



Extended taxon sampling reveals conflicting mitochondrial phylogenomics in the holarctic *Reticulitermes* termites (Blattodea: Termitoidae)

Giobbe Forni^{*,}, Simona Corneti, and Andrea Luchetti^{}

Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy

*Corresponding author. Department of Biological, Geological and Environmental Sciences, University of Bologna, via Selmi 3, 40126 Bologna, Italy
(Email: giobbe.forni2@unibo.it).

Subject Editor: Kim Barão

Reticulitermes is a genus of subterranean termites that experienced a rapid radiation during the late Miocene, resulting in a widespread distribution across the Holarctic region. Despite decades of phylogenetic research that has delineated some well-supported lineages, relationships among them remain unresolved. Previous studies based on a handful of mitochondrial or nuclear markers have produced conflicting topologies and failed to robustly resolve deep nodes within the genus, particularly among the East European–West Asian (EE + WA) and Western Europe lineages. More recently, despite improved resolution achieved through mitochondrial phylogenomics, a major limitation in *Reticulitermes* studies has been the complete absence of mitochondrial genomes for the EE+WA lineage. To address this, we sequenced 6 new mitochondrial genomes from 5 *Reticulitermes* taxa, including the first representatives of the EE + WA clade, and analyzed them alongside previously published data. Phylogenies were inferred across a range of analytical conditions. All analyses supported the monophyly of *Reticulitermes*, though the placement of EE + WA varied depending on the data and model. Analysis of the nucleotide dataset recovered EE + WA as sister to all other clades. In contrast, the amino acid dataset found it nested within the genus, placing East Asia as the earliest diverging lineage. Our results underscore the sensitivity of phylogenetic inference to dataset and model choice in the context of *Reticulitermes* rapid radiation and provide new insights for resolving the evolutionary history of this genus. The inclusion of EE + WA fills a key geographic gap and suggests a complex history of diversification, consistent with either an ancient refugial split in southeastern Europe–Western Asia or an extraordinarily rapid westward expansion.

Keywords: termite, biogeography, Isoptera, systematics, phylomitogenomics

Introduction

The genus *Reticulitermes* comprises a diverse group of subterranean termites distributed across the Holarctic region and is well known for its taxonomic complexity and history of conflicting phylogenetic hypotheses. Following Dedeine et al. (2016), the most taxonomically complete analysis so far, *Reticulitermes* species can be subdivided into four clades: East European–West Asian (EE+WA), West European (WE), Eastern Asia (EA), and North American (NA). Despite multiple efforts to reconstruct their evolutionary history using mitochondrial and nuclear markers, the relationships among major *Reticulitermes* lineages remain unresolved, particularly those involving the EE+WA and WE clades. Previous studies have yielded topologies that vary significantly (Luchetti et al. 2004, Uva et al. 2004, Dedeine et al. 2016, Forni et al. 2019; Fig. 1), likely because all these papers leveraged different taxa, markers, and analytical frameworks. Some suggest the paraphyly of the NA clade, and others propose alternative placements for the EE+WA clade and alternative relationships for the WE clade. These differences in phylogenetic inferences could be due to factors of different nature, such as undetected natural selection, differential phylogenetic informativity of selected loci, biases linked to tree reconstruction methods, incomplete or

insufficient taxon sampling, evolutionary model violations or misspecification, or an incorrect rooting strategy (Som 2015, Steenwyk et al. 2023).

The adoption of whole mitogenomes for insect systematics (Bourguignon et al. 2018, Righetti et al. 2024), also including termites (Cameron et al. 2012, Bourguignon et al. 2016a, 2016b, 2017, Forni et al. 2019, Wang et al. 2019, Buček et al. 2022, Forni et al. 2025), has provided more powerful tools to revisit these questions. Concerning *Reticulitermes*, however, the EE+WA clade has remained entirely unsampled at the mitogenome level, limiting a comprehensive reconstruction.

In this study, we sequence six new mitogenomes from five *Reticulitermes* taxa—including previously unsampled lineages from southern Europe and the Aegean—and re-evaluate the genus's phylogeny using a varied range of datasets, partitioning strategies, and site-heterogeneous models. Moreover, we estimated divergence times to check whether the addition of mitochondrial genomes of previously unsampled species would alter the already known age estimates (Bourguignon et al. 2016a, Dedeine et al. 2016). We aim to assess how methodological choices affect phylogenetic outcomes and clarify the position of the EE+WA clade within the genus. Resolving this node is critical for reconstructing the Holarctic dispersal history of *Reticulitermes*.

Received 7 July 2025. Revised: 28 October 2025. Accepted: 24 November 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of Entomological Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

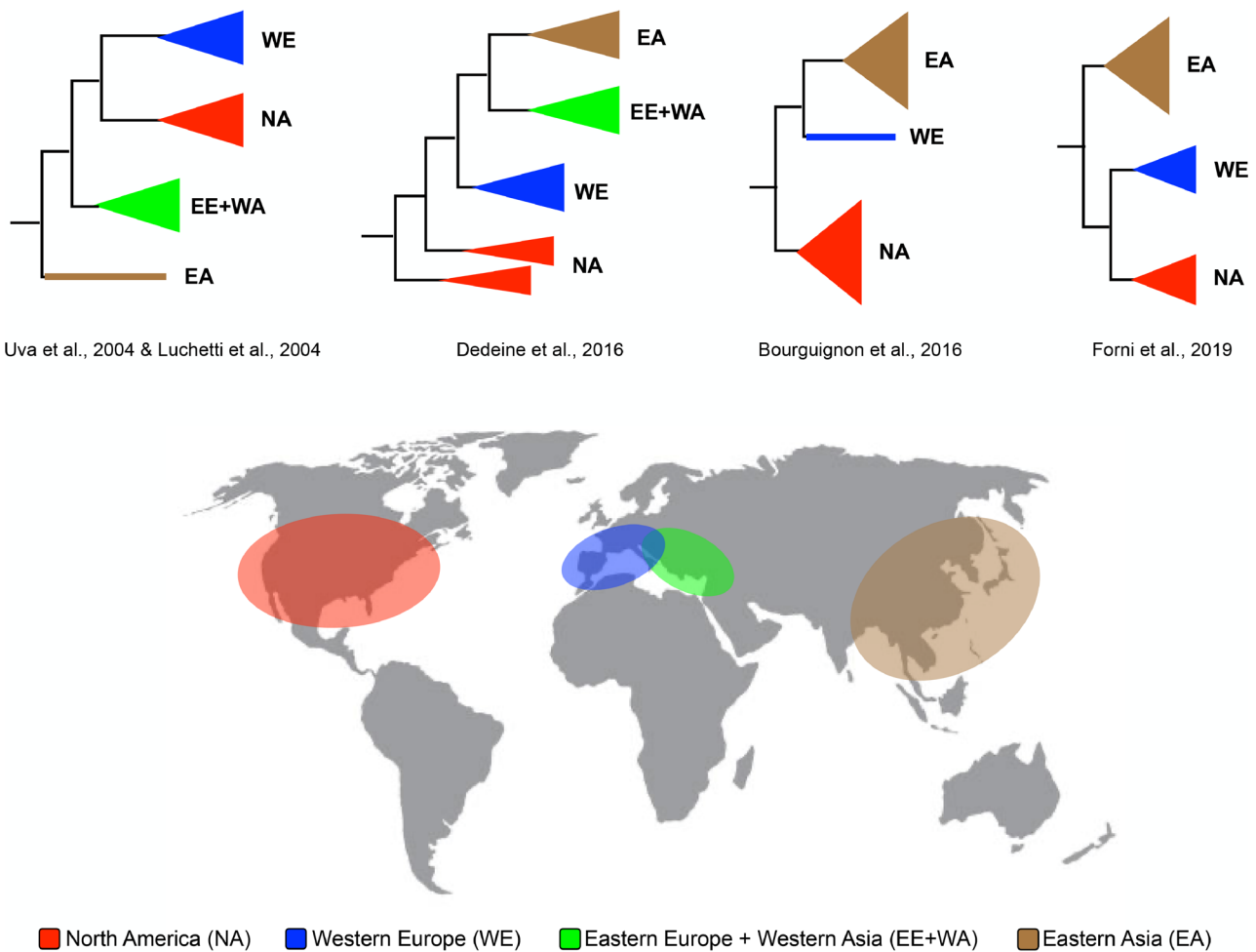


Fig. 1. Topological inconsistency of published phylogenetic hypotheses for *Reticulitermes* termites. Western Europe (WE, blue), North America (NA, red), Eastern Europe + Western Asia (EE+WA, green), and Eastern Asia (EA, brown). Topologies are redrawn from Uva et al. (2004), Luchetti et al. (2004), Dedeine et al. (2016), Bourguignon et al. (2016a), and Forni et al. (2019).

Materials and Methods

Genomic DNA was extracted from field-collected specimens of *Reticulitermes urbis* (Mercadante, Italy), *Reticulitermes lucifugus siculus* (Giarre, Sicily, Italy), *Reticulitermes lucifugus corsicus* (Siniscola, Sardinia, Italy), *Reticulitermes balkanensis* (Nea Artaki and Mount Penteli, Greece), and *Reticulitermes aegeus* (Cyprus) (Supplementary Table S1). To minimize contamination from gut symbionts, DNA was isolated exclusively from the head tissue using the E.Z.N.A. Mollusc & Insect DNA Kit (Omega Bio-Tek). Purified DNA samples were then submitted to Macrogen Europe (Amsterdam, Netherlands) for high-throughput sequencing using the Illumina NovaSeq X platform, generating 150 bp paired-end reads.

Raw reads were quality-checked, cleaned, and trimmed using Trimmomatic v0.36 (Bolger et al. 2014) with default parameters. Genome assemblies were performed using SPAdes v4.0 (Prjibelski et al. 2020), also with default settings. Mitochondrial gene annotation was conducted using MITOS2 (Donath et al. 2019), with tRNA genes further refined using ARWEN (Laslett and Canbäck 2008). Mitogenomes were deposited on NCBI GeneBank under the accessions PV787246 to PV787251

(Supplementary Table S1). Additional mitochondrial genomes from *Reticulitermes* species and from other termites available in NCBI GenBank (Supplementary Table S2) were incorporated into the phylogenetic analyses.

The maximum likelihood phylogenetic inference was performed using IQ-TREE v. 2.3 (Minh et al. 2020), with branch support estimated from 1,000 ultrafast bootstrap replicates (Hoang et al. 2018). For the nucleotide dataset, analyses were conducted both with and without third codon positions in protein-coding genes to assess the potential impact of substitutions saturation. Model selection and partitioning schemes were optimized using ModelFinder (Kalyaanamoorthy et al. 2017), implemented in IQ-TREE.

The Bayesian inference was obtained using MrBayes v. 3.2 (Ronquist et al. 2012), running two independent *mcmc* searches for up to 10,000,000 steps. Runs' convergence was assessed when the average standard deviation of split frequencies was <0.01 and the potential scale reduction factor = 1.0; goodness of parameter mixing was obtained when the estimated sample sizes (ESSs) became >200. The model selection and partitioning schemes optimization were carried out using PartitionFinder v. 1.1 (Lanfear et al. 2012). Since the maximum likelihood analysis showed that the effect of including/excluding the third

codon position was negligible, the Bayesian inference was performed only on the full dataset.

Several partitioning strategies were explored. Initial partitions were defined by genes (for both nucleotide and amino acid datasets) or codon positions (nucleotide dataset only). For each initial partitioning scheme, we applied (i) a fixed substitution model (GTR for nucleotides, LG for amino acids) for each initial partition; (ii) a fixed partition scheme, with best-fitting models selected independently for each initial partition (MF option in IQ-TREE; *scheme search* = *user* in PartitionFinder); and (iii) an algorithmic partition scheme selection, which allows initial partition merging based on model fit (MFP + MERGE option in IQ-TREE; *scheme search* = *greedy* in PartitionFinder).

Additionally, we performed further maximum likelihood analyses using site-heterogeneous mixture models: for the nucleotide dataset, this was done using the MFP + MIX option; for amino acids, model selection was conducted among the C10-C60, LG4M, LG4X, and CF4 model series. Bayesian phylogenetic inference under site-heterogeneous mixture models was performed with PhyloBayes v. 3.3 (Lartillot et al. 2013). Nucleotide alignments were analyzed with GTR + CAT, and amino acid alignments with LG + CAT (Lartillot and Philippe 2004). Two chains of 15,000 cycles were run for each data type. Convergence (maxdiff < 0.1) and mixing (ESSs > 200) among each chain pair were assessed with the commands *tracecomp* and *bpcomp*.

A time estimate for the origin of the *Reticulitermes* clade was obtained using the LSD2 algorithm (To et al. 2016), as implemented in IQ-TREE, applied to both the complete nucleotide and amino acid datasets. Fossil calibrations were taken from Bourguignon et al. (2016a, 2016b) and are detailed in Supplementary Table S3. Confidence intervals for node ages were estimated by resampling branch lengths 1,000 times.

Topology tests were conducted in IQ-TREE to compare alternative phylogenetic hypotheses derived from analyses using mixture models. Log-likelihood scores were calculated for both nucleotide and amino acid datasets. The following tests were performed using 10,000 RELL resampling: the bootstrap proportion using the RELL method (Kishino et al. 1990), the Kishino–Hasegawa test (Kishino and Hasegawa 1989), the Shimodaira–Hasegawa test (Shimodaira and Hasegawa 1999), the Expected Likelihood Weight test (Strimmer and Rambaut 2002), and the approximately unbiased test (Shimodaira 2002).

Results

We successfully assembled six new mitochondrial genomes: one each for *R. lucifugus corsicus*, *R. lucifugus siculus*, *R. urbis*, and *R. aegeus*, and two for *R. balkanensis*. The assembled contigs ranged in length from 15,227 to 15,820 bp, and the genes' content and order were consistent with previously published mitogenomes of subterranean termites. The nucleotide composition does not differ from that of other *Reticulitermes* mitogenomes, both in terms of nucleotide frequencies and AT- or GC-skew (Supplementary Table S4). The 13 protein-coding genes and the two ribosomal RNA genes were aligned with homologous sequences from other Termitidae and Rhinotermitidae species (Supplementary Table S2), producing a final alignment comprising 53 taxa and 13,790 bp.

We performed multiple maximum likelihood and Bayesian inference analyses on both nucleotide (with and without third codon positions) and amino acid datasets to assess the effect of partitioning schemes and substitution models on the tree

topology. Across all analyses, the Termitidae clade was consistently recovered as sister to a monophyletic group comprising the *Coptotermes* + *Heterotermes* clade (with *Heterotermes* rendered paraphyletic) and a monophyletic cluster of all *Reticulitermes* mitogenomes. These relationships were strongly supported, with bootstrap values ranging from 97% to 100% and posterior probabilities of 1.0 (Supplementary Figs S1–S28). All *Reticulitermes* geographic groups from Dedeine et al. (2016) were strongly supported.

Concerning the nucleotide sequences (Fig. 2A; Supplementary Figs S1–S12), the use of a simple GTR model resulted in a topology where the NA clade is paraphyletic and the EE+WA clade is sister to EA, both under by-gene partitioning (either including or excluding third codon positions) and under by-codon partitioning (only when excluding third codon positions). When best-fitting substitution models were implemented, the EA clade was recovered as sister to a clade in which EE+WA is sister to NA + WE only under the by-gene partitioning scheme, including third codon positions. A consistent topology was recovered in all other cases, where EE+WA is sister to all remaining groups, and EA is sister to NA + WE (Fig. 2A; Supplementary Figs S1–S12). Across analyses, the support for the deepest nodes within *Reticulitermes* was often weak. The maximum likelihood tree with the highest overall support was obtained using the by-codon algorithmic partitioning scheme, including third codon positions, where EE+WA was sister to all others, and WA was sister to NA + WE with 71% bootstrap support. The same variation in tree topology and support can be observed for the Bayesian inference on nucleotides, with the only difference of the GTR-based topology, which reported maximum support at all nodes (Fig. 3A; Supplementary Figs S18–S23). The topology obtained using mixture models, both maximum likelihood and Bayesian inference, fully overlaps the one obtained with the algorithmic partitioning scheme, showing only marginally higher average nodal support (Fig. 4A; Supplementary Figs S16 and S27).

For the amino acid dataset (Fig. 2B; Supplementary Figs S13–S15), the use of the simple LG model again recovered a topology with NA paraphyletic and EE+WA sister to EA. Analyses using best-fitting models, with or without the use of algorithmic partition schemes, consistently returned EA as sister to all other clades, with EE+WA sister to NA + WE. However, support for these deeper nodes remained low, and both the monophyly of NA and its relationship with WE received limited support (Fig. 2B; Supplementary Figs S13–S15). The Bayesian inference fully overlaps the maximum likelihood one, and the algorithmic partitioning scheme resulted in all nodes being fully supported (Fig. 3B; Supplementary Figs S24–S26). The application of mixture models, both maximum likelihood and Bayesian inference, increased nodal support for the NA and EA clades, as well as for the NA + WE grouping, but slightly reduced support for the sister relationship between EE+WA and NA + WE (Fig. 4B; Supplementary Figs S17 and S28).

Topology tests comparing the two main topologies shown in Fig. 4 found no significant differences for the nucleotide dataset (Table 1). However, when using the amino acid dataset, the topology shown in Fig. 4A was significantly rejected (Table 1).

The age of the *Reticulitermes* crown group was estimated to lie between 10.3 million years ago (Mya; 95% CI=8.7 to 12.1 Mya) and 11.1 Mya (95% CI=8.8 to 13.4 Mya) using nucleotides and amino acids, respectively (Supplementary Figs S29 and S30).

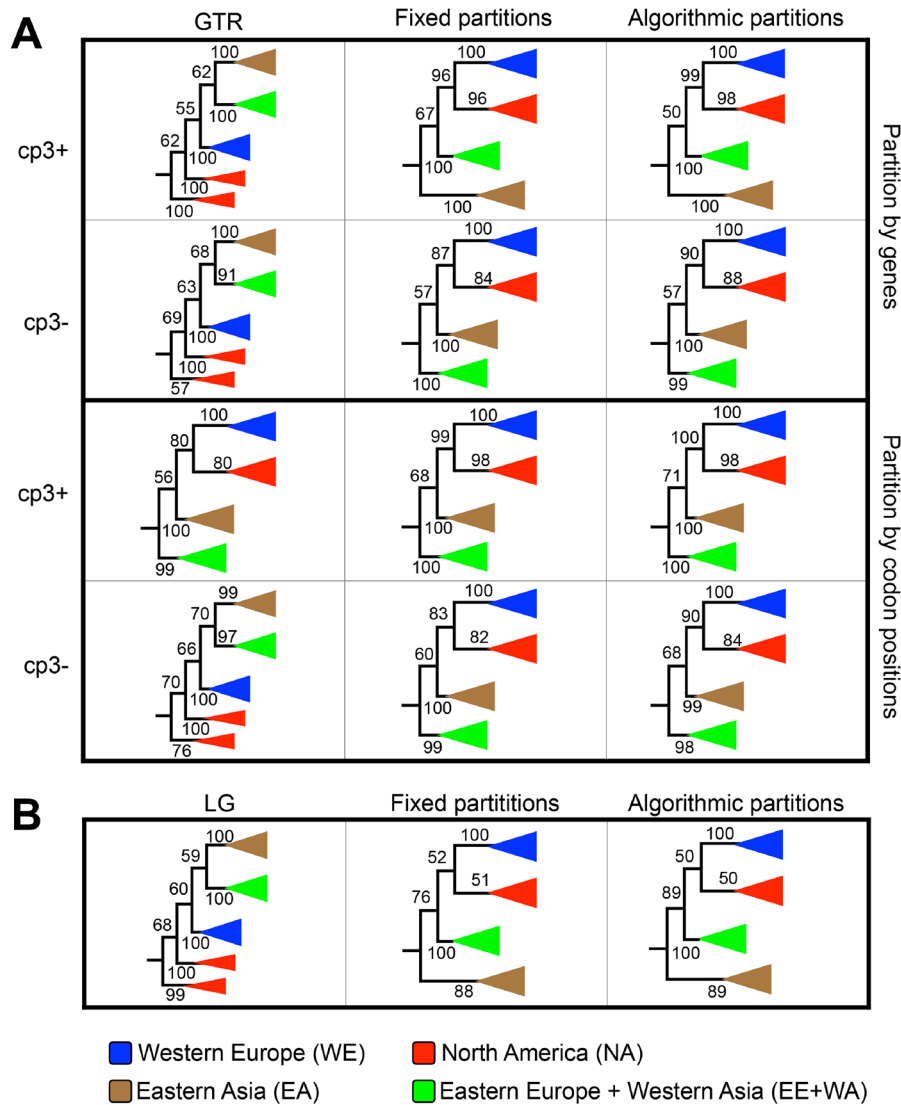


Fig. 2. Maximum likelihood topological variation in *Reticulitermes* phylogenies depending on dataset type, partitioning scheme, and model selection. A) Maximum likelihood trees inferred from nucleotide datasets with either inclusion (cp3+) or exclusion (cp3-) of third codon positions. Analyses were run using 3 strategies: GTR (a single General Time Reversible model), Fixed Partitions (best-fitting models per each partition defined by the user), and Algorithmic partitions (model selection with automatic merging of initial partitions). Trees were inferred under 2 partitioning schemes: by gene (top half) and by codon position within genes (bottom half). B) Maximum likelihood trees based on amino acid translations, using LG, Fixed and Algorithmic partitions model strategies. Clades are color-coded as follows: Western Europe (WE, blue), North America (NA, red), Eastern Europe + Western Asia (EE+WA, green), and Eastern Asia (EA, brown). Values at nodes represent bootstrap supports. Outgroups omitted for graphical purposes.

Discussion

In this study, we report on the complete mitochondrial genomes of 5 *Reticulitermes* taxa belonging to the West European (WE) and EE+WA clades as defined by Dedeine et al. (2016), the latter being the only major clade of the genus lacking mitogenomic data. All mitogenomes share the canonical gene order, content, and overall architecture characteristic of hexapods (Cameron 2014). Moreover, all *Reticulitermes* mitogenomes share the same nucleotide content. The phylogenetic analyses based on all available *Reticulitermes* mitogenomes, and representatives of other Heretotermittinae (*Coptotermes* and *Heterotermes*) and Termitidae, recovered *Reticulitermes* as a well-supported, monophyletic group. Though relationships among major clades within this group remained unstable, mainly depending on the model selection and the partitioning

scheme adopted, as well as the dataset considered (ie nucleotides vs amino acids).

The simplest models we used—those without selection of best-fitting models and non-algorithmic partitioning—yielded a topology in which the NA lineage is paraphyletic to the remaining clades (Figs 2 and 3). This topology is similar to the one presented in the most taxonomically comprehensive study so far (Dedeine et al. 2016); their inference included 29 taxa and three mitochondrial (COI, COII, 16S) markers (the two nuclear markers were found not to be resolutive), leveraging only by-gene partitions without distinguishing among codon positions. Though, when analyzing our mitogenomic dataset with the same methods as Dedeine et al. (2016), i.e. non-algorithmic, by-gene partitions, and best-fitting substitution models, we obtained a different topology, which resulted in overlapping with those obtained by either algorithmic

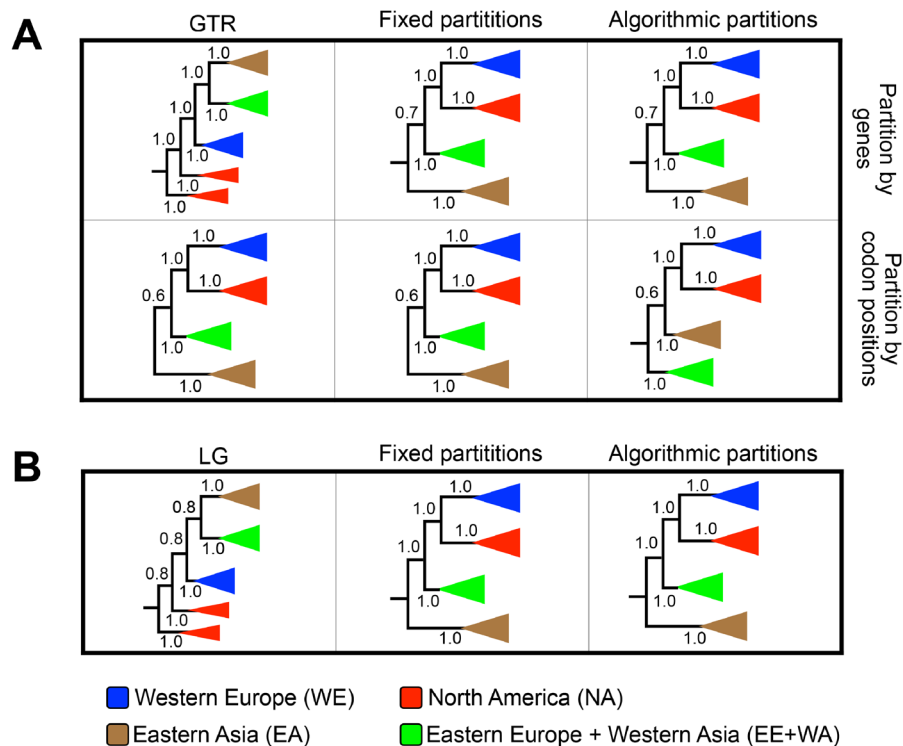


Fig. 3. Bayesian inference topological variation in *Reticulitermes* phylogenies depending on dataset type, partitioning scheme, and model selection. A) Bayesian inference trees inferred from nucleotide datasets. Analyses were run using 3 strategies: GTR (a single General Time Reversible model), Fixed Partitions (best-fitting models per each partition defined by the user), and Algorithmic partitions (model selection with automatic merging of initial partitions). Trees were inferred under 2 partitioning schemes: by gene (top half) and by codon position within genes (bottom half). B) Bayesian inference trees based on amino acid translations, using LG, Fixed and Algorithmic partitions model strategies. Clades are color-coded as follows: Western Europe (WE, blue), North America (NA, red), Eastern Europe + Western Asia (EE+WA, green), and Eastern Asia (EA, brown). Values at nodes represent posterior probability supports. Outgroups omitted for graphical purposes.

partitioning or with amino acids. This may suggest that the observed lack of congruence may rely on the different phylogenetic informativity of the three mitochondrial markers compared to that of all 15 mitochondrial genes. Interestingly, when we excluded the third codon position from this analysis, we got a tree topology that is identical to the one obtained when considering the by-codon partitions (either with algorithmic partitioning or not) or when mixture models are implemented. Although these results do not suggest that third codon positions underwent substitutions saturation (their inclusion/exclusion in this analysis does not significantly affect the inference), they clearly point out their significantly different substitution rate and thus, as it is generally acknowledged, they need to be considered as a separate partition. It is worth noting that analyses based on either algorithmic or non-algorithmic partitioning scheme selection do not give different tree inferences; on the contrary, the implementation of non-best-fitting models produced major incongruities. Our results are in agreement with a survey on the proper use of partitioning scheme selection, which evidenced marginal or no bias at all in topology, branch lengths, and bootstrap supports when using non-algorithmic partitioning with respect to the algorithmic partitioning scheme selection (Kainer and Lanfear 2015).

Early mitochondrial phylogenetic reconstructions, based on a few markers like 16S + ND1 (Uva et al. 2004) or COII + 16S (Luchetti et al. 2004), yielded topologies congruent with our analyses based on amino acids, while more recent mitochondrial phylogenomics showed inconsistencies. However, these latter are likely hampered by incomplete taxon sampling, one

of the most important analytical errors in phylogenetics (Zwickl and Hillis 2002, Hillis et al. 2003, Bernot et al. 2023). The analysis shown by Bourguignon et al. (2016a) did not recover WE as a sister group to NA, though it only included one WE taxon (the Iberian species *R. grassei*). The analysis of Forni et al. (2019), on the other hand, produced topologies compatible with both of ours. Yet, both analyses entirely lacked representatives of the EE+WA clade. As a matter of fact, most of the lack of consensus across trees, both from literature and within our analyses, is due to the placement of this clade: indeed, it consistently emerges as sister to all other *Reticulitermes*, using nucleotides, or as sister to the NA and WE clade, using amino acids or a less thorough initial partitioning scheme (by-gene and including the third codon position), even though with weaker bootstrap supports. The tree topology tests, when assessing the nucleotide dataset, do not significantly reject any of the two obtained topologies represented in Fig. 4; though, this does not hold when assessing topologies using the amino acid data. Of course, this could imply a different phylogenetic informativity of the two datasets, which is known to vary for both the rate of evolution and the possibility of evolutionary convergence (Simmons et al. 2002, Townsend et al. 2008, Philippe et al. 2011), but it points to a general incongruent phylogenetic inference. A possible explanation for this incongruence could be found in the relatively short time of diversification of the *Reticulitermes* crown group, which has been inferred to be approximately between 18 and 16 Mya (Bourguignon et al. 2016a, Dedeine et al. 2016) and presently estimated to be even more recently at around 11 Mya. Considering

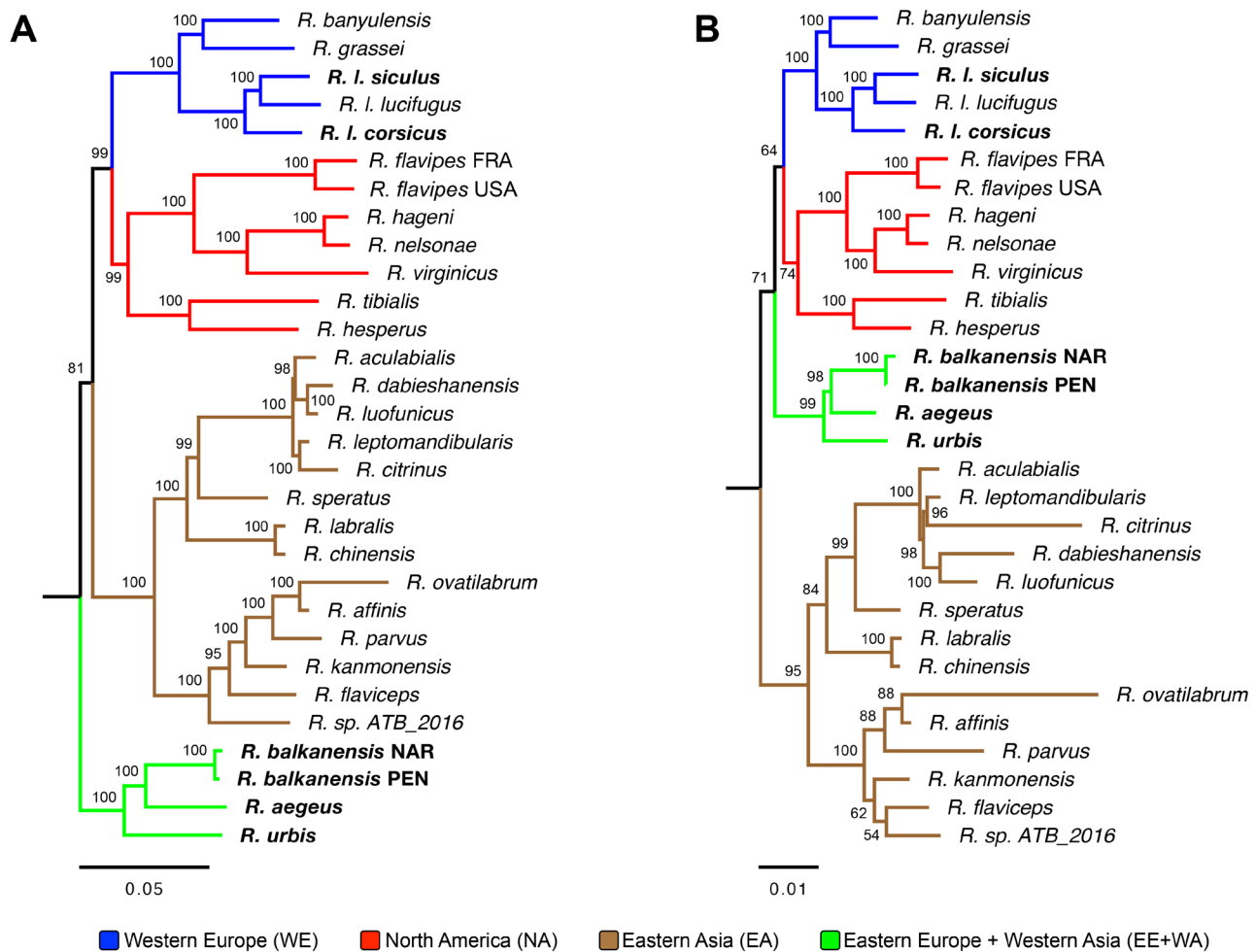


Fig. 4. Phylogenetic inferences using site-heterogeneous mixture models in a maximum likelihood framework. A) Tree inferred using the nucleotide dataset, with codon position partitioning including third positions (cp3+). B) Tree inferred using the amino acid dataset. Color coding of clades corresponds to biogeographic groups: Western Europe (WE, blue), North America (NA, red), Eastern Europe + Western Asia (EE+WA, green), and Eastern Asia (EA, brown). Parametric bootstrap values of 1,000 replicates are shown at nodes. Both analyses recovered *Reticulitermes* as monophyletic but differed in the placement of the EE+WA clade. Although topological differences between trees were not statistically significant for nucleotide data (Table 1), the amino acid dataset significantly rejected the topology shown in panel A. Outgroups omitted for graphical purposes. Phylogenetic inferences using site-heterogeneous mixture models in a Bayesian inference framework are available in [Supplementary Fig. S27](#) for the nucleotide dataset and [Supplementary Fig. S28](#) for the amino acid dataset.

Table 1. Topology tests comparing alternative phylogenetic hypotheses for the placement of the EE + WA clade within *Reticulitermes*, based on nucleotide and amino acid datasets

Dataset	Tree topology	−lnL	ΔlnL	bp-RELL	KH	SH	ELW	AU
Nucleotides	((EE+WA), (EA, (WE, NA)))	−157284.6606	–	0.602	0.628	1	0.602	0.615
	(EA, (EE+WA, (WE, NA)))	−157288.8329	4.1723	0.398	0.372	0.372	0.398	0.385
Amino acids	((EE+WA), (EA, (WE, NA)))	−44034.96666	23.738	0.032*	0.041*	0.041*	0.033*	0.028*
	(EA, (EE+WA, (WE, NA)))	−44011.22869	–	0.968	0.959	1	0.959	0.972

The two alternative tree topologies tested for each dataset are: ((EE+WA), (EA, (WE, NA)))—corresponding to the topology in [Fig. 4A](#), and (EA, (EE+WA, (WE, NA)))—corresponding to the topology in [Fig. 4B](#). Log-likelihood scores (−lnL), log-likelihood differences (ΔlnL), and significance values from several topology tests are reported: bootstrap proportion using REll (bp-RELL), Kishino–Hasegawa (KH), Shimodaira–Hasegawa (SH), expected likelihood weight (ELW), and approximately unbiased (AU) tests. Asterisks (*) indicate statistically significant rejection of a topology ($P < 0.05$). While none of the tests showed significant differences between topologies for the nucleotide dataset, all tests significantly rejected the topology shown in [Fig. 4A](#) for the amino acid dataset.

the high diversity and the widespread distribution of these termites, these data point to a fast diversification of *Reticulitermes* lineages. In our phylogenetic analysis, in fact, it is possible to find the footprint of a sudden radiation in the short, weakly supported branches at deep relationships: the short coalescence time at the origin of extant *Reticulitermes* may be responsible

for the lack of resolution in this part of inferred trees (Rokas et al. 2005, Whitfield and Lockhart 2007).

It is interesting to note that, at variance of maximum likelihood analyses, the Bayesian inference returned fully supported topologies (i.e., with nodes with posterior probabilities close to or equal to 1.0), showing the EA clade as sister to all other

Reticulitermes clades: on nucleotides, when using GTR model on all gene partitions, and on amino acids when using both fixed and algorithmic partitioning schemes (Fig. 3). The tendency of the Bayesian support to be biased by priors and search parameters is long known and well analyzed (Lemmon and Moriarty 2004, Simmons et al. 2004, Lewis et al. 2005, Ekman and Blaalid 2011). In particular, empirical and simulated analyses evidenced an overestimation of Bayesian support for artifactual bipartitions in the case of underparametrization (i.e., too simplistic models) or in the case of hard polytomies, supporting topologies minimally sustained by data or defined by very short branches (Lewis et al. 2005). This clearly suggests a bias in our Bayesian inference due to a too simplistic model, such as the choice of GTR for each gene partition, or for the very short branch lengths observed at bipartitions next to the *Reticulitermes* root.

Determining the position of the EE+WA clade is essential for reconstructing *Reticulitermes* evolutionary history and biogeography. Late Eocene fossils show that early *Reticulitermes* were already widespread across Laurasia well before the complete separation of the North American and Eurasian plates, before 50 Mya (Engel et al. 2009). In the past and present analyses, the origin of extant *Reticulitermes* lineages was, however, estimated to have occurred during the late Miocene, later than the minimum age of the oldest fossils. This implies that vicariance can hardly explain *Reticulitermes* extant distribution and also that much of the earlier diversity within the genus was erased by some major paleoclimatic change, such as the Oligocene cooling (a sudden and deep temperature decline that occurred 34 Mya; Eldrett et al. 2009). Against this backdrop, two biogeographic scenarios can be proposed to explain the evolutionary history of *Reticulitermes*. Analyses based on nucleotide sequences generally identify the EE+WA clade as the earliest divergence within the genus. This scenario suggests that an ancestral *Reticulitermes* lineage found refuge in southeastern Europe and western Asia, potentially serving as a biogeographic reservoir that facilitated subsequent recolonization events into other parts of Europe, Asia, and possibly North America (Dedeine et al. 2016). In contrast, amino acid-based analyses typically indicate that the EA lineage diverged first, aligning with hypotheses proposing a mid-Miocene (~15 to 14 Ma) dispersal event possibly originating from eastern Asia. Under this scenario, dispersal would have occurred subsequently, colonizing Europe and later, through either the Bering land bridge or transient North Atlantic land connections, reaching North America (Sanmartín et al. 2001, Engel et al. 2009). However, the weak support at the deepest nodes prevents ruling out either scenario.

Including EE+WA mitogenomes into the phylogenetic dataset closes a major taxonomic and geographic gap but highlights the instability of the *Reticulitermes* tree inference. It is likely that only dense nuclear-genome datasets—encompassing thousands of independent markers—will carry sufficient resolution to untangle the spatial and temporal dynamics of this rapid termite radiation.

Author Contributions

Giobbe Forni (Formal analysis [equal], Writing—review & editing [equal]), Simona Corneti (Data curation [equal], Formal analysis [equal], Writing—review & editing [equal]), and

Andrea Luchetti (Conceptualization [equal], Formal analysis [equal], Supervision [equal], Writing—original draft [equal])

Supplementary Material

Supplementary material is available at *Journal of Insect Science* online.

Funding

This work has been supported by Canziani funding to A. Luchetti.

Conflicts of interest

None declared.

Data availability

Raw sequencing reads have been deposited on NCBI SRA under the accession numbers: SRR28393303-SRR28393304, SRR32170788-SRR32170791

References

- Bernot JP, Owen CL, Wolfe JM, et al. 2023. Major revisions in pancrustacean phylogeny and evidence of sensitivity to taxon sampling. *Mol. Biol. Evol.* 40:msad175. <https://doi.org/10.1093/molbev/msad175>
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30:2114–2120.
- Bourguignon T, Lo N, Šobotník J, et al. 2016a. Oceanic dispersal, vicariance and human introduction shaped the modern distribution of the termites *Reticulitermes*, *Heterotermes* and *Coptotermes*. *Proc. Roy. Soc.* 283:20160179.
- Bourguignon T, Lo N, Cameron SL, et al. 2016b. The evolutionary history of termites as inferred from 66 mitochondrial genomes. *Mol. Biol. Evol.* 32:406–421.
- Bourguignon T, Lo N, Šobotník J, et al. 2017. Mitochondrial phylogenomics resolves the global spread of higher termites, ecosystem engineers of the tropics. *Mol. Biol. Evol.* 34:589–597.
- Bourguignon T, Tang Q, Ho SY, et al. 2018. Transoceanic dispersal and plate tectonics shaped global cockroach distributions: evidence from mitochondrial phylogenomics. *Mol. Biol. Evol.* 35:970–983.
- Buček A, Wang M, Šobotník J, et al. 2022. Molecular phylogeny reveals the past transoceanic voyages of drywood termites (Isoptera, Kalotermitidae). *Mol. Biol. Evol.* 39:msac093.
- Cameron SL. 2014. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annu. Rev. Entomol.* 59:95–117. <https://doi.org/10.1146/annurev-ento-011613-162007>
- Cameron SL, Lo N, Bourguignon T, et al. 2012. A mitochondrial genome phylogeny of termites (Blattodea: Termitidae): robust support for interfamilial relationships and molecular synapomorphies define major clades. *Mol. Phylogenet. Evol.* 65:163–173.
- Dedeine F, Dupont S, Guyot S, et al. 2016. Historical biogeography of *Reticulitermes* termites (Isoptera: Rhinotermitidae) inferred from analyses of mitochondrial and nuclear loci. *Mol. Phylogenet. Evol.* 94:778–790.
- Donath A, Jühling F, Al-Arab M, et al. 2019. Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes. *Mol. Ecol. Resources.* 19:609–615.
- Ekman S, Blaalid R. 2011. The devil in the details: interactions between the branch-length prior and likelihood model affect node support and branch lengths in the phylogeny of the psoraceae. *Syst. Biol.* 60:541–561.
- Eldrett JS, Greenwood DR, Harding IC, et al. 2009. Increased seasonality through the Eocene to Oligocene transition in northern high latitudes. *Nature* 459:969–973.

- Engel MS, Grimaldi DA, Krishna K. 2009. Termites (Isoptera): their phylogeny, classification, and rise to ecological dominance. *Am. Mus. Novitates* 2009:1–27.
- Forni G, Puccio G, Bourguignon T, et al. 2019. Complete mitochondrial genomes from transcriptomes: assessing pros and cons of data mining for assembling new mitogenomes. *Sci. Rep.* 9:14806.
- Forni G, Corneti S, Luchetti A. 2025. Genomic divergence of the Italian drywood termites *Kaloterme italicus* and *K. flavicollis* (Blattodea: Kalotermitidae): insights from genome skimming. *Ethol. Ecol. Evol.* in Press 1. <https://doi.org/10.1080/03949370.2025.2525081>
- Hillis DM, Pollock DD, McGuire JA, et al. 2003. Is sparse taxon sampling a problem for phylogenetic inference? *Syst. Biol.* 52:124–126. <https://doi.org/10.1080/10635150390132911>
- Hoang DT, Chernomor O, von Haeseler A, et al. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35:518–522.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, et al. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods*. 14:587–589.
- Kainer D, Lanfear R. 2015. The effects of partitioning on phylogenetic inference. *Mol. Biol. Evol.* 32:1611–1627. <https://doi.org/10.1093/molbev/msv026>
- 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.
- Kishino H, Miyata T, Hasegawa M. 1990. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J. Mol. Evol.* 31:215–160. <https://doi.org/10.1007/BF02109483>
- Lanfear R, Calcott B, Ho SYW, et al. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29:1695–1701.
- Lartillot N, Philippe H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* 21:1095–1109.
- Lartillot N, Rodrigue N, Stubbs D, et al. 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst. Biol.* 62:611–615.
- Laslett D, Canbäck B. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* 24:172–175.
- Lemmon AR, Moriarty EC. 2004. The importance of proper model assumption in Bayesian phylogenetics. *Syst. Biol.* 53:265–277.
- Lewis PO, Holder MT, Holsinger KE. 2005. Polytomies and Bayesian phylogenetic inference. *Syst. Biol.* 54:241–253.
- Luchetti A, Trenta M, Mantovani B, et al. 2004. Taxonomy and phylogeny of north mediterranean *Reticulitermes* termites (Isoptera, Rhinotermitidae): a new insight. *Ins. Soc.* 51:117–122.
- Minh BQ, Schmidt HA, Chernomor O, et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37:1530–1534.
- Philippe H, Brinkmann H, Lavrov DV, et al. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS Biol.* 9:e1000602. <https://doi.org/10.1371/journal.pbio.1000602>
- Prjibelski A, Antipov D, Meleshko D, et al. 2020. Using SPAdes de novo assembler. *Curr. Prot. Bioinformatics* 70:e102.
- Rokas A, Krüger D, Carroll SB. 2005. Animal evolution and the molecular signature of radiations compressed in time. *Science* 310:1933–1938. <https://doi.org/10.1126/science.111675>
- Ronquist F, Teslenko M, van der Mark P, et al. 2012. MRBAYES 3.2: efficient Bayesian phylogenetic inference and model selection across a large model space. *Syst. Biol.* 61:539–542.
- Righetti N, Forni G, Luchetti A. 2024. Mitochondrial phylogenomics supports a Carboniferous origin of Xenonomia. *Eur. Zool. J.* 91:1139–1146.
- Sanmartín I, Enghoff H, Ronquist F. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biol. J. Linnean Soc.* 73:345–390.
- Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* 51:492–508.
- Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16:1114–1116.
- Simmons MP, Pickett KM, Miya M. 2004. How meaningful are Bayesian support values? *Mol. Biol. Evol.* 21:188–199.
- Simmons MP, Ochoterena H, Freudenstein JV. 2002. Conflict between amino acid and nucleotide characters. *Cladistics*. 18:200–206.
- Som A. 2015. Causes, consequences and solutions of phylogenetic incongruence. *Brief. Bioinformatics* 16:536–548. <https://doi.org/10.1093/bib/bbu015>
- Steenwyk JL, Li Y, Zhou X, et al. 2023. Incongruence in the phylogenomics era. *Nat. Rev. Genet.* 24:834–850. <https://doi.org/10.1038/s41576-023-00620-x>
- Strimmer K, Rambaut A. 2002. Inferring confidence sets of possibly misspecified gene trees. *Proc. Biol. Sci.* 269:137–142.
- To T-H, Jung M, Lycett S, et al. 2016. Fast dating using least-squares criteria and algorithms. *Syst. Biol.* 65:82–97.
- Townsend JP, López-Giráldez F, Friedman R. 2008. The phylogenetic informativeness of nucleotide and amino acid sequences for reconstructing the vertebrate tree. *J. Mol. Evol.* 67:437–447. <https://doi.org/10.1007/s00239-008-9142-0>
- Uva P, Clément JL, Austin JW, et al. 2004. Origin of a new reticulitermes termite (Isoptera, Rhinotermitidae) inferred from mitochondrial and nuclear DNA data. *Mol. Phylogenet. Evol.* 30:344–353.
- Wang M, Buček A, Šobotník J, et al. 2019. Historical biogeography of the termite clade Rhinotermitinae (Blattodea: Isoptera). *Mol. Phylogenet. Evol.* 132:100–104.
- Whitfield JB, Lockhart PJ. 2007. Deciphering ancient rapid radiations. *Trends Ecol. Evol.* 22:258–265.
- Zwickl DJ, Hillis DM. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Syst. Biol.* 51:588–598.