Molecular Phylogeny Reveals the Past Transoceanic Voyages of Drywood Termites (Isoptera, Kalotermitidae)

Aleš Buček ©,*,[1] Menglin Wang ©,[1] Jan Šobotník ©,[2] Simon Hellemans ©,[1] David Sillam-Dussès ©,[2,3] Nobuaki Mizumoto ©,[1] Petr Stiblík ©,[2] Crystal Clitheroe ©,[1] Tomer Lu,[4] Juan José González Plaza ©,[5] Alma Mohagan,[6,7] Jean-Jacques Rafanomezantsoa,[8] Brian Fisher ©,[8,9] Michael S. Engel ©,[10,11] Yves Roisin ©,[12] Theodore A. Evans ©,[13] Rudolf Scheffrahn,[14] and Thomas Bourguignon ©*,[1,2]

[1]Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna-son, Okinawa 904-0495, Japan

[2]Faculty of Tropical AgriSciences, Czech University of Life Sciences, Prague, Czech Republic

[3]Laboratory of Experimental and Comparative Ethology, UR 4443, University Sorbonne Paris Nord, Villetaneuse, France

[4]Total Hadbara, Gedera, Israel

[5]International Research Centre in Critical Raw Materials-ICCRAM, University of Burgos, Plaza Misael Bañuelos s/n, 09001 Burgos, Spain

[6]Center for Biodiversity Research and Extension in Mindanao, Central Mindanao University, Musuan, Maramag, Bukidnon 8710, Philippines

[7]Department of Biology, College of Arts and Sciences, Central Mindanao University, Musuan, Maramag, Bukidnon 8710, Philippines

[8]Madagascar Biodiversity Center, Parc Botanique et Zoologique de Tsimbazaza, Antananarivo, Madagascar

[9]California Academy of the Sciences, San Francisco, CA, USA

[10]Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, KS, USA

[11]Division of Entomology, Natural History Museum, University of Kansas, Lawrence, KS, USA

[12]Evolutionary Biology and Ecology, Université Libre de Bruxelles, Bruxelles, Belgium

[13]School of Biological Sciences, University of Western Australia, Perth, WA 6009, Australia

[14]Fort Lauderdale Research and Education Center, Institute for Food and Agricultural Sciences, 3205 College Avenue, Davie, FL 33314, USA

*Corresponding authors: E-mails: Bucek.Ales@gmail.com; thomas.bourguignon@oist.jp.
Associate editor: Koichiro Tamura

Abstract

Termites are major decomposers in terrestrial ecosystems and the second most diverse lineage of social insects. The Kalotermitidae form the second-largest termite family and are distributed across tropical and subtropical ecosystems, where they typically live in small colonies confined to single wood items inhabited by individuals with no foraging abilities. How the Kalotermitidae have acquired their global distribution patterns remains unresolved. Similarly, it is unclear whether foraging is ancestral to Kalotermitidae or was secondarily acquired in a few species. These questions can be addressed in a phylogenetic framework. We inferred time-calibrated phylogenetic trees of Kalotermitidae using mitochondrial genomes of ~120 species, about 27% of kalotermitid diversity, including representatives of 21 of the 23 kalotermitid genera. Our mitochondrial genome phylogenetic trees were corroborated by phylogenies inferred from nuclear ultraconserved elements derived from a subset of 28 species. We found that extant kalotermitids shared a common ancestor 84 Ma (75–93 Ma 95% highest posterior density), indicating that a few disjunctions among early-diverging kalotermitid lineages may predate Gondwana breakup. However, most of the ~40 disjunctions among biogeographic realms were dated at <50 Ma, indicating that transoceanic dispersals, and more recently human-mediated dispersals, have been the major drivers of the global distribution of Kalotermitidae. Our phylogeny also revealed that the capacity to forage is often found in early-diverging kalotermitid lineages, implying the ancestors of Kalotermitidae were able to forage among multiple wood pieces. Our phylogenetic estimates provide a platform for critical taxonomic revision and future comparative analyses of Kalotermitidae.

Key words: time-calibrated phylogenetic tree, historical biogeography, social evolution, long distance dispersal, insects, molecular clock.

Introduction

Termites are eusocial cockroaches comprising ~3,000 described species classified into nine extant families (Krishna et al. 2013). The Termitidae, also called higher termites, are the most diverse family, including ~2,100 described species, or about three-fourths of termite diversity (Krishna et al. 2013). By comparison, the Kalotermitidae, the second-largest family, are substantially less speciose, including ~450 described living species classified into 23 genera (Engel, Grimaldi, et al. 2009; Krishna et al. 2013). However, the modest diversity of Kalotermitidae belies their considerable economic significance as one-third of invasive termite species are kalotermitids, including some of the most destructive pests, such as *Cryptotermes brevis* (Evans et al. 2013). Kalotermitidae are also among the most widely distributed termite lineages, occurring worldwide between 45°N and 45°S latitude (Emerson 1969; Jones and Eggleton 2011). Despite their economic importance, there has been no attempt to reconstruct a comprehensive phylogenetic hypothesis for Kalotermitidae. A robust phylogeny for Kalotermitidae is needed to shed light on their evolutionary history.

Among extant families, Kalotermitidae are one of the oldest (Inward et al. 2007; Legendre et al. 2008; Engel, Grimaldi, et al. 2009; Engel et al. 2016), with fossil records dating back to the early Late Cretaceous, 99 Ma (Engel et al. 2007b; Barden and Engel 2021). Time-calibrated phylogenies based on transcriptomes and mitochondrial genomes have estimated that the kalotermitid lineage diverged from their common ancestor with Neoisoptera ~127 Ma (96–148 Ma, combined 95% highest posterior density [HPD] of multiple analyses) (Bourguignon et al. 2015; Bušek et al. 2019). These studies also estimated the age of extant (i.e., crown) Kalotermitidae at 69–81 Ma (54–90 Ma, combined 95% HPD of multiple analyses), albeit with limited taxon sampling within Kalotermitidae (Bourguignon et al. 2015; Bušek et al. 2019). Therefore, the age of early-diverging Kalotermitidae lineages falls within the time range estimations of the final stage of Gondwanan breakup, before the disappearance of land bridges connecting nascent continents (Upchurch 2008). Although vicariance through continental drift may explain the distribution patterns of some early-branching kalotermitid lineages, many genera distributed across almost all terrestrial biogeographic realms, such as *Neotermes*, *Glyptotermes*, and *Cryptotermes* (Krishna et al. 2013), are <60 million years old (Bušek et al. 2019), postdating the breakup of Gondwana (Scotese 2004). Therefore, both vicariance and dispersal processes may have contributed to the geographical distribution and evolution of extant Kalotermitidae.

Termites have winged alates, yet they are poor flyers, unable to actively perform long-range dispersals (Nutting 1969). However, the Kalotermitidae are particularly well-suited to an alternative passive dispersal strategy: transoceanic rafting. The ability of Kalotermitidae to disperse overwater is well illustrated by studies of the fauna of the Krakatau islands, which were entirely defaunated by a volcanic eruption in 1883 and then recolonized by multiple species of Kalotermitidae within 100 years (Abe 1984; Gathorne-Hardy and Jones 2000). The dispersal ability of Kalotermitidae stems from their lifestyle, as they usually nest in and feed on single pieces of wood (Abe 1987; Scheffrahn et al. 2006; Scheffrahn and Postle 2013), which can float across oceans as rafts (Thiel and Haye 2006). The Kalotermitidae are also able to produce secondary reproductives (Myles 1999), increasing the chance of small colony fragments rafting across oceans in wood pieces to reproduce upon arrival to their new destination (Evans et al. 2013). These traits also predispose some species of Kalotermitidae to become invasive, spreading with the help of anthropogenic global material transport (Evans et al. 2013).

Most extant species of Kalotermitidae are unable to forage outside their nesting wood piece. Instead, they make small colonies in wood items like dead branches on living trees (Abe 1987). In these colonies, all members retain their reproductive potential and the ability to develop wings to disperse and found new colonies, except for the soldiers that remain permanently sterile and wingless (Roisin and Korb 2011). These life traits were intuitively interpreted as representative of the ancestral condition of modern termites (Abe 1987; Thorne 1997; Inward et al. 2007). By contrast, some studies have hypothesized that a complex colony organization, with true workers foraging outside their nests, is ancestral to extant termites, in which case foraging ability was secondarily lost in certain termite lineages (Morgan 1959; Watson and Sewell 1985; Bordereau and Pasteels 2011; Bourguignon, Chisholm, et al. 2016; Mizumoto and Bourguignon 2020). Whether foraging is ancestral to Kalotermitidae or was secondarily acquired remains unresolved.

The most comprehensive effort to reconstruct the evolutionary history of Kalotermitidae dates to the 1960s, with the taxonomic revision of Kalotermitidae by Krishna (1961). Krishna's generic classification was based on soldier and imago morphology and has been remarkably stable since its inception, with only two new genera added (Ghesini et al. 2014; Scheffrahn et al. 2018). However, Krishna's generic classification of Kalotermitidae has not been adequately tested by modern phylogenetic methods. A handful of studies have investigated kalotermitid phylogeny using a few genetic markers obtained from samples not representative of the taxonomic and geographic diversity of kalotermitids (e.g., Thompson et al. 2000; Ghesini et al. 2014; Bourguignon et al. 2015; Scheffrahn et al. 2018; Bušek et al. 2019). Here, we used mitochondrial genomes of 172 kalotermitid samples collected worldwide, which includes species of 21 kalotermitid genera, to reconstruct a robust phylogenetic tree. We also reconstructed a phylogenetic tree using nuclear ultraconserved elements (UCEs) derived from a subset of 18 Kalotermitidae species. Using these comprehensive phylogenies, we reconstructed the global historical biogeography of Kalotermitidae,

identifying past natural and human-mediated dispersal events and distribution patterns that bear the signature of potential events of vicariance through continental drift. Finally, we created a list of species of Kalotermitidae with available foraging data and discussed the evolution of such abilities according to patterns implied by the phylogenetic framework.

Results and Discussion

Evolution of Kalotermitidae Inferred from Taxonomically Representative Phylogenetic Trees

We gathered two sequence datasets. The first dataset comprised mitochondrial genomes of 172 samples of Kalotermitidae collected across all biogeographic realms (fig 1, supplementary fig. S1 and table S1, Supplementary Material online). The second dataset included 82–691 UCE loci for a representative subset of 18 species belonging to 14 kalotermitid genera. We reconstructed maximum likelihood and Bayesian time-calibrated trees using these datasets combined with nonkalotermitid outgroups (see Materials and Methods and supplementary table S2, Supplementary Material online for details). The topologies and branch lengths of maximum likelihood trees based on 82 and 691 UCEs were highly congruent (supplementary fig. S2, Supplementary Material online). For this reason, and because of computational power limitation, we used 82 UCEs to reconstruct Bayesian time-calibrated trees. To determine the role of sampling effort and sequence type on the variation of node age estimates, we compared phylogenetic trees reconstructed with 230 mitochondrial genomes (including 172 samples of Kalotermitidae), 28 mitochondrial genomes, and 28 UCEs (including 18 samples of Kalotermitidae). We found that the time-calibrated trees inferred with the UCE and mitochondrial data of 28 species were largely congruent in terms of topology and node age estimates, indicating that our phylogenetic inferences were not influenced by the type of genetic markers used (fig. 2, supplementary fig. S3, Supplementary Material online). The topology of the tree inferred with 230 mitochondrial genomes was also highly congruent with the trees inferred with the UCEs and mitochondrial genomes of 28 species. However, the trees reconstructed with 230 mitochondrial genomes yielded mean age estimates up to 20 Ma older than those of the trees reconstructed with data from 28 species, although their 95% HPD intervals overlapped for most nodes (fig. 2). Next, we investigated the effect of the substitution model by comparing time-calibrated trees reconstructed from the dataset composed of 230 mitochondrial genomes. The analyses performed with a nucleotide substitution model for all mitochondrial genes were highly congruent in terms of topology with the analyses performed with an amino acid substitution model for protein-coding genes and a nucleotide substitution model for rRNA and tRNA genes (fig. 2). Both analyses yielded time-calibrated trees with overlapping 95% HPD intervals for all nodes outside some of

the youngest nodes. Therefore, the sampling effort had more influence on age estimates than the choice of genetic markers or evolution models, and we used the most comprehensive dataset composed of 230 mitochondrial genomes to reconstruct the natural history of Kalotermitidae.

We performed multiple maximum likelihood and Bayesian phylogenetic reconstructions using combinations of protein and nucleotide substitution models, including or excluding the third codon positions (supplementary table S2, Supplementary Material online). The node support statistics of all trees are summarized in a reduced summary evidence tree (RSE-tree) (fig. 1). The RSE-tree was obtained by pruning 35 tips that diverged <1 Ma and were represented by samples collected in the same country. Our RSE-tree was composed of 137 tips representing 21 of the 23 genera of Kalotermitidae and including more than 120 species, which is equivalent to ~27% of described kalotermitid species-level diversity (Krishna et al. 2013) (fig. 1). The RSE-tree was composed of 136 branches, 116 recovered in all four analyses and 20 presenting at least one topological conflict among analyses. The topology was congruent at the genus level among phylogenetic analyses, except for the positions of genus-level branches within the clade *Incisitermes* group C + *Neotermes* group B + *Incisitermes* group B + *Cryptotermes* + *Procryptotermes* that varied among analyses (figs. 1 and 3).

Time-calibrated phylogenetic trees were inferred using termite fossils as node calibrations. We ran Bayesian Binary MCMC ancestral range reconstructions with four biogeographic models on every four phylogenetic estimates with full taxon sampling (supplementary table S2, Supplementary Material online). The reconstructed ancestral range distributions were congruent among the 16 analyses (see fig. 1, supplementary figs. S4–S7, Supplementary Material online). These results were further corroborated by ancestral range reconstructions performed with the dispersal-extinction-cladogenesis (DEC) model (supplementary fig. S8, Supplementary Material online).

Taxonomic Note

Our phylogenetic trees recovered many of the currently recognized genera of Kalotermitidae as nonmonophyletic. These nonnatural divisions previously led to misinformed hypotheses on dispersal routes and ancestral distributions of Kalotermitidae (e.g., Eggleton and Davies 2003). Therefore, our phylogenetic estimates highlight that the Kalotermitidae are in need of a generic revision. Although a comprehensive revision of Kalotermitidae is outside the scope of the present paper, we provide a provisional list of genus-level taxonomic revisions that are required.

Bifiditermes was retrieved as polyphyletic, forming a clade within which *Allotermes*, *Roisinitermes*, and *Epikalotermes* were nested. This clade needs a substantial taxonomic revision. *Procryptotermes* formed two distinct lineages nested within *Cryptotermes*, a result previously found by Thompson et al. (2000) based on a
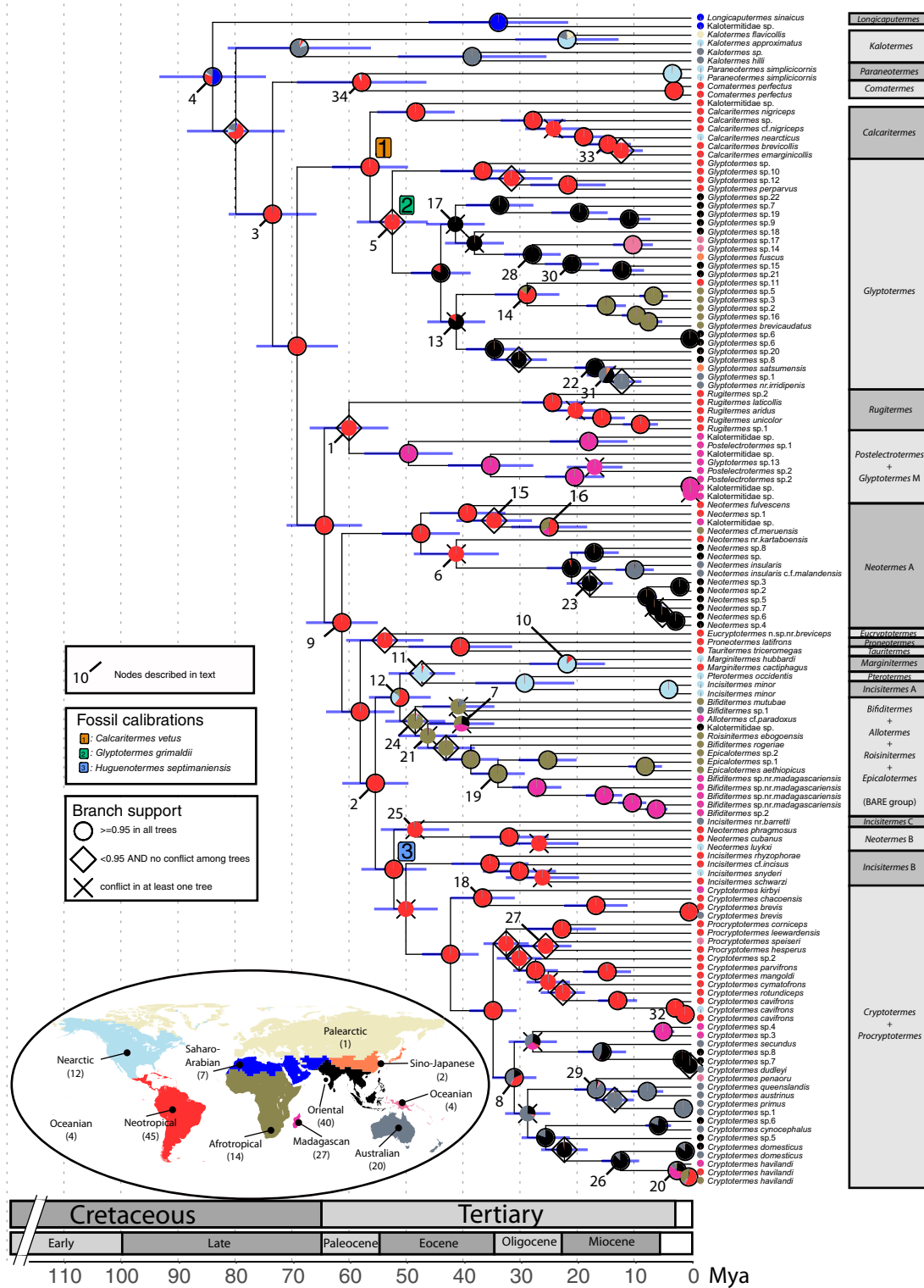
Fig. 1. Time-calibrated phylogeny of Kalotermitidae. Bayesian tree inferred from noncoding nucleotide sequences and protein sequences derived from mitochondrial genome sequences. Internal node shapes summarize the congruence of all other phylogenetic analyses with the backbone tree topology. Node bars represent 95% HPD age intervals. Node pie charts represent the average probabilities of ancestral ranges inferred with four Bayesian Binary MCMC models run on the displayed tree topology. The world map indicates the biogeographic realms recognized in this study and the number of kalotermitid samples analyzed for every realm. Colored numerical node labels indicate fossil calibrations. Tree nodes labeled with numbers 1–34 are referred to in the main text. Nonkalotermitid outgroups were trimmed from the tree. A single tip was kept for every species in every sampling location. For the phylogeny including all samples used in this study and their collection codes, see supplementary figure S1, Supplementary Material online. For ancestral geographic distributions inferred with alternative tree topologies, see supplementary figures S5–S7, Supplementary Material online.
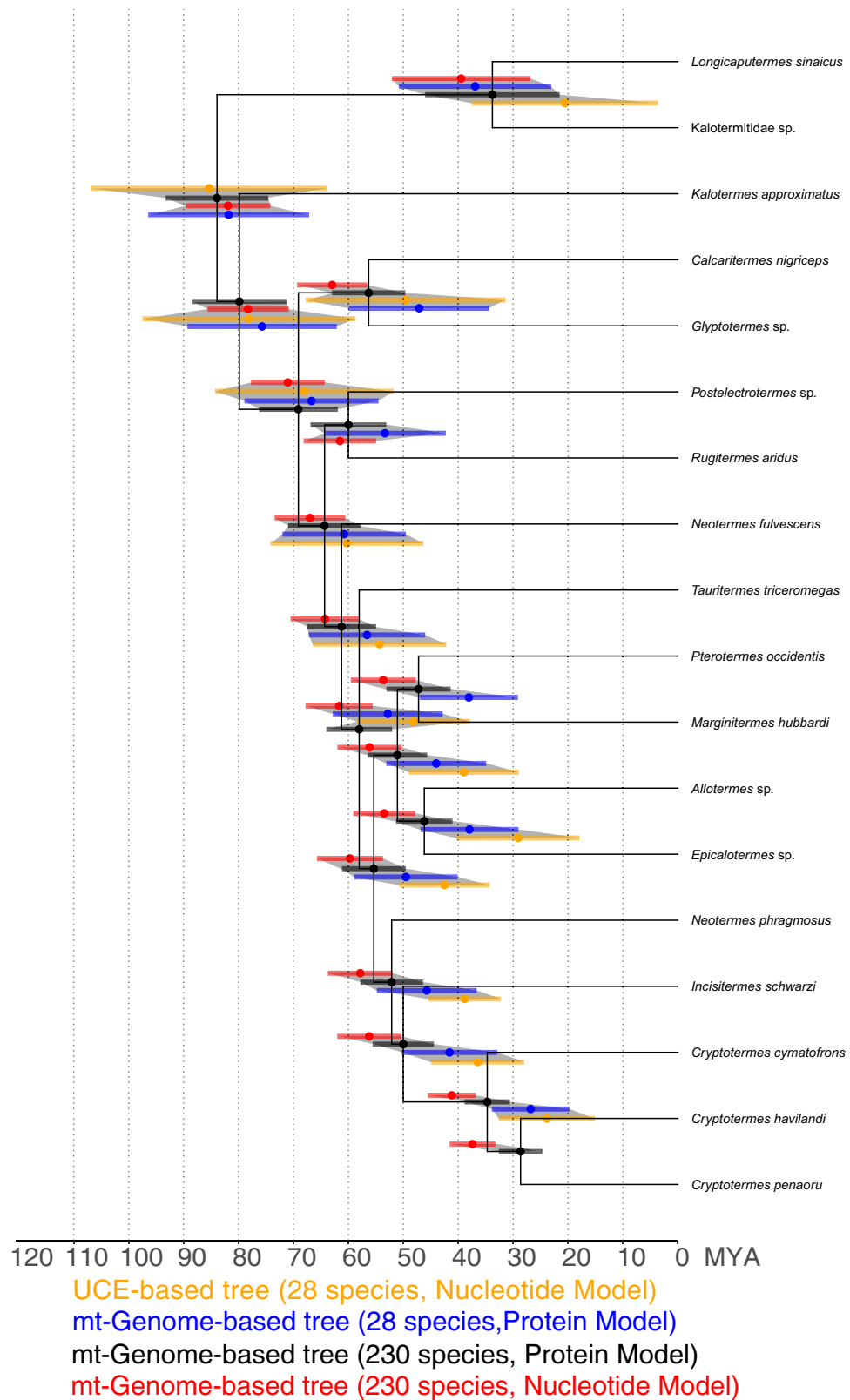
FIG. 2. Comparison of age estimates across phylogenetic analyses performed on a subset of 28 termite species. Node ages were inferred from four different Bayesian inference analyses based on: a UCE dataset including 28 termite species (in orange); a mitochondrial protein sequence and noncoding nucleotide sequence dataset including 28 termite species (in blue); a mitochondrial protein sequence and noncoding nucleotide sequence dataset including 230 termite species (in black, used as the backbone tree here and identical to the tree displayed in supplementary fig. 1, Supplementary Material online); and a mitochondrial nucleotide sequence dataset including 230 termite species (in red). All phylogenies included a subset of 18 species of Kalotermitidae. Mean and 95% HPD age intervals inferred from the four analyses are displayed as colored clouds. Missing age intervals for individual analyses indicate incongruence with the reference tree. See supplementary figure S3, Supplementary Material online for individual trees with outgroups and sample IDs.
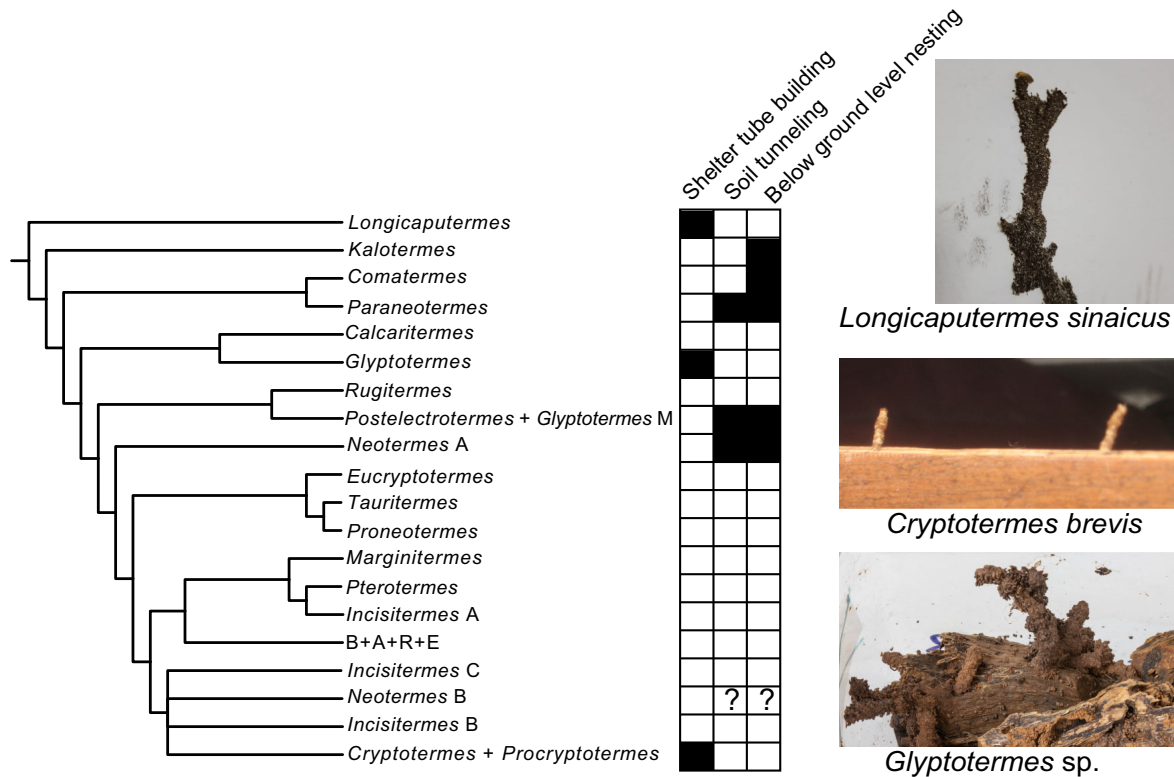
Fig. 3. Construction and foraging abilities of Kalotermitidae. The phylogeny from supplementary figure 1, Supplementary Material online was collapsed at the genus level. Branches showing conflicting topology among our phylogenetic analyses are represented as polytomies. The heat-map indicates the presence of behavioral records (black) for at least one species of the genus. Note that the absence of evidence of a behavior (white) either reflects its actual absence or the absence of its observation. "Below-ground level" indicates the presence of termites in wood items that are below-ground level, such as tree roots. Interrogation marks indicate observations that cannot be unambiguously mapped onto the phylogeny due to the uncertain phylogenetic position of observed species. The photos show the shelter tubes of *Longicaputermes sinaicus* attacking timber in urban environments, the shelter tubes built by *Cryptotermes brevis* under laboratory conditions, and the shelter tubes built by an unidentified *Glyptotermes* species maintained in a laboratory and collected in Papua New Guinea. For details on the literature survey and authors' observations of kalotermitid behavior, see supplementary table S3, Supplementary Material online.

mitochondrial COII phylogeny of Australian Kalotermitidae. *Procryptotermes* differs from *Cryptotermes* by the morphology of its soldiers—possessing slight differences in serration of the anterior margin of the pronotum, longer mandibles, and a less phragmotic head—whereas the alate imagoes of both genera are similar in all respects (Krishna 1961). Given their considerable morphological similarities, the variability of the diagnostic characters, the existence of species with intermediary morphology (Scheffrahn and Krecek 1999), and that both genera are mutually paraphyletic, *Procryptotermes* Holmgren 1910 should be formally synonymized with *Cryptotermes* Banks 1906. The monospecific genus *Ceratokalotermes* is not represented in our summary evidence tree, but we explored its phylogenetic position using a COII sequence previously published by Thompson et al. (2000). We found *Ceratokalotermes* to be nested within a paraphyletic *Kalotermes* (supplementary fig. S9, Supplementary Material online), as originally suggested based on morphology (Hill 1942). These results are in contrast to a previous phylogenetic reconstruction based on COII sequences in which *Ceratokalotermes* was retrieved as sister to *Kalotermes*, albeit with low support and without

the inclusion of species of *Kalotermes* from outside the Australian region (Thompson et al. 2000). Although our phylogenetic estimates strongly support the paraphyly of *Kalotermes* with respect to *Ceratokalotermes*, we did not sequence any specimen of *Ceratokalotermes* and consider it premature to synonymize *Ceratokalotermes* with *Kalotermes*. The genus *Postelectrotermes* is defined by the presence of a spine on the outer margin of the mesotibia, near the outer apical spur (Krishna 1961). We recovered a group of Madagascan kalotermitids that included species with the diagnostic spine and species that do not have this additional spine as monophyletic. The name *Postelectrotermes* would be available for this broader combined group, although a revision is needed to determine what morphological traits serve to circumscribe this clade of species, currently comprising *Postelectrotermes* and Malagasy species erroneously assigned to *Glyptotermes* and termed *Glyptotermes* group M herein. *Incisitermes* is a polyphyletic genus containing at least two distantly related lineages: *Incisitermes* group A includes the Nearctic *Incisitermes minor*; *Incisitermes* group B includes Nearctic and Neotropical species. The phylogenetic position of the

tentative *Incisitermes* group C, consisting of the Australian *Incisitermes* nr. *barretti*, is not well resolved. The *Incisitermes* groups A and B diverged from their sister group 30–50 Ma (fig. 1) and should be classified into separate genera. We did not sequence the type species of the genus, *I. schwarzi*, but morphological characters should be sufficient to properly identify the groups and determine which clade comprises *Incisitermes* proper and which represents a new genus. For example, *I. schwarzi*, *I. minor*, and *I. marginipennis* all have soldiers with pigmented compound eyes, at least suggesting that *Incisitermes* A represents that clade to which the generic name *Incisitermes* should be restricted (with *Incisitermes* B therefore likely representing a new genus). *Incisitermes peritus* in Dominican amber likely represents clade A, whereas *I. krishnai* in Mexican amber could represent clade B, but these will require critical testing once these clades are formally revised. *Neotermes* is also a polyphyletic assemblage represented in our phylogenies by two distant groups, *Neotermes* A and *Neotermes* B. Both *Neotermes* groups share their last common ancestor 61 Ma (55–68 Ma 95% HPD, node 9 in fig. 1). The two *Neotermes* groups should be classified into two separate genera.

Historical Biogeography of Kalotermitids

Since the age estimates of the tree inferred with a nucleotide substitution model for all genes were very similar to that of the tree inferred with an amino acid substitution model for protein-coding genes and a nucleotide substitution model for rRNA and tRNA genes, we will only provide hereunder the age estimates of the latter tree. We estimated the divergence between the lineages comprising Kalotermitidae and Neoisoptera at 115 Ma (104–126 Ma 95% HPD, supplementary fig. S1, Supplementary Material online). Although this divergence age estimate is younger than that estimated by previous time-calibrated phylogenetic trees using mitochondrial genomes and transcriptome data, HPD intervals largely overlapped (Bourguignon et al. 2015; Buček et al. 2019). This age estimate was corroborated by a Bayesian tree inferred with the UCE dataset that yielded 119 Ma (96–145 Ma 95% HPD) as an age estimate for the divergence between Kalotermitidae and Neoisoptera (supplementary fig. S3, Supplementary Material online). The differences with previous studies may result from using different fossil calibrations among studies, changes of fossil age estimates based on revised or refined stratigraphic and radiometric dating for fossil deposits, or different taxonomic sampling among studies. Regardless, all studies estimated the origin of Kalotermitidae to predate the first kalotermitid fossils, the 99-million-year-old species of *Proelectrotermes* (Engel et al. 2007b), by 6–28 million years. Crown-group Kalotermitidae share a common ancestor estimated to live 84 Ma (75–93 Ma 95% HPD, node 4 in fig.1).

Since the most recent ancestor of extant Kalotermitidae predates the final stages of Gondwana breakup (Reguero et al. 2014), vicariance may have contributed to the distribution of early-diverging kalotermitid lineages. Our analyses did not allow us to identify the ancestral range of the last common ancestor of kalotermitids. However, four biogeographic realms, the Saharo-Arabian, Neotropical, Australian, and Nearctic, were inferred as possible ancestral ranges (node 4 in fig. 1). There is also a possibility that the ancestor of modern Kalotermitidae was distributed across several biogeographic realms, such as the Neotropical and Australian realms that were connected via land bridges, such as the Weddellian Isthmus, until at least ~56 Ma (Reguero et al. 2014). Although the ancestral range of the last common ancestor of modern Kalotermitidae remains speculative, we found strong support for a Neotropical origin of the clade composed of all modern Kalotermitidae with the exclusion of the species-poor genera *Kalotermes* (plus the genus *Ceratokalotermes* not sequenced in this study) and *Longicaputermes*. The age of this clade was estimated at 73 Ma (66–82 Ma 95% HPD, node 3 in fig. 1), which predates the separation of Antarctica from South America and Australia (Reguero et al. 2014). Admittedly, the distinction of a Neotropical region as unique from the remainder of the Gondwanan landmasses was not entirely meaningful owing to the lack of paleogeographical differentiation. In addition, it remains unknown to what degree extinction has impacted crown-Kalotermitidae. Nonetheless, the Neotropical region hosts more relict taxa than other realms (fig. 1). Other disjunctions between biogeographic realms postdated the breakup of Gondwana and are therefore consistent with transoceanic and land dispersal modes (fig. 1).

Our results indicate that the Neotropical realm is likely a cradle of modern kalotermitid diversity and the source of repeated dispersals to other biogeographic realms (fig. 1). These results were not a consequence of our more intensive sampling of the Neotropical realm because our ancestral range reconstructions were robust to removing up to two-thirds of the Neotropical samples (supplementary figs. S10 and S11, Supplementary Material online). We found support for two natural long-distance overwater dispersal events to the Neotropical realm: (1) one dispersal of stem- or crown-*Marginitermes* from the Nearctic realm 22–47 Ma (15–53 Ma combined 95% HPD, nodes 10–11 in fig. 1), and one dispersal of a *Glyptotermes* lineage from the Afrotropical or Oriental realm 29–41 Ma (23–46 Ma combined 95% HPD, nodes 13–14 in fig. 1). However, a Neotropical origin of the latter *Glyptotermes* lineage combined with multiple dispersals out of the Neotropical realm could not be excluded (fig. 1, supplementary figs. S4–S7, Supplementary Material online). Note that, throughout this work, we used as dispersal age the age of the most recent ancestor that had, with high probability (>90%), a different range than its descending lineage. This age represents a maximum dispersal age estimate. For disjunctions between ancestral and modern ranges supported by 10–90% probabilities, we also considered alternative dispersal hypotheses. In such case, we provide

the dispersal age as an age interval delimited by the maximum dispersal age and alternative dispersal ages. Other kalotermitid dispersal events to the Neotropics were mediated by human activities, with notable introductions of *Cryptotermes havilandi* from the Afrotropical and *C. dudleyi* from the Oriental realm (Evans et al. 2013). One possible explanation for this observed low dispersal rate to the Neotropics is the presence of a diverse neotropical kalotermitid fauna, which presumably filled most niches available to Kalotermitidae, hindering colonization by foreign kalotermitids.

The Oriental realm was colonized by at least four dispersal events. Our ancestral range reconstructions indicate that *Glyptotermes* was the first lineage dispersing to the Oriental realm, possibly as early as in the Early Eocene, 52 Ma (47–59 Ma 95% HPD, node 5 in fig. 1), and most likely before 38 Ma (33–43 Ma 95% HPD, node 17 in fig. 1). Subsequent dispersal events to the Oriental realm took place 35–45 Ma (31–51 Ma combined 95% HPD), before the Eocene-Oligocene climatic shift and extinction events. These dispersal events included one dispersal of *Neotermes* group A (node 6 in fig. 1) and one or several dispersals of *Cryptotermes* (node 8 in fig. 1), all of which originated from the Neotropical realm. The Oriental realm was also colonized by one additional lineage, wrongly identified as *Neotermes* by Bourguignon et al. (2015), and allied to the *Bifiditermes* + *Allotermes* + *Roisinitermes* + *Epicalotermes* group, which most likely dispersed from the Afrotropics (node 7 in fig. 1). Although our time-calibrated trees suggest the first Oriental kalotermitids date from the early Eocene, the fossils of *Proelectrotermes* provide evidence that the Oriental realm was already occupied by kalotermitids >99 Ma (Engel et al. 2007b). Our ancestral range reconstructions do not support an Oriental origin for crown-kalotermitids, and analyses with the last common kalotermitid ancestor constrained to be Oriental supported that the modern Oriental kalotermitid fauna colonized the Oriental realm through at least four dispersals, including at least three long-distance overwater dispersals (supplementary fig. S8, Supplementary Material online). Early kalotermitid fossils unearthed from the Oriental realm, such as the 99-million-year-old Kachin amber fossils of *Proelectrotermes*, therefore belong to ancient lineages that went extinct in the Oriental realm. If *Proelectrotermes*, perhaps along with concurrent *Kachinitermes* and *Kachinitermopsis*, is representative of stem-kalotermitids, the broader lineage encompassing Kalotermitidae could have originated outside of the Neotropical realm, with crown-kalotermitids representing an originally "western Gondwana" subordinate clade of this broader group. Refined testing of such a hypothesis will require significantly more Cretaceous fossils of stem-Kalotermitidae and a phylogeny ascertaining their relationships to one another as well as to the crown group. At the least, there is no current phylogenetic or paleontological evidence to support an eastern Laurasian (i.e., north of Tethys) origin for the original divergence from the common ancestor of Neoisoptera.

Madagascar was colonized by kalotermitids at least seven times independently (fig. 1). The first dispersal event to Madagascar was that of the ancestor of the Madagascan *Postelectrotermes* + *Glyptotermes* group M, which diverged from the common ancestor of Neotropical *Rugitermes* 60 Ma (53–67 Ma 95% HPD, node 1 in fig. 1). Subsequently, the lineage comprising *Cryptotermes kirbyi*, which diverged from other *Cryptotermes* 37 Ma (31–42 Ma 95% HPD, node 18 in fig. 1), dispersed to Madagascar from the Neotropical realm. *Neotermes* group A represents another colonization event of Neotropical, or possibly Afrotropical, origin 35 Ma (28–41 Ma 95% HPD, node 15 in fig. 1). The monophyletic lineage of Madagascan *Bifiditermes* originated from the Afrotropical realm and diverged from their sister group, the Afrotropical *Epicalotermes*, 34 Ma (29–39 Ma 95% HPD, node 19 in fig. 1). The origins of the three other dispersal events to Madagascar are uncertain. *Cryptotermes havilandi* is presumably invasive in Madagascar (Evans et al. 2013). But the divergence of Madagascan *C. havilandi* and non-Madagascan *C. havilandi* dating to 2.5 Ma (1.5–3.5 Ma 95% HPD, node 20 in fig. 1) suggests dispersal to Madagascar might predate human voyages. More thorough sampling in the Afrotropical realm will be necessary to resolve the native range of *C. havilandi*. The sister lineage of *Allotermes* is unclear from our analyses and could either be the Australian *Bifiditermes* sp. or an unidentified Kalotermitidae sp. from the Oriental realm, from which *Allotermes* diverged 40–45 Ma (35–51 Ma combined 95% HPD; fig. 1, supplementary figs. S4–S7, Supplementary Material online). The phylogenetic position of the third *Cryptotermes* lineage that colonized Madagascar is also not congruent among our analyses, but it likely dispersed from either the Oriental or Australian realm not later than 31 Ma (34–43 Ma 95% HPD, node 8 in fig. 1). Madagascar and India broke away from Africa ∼140 Ma, and India broke apart from Madagascar and started drifting northward ∼85 Ma (Gibbons et al. 2013). The long isolation of Madagascar from other biogeographic realms implies that the seven independent colonization events of Madagascar by Kalotermitidae occurred through long-distance overwater dispersals. The numerous transoceanic dispersals to Madagascar provide evidence for the high-dispersal abilities of Kalotermitidae.

Our ancestral range reconstructions suggest the last common ancestor of Kalotermitidae could be Australian. If so, it is unclear whether the clade composed of *Kalotermes* (plus *Ceratokalotermes*) dispersed to Australia or whether it is relictual in Australia. The timing of other disjunctions between the Australian and other realms correspond to natural dispersal events or human introductions. Our ancestral state reconstructions retrieved four colonization events of Australia by non-*Cryptotermes* Kalotermitidae and up to six colonization events by *Cryptotermes* (fig. 1). *Glyptotermes* and *Neotermes* group A dispersed overwater to Australia from the Oriental realm 17 Ma (13–20 Ma 95% HPD, node 22 in fig. 1) and 18 Ma (14–22 Ma 95% HPD, node

23 in fig. 1), respectively. This timing coincides with the colonization of Australia by higher termites, which possibly benefited from the Miocene expansion of grasslands and growing drier climates in Australia (Bourguignon et al. 2017). The other two colonization events by non-*Cryptotermes* Kalotermitidae were more ancient. *Bifiditermes* sp., which was incorrectly identified by Cameron et al. (2012) as *Neotermes insularis*, has an unresolved phylogenetic position and possibly colonized Australia from Africa or Madagascar through long-distance overwater dispersal 48 Ma (43–54 Ma 95% HPD, node 24 in fig. 1). The phylogenetic position of the Australian *Incisitermes* nr. *barretti* is unresolved; however, it is unambiguously inferred to share its last common ancestor with Neotropical Kalotermitidae 48 Ma (42–54 Ma 95% HPD, node 25 in fig. 1 and supplementary figs. S5–S7, Supplementary Material online). Both the Australian *Incisitermes* and *Bifiditermes* lied on long branches, and the resolution of their phylogenetic position and historical biogeography require additional samples, such as the Oriental "*Incisitermes*", which was missing in our sampling. South America and Australia were connected by land through Antarctica until at least ~56 Ma (Reguero et al. 2014). The formation of deep-water isolation by the Eocene-Oligocene transition, when the Antarctic circumpolar current was established, resulted in an eventual permanent glaciation of Antarctica and global cooling and drying during the latest Paleogene and early Neogene (Zachos 2001). *Incisitermes* nr. *barretti*, therefore, dispersed either overwater or by land, as was previously suggested for many animal taxa, including insects (Sanmartín and Ronquist 2004), which possibly dispersed from South America to Australia through an Antarctic landmass covered by temperate forests (Dawson et al. 1976; Mörs et al. 2020). In addition, *Marginitermes absitus*, which was absent from our phylogenetic inferences, presumably colonized northern Australia via a long-distance overwater dispersal event from Central or North America (Scheffrahn and Postle 2013). Our ancestral range reconstructions indicate up to six colonization events of Australia by *Cryptotermes*, although this number was uncertain because of ambiguous ancestral range reconstructions (fig. 1). Four of these six colonization events correspond to human introductions, from the Oriental realm in the case of *C. domesticus*, *C. dudleyi*, and *C. cynocephalus*, and from the Neotropical realm in the case of *C. brevis* (Evans et al. 2013). The origin of the two other colonization events, that of *C. secundus* and the clade of *Cryptotermes* including *C. austrinus*, is likely the Oriental or Neotropical realm. Future studies are required to refine these patterns of dispersals, especially for *Cryptotermes*.

We identified four likely colonization events of the Afrotropical realm by Kalotermitidae (fig. 1). The most ancient dispersal event was that of *Bifiditermes* + *Epicalotermes* + *Roisinitermes* + *Allotermes* (or possibly of their common ancestors with *Incisitermes* A, *Pterotermes*, and *Marginitermes*), which colonized the Afrotropical realm through long-distance overwater dispersal from

the Neotropical or Nearctic realms 51–55 Ma (46–61 Ma combined 95% HPD, nodes 2 and 12 in fig. 1). Two more long-distance overwater dispersals to the Afrotropical realm involved *Glyptotermes* and *Neotermes*, the former originated from the Oriental or Neotropical realm and dispersed 29–41 Ma (23–46 Ma combined 95% HPD, nodes 13 and 14 in fig. 1), whereas the latter originated from the Neotropical or Madagascan realm and dispersed 25–35 Ma (18–41 Ma combined 95% HPD, nodes 15 and 16 in fig. 1). Lastly, *C. havilandi* is an invasive species believed to be native to West Africa (Evans et al. 2013), in which case its lineage colonized the Afrotropical realm from the Oriental realm 12 Ma (9–16 Ma 95% HPD, node 26 in fig. 1). Alternatively, *C. havilandi* was introduced to the Afrotropical realm by human activities from an unknown location. Therefore, the Afrotropics might have been naturally colonized by Kalotermitidae only thrice independently. This could reflect the abundance of higher termites, whose center of origin is the Afrotropical realm (Bourguignon et al. 2017), and which perhaps hindered the colonization of the realm by kalotermitids.

Our sampling of the Oceanian, Sino-Japanese, Saharo-Arabian, and Palearctic realms was limited, although it somewhat reflects the low termite diversity in these realms. Most notably, the Saharo-Arabian realm is inhabited by one of the earliest-branching kalotermitid lineages, represented in this study by *Longicaputermes sinaicus* and an unidentified species from which we only collected immature individuals. Therefore, despite its low termite diversity, the Saharo-Arabian realm is either a potential center of kalotermitid origin or was invaded by early-diverging kalotermitids. The modern distributions of organisms do not necessarily reflect the ancient distribution of a clade, as evidenced by many fossils that demonstrate the historical ancestral distribution of a clade was elsewhere from surviving species of that same clade (e.g., honey bees: Engel, Hinojosa-Diaz, et al. 2009; Kotthoff et al. 2013; megalyrid wasps: Vilhelmsen et al. 2010; thorny lacewings: Nakamine et al. 2020; mastotermitid termites: Krishna et al. 2013; and many other examples in Grimaldi and Engel 2005). Thus, additional sampling from the sparsely surveyed Saharo-Arabian realm is necessary to resolve the early historical biogeography of kalotermitids since it clearly includes relict species. Our ancestral state reconstructions indicate that the Oceanian realm was colonized at least thrice: once from the Neotropics by a lineage survived by modern *Procryptotermes speiseri* 25 Ma (21–30 Ma 95% HPD, node 27 in fig. 1), once from the Oriental realm by *Glyptotermes* 28 Ma (23–33 Ma 95% HPD, node 28 in fig. 1), and once by a lineage survived by *Cryptotermes penaoru* 17 Ma (13–20 Ma 95% HPD, node 29 in fig. 1) from the Australian realm (fig. 1). The Sino-Japanese realm was colonized from the Oriental region by the lineages including *Glyptotermes fuscus* 24 Ma (18–29 Ma 95% HPD, node 30 in fig. 1) and by the lineage including *G. satsumensis*, whose phylogenetic position was unresolved, but which likely originated from the Oriental or Australian realm 15–17 Ma

(11–17 Ma combined 95% HPD, nodes 22 and 31 in fig. 1). Additional sampling from these poorly sampled realms is required to draw a more accurate picture of these aspects of kalotermitid historical biogeography.

The Palearctic realm only hosts 18 described species (Krishna et al. 2013), but its rich fossil record indicates kalotermitid presence by at least the Early Eocene (supplementary fig. S4, Supplementary Material online). The absence of early-branching Palearctic kalotermitid lineages other than *Kalotermes* in our sampling (figs. 1 and 2) either reflects an undersampling of extant Palearctic species or, more likely, the extinction of lineages once abundant during the past global climatic optimum.

We inferred 6–7 dispersal events to the Nearctic realm, six from the Neotropics (fig. 1), which are consistent with the geographic vicinity of both realms since at least the Paleocene, ~66–56 Ma (Buchs et al. 2010). The most recent colonization event of the Nearctic realm by a Neotropical kalotermitid was that of *C. cavifrons* 1.0 Ma (0.5–1.5 Ma 95% HPD, node 32 in fig. 1). The five other dispersal events, involving lineages including *I. snyderi*, *Neotermes luykxi*, *Pterotermes* + *Incisitermes* group A, *Calcaritermes nearticus*, and *Paraneotermes simplicornis*, took place between 15 Ma (11–19 Ma 95% HPD, node 33 in fig. 1) and 58 Ma (45–68 Ma 95% HPD, node 34 in fig. 1), and therefore involved overwater dispersals. *Marginitermes* either diversified within the Nearctic realm from a common ancestor with *Pterotermes* and *Incisitermes* group A, or it dispersed from the Neotropical realm, thus representing a possible seventh dispersal to the Nearctic realm, the age of which is uncertain (nodes 2, 10, 11, and 12 in fig. 1). The origin of Nearctic *Kalotermes* is challenging to trace. *Kalotermes* either dispersed to the Nearctic realm from an unidentified location, the lineage has experienced considerable extinction, or the Nearctic realm was the center of kalotermitid origin. If the last hypothesis were the case, then the presence of the lineage that gave rise to *Kalotermes* in the Nearctic realm dates back to the earliest history of Kalotermitidae. Note that this scenario received a low probability by our ancestral range reconstructions (fig. 1). The modern kalotermitid fauna of the Nearctic realm is overwhelmingly of Neotropical origin.

Reconciliation of Extant and Extinct Kalotermitid Evidence

Molecular phylogenies are reconstructed using extant taxa or, rarely, using well-preserved fossils younger than ~1 Ma (van der Valk et al. 2021). The exclusion of extinct species can lead to spurious historical biogeographic reconstructions (Lieberman 2002). Although fossil evidence has been integrated into biogeographic reconstructions inferred from molecular phylogenies by previous studies (e.g., Meseguer et al. 2015), these methods are hampered by an incomplete fossil record and uncertain taxonomic affiliations, as is the case of Kalotermitidae. Kalotermitids are represented by 74 fossils belonging to 12 extinct and six

extant genera in the paleontological database PaleoDB (supplementary fig. S4, Supplementary Material online). Despite the paucity of available fossils, the age and biogeographic origin of kalotermitid fossils are somewhat consistent with our phylogenetic inferences, although the earliest kalotermitids are from the Oriental region. The available fossil record suggests a more intricate historical biogeography for Kalotermitidae, characterized by considerable past extinctions, consistent with the fact that most organisms that have ever lived are today extinct, and so historical biogeographic patterns must always be tempered by a consideration of significant extinction. Extinction events are at minimum evidenced by at least four groups of Kalotermitidae: (1) *Electrotermes* is the first unambiguous kalotermitid fossil from the Palearctic realm represented by ~49–41 million-year-old fossils (Nel and Bourguet 2006), and predating our oldest estimates of Palearctic kalotermitids (fig. 1, supplementary fig. S12, Supplementary Material online); (2) *Huguenotermes*, an extinct genus allied to modern *Cryptotermes*, which appeared in the Palearctic at least by 35 Ma (Engel and Nel 2015), whereas there are no modern native species of *Cryptotermes* or related genera in the Palearctic realm; (3) a wholly unique fauna of kalotermitids from the Miocene of Zealandia which may have diverged from other Kalotermitidae when Zealandia separated from Australia in the Late Cretaceous, and then become extinct as the landmasses of Zealandia became submerged during the Neogene, necessitating more recent recolonization of modern New Zealand (Engel and Kaulfuss 2017); and (4) the oldest kalotermitid fossils of *Proelectrotermes* (a genus also represented in 41 Ma Baltic amber of the Palearctic realm), *Kachinitermes*, and *Kachinitermopsis* from the ~99 million year old Kachin amber (Engel et al. 2007b), substantially predating our estimates of the oldest Oriental kalotermitids (52 Ma, 47–59 Ma 95% HPD, node 5 in fig. 1). These last three are possible representative of the kalotermitid stem group, but the others are crown-Kalotermitidae. These fossils imply the extinction of ancient kalotermitid faunas, or subsets of those faunas, in the Palearctic and Oriental realms, which, in the case of at least *Proelectrotermes*, *Kachinitermes*, and *Kachinitermopsis*, involved the extinction of stem-kalotermitids (Engel, Grimaldi, et al. 2009, 2016). Furthermore, extinct representatives of many other kalotermitid genera (Krishna et al. 2013; supplementary fig. S12, Supplementary Material online) provide a small glimpse of localized extinctions throughout the phylogeny of Kalotermitidae.

Dispersal Potential and Invasive Traits

Herein, we explored kalotermitid phylogeny based on the most comprehensive taxonomic sampling yet undertaken, including most major lineages sampled across their distribution. We did not attempt to include samples of invasive species from across their introduced range and, therefore, many recent human-assisted kalotermitid dispersals are not represented in our ancestral range reconstructions

(Evans et al. 2013). Altogether, we inferred ∼40 kalotermitid dispersals among biogeographic realms, most of which were transoceanic dispersals. In comparison, our previous study on the historical biogeography of Termitidae based on a representative taxonomic sampling of 415 species (∼18% of termitid species diversity) revealed about 30 dispersals among realms (Bourguignon et al. 2017). Our previous phylogeny of the clade *Heterotermes* + *Coptotermes* + *Reticulitermes*, based on 37 of the 258 extant described species, inferred 14 dispersals among biogeographic realms (Bourguignon, Lo, et al. 2016), and our phylogeny of Rhinotermitinae, based on 27 of 66 described species, retrieved four dispersals among biogeographic realms (Wang et al. 2019). The high number of overwater dispersals inferred for kalotermitids is consistent with their lifestyle, suitable for transoceanic voyages in rafting wood.

Not all branches of the kalotermitid tree had comparable dispersal success. One branch with high-dispersal success is that of *Cryptotermes* (fig. 1), a genus that also includes some of the most successful invasive termite species and the most successful invasive species of Kalotermitidae (Evans et al. 2013). *Cryptotermes* includes five invasive species, *C. havilandi*, *C. cynocephalus*, *C. dudleyi*, *C. brevis*, and *C. domesticus*, making it one of the most cosmopolitan termite genera (Evans et al. 2013). Species of *Cryptotermes* possess three biological traits that enhanced their invasiveness and have allowed them to disperse around the world with human assistance: they are wood-feeders, nesting in wood pieces, and easily produce secondary reproductives (Evans et al. 2013). In addition, *C. domesticus* is tolerant to saline water, a biological trait that presumably increases its ability to survive long-distance oceanic dispersal in wood pieces and possibly human-assisted journeys (Chiu et al. 2021). Additional life-history traits might contribute to the invasiveness of *Cryptotermes*, but the physiology and ecology of Kalotermitidae have not been fully elucidated, preventing us from comprehending what may also aid their suitability for such passive dispersal. This question could be addressed using a taxonomically representative sampling of kalotermitids. The phylogenies of Kalotermitidae reconstructed here provide a framework to guide such sampling.

Evolution of Social Organization

Most extant species of Kalotermitidae are unable to forage outside their nests. This life trait has been alternatively considered as ancestral for termites (Abe 1987; Thorne 1997; Inward et al. 2007) or as a derived condition stemming from the secondary loss of foraging abilities in some lineages such as kalotermitids (Morgan 1959; Watson and Sewell 1985; Bordereau and Pasteels 2011; Bourguignon, Chisholm, et al. 2016; Mizumoto and Bourguignon 2020). Our taxonomically representative phylogeny of Kalotermitidae represents a starting point to disentangle the evolution of social organization in Kalotermitidae.

Due to the scarcity of behavioral observations, we did not attempt to reconstruct the ancestral behavior of Kalotermitidae quantitatively. However, some pieces of information on the behavior of Kalotermitidae are available for several species. Some Kalotermitidae are known to have the ability to build underground tunnels and shelter tubes covering foraging trails. This is notably the case of the two early-branching kalotermitid lineages (fig. 3, supplementary table S3, Supplementary Material online). Indeed, *Longicaputermes* can build shelter tubes, and *Paraneotermes* build underground tunnels to colonize new wood resources (Light 1937) (fig. 3). These foraging abilities are not unique to early-branching Kalotermitidae and have also been observed in other kalotermitids. For example, we observed *C. brevis* building shelter tubes under laboratory conditions (fig. 3), and *Neotermes* has been reported to build below-ground tunnels (Thompson 1934; Goetsch 1936; Waterhouse 1993). Moreover, all termite species investigated so far are physiologically equipped for foraging as they produce and follow trail pheromones (Sillam-Dussès et al. 2009). Notably, the chemical composition of trail pheromones in Kalotermitidae is identical to those of extensively foraging fungus-growing termites (Macrotermitinae) (Sillam-Dussès et al. 2009). The occurrence of foraging behavior in early-branching kalotermitid lineages and in species spread across the kalotermitid phylogeny initially suggests that the last common ancestor of Kalotermitidae was capable of foraging among multiple wood pieces.

Although several species of Kalotermitidae have foraging abilities, especially those belonging to early-branching lineages, most modern species of Kalotermitidae appear to lack foraging abilities or exhibit such behaviors only under specific circumstances, as in laboratory colonies. Such facultative foraging would be consistent with an ancestral presence of foraging ability that is not utilized in most lineages under natural environments, perhaps only expressing such behaviors under extreme conditions, much like facultative sociality in halictine bees (Eickwort et al. 1996). We hypothesize that the shift in kalotermitid social organization and foraging abilities was driven by competition. In particular, two lineages of Neoisoptera, *Heterotermes* + *Coptotermes* + *Reticulitermes* and the higher termites (Termitidae), are major competitors of kalotermitids owing to the overlap of their niches, high abundance, mutually high aggression, and the existence of adaptive avoidance mechanisms (Thorne and Haverty 1991; Evans et al. 2009). Biotic factors, such as species competition, can shape species spatial distribution at a global scale (Wisz et al. 2013). Kalotermitids expressing foraging abilities often inhabit biomes hosting few or no other termite species, such as arid regions (*Longicaputermes*, *Paraneotermes*, and *Neotermes chilensis*), remote islands (*Neotermes rainbowii*) (Waterhouse 1993), and high-elevation mountains (*Postelectrotermes militaris* and *Comatermes perfectus*) (Hemachandra et al. 2014; Scheffrahn 2014; Gnanapragasam 2018). This geographical

distribution is consistent with foraging kalotermitids being competitively excluded from regions where wood-feeding Neoisoptera are abundant, as well as with the notion that expression of foraging is only under atypical environmental conditions. As noted, kalotermitids exhibit behavioral flexibility with respect to foraging behavior, as demonstrated by *Cryptotermes*, which is capable of building galleries in laboratory conditions (fig. 3). Foraging in Kalotermitidae is likely a complex interplay between extreme conditions, local competition, and perhaps yet unidentified factors. It is unclear to what extent the foraging behavior of kalotermitids could decrease their fitness in biomes occupied by competing wood-feeding Neoisoptera. Kalotermitidae are generally rare at the ground level of tropical forests (e.g., Eggleton et al. 1996), although their abundance and diversity might be underestimated owing to their ability to live in microhabitats hardly accessible to Neoisoptera and termitologists, such as very hard wood or trees in partially flooded forests (authors' personal observations). The scarcity of Kalotermitidae in tropical forests indicates that single-piece nesting is not an efficient strategy to compete against wood-feeding Neoisoptera. Ants represent another hyper-diverse and abundant organismal group that interacts with termites, primarily through predation (Wilson 1971). Ant predation thus likely represents another factor favoring confinement of termites into a single piece of wood. Additional work is needed to quantify competition among termites and reconcile kalotermitid historical biogeography with that of their presumed wood-feeding termite competitors and ant predators.

Conclusion

We present the most comprehensive molecular phylogenetic reconstruction of Kalotermitidae, based on a taxonomic sampling of global kalotermitid diversity and distribution. This study is also the first to provide an estimated timeline of kalotermitid evolution and to reconstruct the historical biogeographic events across the family. Our ancestral range reconstructions imply that the Neotropical realm is a cradle for crown-group kalotermitid diversity, with the exception of the two earliest-diverging lineages and putative stem-group fossils from the Oriental region. We show that the distribution of modern Kalotermitidae includes upwards of 40 disjunctions between biogeographic realms, most of which can be explained by long-distance transoceanic dispersal events. Although the single-piece nesting lifestyle has been sometimes considered as representative of the proto-termite colony, or at least ancestral to Kalotermitidae, we found that many kalotermitids, including early-diverging species-poor genera, are capable of foraging outside their nesting piece of wood. These results suggest that the social organization of the colonies of the common ancestor to all modern Kalotermitidae was more complex than previously appreciated, and perhaps includes some degree of facultative expression. We postulate that single-piece nesting is

a derived trait that evolved as a specialization to minimize niche overlap with competitors, such as species of Neoisoptera. Finally, our molecular phylogeny reveals that several kalotermitid genera are nonmonophyletic and provides a framework for a future taxonomic revision of the family.

Materials and Methods

Sample Collection
The termite samples used in this study were collected during the last three decades by the authors. Two sampling procedures were employed. Samples were either collected in 80% ethanol and stored at room temperature or collected in RNA-later® and stored at −80 °C until DNA extraction.

Mitochondrial Genome Sequencing, Assembly, and Annotation
We used different approaches for the samples preserved in RNA-later® and 80% ethanol. For the 108 samples preserved in RNA-later®, we extracted total genomic DNA from one to ten termite immature (pseudergate) heads using the DNeasy Blood & Tissue Kit (Qiagen). The total extracted DNA was used to amplify mitochondrial genomes with TaKaRa LA Taq in two long-range PCR reactions. The primers and conditions of the reactions were as described by Bourguignon et al. (2015). The concentration of the two long-range PCR amplicons was measured with Qubit 3.0 fluorometer, and both amplicons were mixed in equimolar concentration. One DNA library was then prepared for each sample separately. Libraries were either prepared in-house, using unique dual index barcodes, or by a sequencing company, using a unique combination of index barcodes. In-house libraries were prepared with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs) using one-fifteenth of the reagent volume indicated in the manufacturer protocol. Libraries were pooled and paired-end sequenced on the Illumina HiSeq platform or the Illumina MiSeq platform. For the 56 samples preserved in 80% ethanol, total genomic DNA was extracted from entire pseudergate bodies using the DNeasy Blood & Tissue Kit (Qiagen). Libraries were prepared for shotgun sequencing, directly from whole pseudergate body DNA, without any long-range PCR step. Because the DNA extracted from these samples was generally highly fragmented, we prepared the libraries using a modified protocol. Libraries were prepared with the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs) without an enzymatic fragmentation step. In addition, library preparation was downsized to one-fifteenth of the volume indicated in the manufacturer protocol for all reagents. Other steps were performed as suggested by the manufacturer protocol. One library with unique dual index barcodes was prepared for each sample separately. A total of 3 µl of each library was pooled and sequenced on the Illumina HiSeq X platform.

We obtained an average of 52 megabases and 2 gigabases of raw sequencing data for samples sequenced with the amplicon-based approach and samples sequenced with the shotgun sequencing approach, respectively. Raw sequencing data was quality-checked with FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimmed and quality-filtered with bbduk (sourceforge.net/projects/bbmap/) using the following parameters: "ktrim = r k = 23 mink = 7 hdist = 1 tpe tbo maq = 10 qtrim = rl trimq = 15 minlength = 35." Multiple mitochondrial genome assembly approaches were used as described below, and the longest mitochondrial genome contigs or scaffolds were selected for each sample. The mitochondrial genomes obtained through long-range PCR amplifications were assembled using SPAdes with default parameters (Bankevich et al. 2012). The mitochondrial genomes obtained from shotgun sequencing data were assembled with one of the three following methods: (1) metaSPAdes with default parameters (Nurk et al. 2017); (2) metaSPAdes with default parameters followed by the iterative mapping-reassembling approach implemented in TCSF and IMRA (Kinjo et al. 2015) and using the mitochondrial genome of *C. brevis* as a reference; (3) Novoplasty organelle genome assembler (Dierckxsens et al. 2016) with default parameters and using *C. brevis* mitochondrial genome as a seed sequence. The mitochondrial genomes assembled in multiple high-coverage contigs were merged using published kalotermitid mitochondrial genomes as references. Assembly gaps were filled with the symbol N (supplementary table S2, Supplementary Material online). All mitochondrial genomes were cyclically shifted using MARS (Ayad and Pissis 2017) to start with the tRNA gene coding for isoleucine and end with the 12S rRNA gene. Each mitochondrial genome linearized with MARS was aligned with MAFFT v.7.475 (Katoh and Standley 2013). The aligned draft mitochondrial genomes were visually inspected for the presence of 5′- and 3′-terminal assembly artifacts, including low-quality stretches of terminal nucleotides and duplicated sequences present at both 5′-end and 3′-end of the original contigs that appeared as sequence insertions or low-sequence-similarity regions in the alignments. All detected artifacts were removed. The contiguity of the mitochondrial genome assembly was confirmed by remapping of sequence reads to the mitochondrial genome assembly using bowtie2 v2.4.2 (Langmead and Salzberg 2012) and visual inspection using Integrative Genomics Viewer v2.11.0 (Robinson et al. 2011). We also removed the noncoding portion of mitochondrial genomes between 16S ribosomal RNA and tRNA-Ile, which consists mostly of the control region, since the sequence repeats in the control region generally lead to artifacts during short-read de novo assembly. The resulting trimmed mitochondrial genomes were annotated with MitoZ (Meng et al. 2019). The genome annotations of previously published mitochondrial genomes were retrieved from GenBank. MitoZ annotations were visually inspected by first aligning mitochondrial genomes with MAFFT as described above. The alignments and the MitoZ-generated feature annotations were inspected with JalView (Clamp et al. 2004). The errors in MitoZ-generated annotations, mostly comprising of omitted tRNA-Val and 5′-terminally truncated 16S rRNA gene, were corrected manually. The final mitochondrial genomes and their annotations are available on GenBank (supplementary table S1, Supplementary Material online). The final mitochondrial genome sequence dataset consisted of 164 mitochondrial genomes. We also used previously published sequences, including the mitochondrial genomes of eight species of Kalotermitidae (Cameron et al. 2012; Bourguignon et al. 2015) (supplementary table S2, Supplementary Material online) and 58 mitochondrial genomes of nonkalotermitid termites that we used as outgroups (Bourguignon et al. 2015). The mitochondrial genome sequence dataset used for comparison with the UCE dataset included an additional ten publicly available nonkalotermitid termite mitochondrial sequences (supplementary table S4, Supplementary Material online).

The rRNA and tRNA sequences were aligned using MAFFT v.7.475 (Katoh and Standley 2013). For protein-coding genes, we first translated the nucleotide sequences into amino acid sequences with the transeq command of the EMBOSS suite of programs (Rice et al. 2000) using the invertebrate mitochondrial genetic code. The protein sequences were then aligned using MAFFT v.7.475. The protein sequences were either directly concatenated with the noncoding nucleotide sequences or were first back-translated into nucleotide sequence alignments using PAL2NAL v.14 (Suyama et al. 2006) before concatenation with the noncoding nucleotide sequences. All alignments were concatenated with FASconCAT v.1.04 (Kück and Meusemann 2010). We also generated alternative alignments, one with the third codon sites of protein-coding genes and one without the third codon sites of protein-coding genes. An additional mitochondrial genome dataset included the COII sequence of *Ceratokalotermes spoliator* (Thompson et al. 2000) to elucidate its phylogenetic position.

UCE Datasets

UCEs were extracted from metaSPAdes assemblies using the PHYLUCE suite of programs (Faircloth 2016) with the baits and parameters used by Hellemans et al. (2021). Since the coverage of the Kalotermitidae whole-genome shotgun sequence data was not uniform among samples, we selected 18 Kalotermitidae species as a trade-off between maximizing the number of species and genera and maximizing the completeness of the UCE matrices. The final UCE datasets included 28 species (18 Kalotermitidae species and 10 non-Kalotermitidae termite species; see supplementary table S4, Supplementary Material online). All extracted UCEs were deposited in Dryad (https://doi.org/10.5061/dryad.5mkkwh77v; https://github.com/oist/TER-UCE-DB/). Two UCE datasets were

used for phylogenetic analyses. The first dataset had a completeness of 70% and included 691 loci, whereas the second dataset was complete at 75% and included 82 loci. Alignments were produced with MAFFT and trimmed using Gblocks (Castresana 2000), as implemented in the PHYLUCE suite of programs.

Maximum Likelihood Phylogeny

Sequence alignments were used to infer maximum likelihood phylogenies with IQtree v.1.6.7 (Nguyen et al. 2015). IQtree was run with the extended nucleotide model selection, including FreeRate models. The best partitioning scheme was determined from a set of a priori defined partitions using ModelFinder (Kalyaanamoorthy et al. 2017) implemented in IQtree (option "-m MFP + MERGE"). The partition scheme searching was constrained by setting the maximum number of partition pairs during the partition merging phase to 2,000 and considering only the top 10% partition schemes (options "rcluster-max 2000 -rcluster 10 -bnni"). The partitions of the mitochondrial genome datasets were defined a priori assuming differences in evolutionary rates among codon positions of protein-coding genes, tRNAs, and rRNAs. Therefore, we set one partition for each codon position of the 13 mitochondrial protein-coding genes, one partition for the 22 tRNA genes, and one partition for the 12S and 16S rRNA genes. For UCE datasets, each genetic locus was predefined as a separate partition. Branch supports were calculated using ultrafast bootstrapping method (option "-bb 1000") (Minh et al. 2013). The reconstructed trees were rooted using *Mastotermes darwiniensis*, the extant sister lineage of all other termites (Kambhampati et al. 1996). See supplementary table S2, Supplementary Material online for an overview of maximum likelihood analyses.

Bayesian Time-Calibrated Phylogeny

We estimated time-calibrated phylogenies of Kalotermitidae using the Bayesian phylogenetic software BEAST2 v.2.6.2 (Bouckaert et al. 2014). We used the uncorrelated lognormal relaxed clock model to account for substitution rate variation among branches (Drummond et al. 2006). Several alternative datasets were used: (1) a UCE dataset including 28 species, (2) a mitochondrial genome dataset including 28 species and consisting of protein sequence alignments for protein-coding genes and nucleotide sequence alignments for rRNA and tRNA genes, (3) a mitochondrial genome dataset including 230 species and consisting of protein sequence alignments for protein-coding sequences and nucleotide sequence alignments for rRNAs and tRNAs, and (4) a mitochondrial genome dataset including 230 species and consisting of concatenated nucleotide sequence alignments for all genes. The partitioning schemes were determined a priori as described for the maximum likelihood analyses. The nucleotide substitution models were inferred using bModelTest from a set of 30 transition/transversion split models (Bouckaert and Drummond 2017). The empirical mtREV model

implemented in BEAST2 was used for the mitochondrial protein sequences. A Yule speciation model was used as a tree prior. We used 15 fossils as node calibrations (supplementary table S6, Supplementary Material online). The fossils were selected based on the criteria described by Parham et al. (2012) (for justification concerning the choice of fossil calibrations, see details below). The fossil ages were used as minimum age constraints, which we implemented as exponential priors on node ages. For each fossil, we determined a soft maximum bound using the phylogenetic bracketing approach described by Ho and Phillips (2009). The MCMC chains were run three times independently for each dataset. We used between 169×10^6 and 337×10^6 generations sampled every 10,000 generations to build the final maximum clade credibility trees. The convergence of the MCMC chains was inspected with Tracer 1.7.1 (Rambaut et al. 2018). Between 10% and 30% of initial generations were discarded for each MCMC chain following visual inspection of trace plots that ensured MCMC chain convergence. The trees sampled from the MCMC runs were summarized as maximum clade credibility tree with TreeAnnotator v2.6.2 (Bouckaert et al. 2014).

Fossil Calibrations

We used 15 termite fossils as minimal age internal calibrations (supplementary table S5, Supplementary Material online). In all cases, we used the youngest fossil age estimates reported on the Paleobiology Database (https://paleobiodb.org/, accessed November 2021). Note that fossil age estimates can be controversial and alternative age estimates exist (e.g., Kasinski et al. 2020 for hypotheses on Baltic amber age). However, the widths of the 95% HPD intervals produced by our Bayesian time-calibrated phylogenies typically exceed the interval of alternative estimates of fossil ages. Therefore, we presume that our inferences of time-calibrated phylogenies are robust with respect to the use of alternative fossil age estimates. For reproducibility, we adhered here to the ages reported in Paleobiology Database although these might not always reflect the latest opinions on the fossil age. We used the age of *Melqartitermes myrrheus* as the minimal age of all termites (Engel et al. 2007b). *Melqartitermes myrrheus* is part of the *Meiatermes*-grade that comprised Cretaceous termite fossils intercalating between Mastotermitidae and other termite families (Engel, Grimaldi, et al. 2009, 2016). We used *Cosmotermes multus* to calibrate the nodes corresponding to Stolotermitidae + Hodotermitidae + Archotermopsidae. All castes of *Cosmotermes* are known from several aggregations in amber and have similar morphology to modern Stolotermitidae (Zhao et al. 2020). We used *Proelectrotermes swinhoei* to calibrate the Kalotermitidae + Neoisoptera group. *Proelectrotermes*, along with the fossil genera *Electrotermes* and *Prokalotermes*, were separated by Emerson (1942) into a distinct extinct kalotermitid subfamily Electrotermitinae and considered sister to all other Kalotermitidae, principally owing to what he intuited

to be the plesiomorphic retention of tetramerous cerci. As noted by Krishna (1961), these fossils have the typical dimerous cerci of Kalotermitidae and the interpretation of additional cercomeres was an error. The diagnostic characters of the alate imago of *Proelectrotermes* are solely plesiomorphies (Engel, pers. obs.), and therefore *Proelectrotermes* could represent a group outside of crown-Kalotermitidae or could be basal to one or more of the earliest-diverging branches. Indeed, in the phylogenetic analyses of Engel, Hinojosa-Diaz, et al. (2009) and Engel et al. (2016), *Proelectrotermes* was found to be sister to crown-Kalotermitidae, although the available characters were limited. Note that another fossil species from the Burmese amber, *Archeorhinotermes rossi* (Archeorhinotermitidae), has alates with a fontanelle and is, therefore, the earliest known representative of Neoisoptera (Krishna and Grimaldi 2003), the extant sister group of Kalotermitidae, and could be used to calibrate the same node.

The phylogenetic placement of *Electrotermes* was intuited by Krishna (1961) to be near *Postelectrotermes* and *Neotermes* based on an assumed transition series in the reduction of mesotibial spines, which were assumed to be symplesiomorphic, and changes in the sclerotization and position of M-vein in the forewing, two characteristics homoplastic across Kalotermitidae. The mesotibial spines are likely independent in the fossil genera *Electrotermes* and *Proelectrotermes*, as well as modern *Postelectrotermes*, rather than stages of a transition series (Engel, pers. obs.). Indeed, the shortened wing R-vein of *Electrotermes* is more similar to that of African *Glyptotermes* or some *Kalotermes* than to either *Proelectrotermes* or *Postelectrotermes*. Although *Electrotermes* fell outside of crown-Kalotermitidae in recent analyses (Engel, Grimaldi, et al. 2009; Engel et al. 2016), the placement is potentially spurious as the genus lacked most codings for the character-state matrix. *Prokalotermes* is a genus of Kalotermitidae known from one species at the Eocene-Oligocene boundary of Florissant, Colorado, and another from the Early Oligocene (32 Ma) of Montana (Krishna et al. 2013). Although certainly a kalotermitid, the genus remains too poorly understood to provide a more refined phylogenetic placement (Engel, pers. obs.). For these reasons, we did not use *Electrotermes* or *Prokalotermes* as a fossil calibration for crown-Kalotermitidae.

The fauna of kalotermitids from the Early Miocene of New Zealand represents a distinctive fauna in ancient Zealandia (Engel and Kaulfuss 2017). These taxa are poorly understood with unknown affinities to crown-kalotermitids. Other Miocene fossil termites, such as those kalotermitids reported from Shandong, China, are poorly preserved and documented, and of entirely uncertain taxonomic affinities. Such species are best-considered *incertae sedis* and cannot be used as definitive evidence of those genera to which they have historically been assigned. Other fossils of Kalotermitidae include two species in mid-Cretaceous (99 Ma) Kachin amber of the genera

Kachinitermes and *Kachinitermopsis* (Engel et al. 2007b; Engel and Delclòs 2010). Both genera, even more so than *Proelectrotermes*, exhibit only plesiomorphies for Kalotermitidae and are likely stem groups. However, both are known from only fragmentary remains and lack a considerable number of critical traits that would help to better refine their phylogenetic affinities. Thus, both have to be considered as *incertae sedis* among stem-group Kalotermitidae. Trace fossils of kalotermitid nests are known from the Cretaceous through the Miocene, but aside from indicating the presence of kalotermitids with their distinctive frass, nothing more can be determined of the particular groups of Kalotermitidae that produced these ichnological remains.

We used the age of *Nanotermes isaacae* as a minimal constraint to calibrate the node corresponding to Termitidae + sister group (*Coptotermes* + *Heterotermes* + *Reticulitermes*). *N. isaacae* is the oldest known Termitidae, with unclear affinities to modern subfamilies, but is clearly affiliated to Termitidae (Engel et al. 2011). Eight of the 11 remaining fossils were used to calibrate internal nodes of Neoisoptera, including three nodes of Rhinotermitidae and five nodes of Termitidae. Within the Rhinotermitidae, *Reticulitermes antiquus* is common in the Baltic amber and is clearly assigned to *Reticulitermes* (Engel et al. 2007a), probably representing a stem group *Reticulitermes*, and we therefore used this fossil to calibrate the node corresponding to *Reticulitermes* + sister group. *Coptotermes sucineus* is known from Mexican amber (Emerson 1971) and was found in the same amber piece with *Heterotermes*, the paraphyletic genus within which *Coptotermes* is nested (Bourguignon, Lo, et al. 2016; Buček et al. 2019). We therefore used the fossil of *C. sucineus* to calibrate *Heterotermes* + *Coptotermes*. Finally, *Dolichorhinotermes dominicanus*, a clear member of the Rhinotermes-complex (Schlemmermeyer and Cancello 2000), was used to calibrate *Dolichorhinotermes* + *Schedorhinotermes*. Note that additional species of *Dolichorhinotermes* are known from Mexican amber (Engel and Krishna 2007b). Within Termitidae, *Macrotermes pristinus*, described by Charpentier (1843) as *Termes pristinus*, and assigned to *Macrotermes* by Snyder (1949), was used as a calibration for the node *Macrotermes* + *Odontotermes* + *Synacanthotermes*. The four other fossils of Termitidae were described from Dominican amber: *Constrictotermes electroconstrictus*, *Microcerotermes insulanus*, *Amitermes lucidus*, and *Anoplotermes* sensu lato. The first three fossil species belong to modern genera (Krishna 1996; Krishna and Grimaldi 2009), and were used as minimal age calibrations for the nodes corresponding to their respective genera + sister groups. Although the 11 species of the *Anoplotermes*-group described from Dominican amber were assigned to the genus *Anoplotermes* (Krishna and Grimaldi 2009), they likely belong to different genera within the *Anoplotermes*-group. The genus *Anoplotermes* is presently a polyphyletic assemblage belonging to the South American *Anoplotermes*-group (Bourguignon, Šobotník, et al. 2016). The group is in great need of a comprehensive

revision, which implies that fossil species of *Anoplotermes* cannot be precisely assigned to a modern lineage within the *Anoplotermes*-group, and are better referred to as *Anoplotermes* sensu lato. We used these fossils to calibrate all soldierless Apicotermitinae. The last three fossils were used to calibrate the internal nodes of Kalotermitidae. *Huguenotermes septimaniensis* is known from a wing impression found in French sediments dated from the late Eocene. The wing fossil presents plesiomorphies of the clade comprising *Cryptotermes* and *Procryptotermes* (Engel and Nel 2015) and was used to calibrate *Cryptotermes* + *Procryptotermes* + sister group. *Glyptotermes grimaldii* was used as a calibration for the node including all species of *Glyptotermes* sensu stricto. *Glyptotermes grimaldii* was found in Dominican amber and closely resembles one extant sympatric species, *Glyptotermes pubescens* (Engel and Krishna 2007a). For this reason, we considered *G. grimaldii* as a crown-*Glyptotermes* and used it to calibrate the node that contains all species of *Glyptotermes*. The last fossil we used, *Calcaritermes vetus*, was assigned to *Calcaritermes* by Emerson (1969), who stated that this assignment is probable. We considered *C. vetus* as a stem-*Calcaritermes* and used it to calibrate *Calcaritermes* + sister clade.

We used the age of five termite fossils as 97.5% soft maximum bounds. The age of one of these five fossils was used as soft maximum bounds for each 15 nodes with fossil calibrations (supplementary table S6, Supplementary Material online). The maximum soft bounds were assigned using a combination of phylogenetic bracketing and absence of fossil evidence (Ho and Phillips 2009). For the nodes calibrated with *M. myrrheus*, *C. multus*, and *P. swinhoei*, the three oldest calibrations used in this study, we used as maximum soft bound *Triassoblatta argentina*, the first representative of Mesoblattinidae (Martins-Neto et al. 2005). We used the first fossil of Neoisoptera, *A. rossi* (Krishna and Grimaldi 2003), as maximum soft bound for the nodes calibrated with the two rhinotermitid fossils *R. antiquus* and *D. dominicanus*, and for the node calibrated with *N. isaacae*, the oldest known fossil of termitids. *P. swinhoei*, one of the oldest undisputed fossils of Kalotermitidae found in Kachin amber, was used as a maximum soft bound for the nodes calibrated with the three crown-kalotermitid fossils, *H. septimaniensis*, *C. vetus*, and *G. grimaldii* (Emerson 1971). The oldest known fossil of *Heterotermes*, *Heterotermes eocenicus* (Engel 2008), was used as a maximum soft bound for the nodes calibrated with *C. sucineus*. This is justified by the paraphyly of *Heterotermes* to *Coptotermes* (Bourguignon, Lo, et al. 2016; Buček et al. 2019). Finally, we used the age of *N. isaacae* as a maximum soft bound for all the nodes calibrated with the crown-termitid fossils, *A. lucidus*, *M. insularis*, *Anoplotermes* sensu lato, *C. electroconstrictus*, and *M. pristinus*.

We excluded several fossils because of their uncertain or revised taxonomic assignments. *Cratokalotermes santanensis* (PaleoDB occurrence numbers 1024939 and 1108956) was included in the Kalotermitidae by Krishna et al. (2013), but Grimaldi et al. (2008) and the analysis of Engel et al. (2016) demonstrated that it is not a

kalotermitid, instead falling within the *Meiatermes* grade of Cretaceous Isoptera.

Eotermes was considered not to be a kalotermitid by Krishna (1961). Several Nearctic fossil specimens were assigned to *Cryptotermes* (PaleoDB collection numbers 930175 and 1134865) but lack formal description (Park and Downing 2001), which precluded their use as fossil calibrations. We did not use *Neotermes grassei* (PaleoDB collection number 995549), from Eocene deposit in France, because its assignment to *Neotermes* was controversial and not accepted by some authors (Nel and Paicheler 1993; Krishna et al. 2013). In addition, our phylogenetic analyses showed that the genus *Neotermes* is a polyphyletic assemblage, hampering the use of fossils of *Neotermes* as calibrations. Lastly, *Kalotermes piacentinii* (PaleoDB collection number 120351), originally described as a termite fossil (Piton and Théobald 1937), is no longer considered to be a termite fossil (Nel and Paicheler 1993). See supplementary table S6, Supplementary Material online for an overview of the Kalotermitidae fossils retrieved from the Paleobiology Database.

Summary Evidence Time-Calibrated Phylogeny

A summary evidence time-calibrated phylogeny (SE-tree) was generated by summarizing all phylogenies onto the time-calibrated tree reconstructed using concatenated protein alignments of mitochondrial protein-coding genes and nucleotide alignments of mitochondrial tRNA and rRNA genes (supplementary table S2, Supplementary Material online). Conflicts and supports among trees were summarized using TreeGraph v2.15.0-887 (Stöver and Müller 2010). Prior to importing maximum likelihood trees into Treegraph, they were rerooted with Dendroscope using *M. darwiniensis* as an outgroup. The SE-tree was further summarized to create a reduced SE-tree (RSE-tree) by pruning 35 tips that diverged <1 Ma and were collected in the same country. The genus-level tree with polytomies displayed in figure 3 was generated by collapsing genus-level clades and using polytomies for genus-level branches having topological conflicts in the RSE-tree. The RASP probabilities of ancestral ranges inferred with RSE-tree topology for each of the four models were averaged (fig. 1, supplementary figs. S4–S7, Supplementary Material online). Final trees were plotted in R programming language.

Historical Biogeography

We reconstructed the historical biogeography of Kalotermitidae using the Bayesian Binary MCMC analysis (Ronquist and Huelsenbeck 2003) implemented in RASP4.2 (Yu et al. 2015, 2020). The maximum number of areas allowed for each node was set to 1. We ran the model on the two maximum likelihood trees and the two Bayesian time-calibrated trees generated in this study. Each tip was assigned to the biogeographic realm corresponding to its sampling location. We recognized ten biogeographic realms. Nine realms were as defined by Holt

et al. (2013): Afrotropical, Australian, Madagascan, Nearctic, Oceanian, Oriental, Palearctic, Saharo-Arabian, and Sino-Japanese. For the tenth realm, the Neotropical realm, we used the original definition of Wallace (1876) instead of that of Holt et al. (2013), which divided it into two realms, the Panamian and Neotropical realms. We justify the use of the Neotropical realm *sensu* Wallace (1876) by our observations of termite fauna that do not substantiate its division into two entities (authors' personal observation). The large-scale geographical division of Earth terrestrial biota of Holt et al. (2013) has several limitations: (1) it was inferred from species distribution and phylogenetic relationships of vertebrates rather than insects and, thus, might not unconditionally apply to the historical biogeography of insects; and, (2) during the ∼90 million years of Kalotermitidae evolutionary history, the geographical delimitation of each realm likely underwent substantial changes. However, comparable large-scale biogeographic divisions proved useful in the description of termite natural history (Bourguignon et al. 2017). The analyses were run on each tree with four models: JC (state frequencies set to fixed and among site variation set to equal), JC + GAMMA (state frequencies set to fixed and among site variation set to gamma), F81 (state frequencies set to estimated and among site variation set to equal), and F81 + GAMMA (state frequencies set to estimated and among site variation set to gamma). Default values were used for other parameters. Therefore, we ran the Bayesian Binary MCMC analyses four times for every four trees, which make a total of 16 ancestral range reconstructions of Kalotermitidae. Ancestral ranges reconstructed for each of four alternative tree topologies were averaged across the models used and were mapped onto the trees (supplementary figs. S4–S7, Supplementary Material online).

The sampling was unequal among biogeographic realms, especially for the Neotropical realm, which was represented by 45 tips in the RSE-tree. The Neotropical realm was reconstructed as the center of origin for many deep nodes in the phylogeny of Kalotermitidae, possibly reflecting a sampling bias toward Neotropical Kalotermitidae. To rule out an effect of this sampling bias on our ancestral range reconstructions, we performed additional RASP analyses using phylogenies of Kalotermitidae for which Neotropical species were randomly subsampled to 22 and 15 tips. The random subsampling was performed thrice for a total of six RASP analyses. RASP was run with the JC model. The node probabilities were plotted onto SE-tree and RSE-tree as pie charts using R. A biogeographic realm map was plotted in R using map data downloaded from https://macroecology.ku.dk/resources/wallace/cmec_regions___realms.zip (Holt et al. 2013).

Since the oldest fossils of kalotermitids were discovered in the Oriental realm, we tested the hypothesis of Oriental distribution of crown-Kalotermitidae by constraining the range of the last common ancestor of extant kalotermitids to include Oriental realm. The analysis was performed using R package BioGeoBEARS v1.1.2 (Matzke 2013) with maximum range size set to 3 and using a DEC model (Ree and Smith 2008).

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

Data Availability

Additional data including all trees generated in this study in Newick tree format, R scripts and input data used to generate figures in this article are available online at https://github.com/AlesBucek/Kalotermitidae_biogeography.

References

Abe T. 1984. Colonization of the Krakatau Islands by termites (Insecta: Isoptera). *Physiol Ecol Japan* **21**:63–88.

Abe T. 1987. Evolution of life types in termites. In: Kawano S, Connell J, Hidaka T, editors. *Evolution and coadaptation in biotic communities*. Tokyo: University of Tokyo Press. p. 125–148.

Ayad LAK, Pissis SP. 2017. MARS: improving multiple circular sequence alignment using refined sequences. *BMC Genomics* **18**:86.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Alexander S, Lesin VM, Nikolenko SI, Pham S, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* **19**:455–477.

Barden P, Engel MS. 2021. Fossil social insects. In: Starr CK, editor. *Encyclopedia of social insects*. Cham: Springer International Publishing. p. 384–403.

Bordereau C, Pasteels JM. 2011. Pheromones and chemical ecology of dispersal and foraging in termites. In: Bignell DE, Roisin Y, Lo N, editors. *Biology of termites: a modern synthesis*. Dordrecht: Springer SBM. p. 279–320.

Bouckaert RR, Drummond AJ. 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evol Biol.* **17**:1–11.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol.* **10**: e1003537.

Bourguignon T, Chisholm RA, Evans TA. 2016. The termite worker phenotype evolved as a dispersal strategy for fertile wingless individuals before eusociality. *Am Nat.* **187**:372–387.

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

Bourguignon T, Lo N, Šobotník J, Ho SYW, Iqbal N, Coissac E, Lee M, Jendryka MM, Sillam-Dusses D, Křížková B, 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, Lo N, Šobotník J, Sillam-Dussès D, Roisin Y, Evans TA. 2016. Oceanic dispersal, vicariance and human introduction shaped the modern distribution of the termites *Reticulitermes*, *Heterotermes* and *Coptotermes*. *Proc R Soc B Biol Sci.* **283**: 20160179.

Bourguignon T, Šobotník J, Dahlsjö CAL, Roisin Y. 2016. The soldierless Apicotermitinae: insights into a poorly known and ecologically dominant tropical taxon. *Insect Soc.* **63**:39–50.

Bucek A, Šobotník J, He S, Shi M, McMahon DP, Holmes EC, Roisin Y, Lo N, Bourguignon T. 2019. Evolution of termite symbiosis informed by transcriptome-based phylogenies. *Curr Biol.* **29**: 3728–3734.e4.

Buchs DM, Arculus RJ, Baumgartner PO, Baumgartner-Mora C, Ulianov A. 2010. Late Cretaceous arc development on the SW margin of the Caribbean Plate: insights from the Golfito, Costa Rica, and Azuero, Panama, complexes. *Geochem Geophys Geosyst.* **11**:Q07S24.

Cameron SL, Lo N, Bourguignon T, Svenson GJ, Evans TA. 2012. A mitochondrial genome phylogeny of termites (Blattodea: Termitoidae): robust support for interfamilial relationships and molecular synapomorphies define major clades. *Mol Phylogenet Evol.* **65**:163–173.

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

Charpentier T. 1843. Über einige fossile Insecten aus Radoboj in Croatien. *Novorum Actorum Acad. Caesareae Leopoldino-Carolinae Naturae Curiosorum* **20**:399–410,Pl. XXI–XXIII.

Chiu C, Mullins AJ, Kuan K, Lin M, Su N, Li H. 2021. Termite salinity tolerance and potential for transoceanic dispersal through rafting. *Ecol Entomol.* **46**:106–116.

Clamp M, Cuff J, Searle SM, Barton GJ. 2004. The Jalview Java alignment editor. *Bioinformatics* **20**:426–427.

Dawson MR, West RM, Langston W, Hutchison JH. 1976. Paleogene terrestrial vertebrates: northernmost occurrence, Ellesmere Island, Canada. *Science* **192**:781–782.

Dierckxsens N, Mardulyn P, Smits G. 2016. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **45**:e18.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**:e88.

Eggleton P, Bignell DE, Sands WA, Mawdsley NA, Lawton JH, Wood TG, Bignell NC. 1996. The diversity, abundance and biomass of termites under differing levels of disturbance in the Mbalmayo Forest Reserve, southern Cameroon. *Philos Trans R Soc London Ser B Biol Sci.* **351**:51–68.

Eggleton P, Davies R. 2003. Isoptera, termites. In: Goodman SM, Benstead JP, editors. *The natural history of Madagascar.* Chicago and London: The University of Chicago Press. p. 654–660.

Eickwort GC, Eickwort JM, Gordon J, Eickwort MA, Wcislo WT. 1996. Solitary behavior in a high-altitude population of the social sweat bee *Halictus rubicundus* (Hymenoptera: Halictidae). *Behav Ecol Sociobiol.* **38**:227–233.

Emerson AE. 1942. The relations of a relict South African Termite (Isoptera, Hodotermitidae, *Stolotermes*). *Am Museum Novit.* **1187**:1–12.

Emerson AE. 1969. A revision of the Tertiary Fossil Species of the Kalotermitidae (Isoptera). *Am Museum Novit.* **2359**:1–57.

Emerson AE. 1971. Tertiary fossil species of the Rhinotermitidae (Isoptera), phylogeny of genera, and reciprocal phylogeny of associated Flagellata (Protozoa) and the Staphylinidae (Coleoptera). *Bull Am Museum Nat Hist.* **146**:243–304.

Engel MS. 2008. Two new termites in Baltic Amber (Isoptera). *J Kansas Entomol Soc.* **81**:194–203.

Engel MS, Barden P, Riccio ML, Grimaldi DA. 2016. Morphologically specialized termite castes and advanced sociality in the early cretaceous. *Curr Biol.* **26**:522–530.

Engel MS, Delclòs X. 2010. Primitive Termites in Cretaceous Amber from Spain and Canada (Isoptera). *J Kansas Entomol Soc.* **83**: 111–128.

Engel MS, Grimaldi DA, Krishna K. 2007a. A synopsis of Baltic amber termites (Isoptera). *Stuttgarter Beiträge Zur Naturkd Ser. B* **372**: 1–20.

Engel MS, Grimaldi DA, Krishna K. 2007b. Primitive termites from the early Cretaceous of Asia (Isoptera). *Stuttgarter Beiträge Naturkd. Ser. B* **371**:1–32.

Engel MS, Grimaldi DA, Krishna K. 2009. Termites (Isoptera): their phylogeny, classification, and rise to ecological dominance. *Am Museum Novit.* **3650**:1–27.

Engel M, Grimaldi D, Nascimbene P, Singh H. 2011. The termites of Early Eocene Cambay amber, with the earliest record of the Termitidae (Isoptera). *Zookeys* **148**:105–123.

Engel MS, Hinojosa-Diaz IA, Rasnitsyn AP. 2009. A honey bee from the Miocene of Nevada and the biogeography of *Apis* (Hymenoptera: Apidae: Apini). *Proc Calif Acad Sci.* **60**:23–38.

Engel MS, Kaulfuss U. 2017. Diverse, primitive termites (Isoptera: Kalotermitidae, incertae sedis) from the early Miocene of New Zealand. *Austral Entomol.* **56**:94–103.

Engel MS, Krishna K. 2007a. Drywood termites in Dominican amber (Isoptera: Kalotermitidae). *Beiträge Zur Entomol.* **57**:263–275.

Engel MS, Krishna K. 2007b. New *Dolichorhinotermes* from Ecuador and in Mexican amber (Isoptera: Rhinotermitidae). *Am Museum Novit.* **2811**:1–8.

Engel MS, Nel A. 2015. A new fossil drywood termite species from the Late Eocene of France allied to *Cryptotermes* and *Procryptotermes* (Isoptera: Kalotermitidae). *Novit Paleoentomologicae.* 1–7.

Evans TA, Forschler BT, Grace JK. 2013. Biology of invasive termites: a worldwide review. *Annu Rev Entomol.* **58**:455–474.

Evans TA, Inta R, Lai JCS, Prueger S, Foo NW, Fu EW, Lenz M. 2009. Termites eavesdrop to avoid competitors. *Proc R Soc B Biol Sci.* **276**:4035–4041.

Faircloth BC. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics* **32**:786–788.

Gathorne-Hardy FJ, Jones DT. 2000. The recolonization of the Krakatau islands by termites (Isoptera), and their biogeographical origins. *Biol J Linn Soc.* **71**:251–267.

Ghesini S, Simon D, Marini M. 2014. *Kalotermes sinaicus* Kemner (Isoptera, Kalotermitidae): new morphological and genetic evidence, and assignment to the new genus *Longicaputermes* gen. nov. *Insectes Soc.* **61**:123–131.

Gibbons AD, Whittaker JM, Müller RD. 2013. The breakup of East Gondwana: assimilating constraints from Cretaceous ocean basins around India into a best-fit tectonic model. *J Geophys Res Solid Earth.* **118**:808–822.

Gnanapragasam NC. 2018. Insect pests of tea: shot hole borers, termites and nematodes. In: *Global tea science.* Cambridge UK & Philadelphia USA: Burleigh Dodds Science Publishing Limited. p. 201–240.

Goetsch W. 1936. Beiträge zur Biologie des Termitenstaates. *Zeitschrift Morphol Ökologie Tiere* **31**:490–560.

Grimaldi DA, Engel MS. 2005. Evolution of the insects. Cambridge: Cambridge University Press.

Grimaldi DA, Engel MS, Krishna K. 2008. The species of Isoptera (Insecta) from the early Cretaceous Crato Formation: a revision. *Am Museum Novit.* **3626**:1.

Hellemans S, Wang M, Hasegawa N, Šobotník J, Scheffrahn RH, Bourguignon T. 2021. Using ultraconserved elements to reconstruct the termite tree of life. *BioRxiv* 12.09.472027. doi:10.1101/2021.12.09.472027

Hemachandra II, Edirisinghe JP, Karunaratne WAIP, Gunatilleke CVS, Fernando RHSS. 2014. Diversity and distribution of termite assemblages in montane forests in the Knuckles Region, Sri Lanka. *Int J Trop Insect Sci.* **34**:41–52.

Hill GF. 1942. Termites (Isoptera) from the Australian Region. Melbourne: CSIR

Ho SYW, Phillips MJ. 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst Biol*. **58**:367–380.

Holt BG, Lessard J-P, Borregaard MK, Fritz SA, Araújo MB, Dimitrov D, Fabre P-H, Graham CH, Graves GR, Jønsson KA, et al. 2013. An update of Wallace's zoogeographic regions of the world. *Science* **339**:74–78.

Inward DJG, Vogler AP, Eggleton P. 2007. A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspects of their evolutionary biology. *Mol Phylogenet Evol*. **44**:953–967.

Jones DT, Eggleton P. 2011. Global biogeography of termites: a compilation of sources. In: Bignell DE, Roisin Y, Lo N, editors. *Biology of termites: a modern synthesis*. Dordrecht: Springer SBM. p. 477–498.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Haeseler Av, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. **14**:587–589.

Kambhampati S, Kjer KM, Thorne BL. 1996. Phylogenetic relationship among termite families based on DNA sequence of mitochondrial 16S ribosomal RNA gene. *Insect Mol Biol*. **5**: 229–238.

Kasinski JR, Kramarska R, Slodkowska B, Sivkov V, Piwocki M. 2020. Paleocene and Eocene deposits on the eastern margin of the Gulf of Gdańsk (Yantarny P-1 borehole, Kaliningrad region, Russia). *Geol Q*. **64**:29–53.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. **30**:772–780.

Kinjo Y, Saitoh S, Tokuda G. 2015. An efficient strategy developed for next-generation sequencing of endosymbiont genomes performed using crude DNA isolated from host tissues: a case study of *Blattabacterium cuenoti* inhabiting the fat bodies of cockroaches. *Microbes Environ*. **30**:208–220.

Kotthoff U, Wappler T, Engel MS. 2013. Greater past disparity and diversity hints at ancient migrations of European honey bee lineages into Africa and Asia. *J Biogeogr*. **40**:1832–1838.

Krishna K. 1961. A generic revision and phylogenetic study of the family Kalotermitidae (Isoptera). *Bull Am Mus Nat Hist*. **122**: 303–408.

Krishna K. 1996. New fossil species of termites of the subfamily Nasutitermitinae from Dominican and Mexican amber (Isoptera, Termitidae). *Am Museum Novit*. **3176**:1–13.

Krishna K, Grimaldi DA. 2003. The first Cretaceous Rhinotermitidae (Isoptera): a new species, genus, and subfamily in Burmese amber. *Am Museum Novit*. **3390**:1–10.

Krishna K, Grimaldi D. 2009. Diverse Rhinotermitidae and Termitidae (Isoptera) in Dominican amber. *Am Museum Novit*. **3640**:1–48.

Krishna K, Grimaldi DA, Krishna V, Engel MS. 2013. Treatise on the Isoptera of the World 1. Introduction. *Bull Am Museum Nat Hist*. **377**:1–200.

Kück P, Meusemann K. 2010. FASconCAT: convenient handling of data matrices. *Mol Phylogenet Evol*. **56**:1115–1118.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**:357–359.

Legendre F, Whiting MF, Bordereau C, Cancello EM, Evans TA, 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.

Lieberman BS. 2002. Phylogenetic biogeography with and without the fossil record: gauging the effects of extinction and paleontological incompleteness. *Palaeogeogr Palaeoclimatol Palaeoecol*. **178**:39–52.

Light SF. 1937. Contributions to the biology and taxonomy of *Kalotermes* (*Paraneotermes*) *simplicicornis* Banks (Isoptera). *Univ Calif Publ Entomol*. **6**:423–464.

Martins-Neto RG, Mancuso A, Gallego OF. 2005. The Triassic insect fauna from Argentina. Blattoptera and Coleoptera from the Ischichuca Formation (Bermejo Basin), la Rioja Province. *Ameghiniana* **42**:705–723.

Matzke NJ. 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Front Biogeogr*. **5**(4): 242–248.

Meng G, Li Y, Yang C, Liu S. 2019. MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Res*. **47**:e63.

Meseguer AS, Lobo JM, Ree R, Beerling DJ, Sanmartín I. 2015. Integrating fossils, phylogenies, and niche models into biogeography to reveal ancient evolutionary history: the case of *Hypericum* (Hypericaceae). *Syst Biol*. **64**:215–232.

Minh BQ, Nguyen MAT, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol*. **30**:1188–1195.

Mizumoto N, Bourguignon T. 2020. Modern termites inherited the potential of collective construction from their common ancestor. *Ecol Evol*. **10**:6775–6784.

Morgan FD. 1959. The ecology and external morphology of *Stolotermes ruficeps* Brauer (Isoptera: Hodotermitidae). *Trans R Soc New Zeal*. **86**:155–195.

Mörs T, Reguero M, Vasilyan D. 2020. First fossil frog from Antarctica: implications for Eocene high latitude climate conditions and Gondwanan cosmopolitanism of Australobatrachia. *Sci Rep*. **10**:5051.

Myles TG. 1999. Review of secondary reproduction in termites (Insecta: Isoptera) with comments on its role in termite ecology and social evolution. *Sociobiology* **333**:1–91.

Nakamine H, Yamamoto S, Takahashi Y. 2020. Hidden diversity of small predators: new thorny lacewings from mid-Cretaceous amber from northern Myanmar (Neuroptera: Rhachiberothidae: Paraberothinae). *Geol Mag*. **157**:1149–1175.

Nel A, Bourguet E. 2006. Termite of the Early Eocene amber of France (Isoptera: Mastotermitidae. Kalotermitidae). *Neues Jahrb Für Geol Und Paläontologie – Monatshefte*. **2006**:101–115.

Nel A, Paicheler J-C. 1993. Les Isoptera fossiles. État actuel des connaissances, implications paléoécologiques et paléoclimatologiques [Insecta. Dictyoptera]. In: Nel A, Martinez-Delclos X, Paicheler J-C, editors. *Essai de révision des Aeschnidoidea [Insecta, Isoptera, Anisoptera] - Les Isoptera fossiles [Insecta, Dictyoptera]. Paris: Cahiers Paléontologie, CNRS Éditions*. p. 101–179.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. **32**:268–274.

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res*. **27**:824–834.

Nutting WL. 1969. Flight and colony foundation. In: *Krishna K, Weesner FM, editors Biology of termites*. Vol. I. New York: Academic Press. p. 233–282.

Parham J, Donoghue PCJ, Bell CJ, Calway TD, Head JJ, Holroyd PA, Inoue JG, Irmis RB, Joyce WG, Ksepka DT, et al. 2012. Best practices for justifying fossil calibrations. *Syst Biol*. **61**:346–359.

Park LE, Downing KF. 2001. Paleoecology of an exceptionally preserved arthropod fauna from lake deposits of the Miocene Barstow Formation, Southern California, U.S.A. *Palaios* **16**: 175–184.

Piton L, Théobald N. 1937. Les Lignites et Schistes bitumineux de Menat (Puy-de-Dôme). II : Les insectes fossiles de Menat. *Rev Sci Nat Auvergne*. **3**:9–21.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol*. **67**:901–904.

Ree RH, Smith SA. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst Biol*. **57**:4–14.

Reguero MA, Gelfo JN, López GM, Bond M, Abello A, Santillana SN, Marenssi SA. 2014. Final Gondwana breakup: the Paleogene South American native ungulates and the demise of the South America–Antarctica land connection. *Glob Planet Change*. **123**: 400–413.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. *Trends Genet*. **16**:276–277.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol*. **29**:24–26.

Roisin Y, Korb J. 2011. Social organisation and the status of workers in termites. In: Bignell DE, Roisin Y, Lo N, editors. *Biology of termites: a modern synthesis*. Dordrecht: Springer SBM. p. 133–164.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**:1572–1574.

Sanmartín I, Ronquist F. 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Syst Biol*. **53**:216–243.

Scheffrahn RH. 2014. *Incisitermes nishimurai*, a new drywood termite species (Isoptera: Kalotermitidae) from the highlands of Central America. *Zootaxa* **3878**:471–478.

Scheffrahn RH, Bourguignon T, Akama PD, Sillam-Dussès D, Šobotník J. 2018. *Roisinitermes ebogoensis* gen. & sp. n., an outstanding drywood termite with snapping soldiers from Cameroon (Isoptera, Kalotermitidae). *Zookeys* **787**:91–105.

Scheffrahn R, Krecek J. 1999. Termites of the Genus *Cryptotermes* Banks (Isoptera: Kalotermitidae) from the West Indies. *Insecta Mundi*. **13**:111–171.

Scheffrahn RH, Křeček J, Chase JA, Maharajh B, Mangold JR. 2006. Taxonomy, biogeography, and notes on termites (Isoptera: Kalotermitidae, Rhinotermitidae, Termitidae) of the Bahamas and Turks and Caicos Islands. *Ann Entomol Soc Am*. **99**:463–486.

Scheffrahn RH, Postle A. 2013. New termite species and newly recorded genus for Australia: *Marginitermes absitus* (Isoptera: Kalotermitidae). *Aust J Entomol*. **52**:199–205.

Schlemmermeyer T, Cancello EM. 2000. New fossil termite species: *Dolichorhinotermes dominicanus* from Dominican amber. *Pap Avulsos Zool*. **41**:303–311.

Scotese RS. 2004. Cenozoic and mesozoic paleogeography: changing terrestrial biogeographic pathways. In: Lomolino MV, Heaney LR, editors. *Frontiers of biogeography: new directions in the geography of nature*. Sunderland, MA: Sinauer Associates. p. 9–26

Sillam-Dussès D, Sémon E, Robert A, Bordereau C. 2009. (Z)-Dodec-3-en-1-ol, a common major component of the trail-following pheromone in the termites Kalotermitidae. *Chemoecology* **19**:103–108.

Snyder TE. 1949. Catalog of the termites (Isoptera) of the world. *Smithson Misc Collect*. **112**:1–490.

Stöver BC, Müller KF. 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinform*. **11**:7.

Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res*. **34**:W609–W612.

Thiel M, Haye P. 2006. The ecology of rafting in the marine environment. III. Biogeographical and evolutionary consequences. In: Gibson RN, Atkinson RJA, Gordon DM, editors. *Oceanography and marine biology: an annual review*. Boca Raton: CRC Press. p. 323–429.

Thompson WL. 1934. Notes on *Neotermes castaneus* Burm. *Florida Entomol*. **18**:33.

Thompson GJ, Miller LR, Lenz M, Crozier RH. 2000. Phylogenetic analysis and trait evolution in Australian lineages of Drywood Termites (Isoptera, Kalotermitidae). *Mol Phylogenet Evol*. **17**: 419–429.

Thorne BL. 1997. Evolution of eusociality in termites. *Annu Rev Ecol Syst*. **28**:27–54.

Thorne BL, Haverty MI. 1991. A review of intercolony, intraspecific and interspecific agonism in termites. *Sociobiology* **19**: 115–145.

Upchurch P. 2008. Gondwanan break-up: legacies of a lost world? *Trends Ecol Evol*. **23**:229–236.

van der Valk T, Pečnerová P, Díez-del-Molino D, Bergström A, Oppenheimer J, Hartmann S, Xenikoudakis G, Thomas JA, Dehasque M, Sağlıcan E, et al. 2021. Million-year-old DNA sheds light on the genomic history of mammoths. *Nature* **591**: 265–269.

Vilhelmsen L, Perrichot V, Shaw SR. 2010. Past and present diversity and distribution in the parasitic wasp family Megalyridae (Hymenoptera). *Syst Entomol*. **35**:658–677.

Wallace AR. 1876. *The geographical distribution of animals*. Cambridge: Cambridge Univ. Press.

Wang M, Bučk A, Šobotník J, Sillam-Dussès D, Evans TA, Roisin Y, Lo N, Bourguignon T. 2019. Historical biogeography of the termite clade Rhinotermitinae (Blattodea: Isoptera). *Mol Phylogenet Evol*. **132**:100–104.

Waterhouse DF. 1993. *Biological control: pacific prospects – Supplement 2*. Canberra: Australian Centre for International Agricultural Research.

Watson JAL, Sewell JJ. 1985. Caste development in *Mastotermes* and *Kalotermes*: which is primitive? In: Watson JAL, Okot-Kotber BM, Noirot C, editors. *Caste differentiation in social insects*. Oxford: Pergamon Press Ltd. p. 27–40.

Wilson EO. 1971. *The insect societies*. Masachusetts: Harvard Unviersity Press.

Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, Dormann CF, Forchhammer MC, Grytnes J-A, Guisan A, et al. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biol Rev*. **88**:15–30.

Yu Y, Blair C, He X. 2020. RASP 4: ancestral state reconstruction tool for multiple genes and characters. *Mol Biol Evol*. **37**:604–606.

Yu Y, Harris AJ, Blair C, He X. 2015. RASP (reconstruct ancestral state in phylogenies): a tool for historical biogeography. *Mol Phylogenet Evol*. **87**:46–49.

Zachos J. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**:686–693.

Zhao Z, Yin X, Shih C, Gao T, Ren D. 2020. Termite colonies from mid-Cretaceous Myanmar demonstrate their early eusocial lifestyle in damp wood. *Natl Sci Rev*. **7**:381–390.