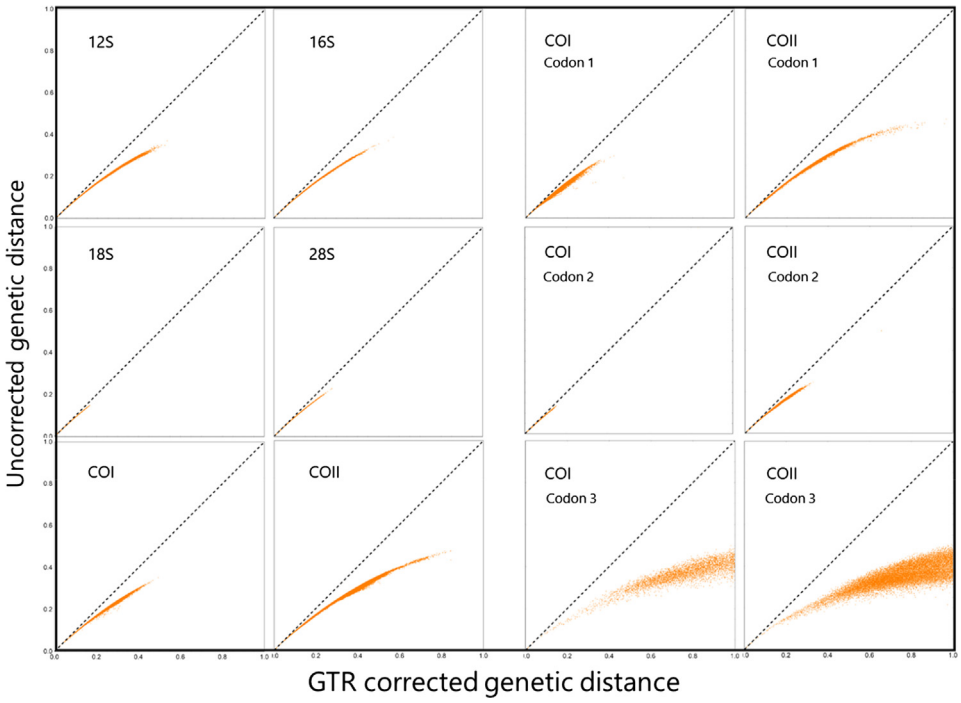

Fig. 4. Overall locus saturation overall. Pairwise genetic distance with GTR model corrected distance on the horizontal axis and uncorrected genetic distances on the vertical axis. A dashed line is included so that the genetic distances can be compared with parity. Parity would indicate that there is no locus saturation (no homoplasy, and high signal) and high deviation from parity indicates that there is high saturation (high homoplasy and noise). For consistency, all plot ranges are given as between 0 and 1. Plots indicate that COI, COII and 12S are the most saturated and 16S is moderately saturated. 18S and 28S are not saturated but have low genetic distances overall, indicating high conservation. For the protein coding genes, COII first codon positions and COI and COII 3rd codon positions are highly saturated.

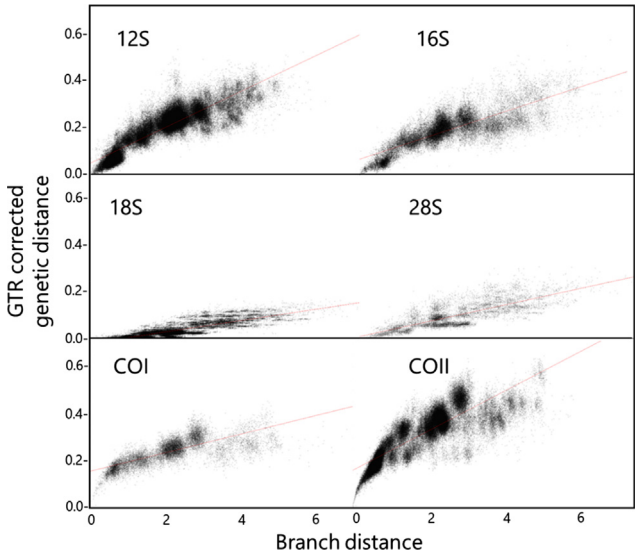

Fig. 5. Locus saturation per time-scale. Pairwise branch distance on the horizontal axis plotted against pairwise GTR model corrected distance on the vertical axis. Trend lines calculated from a best fit linear regression. Equations for each trend line are as follows: 12S, 0.048 + 0.077x; 16S, 0.063 + 0.053x; 18S, −0.019 + 0.023x; 28 s, 0.01 + 0.035x; COI, 0.155 + 0.040x; COII, 0.160 + 0.084x. Plots show that 18S and 28S are more conserved at all ages and have phylogenetic utility at deep phylogenetic scales. 16S and COI are moderately fast evolving. 12S and COII are the least conserved and may be limited in their use at deep phylogenetic scales.

(Narechania et al., 2012). The reasons for this pattern are unclear, however we suspect it might have to do with internal non-independence of rRNA's (due to paired-mutations in self-bonded regions), patterns of missing data (Dell'Ampio et al., 2014), or heterogeneous partition quality (Table 1). The low scores overall and the heterogeneity among pseudoreplicate trees indicates that: more markers are needed to stabilize the tree topology (Narechania et al., 2012) and dilute potential biases; better modelling strategies are needed (e.g. Letsch and Kjer, 2011; Zhang et al., 2015); or missing data patterns should be filled in through new sequencing. In parallel, signal analyses must be conducted to differentiate the markers supporting the relationships from those with highly conflicting signal.

4.2. Signal and conflicts in molecular markers

Lento-plots (Fig. 3; Appendix E) give insights into three classes of long-branch effects (Wägele and Mayer, 2007): symplesiomorphy driven (class 1), signal-erosion driven (class 2), and homoplasy driven (class 3). We did not find evidence of class 1 in the data because there are few splits where large clades are lumped with outgroups. When ingroup taxa matched with outgroups they also tended to group incorrectly with a few other ingroup taxa supported by long branches. These taxa were: Latindiinae spp., Nocticolidae sp., *Anaplecta* sp., *Supella longipalpa*, *Dendroblatta* sp. and *Isoldaia* sp. Also, the taxa clustering with outgroups were individual species or members of the same genus, suggesting that the states driving the long branch effect are derived. These taxa are probably clustering under a class 2 long branch effect derived from saturated nucleotide positions (Figs. 4 and 5; Test 6; as in Omilian and Taylor, 2001). Lento plots for 3rd codon positions clearly indicate class 2 long branch effects and all splits are polyphyletic when mapped onto the "focal tree". Class 3 effects would result from homoplastic sites contributing to conflicting signal (as in the data of Crandall et al., 2001), but conflict was low in the combined data Lento plots (Fig. 3) so presence of a class 3 effect seems less likely here.

Interpretation of splits analysis methods (such as Lento plots; e.g. Wägele et al., 2009), and integration with other analytics, can also give insight into the phylogenetic value of specific molecular markers. 18S and 28S provided the highest support to splits (Fig. 3) despite being conserved overall (Fig. 5) and having the lowest density of unique patterns (Table 1). Yet, including mtDNA markers, which had low phylogenetic signal on their own (Fig. 3), did improve split support and lessened internal conflict (Fig. 3). Given that many of the nodes on the final tree have low support (Appendix H) it is unclear how the addition of the mtDNA markers, which were comparatively more conflicting, affected them. In fact, given the much stronger nucleotide compositional bias in the mtDNA (Table 1) it is possible that nodes supported ambiguously by nucDNA markers would have been swayed by mtDNA compositional bias. Additional tests are needed to further assess these effects (i.e. site rate variation, compositional heterogeneity and violation of sequence evolution assumptions) from individual markers (Song et al., 2010).

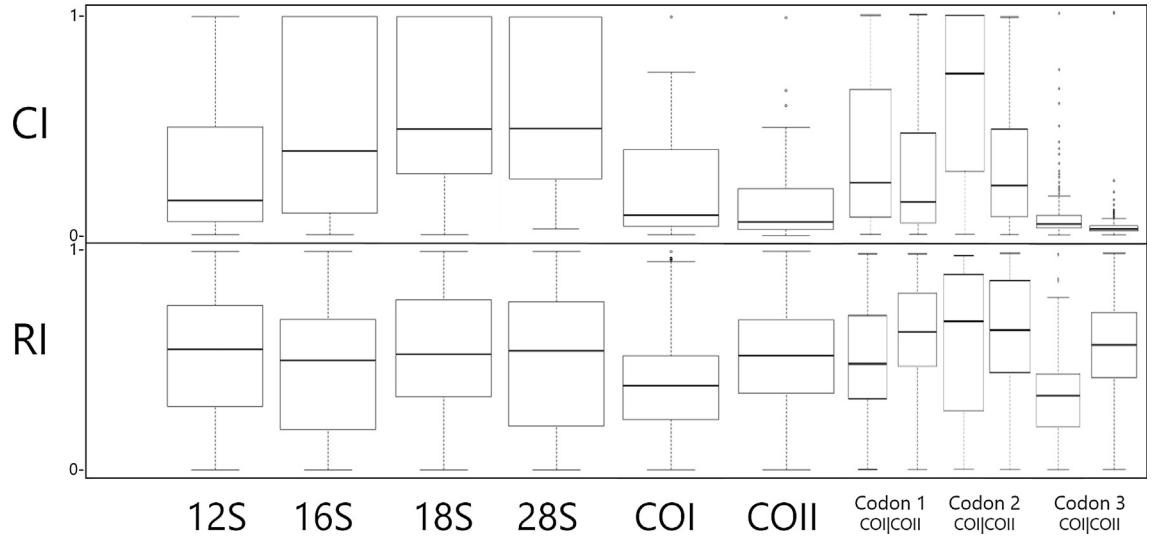

Fig. 6. Homoplasy among loci. Plots of consistency index (CI) and retention index (RI) for the six loci considered in this study. Protein coding genes are also assessed for each codon position. CI scores show greater differences than RI scores. CI scores for 16S, 18S, and 28S show that these loci have low levels of homoplasy. Contrarily, 12S, COI and COII are highly homoplasious. However, COI and COII second codon positions and COI first codon position appear less homoplastic, while third codon positions for both and first codon positions for COII appear highly homoplastic.

	Test 1 Taxon reductions	2 Gene reductions	3 Weighted Bootstrap	3 Weighted IC	4 AU test
Corydioidea sister to Blattoidea	100%	66%	41	0.09	Not unique
Lamproblattidae sister to Cryptocercidae + Termites	75%	NA	70	0.57	Not unique
Tryonicidae sister to Blattidae	66%	NA	72	0.6	Not unique
Blattellinae + Nyctiborinae sister to remaining Blaberoidea	100%	NA	17	-0.06	Not unique
Ectobiinae sister to Pseudophyllodromiinae	100%	NA	39	0.29	Not unique
Panchlora sister to all other Blaberidae	33%	NA	44	0.23	Not unique

Fig. 7. Summary of Tests 1–4 showing that multiple tests are needed to assess phylogenetic support. Results from Tests 3 to 4 taken from "focal tree". Tests 1 and 2 results are the percentage of relevant trees in which the relationship was recovered (trees lacking the taxa in question were omitted from this calculation). "NA" indicates the relationship could not be tested (see Appendix H). Test 3 values are weighted averages. See Section 2.2 for explanation. Test 4 results indicate if a relationship was uniquely supported by an AU test, or if other relationships were also plausible (not unique). Black cells indicate results supporting the relationship in question, white cells are results showing lack of support, and grey cells are ambiguous.

4.3. Combining several tests for a better understanding of phylogenetic hypotheses

The assays employed here proved useful for elucidating some dark regions in the Blattodea tree and for understanding the discordance among previous studies (as illustrated in Evangelista et al., 2017). However, no single test or metric was sufficient to identify either all areas of the tree that are in need of new data (Fig. 7) or the relative contribution of molecular markers to tree instability (Chiapella et al., 2014; Vogler et al., 2005). This is unsurprising given that phylogenetic uncertainty comes from a multitude of sources and that no single test or metric is able to identify them all, which underline the importance of confronting them.

A dark region identified using multiple tests is the position of Corydioidea (= Corydiidae *s.l.* = Polyphagoidea). On one hand, alignment permutation tests (Tests 1 and 2) consistently recovered Corydioidea as sister to Blattoidea (Appendix H). There were two exceptions (-92Blocks and -460Blocks trees) but this was still one of the more stable relationships and it had moderately high support (weighted bootstrap in the "focal tree" = 41). On the other hand, weighted ICA for this relationship was near 0 (Test 3), and AU tests (Appendices H and I) showed that the alternative of Corydioidea as sister to all of Blattodea is also a plausible topology (Test 4). Wang et al. (2017) claimed to have recovered Corydioidea + all other Blattodea with strong support, but their assessment of support was limited to bootstrap frequency and Bayesian posterior probability. We have not reanalyzed their data so we do not know if other tests would agree with their topology. Given the discordance between our tests, we doubt the certainty of either proposed relationship. To summarize, we found the phylogenetic position of Corydioidea as stable and with moderate bootstrap support, but with very low ICA scores and no statistical support in the AU test. Using only one of these tests would have resulted in drastic conclusions (in one way or another), which would have been premature.

Similarly, the set of analyses conducted here brought mixed results for the position of Tryonicidae. This lineage is thought to be important in understanding the evolutionary past of termites because Murienne (2009) and Djernæs et al. (2015) recovered Tryonicidae as sister to (Cryptocercidae + termites). This relationship was never recovered in any of the alignment permutations (Tests 1 and 2), and IC values were relatively high in the 1GeneMin tree (Test 3). This set of results is congruent with Legendre et al. (2015). AU-tests, however, did reject some alternative topologies but not all (Test 4; Appendix I) and Lento plots gave no indication of a potential long branch effect with this taxon (Test 5; Appendix E). The different positions of Tryonicidae in past studies could be due to: missing Anaplectinae and Lamproblattidae and a relatively limited taxon sampling in Murienne (2009); inclusion of morphological data, different alignment partitioning schemes or tree reconstruction methods in Djernæs (2015).

4.4. Suggestions and future directions

The tree of life will undoubtedly become more fully resolved as transcriptomic and genomic data are increasingly collected from diverse taxa (e.g. Blaimer et al., 2015). "Big data" are not yet available for every group of organisms and some progress could still be made with pre-omics datasets, such as the six molecular markers studied here. However, we must analyze how these molecular data contribute to phylogenetic inferences using tests that will help identify both current dark areas in phylogenetic trees and the target profile of future markers to decipher these unresolved areas. An effective strategy would be to mine genomic studies (e.g. Granados Mendoza et al., 2015; Misof et al., 2013, 2014) for new, more informative, and more independent markers (as in Chen et al., 2015). Some of the tests discussed here should be used to tailor locus choice to fit the needs of a particular phylogenetic question. For instance, our evidence of class 2 or 3 long branch effects could be remedied by targeting characters more conserved than the saturated or low consistency ones found here. This aim is not easily achieved though, as the interaction of mutation rate, timescale and phylogenetic context is rarely straightforward. The phylogeny of Blattodea illustrates how a given set of molecular markers provide excellent resolution at a certain timescale for one clade (i.e. termites) but low resolution for similar time-scales on nearby clades (i.e. Blaberoidea; considering dates from Wang et al., 2017). Though, it is good to keep in mind that some nodes in phylogenies are inherently difficult to infer (Narechania et al., 2012; Salichos and Rokas, 2013) and identifying which nodes these are is key to our understanding of evolutionary histories (Whitfield and Lockhart, 2007).

For numerous taxa, pre-omics datasets are still useful and, to benefit

even further from them, a number of steps could be considered prior to tree reconstruction (Misof et al., 2014b). With our dataset, structural alignment of rRNA had minimal effect on tree reconstruction. Still, not aligning with respect to structural homology (as in Bourguignon et al., 2015; Djernæs et al., 2015; Legendre et al., 2015) could introduce error associated with phenetic inference (Kjer et al., 2007), which could reduce tree accuracy and artificially inflate support values. Further improvements could be made by down-weighting rRNA stems, and modelling stem and loop regions separately, but only if these regions do not violate model assumptions (Letsch and Kjer, 2011). Finally, omitting saturated nucleotide regions could improve tree reconstruction (Breinholt and Kawahara, 2013; Wenzel and Siddall, 1999), although sometimes highly homoplastic sites provide "hidden support" to nodes, even deep ones (Vogler et al., 2005; Wenzel and Siddall, 1999). Saturated sites could be removed by alignment masking software (e.g. Castresana, 2000; Wu et al., 2012) or in protein coding genes analyzed with a "R-Y" coding. Running multiple tests prior to tree reconstruction might help in making educated guesses as to the best strategy to follow for a given dataset. Relationships recovered consistently in taxon reduction tests but that also have low node support (such as the position of Corydioidea) could be better resolved by improving alignment completeness with the current molecular markers. Conversely, highly volatile taxa (like Blaberidae lineages) will probably not benefit from additional sequencing to improve completeness of these six markers, since multiple individuals representing major lineages have highly complete data but are still very volatile. Anaplectidae, although volatile, currently has such a low data completeness that further sequencing might provide resolution.

Our results show that assessing the strength of a phylogenetic hypothesis requires integration of tests and metrics. Tree independent methods of visualizing support, such as Lento Plots (via SAMS; Wägele and Mayer, 2007) or split-networks (Bryant and Moulton, 2004; Wägele et al., 2009), directly measure support for relationships and can be used to identify data features causing long branch attraction (Wägele and Mayer, 2007). Bootstrap resampling can give estimates of conflicting signal for clades (Kobert et al., 2016; Soltis and Soltis, 2003) but bootstrap frequencies alone may not be fully indicate the extent of conflict and can be complemented with certainty scores (Kobert et al., 2016). Similarly, statistical tests of topology (i.e. AU test) can test alternatives not recovered in the most likely tree (Shimodaira, 2002) but their power is limited by data homogeneity. These multifaceted approaches improve understanding of the dependence of the results on sampling (Narechania et al., 2012; Ware et al., 2008), modelling (Djernæs et al., 2015), alignment bias (Dell'Ampio et al., 2014), or inference method (Djernæs et al., 2015; Zhang et al., 2015).

5. Conclusions

Molecular datasets are ever-increasing in size so that interpreting the relationship between input and output is becoming more difficult. However, comprehensive tests can disentangle dataset weaknesses and strengths. Assessments of data quality can be done both before and after tree inference but combining the results gives the most meaningful assessment. Multiple lines of evidence are needed to fully identify tree support and data robustness.

In the phylogeny of Blattodea, only AU tests and ICA scores, and not bootstrap scores or alignment permutations, illustrated the lack of support for the placement of Corydioidea. Split support analysis showed that few deep relationships in the ingroup had visible character support. As such, tree inference and the calculation of support values are vulnerable to data biases, signal saturation, conflicting signal and other issues. Indeed, our contradictory results and alternative topologies could not be rejected using the information in the alignment. More optimistically, although there was evidence of long-branch attraction in the split support plots, long-branch taxa appear correctly placed in the "focal tree". Yet, the long-branch signal may be contributing to low tree support.

Acknowledgements

Funding: This work was supported by the National Science Foundation [award#: 1608559] and Labex BCDiv for France Thouzé.

The authors would like to acknowledge Jessica L. Ware and others for assistance with rRNA secondary structure alignment. Also, thanks to Karen Meusemann and Christoph Mayer for advice on software and analyses. Finally, a deep thank you to the editors handling this paper, and the reviewers as well.

Author contributions: FT and FL conceived the study and laid the groundwork; FT did preliminary analyses, DE and MKK composed the final alignment and did rRNA structural alignment; DE completed the final analyses, and composed the figures; DE and FL wrote the paper; FT, and MKK also contributed to the text of the paper and figures.

Appendices. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.05.007.

References

Beaulieu, J.M., Ree, R.H., Cavender-Bares, J., Weiblen, G.D., Donoghue, M.J., 2012. Synthesizing phylogenetic knowledge for ecological research. Ecology 93, S4–S13.

Blaimer, B.B., Brady, S.G., Schultz, T.R., Lloyd, M.W., Fisher, B.L., Ward, P.S., 2015. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. BMC Evol. Biol. 15, 271.

Borowiec, M.L., 2017. Convergent evolution of the army ant syndrome and congruence in big-data phylogenetics. bioRxiv (preprint).

Borowiec, M.L., Lee, E.K., Chiu, J.C., Plachetzki, D.C., 2015. Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa. BMC Genom. 16, 987.

Bourguignon, T., Lo, N., Cameron, S.L., Sobotnik, J., Hayashi, Y., Shigenobu, S., Watanabe, D., Roisin, Y., Miura, T., Evans, T.A., 2015. The evolutionary history of termites as inferred from 66 mitochondrial genomes. Mol. Biol. Evol. 32, 406–421.

Bradler, S., Robertson, J.A., Whiting, M.F., 2014. A molecular phylogeny of Phasmatodea with emphasis on Necrosciinae, the most species-rich subfamily of stick insects. Syst. Entomol. 39, 205–222.

Breinholt, J.W., Kawahara, A.Y., 2013. Phylotranscriptomics: saturated third codon positions radically influence the estimation of trees based on next-gen data. Genome Biol. Evol. 5, 2082–2092.

Bremer, K., Friis, E.M., Bremer, B., Linder, P., 2004. Molecular phylogenetic dating of asterid flowering plants shows early cretaceous diversification. Syst. Biol. 53, 496–505.

Bryant, D., Moulton, V., 2004. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. Mol. Biol. Evol. 21, 255–265.

Buddenhagen, C., Lemmon, A.R., Lemmon, E.M., Bruhl, J., Cappa, J., Clement, W.L., Donoghue, M., Edwards, E.J., Hipp, A.L., Kortyna, M., Mitchell, N., Moore, A., Prychid, C.J., Segovia-Salcedo, M.C., Simmons, M.P., Soltis, P.S., Wanke, S., Mast, A., 2016. Anchored Phylogenomics of Angiosperms I: Assessing the Robustness of Phylogenetic Estimates. bioRxiv (preprint).

Cameron, S.L., 2013. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu. Rev. Entomol. 59, 95–117.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552.

Chaboo, C.S., Frieiro-Costa, F.A., Gómez-Zurita, J., Westerduijn, R., 2014. Origins and diversification of subsociality in leaf beetles (Coleoptera: Chrysomelidae: Cassidinae: Chrysomelinae). J. Nat. History 1–43.

Chen, M.-Y., Liang, D., Zhang, P., 2015. Selecting question-specific genes to reduce incongruence in phylogenomics: a case study of jawed vertebrate backbone phylogeny. Syst. Biol. 64, 1104–1120.

Chiapella, J., Kuhl, J., Demaio, P., Amarilla, L., 2014. Fishing for significance in phylogenies: too many alternatives for the same outcome, or an appeal to journal editors. Ideas Ecol. Evol. 7, 3–7.

Crandall, K., Fetzner, J., Jara, C., Buckup, L., 2001. On the phylogenetic positioning of the South American freshwater crayfish genera (Decapoda: Parastacidae). J. Crustacean Biol. 20, 530–540.

Deitz, L., Nalepa, C.A., Klass, K.D., 2003. Phylogeny of the Dictyoptera Re-examined (Insecta). Entomol. Abhandl. 61, 69–91.

Dell'Ampio, E., Meusemann, K., Szucsich, N.U., Peters, R.S., Meyer, B., Borner, J., Petersen, M., Aberer, A.J., Stamatakis, A., Walzl, M.G., Minh, B.Q., von Haeseler, A., Ebersberger, I., Pass, G., Misof, B., 2014. Decisive data sets in phylogenomics: lessons from studies on the phylogenetic relationships of primarily wingless insects. Mol. Biol. Evol. 31, 239–249.

Djernæs, M., Klass, K.-D., Picker, M.D., Damgaard, J., 2012. Phylogeny of cockroaches (Insecta, Dictyoptera, Blattodea), with placement of aberrant taxa and exploration of out-group sampling. Syst. Entomol. 37, 65–83.
Djernæs, M., Klass, K.D., Eggleton, P., 2015. Identifying possible sister groups of Cryptocercidae + Isoptera: a combined molecular and morphological phylogeny of Dictyoptera. Mol. Phylogenet. Evol. 84, 284–303.
Evangelista, D.A., Djernæs, M., Kohli, M.K., 2017. Fossil calibrations for the cockroach phylogeny (Insecta, Dictyoptera, Blattodea), comments on the use of wings for their identification, and a redescription of the oldest Blaberidae. Palaeontol. Electron. 20 (3), 1–23.
Farris, J.S., 1989. The retention index and homoplasy excess. Syst. Zool. 38, 406–407.
Fisher-Reid, M.C., Wiens, J.J., 2011. What are the consequences of combining nuclear and mitochondrial data for phylogenetic analysis? Lessons from Plethodon salamanders and 13 other vertebrate clades. BMC Evol. Biol. 11, 1–20.
Garrison, N.L., Rodriguez, J., Agnarsson, I., Coddington, J.A., Griswold, C.E., Hamilton, C.A., Hedin, M., Kocot, K.M., Ledford, J.M., Bond, J.E., 2016. Spider phylogenomics: untangling the Spider Tree of Life. PeerJ 4, e1719.
Granados Mendoza, C., Naumann, J., Samain, M.S., Goetghebeur, P., De Smet, Y., Wanke, S., 2015. A genome-scale mining strategy for recovering novel rapidly-evolving nuclear single-copy genes for addressing shallow-scale phylogenetics in Hydrangea. BMC Evol. Biol. 15, 1–13.
Grandcolas, P., 1996. The phylogeny of cockroach families: a cladistic appraisal of morpho-anatomical data. Can. J. Zool. 74, 508–527.
Grandcolas, P., 1997. Systematique phylogenetique de la sous-famille des Tryonicinae (Dictyoptera, Blattaria, Blattidae). Mem. Mus. Nat. Hist. Naturelle 171, 91–124.
Grandcolas, P., 1999. Reconstructing the past of Cryptocercus (Blattaria: Polyphagidae): phylogenetic histories and stories. Ann. Entomol. Soc. Am. 92, 303–307.
Heath, T.A., Hedtke, S.M., Hillis, D.M., 2008. Taxon sampling and the accuracy of phylogenetic analyses. J. Syst. Evol. 46, 239–257.
Hillis, D.M., Dixon, M.T., 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. Quart. Rev. Biol. 66, 411–453.
Inward, D., Beccaloni, G., Eggleton, P., 2007. Death of an order: a comprehensive molecular phylogenetic study confirms that termites are eusocial cockroaches. Biol. Lett. 3, 331–335.
Johnson, B.R., Borowiec, M.L., Chiu, J.C., Lee, E.K., Atallah, J., Ward, P.S., 2013. Phylogenomics resolves evolutionary relationships among ants, bees, and wasps. Curr. Biol. 23, 2058–2062.
Kambhampati, S., 1996. Phylogenetic relationship among cockroach families inferred from mitochondrial12S rRNA gene sequence. Syst. Entomol. 21, 89–98.
Kambhampati, S., Luykx, P., Nalepa, C.A., 1996. Evidence for sibling species in *Cryptocercus punctulatus*, the wood roach, from variation in mitochondrial DNA and karyotype. Heredity (Edinb.) 76, 485–496.
Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29, 170–179.
Kjer, K.M., Gillespie, J.J., Ober, K.A., 2007. Opinions on multiple sequence alignment, and an empirical comparison of repeatability and accuracy between POY and structural alignment. Syst. Biol. 56, 133–146.
Kjer, K.M., Roshan, U., Gillespie, J.J., 2009. Structural and evolutionary considerations for multiple sequence alignment of RNA, and the challenges for algorithms that ignore them. In: Rosenberg, M. (Ed.), Perspectives on Biological Sequence Alignment: Where, How, and Why It Matters. University of California Press, USA, pp. 105–332.
Kjer, K.M., Simon, C., Yavorskaya, M., Beutel, R.G., 2016. Progress, pitfalls and parallel universes: a history of insect phylogenetics. J. R. Soc. Interface 13, 1–29.
Klass, K.D., Meier, R., 2006. A phylogenetic Analysis of Dictyoptera (Insecta) based on morphological characters. Entomol. Abhandl. 63, 3–50.
Klass, K.D., 2001. Morphological evidence on Blattarian phylogeny: “phylogenetic histories and stories” (Insecta, Dictyoptera). Berliner Entomol. Z. 48, 223–265.
Klassen, G.J., Mooi, R.D., Locke, A., 1991. Consistency indices and random data. Syst. Biol. 40, 446–457.
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.
Kodandaramaiah, U., Simonsen, T.J., Bromilow, S., Wahlberg, N., Sperling, F., 2013. Deceptive single-locus taxonomy and phylogeography: Wolbachia-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. Ecol. Evol. 3, 5167–5176.
Kozak, K.H., Wiens, J.J., 2012. Phylogeny, ecology, and the origins of climate–richness relationships. Ecology 93, S167–S181.
Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701.
Lanier, H.C., Knowles, L.L., 2015. Applying species-tree analyses to deep phylogenetic histories: challenges and potential suggested from a survey of empirical phylogenetic studies. Mol. Phylogenet. Evol. 83, 191–199.
Legendre, F., Grandcolas, P., Thouzé, F., 2017. Molecular phylogeny of Blaberidae (Dictyoptera, Blattodea) with implications for taxonomy and evolutionary studies. Eur. J. Taxon. 291, 1–13.
Legendre, F., Nel, A., Svenson, G.J., Robillard, T., Pellens, R., Grandcolas, P., 2015. Phylogeny of Dictyoptera: dating the origin of cockroaches, praying mantises and termites with molecular data and controlled fossil evidence. PLoS One 10, e0130127.
Legendre, F., Robillard, T., Song, H., Whiting, M.F., Desutter-Grandcolas, L., 2010. One hundred years of instability in ensiferan relationships. Syst. Entomol. 35, 475–488.
Legendre, F., Whiting, M.F., Bordereau, C., Cancello, E.M., Evans, T.A., Grandcolas, P., 2008. The phylogeny of termites (Dictyoptera: Isoptera) based on mitochondrial and nuclear markers: Implications for the evolution of the worker and pseudergate castes, and foraging behaviors. Mol. Phylogenet. Evol. 48, 615–627.
Letsch, H.O., Kjer, K.M., 2011. Potential pitfalls of modelling ribosomal RNA data in phylogenetic tree reconstruction: evidence from case studies in the Metazoa. BMC Evol. Biol. 11, 146.
Lo, N., Bandi, C., Watanabe, H., Nalepa, C., Beninati, T., 2003. Evidence for cocladogenesis between diverse dictyopteran lineages and their intracellular endosymbionts. Mol. Biol. Evol. 20, 907–913.
Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J., 2000. Molecular Cell Biology. Freeman, New York, W.H.
Maddison, W.P., Maddison, D.R., 2017. Mesquite: a modular system for evolutionary analysis.
Maekawa, K., Lo, N., Rose, H.A., Matsumoto, T., 2003. The evolution of soil-burrowing cockroaches (Blattaria: Blaberidae) from wood-burrowing ancestors following an invasion of the latter from Asia into Australia. Proc. Biol. Sci. 270, 1301–1307.
Maganga, G.D., Kapetshi, J., Berthet, N., Ilunga, B.K., Kabange, F., Kingebeni, P.M., Mondonge, V., Muyembe, J.-J.T., Bertherat, E., Briand, S., Cabore, J., Epelboin, A., Formenty, P., Kobinger, G., González-Angulo, L., Labouba, I., Manuguerra, J.-C., Okwo-Bele, J.-M., Dye, C., Leroy, E.M., 2014. Ebola virus disease in the democratic Republic of Congo. New Engl. J. Med. 371, 2083–2091.
Mandal, S.D., Chhakchhuak, L., Gurusubramanian, G., Kumar, N.S., 2014. Mitochondrial markers for identification and phylogenetic studies in insects – a review. DNA Barcodes 2, 1–9.
Marvaldi, A.E., Duckett, C.N., Kjer, K.M., Gillespie, J.J., 2009. Structural alignment of 18S and 28S rDNA sequences provides insights into phylogeny of Phytophaga (Coleoptera: Curculionoidea and Chrysomeloidea). Zool. Scr. 38, 63–77.
McKittrick, F.A., 1965. A contribution to the understanding of cockroach-termite affinities. Ann. Entomol. Soc. Am. 58, 18–22.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for Inference of. Large Phylogenetic Trees.
Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A.J., Aspock, U., Aspock, H., Bartel, D., Blanke, A., Berger, S., Bohm, A., Buckley, T.R., Calcott, B., Chen, J., Friedrich, F., Fukui, M., Fujita, M., Greve, C., Grobe, P., Gu, S., Huang, Y., Jermiin, L.S., Kawahara, A.Y., Krogmann, L., Kubiak, M., Lanfear, R., Letsch, H., Li, Y., Li, Z., Li, J., Lu, H., Machida, R., Mashimo, Y., Kapli, P., McKenna, D.D., Meng, G., Nakagaki, Y., Navarrete-Heredia, J.L., Ott, M., Ou, Y., Pass, G., Podsiadlowski, L., Pohl, H., von Reumont, B.M., Schutte, K., Sekiya, K., Shimizu, S., Slipinski, A., Stamatakis, A., Song, W., Su, X., Szucsich, N.U., Tan, M., Tan, X., Tang, M., Tang, J., Timelthaler, G., Tomizuka, S., Trautwein, M., Tong, X., Uchifune, T., Walzl, M.G., Wiegmann, B.M., Wilbrandt, J., Wipfler, B., Wong, T.K., Wu, Q., Wu, G., Xie, Y., Yang, S., Yang, Q., Yeates, D.K., Yoshizawa, K., Zhang, Q., Zhang, R., Zhang, W., Zhang, Y., Zhao, J., Zhou, C., Zhou, L., Ziesmann, T., Zou, S., Li, Y., Xu, X., Zhang, Y., Yang, H., Wang, J., Wang, J., Kjer, K.M., Zhou, X., 2014a. Phylogenomics resolves the timing and pattern of insect evolution. Science 346, 763–767.
Misof, B., Meyer, B., von Reumont, B.M., Kück, P., Misof, K., Meusemann, K., 2013. Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. BMC Bioinform. 14, 1–13.
Misof, B., Meusemann, K., Reumont, B.M.v., Kück, P., Prohaska, S.J., Stadler, P.F., 2014b. A priori assessment of data quality in molecular phylogenetics. Algorithms Mol. Biol. 9, 1–8.
Muraji, M., Tachikawa, S., 2000. Phylogenetic analysis of water striders (Hemiptera: Gerroidea) based on partial sequences of mitochondrial and nuclear ribosomal RNA genes. Entomol. Sci. 3, 615–626.
Murienne, J., 2009. Molecular data confirm family status for the *Tryonicus-Lauraesilpha* group (Insecta: Blattodea: Tryonicidae). Organ. Divers. Evol. 9, 44–51.
Narechania, A., Baker, R.H., Sit, R., Kolokotronis, S.O., DeSalle, R., Planet, P.J., 2012. Random Addition Concatenation Analysis: a novel approach to the exploration of phylogenomic signal reveals strong agreement between core and shell genomic partitions in the cyanobacteria. Genome Biol. Evol. 4, 30–43.
Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.
Omilian, A., Taylor, D., 2001. Rate acceleration and long-branch attraction in a conserved gene of cryptic daphniid (Crustacea) species. Mol. Biol. Evol. 18, 2201–2212.
Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R.P., Moret, B.M.E., Stamatakis, A., 2009. How many bootstrap replicates are necessary? In: Batzoglou, S. (Ed.), LNCS, pp. 184–200.
Roth, L.M., 1996. The cockroach genera *Anaplecta*, *Anaplectella*, *Anaplectoidea*, and *Malaccina* (Blattaria, Blattellidae: Anaplectinae and Blattellinae). Oriental Insects 30, 301–372.
Salichos, L., Rokas, A., 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature 497, 327–331.
Schliep, K.P., 2011. phangorn: phylogenetic analysis in R. Bioinformatics 27, 592–593.
Schulte, K., Barfuss, M.H., Zizka, G., 2009. Phylogeny of Bromelioideae (Bromeliaceae) inferred from nuclear and plastid DNA loci reveals the evolution of the tank habit within the subfamily. Mol. Phylogenet. Evol. 51, 327–339.
Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.
Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.
Sneath, P.H.A., Sokal, R.R., 1973. Numerical Taxonomy: The Principles and Practice of Numerical Classification. Freeman, San Fransisco, USA, pp. 278.
Soltis, P.S., Soltis, D.E., 2003. Applying the bootstrap in phylogeny reconstruction. Stat. Sci. 18, 256–267.

Song, H., Amedegnatoc, C., Ciglianod, M.M., Desutter-Grandcolas, L., Heads, S.W., Huang, Y., Otte, D., Whiting, M.F., 2015. 300 million years of diversification: elucidating the patterns of Orthopteran evolution based on comprehensive taxon and gene sampling. Cladistics 31, 621–651.
Song, H., Sheffield, N.C., Cameron, S.L., Miller, K.B., Whiting, M.F., 2010. When phylogenetic assumptions are violated: base compositional heterogeneity and among-site rate variation in beetle mitochondrial phylogenomics. Syst. Entomol. 35, 429–448.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, Massachusetts.
Tang, C.Q., Humphreys, A.M., Fontaneto, D., Barraclough, T.G., Paradis, E., 2014. Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-locus data. Methods Ecol. Evol. 5, 1086–1094.
Thorne, B.L., Carpenter, J.M., 1992. Phylogeny of Dictyoptera. Syst. Entomol. 17, 253–268.
Trautwein, M.D., Wiegmann, B.M., Beutel, R., Kjer, K.M., Yeates, D.K., 2012. Advances in insect phylogeny at the dawn of the postgenomic era. Annu. Rev. Entomol. 57, 449–468.
Vogler, A., Cardoso, A., Barraclough, T., 2005. Exploring rate variation among and within sites in a densely sampled tree: species level phylogenetics of north american tiger beetles (genus cicindela). Syst. Biol. 54, 4–20.
Wägele, J.W., Letsch, H., Klussmann-Kolb, A., Mayer, C., Misof, B., Wägele, H., 2009. Phylogenetic support values are not necessarily informative: the case of the Serialia hypothesis (a mollusk phylogeny). Front. Zool. 6, 1–15.
Wägele, J.W., Mayer, C., 2007. Visualizing differences in phylogenetic information content of alignments and distinction of three classes of long-branch effects. BMC Evol. Biol. 7, 1–24.
Wang, Z., Shi, Y., Qiu, Z., Che, Y., Lo, N., 2017. Reconstructing the phylogeny of Blattodea: robust support for interfamilial relationships and major clades. Sci. Rep. 7, 1–8.
Ware, J.L., Litman, J., Klass, K.-D., Spearman, L.A., 2008. Relationships among the major lineages of Dictyoptera: the effect of outgroup selection on dictyopteran tree topology. Syst. Entomol. 33, 429–450.
Wenzel, J.W., Siddall, M.E., 1999. Noise. Cladistics 15, 51–64.
Whitfield, J.B., Lockhart, P.J., 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22, 258–265.
Wiens, J.J., Kuczynski, C.A., Stephens, P.R., 2010. Discordant mitochondrial and nuclear gene phylogenies in emydid turtles: implications for speciation and conservation. Biol. J. Linnean Soc. 99, 445–461.
Wolfram Research, I., 2012. Mathematica Edition: Version 9.1. Wolfram Research Inc., Champaign, Illinois, United States.
Wu, M., Chatterji, S., Eisen, J.A., 2012. Accounting for alignment uncertainty in phylogenomics. PLoS One 7, e30288.
Zhang, L., Wu, W., Yan, H.F., Ge, X.J., 2015. Phylotranscriptomic analysis based on coalescence was less influenced by the evolving rates and the number of genes: a case study in ericales. Evol. Bioinform. Online 11, 81–91.
Zwickl, D.J., Hillis, D.M., 2002. Increased taxon sampling greatly reduces phylogenetic error. Syst. Biol. 51, 588–598.