Phylogenetic relationships of nearctic *Reticulitermes* species (Isoptera: Rhinotermitidae) with particular reference to *Reticulitermes arenincola* Goellner

Weimin Ye, a Chow-Yang Lee, b Rudolf H. Scheffrahn, c Jody M. Aleong, a Nan-Yao Su, c Gary W. Bennett, a and Michael E. Scharf a, b

a Department of Entomology and Center for Urban and Industrial Pest Management, Purdue University, West Lafayette, IN 47907-2089, USA

b School of Biological Sciences, Universiti Sains Malaysia, Penang 11800, Malaysia

c Ft. Lauderdale Research and Education Center, University of Florida, 3205 College Avenue, Ft. Lauderdale, FL 33314, USA

Received 19 February 2003; revised 3 June 2003

Abstract

DNA sequence comparisons of the mitochondrial COI, 16S, and 12S rRNA genes were used to infer phylogenetic relationships among the six known US *Reticulitermes* species (*Reticulitermes flavipes*, *Reticulitermes arenincola*, *Reticulitermes tibialis*, *Reticulitermes hageni*, *Reticulitermes virginicus*, and *Reticulitermes hesperus*) and the closely related European species *Reticulitermes santonensis*. The interspecific pairwise sequence divergence, based on uncorrected "p" distance, varied up to 10% across the COI, 4.9% across the 16S, and 3% across the 12S fragments. Phylogenetic trees were constructed using maximum parsimony, likelihood, and distance methods. The combined results suggest several phylogenetic relationships including: (i) *R. flavipes*, *R. arenincola*, and European *R. santonensis* are possibly conspecific; (ii) *R. virginicus* and *R. hageni* are closely related species; and (iii) *R. tibialis* and *R. hesperus* are closely related species. Interestingly, while there is apparent synonymy between *R. flavipes* and *R. arenincola* by DNA sequence, there are clear morphological differences in the soldier caste. This finding suggests a combination of molecular and morphological approaches are necessary for accurate species identification. These data lend resolution to the complex problem of *Reticulitermes* systematics, and will assist future efforts directed toward characterizing species distribution and ecology.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Subterranean termite; mtDNA; COI; 12S; 16S; Phylogenetics; Molecular systematics

1. Introduction

The termites (order Isoptera) are important insects from a number of economic and ecological perspectives. Economically, the damage termites cause to structures and buildings accounts for greater than US$20 billion annually worldwide (Su, 2002). Ecologically, termites fill important niches in natural ecosystems by their breaking down of cellulose and their role in the decomposition of organic matter (Sugimoto et al., 2000). Despite their importance, our understanding of a number of basic biological processes in the most common North American termites (family Rhinotermitidae, genus *Reticulitermes*) are extremely limited (e.g., reproduction, caste differentiation, foraging ecology, species distribution, etc.). Developing a better understanding of termite biology is closely dependent upon reliable species identification.

Subterranean termites of the genus *Reticulitermes* are the dominant termites in North America, where six species phenotypes have been described based on alate and soldier characters: *Reticulitermes flavipes* Kollar, *Reticulitermes tibialis* Banks, *Reticulitermes hesperus* Banks, *Reticulitermes virginicus* Banks, *Reticulitermes hageni* Banks, and *Reticulitermes arenincola* Goellner
(Goellner, 1931; Nutting, 1990; Snyder, 1954; Weesner, 1965). The geographical distribution and habitat preferences of the North American Reticulitermes are only partially understood. For example, the distribution of R. hesperus (the western subterranean termite) is thought to be limited to the extreme western US and Pacific Coast; R. tibialis (the dry land subterranean termite) is considered to range across the desert Southwest through the Great Plains to the eastern Midwest, and R. flavipes (the eastern subterranean termite) is believed to occur over the entire US East of the Missouri River, from southern Florida northward into southern Ontario, Canada (Nutting, 1990; Snyder, 1954; Weesner, 1965). Interestingly, the species R. arenicola has been reported to occur exclusively in two very limited areas: the sand dunes on the South end of Lake Michigan in northwest Indiana and southwest Michigan, and in the vicinity of Boston, Massachusetts (Goellner, 1931; Nutting, 1990; Snyder, 1954; Weesner, 1965). R. arenicola has not been extensively studied in the past 50 years, and therefore, the biology and ecology of this species remains mostly unknown. In Indiana, all of the above species, with the exception of R. hesperus, have been reported.

Systematic research on the Reticulitermes genus by morphology alone has been difficult (Nutting, 1990). Recent studies have demonstrated a great potential for DNA sequence analysis in Reticulitermes species identification, phylogeny, and gene flow (Austin et al., 2002; Jenkins et al., 1999, 2000, 2001; Marini and Mantovani, 2002; Miura et al., 1998). Sequences of the mitochondrial genes cytochrome oxidase subunit II (COII), rRNA large subunit (16S), and rRNA small subunit (12S) have been extensively applied in phylogenetic reconstructions for a variety of taxa (Frati et al., 1997; Kambhampati, 1995; Kambhampati et al., 1996; Liu and Beckenbach, 1992). In this study, our objectives were to: (1) obtain nucleotide sequences of the COII, 16S and 12S mitochondrial genes from Reticulitermes termites as a basis for further species identification, (2) conduct phylogenetic analyses using these nucleotide sequences (alone and in combination), (3) examine the species status of R. arenicola, Reticulitermes santonensis, and R. flavipes by molecular phylogeny, and (4) compare molecular findings with morphological assessments of the soldier caste.

2. Materials and methods

2.1. Termite samples, DNA extraction, and GenBank sequences

Reticulitermes samples and one outgroup species (Coptotermes formosanus Shiraki) were collected from the field or procured from various sources (Table 1), and were preserved in 85–100% ethanol prior to DNA extraction. Total genomic DNA was extracted from single termites using the Wizard genomic DNA purification kit (Promega; Madison, WI, USA) with slight modifications. Extracted genomic DNA was stored at −20°C until it was used as polymerase chain reaction (PCR) template. Published sequences from R. santonensis, R. flavipes, and R. hesperus were also included in our phylogenetic analyses (see Table 1 for details).

2.2. PCR

Fragments of the mitochondrial genes COII, 16S, and 12S were amplified by PCR, using established primer pairs and reaction conditions (Kambhampati, 1995; Liu and Beckenbach, 1992; Marini and Mantovani, 2002; Miura et al., 1998; Simon et al., 1994). See supplementary materials for information on primers and PCR conditions.

2.3. Sequencing

PCR products were cleaned using either a PCR purification kit (Qiagen; Valencia, CA, USA) or Microcon-PCR filter unit (Millipore; Bedford, MA, USA). Direct sequencing of PCR products was performed by the Purdue University Genomics Core Facility. Gene sequences were obtained for at least three individuals per termite collection site. Representative sequences were deposited into the GenBank database under accession numbers shown in Table 1.

2.4. Sequence alignments and phylogenetic inferences

DNA sequences were aligned using C. formosanus as the outgroup taxon. Sequences were aligned using GCG (Genetics Computer Group) PILEUP program (with a gap weight of 5 and a gap length weight of 1). The distances were calculated as uncorrected percentages of divergence according to the Kimura two-parameter model (Kimura, 1980) using PAUP* 4.0b10 (Swofford, 2002). Three different analyses were carried out in PAUP* 4.0b10. The maximum parsimony method was applied, using the branch-and-bound search with stepwise-addition options to determine the most parsimonious tree. Analysis of nucleotide sequences was also performed using the neighbor-joining algorithm with Kimura’s (1980) two-parameter model and maximum likelihood (Felsenstein, 1981) using substitution model and HKY85 as DNA distance. Sites with missing data or gaps were treated as missing characters for all analyses. Parsimony analyses were done using a 2:1 transition/transversion weighting step matrix. The robustness of the trees was tested using the bootstrap method. All bootstrap values (Felsenstein, 1985) are based on 1000 replicates.
Table 1
Termite species and populations included in the present study

<table>
<thead>
<tr>
<th>Species</th>
<th>Location</th>
<th>GenBank Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>R. arenuncola</em> Sd</td>
<td>Dune Acres, IN, USA</td>
<td>AY168209</td>
</tr>
<tr>
<td><em>R. flavipes</em> Indiana</td>
<td>Entomology Field Office Building, Purdue Research Station, W. Lafayette, IN, USA</td>
<td>AY168203</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Alachua, FL, USA</td>
<td>AF525321</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Bahamas</td>
<td>AF525322</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Hamburg, Germany</td>
<td>AF525323</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Lincoln, NE, USA</td>
<td>AF525325</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Toronto, Ontario, Canada</td>
<td>AF525326</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>USA</td>
<td>AY027477</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Sapelo Island, GA, USA</td>
<td>AY107479</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Sapelo Island, GA, USA</td>
<td>AY107480</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Sapelo Island, GA, USA</td>
<td>AY107481</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Sapelo Island, GA, USA</td>
<td>AY107482</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Sapelo Island, GA, USA</td>
<td>AY107484</td>
</tr>
<tr>
<td><em>R. flavipes</em></td>
<td>Spalding Co., GA, USA</td>
<td>AY107489</td>
</tr>
<tr>
<td><em>R. flavipes</em> Hh</td>
<td>Happy Hollow Park (Sandy Hill), W. Lafayette, IN, USA</td>
<td>AY168219</td>
</tr>
<tr>
<td><em>R. flavipes</em> Rh</td>
<td>Ross Hills Park, Tippecannoe Co., IN, USA</td>
<td>AY168213</td>
</tr>
<tr>
<td><em>R. flavipes</em> Rw</td>
<td>Robinson Woods, Tippecannoe Co., IN, USA</td>
<td>AY168214</td>
</tr>
<tr>
<td><em>R. flavipes</em> Sm</td>
<td>Marion, IN, USA</td>
<td>AY168211</td>
</tr>
<tr>
<td><em>R. flavipes</em> Wi</td>
<td>Soldiers Home Rd., W. Lafayette, IN, USA</td>
<td>AY168210</td>
</tr>
<tr>
<td><em>R. hageni</em></td>
<td>Conway, AR, USA</td>
<td>AY168208</td>
</tr>
<tr>
<td><em>R. hesperus</em></td>
<td>Los Angeles, CA, USA</td>
<td>AF525329</td>
</tr>
<tr>
<td><em>R. santonensis</em></td>
<td>Charente, France</td>
<td>AF525343</td>
</tr>
<tr>
<td><em>R. santonensis</em></td>
<td>La Rochelle, France</td>
<td>AF262607</td>
</tr>
<tr>
<td><em>R. santonensis</em></td>
<td>France</td>
<td>AF291742</td>
</tr>
<tr>
<td><em>R. santonensis</em></td>
<td>France</td>
<td>AF291743</td>
</tr>
<tr>
<td><em>R. santonensis</em></td>
<td>Bordeaux, France</td>
<td>AY027473</td>
</tr>
<tr>
<td><em>R. santonensis</em></td>
<td>Bordeaux, France</td>
<td>AY027474</td>
</tr>
<tr>
<td><em>R. santonensis</em></td>
<td>Bordeaux, France</td>
<td>AY027475</td>
</tr>
<tr>
<td><em>R. santonensis</em></td>
<td>Bordeaux, France</td>
<td>AY027476</td>
</tr>
<tr>
<td><em>R. tibialis</em> Ag</td>
<td>Agricultural Administration Building, Purdue University, W. Lafayette, IN, USA</td>
<td>AY168207</td>
</tr>
<tr>
<td><em>R. tibialis</em> Fs</td>
<td>Forestry Building, Purdue University, W. Lafayette, IN, USA</td>
<td>AY168206</td>
</tr>
<tr>
<td><em>R. virginicus</em></td>
<td>Area 7, northwest side, Jasper-Pulaski Fish &amp; Wildlife Area, Wheatfield, IN, USA</td>
<td>AY168205</td>
</tr>
<tr>
<td><em>R. virginicus</em> Jp</td>
<td>Area 7, south-central side, Jasper-Pulaski Fish &amp; Wildlife Area, Wheatfield, IN, USA</td>
<td>AY168223</td>
</tr>
<tr>
<td><em>Coptotermes formosanus</em></td>
<td>Golden Beach, FL, USA</td>
<td>AY168204</td>
</tr>
</tbody>
</table>

* Austin et al. (2002).
* Thompson et al. (2000).
* Marini and Mantovani (2002).
* Jenkins et al. (2001).

3. Results

3.1. Nucleotide analyses

The PCR products of COI, 16S, and 12S are approx. 780, 550, and 450 bp, respectively for the *Reticulitermes* species examined. For the COI gene, the multiple sequence alignment (including outgroup) has 684 characters, of which 497 are constant and 91 are parsimony-informative. Excluding the outgroup and sequences obtained from GenBank, there are 112 variable COI sequence sites among all the *Reticulitermes* species. Most of these variations correspond to the third codon positions (77.68%), followed by the first (15.18%) then second (7.14%) position. For 16S, multiple sequence alignments including outgroup had 459 characters, of which 388 are constant and 25 are parsimony-informative. For 12S, there are 447 characters including the forward and reverse priming regions, of which 394 are constant and 15 are parsimony-informative. For combined analyses of the three genes, the multiple sequence alignment has 1590 characters, of which 1316 are constant and 114 are parsimony-informative.
Fig. 1. Neighbor-joining trees inferred from (A) COII mtDNA (parsimony tree length [TL] = 294; consistency index [CI] = 0.7279; retention index [RI] = 0.7727; Ln likelihood = -2457.1980), (B) 16S rDNA (TL = 89; CI = 0.8876; RI = 0.8333), (C) 12S rDNA (TL = 65; CI = 0.8485; RI = 0.7826; Ln likelihood = -934.2792), and (D) combined data on COII, 16S, and 12S (TL = 352; CI = 0.8551; RI = 0.7560). Bootstrap percentages are shown above branches supported in at least 50% of 1000 replicates in parsimony analysis during a branch and bound search using Paup.
Average nucleotide composition among *Reticulitermes* species (excluding GenBank sequences) is as follows: COII [39.23\% (A), 23.51\% (T), 13.76\% (G), and 23.50\% (C)]; 16S [40.08\% (A), 22.33\% (T), 13.11\% (G), and 24.50\% (C)]; and 12S [43.98\% (A), 22.29\% (T), 12.45\% (G), and 21.27\% (C)]. The interspecific pairwise sequence divergence, based on uncorrected “p” distance, ranges from 0 to 10\% across the entire COII gene, from 0 to 4.9\% across the 16S fragment, and from 0 to 3\% across the 12S fragment. Thus, COII is considered as the fastest evolving gene of those examined, while 12S is the most conserved. Alternatively, the divergence values between *Reticulitermes* species and the outgroup vary from 13.7 to 15.6\% in COII, 10.7 to 13.1\% in 16S, and 8.3 to 9.2\% in 12S.

The aligned DNA data matrices are available at TreeBASE (http://www.treebase.org, submission number SN1460). Identical sequences were found between two populations of *R. tibialis* for COII, 16S, and 12S; three populations of *R. santonensis* from France for COII, two populations of *R. flavipes* from the US for COII; one population of *R. flavipes* from Indiana, one population of *R. arenicola* from Indiana, and one population of *R. santonensis* from France for COII and 16S; two populations of *R. flavipes* from Indiana for 12S; one population of *R. flavipes* and one population of *R. arenicola* from Indiana for 12S. Intraspecific sequence divergences were found for sixteen populations of *R. flavipes* for COII (uncorrected “p” distance = 0–3.5\%), eight populations of *R. santonensis* for COII (0–1.5\%); two populations of *R. virginicus* from Indiana for 12S (1.1\%); three populations of *R. flavipes* from Indiana for 16S (0.4–2\%); and six populations of *R. flavipes* from Indiana for 12S (0.2–0.9\%).

3.2. Phylogenetic relationships inferred from COII, 16S, and 12S

Phylogenies derived using distance, maximum parsimony, and maximum likelihood methods were generally congruent, therefore, only distance neighbor-joining trees are shown in the results. Based on the COII gene sequence with *C. formosanus* as the outgroup, two major clades within the *Reticulitermes* genus are apparent (Fig. 1A). One clade includes *R. flavipes*, *R. arenicola*, and *R. santonensis* with strong bootstrap support (98\%). Within this clade, there are two sub-clades: one with *R. flavipes* from Indiana and Georgia, and the other includes various populations of *R. flavipes*, *R. arenicola*, and *R. santonensis*. Here, it is also noteworthy that one Indiana population of *R. arenicola* shares identical DNA sequences with one *R. santonensis* population from France and one *R. flavipes* population from Indiana. The second *Reticulitermes* clade is composed of two sub-clades. These two sub-clades include *R. tibialis* and *R. hesperus* in one sister group, and *R. virginicus* and *R. hageni* in another. These results for COII are in general agreement with the phylogenetic trees inferred from the 16S and 12S partial sequences (Figs. 1B and C), except that *R. virginicus* has a closer relationship to *R. tibialis* than *R. hageni* in the 12S tree. However, this relationship is not supported by high bootstrap values. Also, *R. arenicola* and one population of *R. flavipes* from Indiana share identical 16S sequences with *R. santonensis* from France. These three species form a monophyletic cluster. Finally, *R. arenicola* and one population of *R. flavipes* from Indiana also share identical 12S sequences.

Maximum parsimony analysis of the COII, 16S, 12S, and combined nucleotide matrices results in a total of 10, 100 (initial MaxTrees setting = 100), 1, and 2 equally most parsimonious trees, respectively. The strict consensus tree of each has the same topology as the neighbor-joining and maximum likelihood trees, thus these trees are not shown.

3.3. Combined phylogenetic analysis inferred from COII, 16S, and 12S

The combined nucleotide matrix of the three genes generates a tree topology similar to that obtained using COII DNA sequences (Fig. 1D). Three clades are suggested: one clade includes two populations of *R. tibialis*, one clade includes *R. hageni* and *R. virginicus*, while the other includes *R. flavipes* and *R. arenicola*.

3.4. Morphological comparisons

Identities of individuals shown in Fig. 2 were verified by DNA sequencing. Soldiers of Indiana *R. tibialis*, *R. virginicus*, and *R. hesperus* can be readily distinguished from *R. flavipes* and *R. arenicola* by head morphology (Figs. 2A and B), and by following characters noted in existing keys such as head pigmentation and gular width ratios (e.g., Goellner, 1931; Nutting, 1990). Morphometric measurements on head + mandible lengths and pronotal widths of *R. arenicola* soldiers were 2.43 and 0.78 mm, respectively (*n = 3*). These values are within the limits of 2.28–2.56 mm (head + mandible) and 0.73–0.87 mm (pronotum) described by Goellner (1931). Other *R. flavipes* samples (*n ≥ 10*) shown in Fig. 2A were within (Rf1) and above (Rf2) the limits, respectively, of 2.65–3.20 mm described by Goellner (1931). Examples of observed size and morphological variation of *R. flavipes* soldiers between different collection locations, and between various *R. flavipes* populations and *R. arenicola* are shown in Fig. 2A.

4. Discussion

From our phylogenetic analyses, three clades within the *Reticulitermes* genus from the Nearctic region are
apparent. These include: (1) *R. hesperus* and *R. tibialis*; (2) *R. virginicus* and *R. hageni*; and (3) *R. flavipes*, *R. arenincola*, and *R. santonensis*. This conclusion is supported by phylogenies reconstructed from three individual mitochondrial gene sequences, and the three combined gene sequences. Previous studies, which employed different approaches have also made similar conclusions (Austin et al., 2002; Jenkins et al., 1999, 2000, 2001; Marini and Mantovani, 2002; Miura et al., 1998). The relationships observed in this study using molecular systematic approaches generally correspond to phylogenies based on morphology, with the exception of the clade including *R. flavipes*, *R. arenincola*, and *R. santonensis*.

*Reticulitermes arenincola* was described as a new species from previously identified *R. flavipes* collected from the sand dunes of Indiana and Michigan, along the shores of Lake Michigan (Goellner, 1931). *R. arenincola* has not been studied in over 50 years. Our *R. arenincola* specimens collected from Indiana sand dune habitats agree with the original descriptions of Goellner (1931). *R. arenincola* can also readily be differentiated from *R. flavipes* according to the species diagnosis presented by Goellner (1931).

The present study includes the first examination of DNA sequences for *R. arenincola*. However, DNA sequencing data presented here clearly indicates that *R. flavipes* and *R. arenincola* from the US and *R. santonensis* from France are conspecific. This result is supported by high degrees of similarity in COII (uncorrected "p" distance = 0–4%), 16S (0–1.9%), 12S (0–1%), from combined analyses of the three genes (0–2%), and translated COII amino acid sequences (0–2.6%).

*Reticulitermes santonensis* has been previously reported to be closely related to the North American species *R. flavipes* based on morphological characters (Clément et al., 2001), and by nucleotide sequences of the mitochondrial NADH dehydrogenase-1 gene (Clément et al.,...
the mitochondrial cytochrome oxidase I and II genes (Austin et al., 2003; Jenkins et al., 2001; Marini and Mantovani, 2002), the mitochondrial 16S ribosomal RNA gene (Marini and Mantovani, 2002), and the nuclear ribosomal intergenic transcribed spacer region II (Jenkins et al., 2001). It is believed that R. santonensis originated from a population(s) of US R. flavipes that was introduced to France during the late 19th century (Bagnères et al., 1990; Feytaud, 1925), and subsequent locations in Europe including Germany (Harris, 1962; Weidner, 1937) and Austria (Hrdy, 1961). Previous investigations have also suggested that R. santonensis is a synonym of R. flavipes (Austin et al., 2002; Jenkins et al., 2001; Marini and Mantovani, 2002; Vieau, 2001). The results of our concurrent investigation of three mitochondrial gene sequences further supports such a classification, but additionally suggests R. arenincola is also a possible synonym of R. flavipes. However, synonymizing these two species would be premature because of the clear morphological separation of R. arenincola and R. flavipes by soldier characters. It should also be noted that we have not assessed morphological features of alate forms and type specimens between R. arenincola and R. flavipes. Making such comparisons would certainly resolve the relationships of these two species.

Earlier studies have shown that termite morphology can be influenced by geography (Weidner, 1970), therefore, the morphological difference between R. arenincola and R. flavipes (and between geographically isolated R. flavipes populations) could be the result of environment-related intraspecific variation. Therefore, it appears imperative that future research efforts are directed at determining whether or not habitat/environmental-associated variation correlates with taxonomy derived from morphological, molecular, biochemical, and behavioral approaches in Reticulitermes termites.

In this study, we: (1) considered three mitochondrial gene sequences concurrently when inferring the phylogenetic relationships of US Reticulitermes species, (2) obtained novel 16S and 12S sequences for the species R. tibialis, R. virginicus, and R. hageni, and (3) obtained novel sequences for the highly under-studied species R. arenincola. Additionally, we considered intraspecific morphological variation in combination with molecular systematic investigations. The results of our comparisons indicate the existence of an apparently intractable system. Our observations, in agreement with previous conclusions (Forschler and Jenkins, 2000; Jenkins et al., 2000), suggest a combination of molecular systematics and morphology to be the most reliable approach for the identification of US Reticulitermes termites. Further investigations will be important for addressing questions relating to the effects of environment on morphological variation across geographic locations, and further, how such differences correlate with phylogenies based on molecular evidence.

Acknowledgments

We thank the Fulbright Scholarship Program and Department of Entomology (Purdue University) for supporting C.Y.L.; Dr. Allen L. Szalanski (University of Arkansas) for R. hageni samples; Dr. Changlu Wang (Purdue University) for R. hageni soldiers; Doug Murphy of the Purdue University Agricultural Genomics Facility for sequencing assistance, and Dow AgroSciences Inc. for partial research support. This is journal article No. 17030 of the agricultural research program of Purdue University, West Lafayette, IN.

References


Kambhampati, S., Kjer, K.M., Thorne, B.L., 1996. Phylogenetic relationship among termite families based on DNA sequence of the