

## Phylogenetic Analysis of Bacterial Communities Associated with Ectoparasitic Chewing Lice of Pocket Gophers: A Culture-Independent Approach

D.L. Reed,<sup>\*</sup> M.S. Hafner

Museum of Natural Science and Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803

Received: 23 April 2001; Accepted: 12 March 2002; Online publication: 20 May 2002

### ABSTRACT

This study identifies the bacteria associated with ectoparasitic chewing lice that live in the fur of pocket gophers. Samples of chewing lice were collected from pocket gopher hosts in Florida, Missouri, New Mexico, and Costa Rica. We used a molecular sampling method whereby total community DNA was extracted from samples of chewing lice, and PCR was used to selectively amplify small-subunit rRNA genes from bacteria. This culture-independent method yielded ca. 35 distinct lineages representing eight widely divergent groups within the domain Bacteria. Phylogenetic analysis of two lineages (*Acinetobacter* and *Staphylococcus*) provides evidence that multiple species of each group are found in chewing lice. Phylogenetic analysis also demonstrated that diversification within chewing lice may be evident in both *Acinetobacter* and *Staphylococcus*. Some clones amplified from chewing louse hosts appeared to be distinct from known species of *Acinetobacter* and *Staphylococcus*. This diversification may be the result of the extreme isolation of populations of both chewing lice and their pocket gopher hosts.

### Introduction

Pocket gophers and their ectoparasitic chewing lice show distinct patterns of cospeciation [for review see 6]. Pocket gophers (Rodentia: Geomyidae) are fossorial rodents whose geographical range extends from southern Canada through northwestern Colombia. These rodents are ex-

tremely asocial and live in isolation in sealed underground burrow systems. Different species of pocket gophers are rarely found in the same region, and then only along narrow zones of parapatry. These natural history characteristics prevent widespread transfer of parasites among individuals in a population, among conspecific populations, or among species of pocket gophers.

Chewing lice of the genera *Geomydoecus* and *Thomomydoecus* (Phthiraptera: Trichodectidae) are restricted to pocket gopher hosts. These lice are small, permanent ectoparasites that complete their entire life cycle on a single host. Chewing lice are highly host specific [17, 27] and

<sup>\*</sup>Present address: Department of Biology, 257 South 1400 East, University of Utah, Salt Lake City, UT 84112

Correspondence to: D.L. Reed; Phone: (801) 585-9130; Fax: (801) 581-4668; E-mail: reed@biology.utah.edu

disperse to new hosts only through direct host-to-host contact [27]. The combination of low parasite vagility and obligate contact transmission of lice limits opportunities for colonization of new host species. The absence of widespread transfer of lice among gopher species has, in part, led to the pattern of cospeciation as documented by Hafner et al. [7], which appears to be common in this host-parasite assemblage [e.g., 3, 5, 7, 16]. The goals of this study are to determine the identity and diversity of the community of bacteria associated with these chewing lice, and to determine whether the host's geographic isolation has led to bacterial diversification within lice.

Numerous relationships have been described between invertebrates and endosymbiotic bacteria [1, 4, 9, 13, 14, 22]. In a review of the symbioses of insects and bacteria, Douglas [4] determined that approximately 10% of known insect species contain nonparasitic microorganisms, whereas recently, Jeyaprakash and Hoy [10] found that 76% of arthropod species tested harbor the bacterium *Wolbachia*. Modern molecular techniques are now being used to identify bacteria associated with a variety of insects, and it is becoming clear that most insects host multiple species of microorganisms. Often these microorganisms are required for specialized functions in the host, such as production of vitamins, synthesis of essential amino acids, digestion of complex foods, and development of offspring.

Ries [20] used microscopy to document the presence of bacteria in a wide variety of insects including species of chewing lice that live among the feathers of birds. He was unsuccessful, however, in finding bacteria associated with the species of chewing lice that live in the fur of mammals. Saxena and Agarwal [23] suggested that the endosymbiotic bacteria of feather lice might aid in the digestion of keratin-based feathers, although the biological interaction between these symbionts and their host is not yet known. None of the endosymbiotic bacteria from feather lice has been cultured outside of its host, and to date, no molecular techniques have been used to characterize the bacteria associated with any species of chewing louse.

We used a molecular sampling method whereby total community DNA was extracted from an environmental sample and polymerase chain reaction (PCR) was used to selectively amplify small-subunit rRNA genes (rDNA) from bacteria. This technique has been used extensively in studies of microbial diversity and evolution [15, 30], and it permits culture-independent detection and phylogenetic placement of unknown bacteria [8, 29].

## Materials and Methods

### Collection of Specimens and DNA extraction

We surveyed chewing lice of the genera *Geomydoecus* and *Thomomydoecus* from seven species of pocket gophers in four genera (*Cratogeomys*, *Geomys*, *Orthogeomys*, and *Thomomys*). The following lice were collected: *Geomydoecus geomydis* from *Geomys bursarius halli* (host number LSUMZ 31463) and *G. b. major* (LSUMZ 31448); *Geomydoecus scleritus* from *Geomys pinetus*; *Geomydoecus expansus* from *Cratogeomys castanops* (LSUMZ 31455); *Geomydoecus panamensis* from *Orthogeomys cavator* (LSUMZ 29253); *Geomydoecus costaricensis* from *O. heterodus* (LSUMZ 29501); *Geomydoecus setzeri* from *O. underwoodi* (LSUMZ 29493); and *Geomydoecus centralis* and *Thomomydoecus minor* from *Thomomys bottae* (LSUMZ 29569). Samples were collected from geographically distant localities in an attempt to maximize the diversity of bacteria found. Chewing louse samples were collected from their pocket gopher hosts, stored temporarily in liquid nitrogen, and later deposited in the LSU Museum of Natural Science Collection of Genetic Resources.

Adult chewing lice were washed twice in a solution of 400  $\mu$ L saline EDTA buffer (containing 150 mM NaCl, 10 mM EDTA, pH 8.0), 10  $\mu$ L of 25% SDS (sodium dodecyl sulfate), and 5  $\mu$ L of 10 mg/mL lysozyme, and were incubated at 37°C and agitated for 1 h. This washing protocol was designed to wash away external sources of bacterial DNA by lysing bacterial cells on the exterior surface of the louse. Whole lice were removed from the wash, placed in 1.5 mL microcentrifuge tubes along with 400  $\mu$ L of saline EDTA buffer and 5  $\mu$ L lysozyme (10 mg/mL), and then crushed with sterile micropestles. Negative controls, which contained all extraction reagents (and no DNA template), were used to ensure that reagents were not contaminated with extraneous sources of bacteria. The extraction solutions were incubated at 37°C for 1 h. Five microliters of proteinase-K (15 mg/mL) and 10  $\mu$ L of SDS (25%) were added and the solutions were incubated at 55°C for 1 h. Genomic DNA was extracted using a standard phenol/chloroform procedure followed by ethanol precipitation. The resulting pellets were resuspended in 50  $\mu$ L of filtered TE buffer.

### PCR, Cloning, and Cycle Sequencing

The polymerase chain reaction was used to amplify copies of bacterial 16S rDNA from total genomic DNA. Amplifications were performed using universal bacteria primers 27-f (5'-GAG TTT GAT CCT GGC TCA G-3') and 1525-r (5'-AGA AAG GAG GTG ATC CAG CC-3'). Genomic DNA (2  $\mu$ L), 3  $\mu$ L of each primer (10  $\mu$ M), 3  $\mu$ L of deoxynucleoside triphosphate (dNTP) mixture (dATP, dGTP, dCTP, and dTTP, each 1  $\mu$ M), 3  $\mu$ L of MgCl<sub>2</sub> (25 mM), and 1 unit of *Taq* DNA polymerase were combined in a 50  $\mu$ L PCR reaction. Negative PCR controls, which contained no DNA template, were used to test for contamination of the PCR reagents. Thirty-five thermal cycles were performed, each with a denaturation period at 94°C for 1 min, an annealing period at

56°C for 1 min, followed by an elongation period at 72°C for 2 min. After 35 cycles, a single extension time of 10 min at 72°C was used to facilitate polymerase activity and extend PCR products. Reactions containing fragments of the appropriate size (ca. 1530 base pairs) were cleaned with Qiaquick spin column PCR purification kits (Qiagen, Valencia, CA) as prescribed by the manufacturer.

Cleaned PCR fragments were ligated into pCR 4-TOPO plasmid vectors as prescribed by the manufacturer (Invitrogen TOPO TA Cloning Kit for Sequencing, Carlsbad, CA). Ligated plasmids were transformed into TOP10 One Shot competent cells as prescribed by the manufacturer. Two methods were used (alternatively) to isolate cloned plasmid DNA. The first method used S.N.A.P. mini-prep kits (Invitrogen, Carlsbad, CA) to remove plasmid DNA from competent cells by centrifugation. The second method employed an additional PCR step that used primers located in plasmid regions that flank the DNA template insert (M13f: 5'-GTA AAA CGA CGG CCA G-3' and M13r: 5'-CAG GAA ACA GCT ATG AC-3'). Both procedures produced high-quality PCR fragments suitable for DNA sequencing.

Cycle sequencing was performed using the ABI PRISM Big Dye kit (PE Applied Biosystems, Foster City, CA) as prescribed by the manufacturer. The number of clones sequenced was limited by the number of chewing lice collected and the efficiency of PCR, cloning, and sequencing. Clones ( $n = 16\text{--}78$  per host) were partially sequenced using two M13 plasmid primers, which produced two nonoverlapping fragments from the 5' and 3' regions of the 16S rRNA gene (each approximately 500 bp in length). These domains contained regions of both conserved and variable sequence.

### Taxonomic Identification

Mixed template PCR can produce chimeric sequences from multiple organisms [11, 21, 24, 28]. Therefore, all sequences were evaluated using the CHIMERA CHECK subroutine at the Ribosomal Database Project Web site (<http://rdp.cme.msu.edu>). Those found to be chimeric ( $n = 11$ ) were not included in further analyses.

Two methods were used to determine the identity of each bacterial sequence. First, partial sequences ( $n = 95$ ; 5' end of the 16S rDNA) were imported into the computer program ARB (version 2.5b; Department of Microbiology, TU Munich, Germany [<http://www.arb-home.de>]), which contains a database of more than 15,000 16S rDNA sequences of Eubacteria. The sequences were aligned with nearest neighbor sequences using the 16S alignment in ARB and several reference sequences were chosen for each clade from these nearest neighbor sequences. Parsimony and maximum-likelihood-based phylogenies were constructed to determine the identity of our bacterial clones. These sequences were deposited in GenBank under the accession numbers AF467307–AF467401 and AF467895. A second method of taxonomic identification was used for all sequence fragments ( $n = 234$ ; both 5' and 3' end of the 16S). These sequences were evaluated using the Basic Local Alignment Search Tool (BLAST)

at the National Center for Biotechnology Information (GenBank; <http://www.ncbi.nlm.nih.gov>). BLAST search results with the highest sequence similarity were used for identification if the expected value (E) was less than  $10^{-5}$ .

### Phylogenetic Analyses

Complete 16S rDNA sequences (~1530 bp) were needed for phylogenetic study of two bacterial lineages; therefore, four internal 16S rDNA primers (536f: 5'-CAG CMG CCG CGG TAA TWC-3', 1114f: 5'-GCA ACG AGC GCA ACC C-3', 519r: 5'-GWA TTA CCG CGG CKG CTG-3', and 960r: 5'-GCT TGT GCG GGY CCC CG-3') were used to generate complete 16S rDNA sequences in both forward and reverse directions for sequences that were identified as *Acinetobacter* and *Staphylococcus* from BLAST searches ( $n = 59$ ). All cycle sequencing products were cleaned using the ethanol/sodium acetate procedure outlined in the ABI Big Dye manual. Automated DNA sequencing was performed with an ABI 377 DNA Sequencer. The computer program Sequencher v. 3.1 (Gene Codes Corporation, Ann Arbor, MI) was used to join contiguous 16S rDNA fragments into a single consensus sequence for each clone. A BLAST search was performed for each complete 16S sequence so that reference sequences could be chosen. When possible we used the same reference sequence as in previous analyses; however, this was not always possible. Our clones were aligned to reference sequences using ClustalX [26]. The cloned sequences for *Acinetobacter* and *Staphylococcus* were deposited in GenBank under the accession numbers AF467293–AF467306 and AF467402–AF467430, respectively.

We used the computer program ModelTest [18] as a guide to determine the best-fit maximum likelihood (ML) model as described by Cunningham et al. [2]. ModelTest examines maximum likelihood models ranging from simple to complex. This method increases the number of parameters in the ML model incrementally until the addition of a new parameter no longer increases significantly the fit between the model and the data. ModelTest calculated likelihood scores for 56 nested ML models and used hierarchical likelihood ratio tests (LRTs) to determine the best-fit model. We performed post-hoc LRTs to examine several ML models that were not evaluated by the program ModelTest. We incorporated the best-fit model of nucleotide evolution in PAUP\* [25] using the maximum likelihood optimality criterion. Multiple outgroups were chosen from well-established phylogenies of 16S rDNA sequences [12, 19].

We used PAUP\* to generate an initial neighbor joining tree (using the default settings) and estimated the ML parameters from that topology. Using these parameter estimates, we performed a heuristic search with random sequence addition ( $n = 1$ ) and tree-bisection-reconnection (TBR) branch swapping. We then re-estimated the ML parameters from the new tree topology. These values were used in another heuristic search with random sequence addition ( $n = 1$ ) and TBR branch swapping. Again we re-estimated the ML parameters from the new topology. This iterative method of refining the parameter estimates was repeated until the estimates remained unchanged in three

successive iterations. A final heuristic search was performed using random sequence addition ( $n = 10$ ) and TBR branch swapping.

## Results

### Identification of Bacterial Clones

The negative controls used during DNA extraction produced no PCR products, indicating that none of the extraction reagents contained external sources of bacterial DNA. The negative controls used during PCR amplification, which contained no DNA template, occasionally amplified 16S rDNA. In such instances all PCR reagents were replaced and PCR reactions were repeated until there

was no amplification for the negative control sample. Contamination was most often attributed to the diluted ( $10 \mu\text{M}$ ) primers used during PCR amplification. Solutions used to wash bacteria from the exterior surface never resulted in successful PCR amplification, cloning, and DNA sequencing, which suggests that the exterior surface of chewing lice harbors few, if any, bacteria. Scanning electron microscopy revealed no bacteria on the external surfaces of chewing lice (unpublished data).

DNA sequence fragments from the 5' end (ca. 500 bp) of the 16S rDNA ( $n = 95$ ) were added to the ARB database and analyzed using parsimony (Fig. 1) and maximum likelihood (Fig. 2) methods. The resulting phylogenies show that the bacteria amplified from chewing lice are

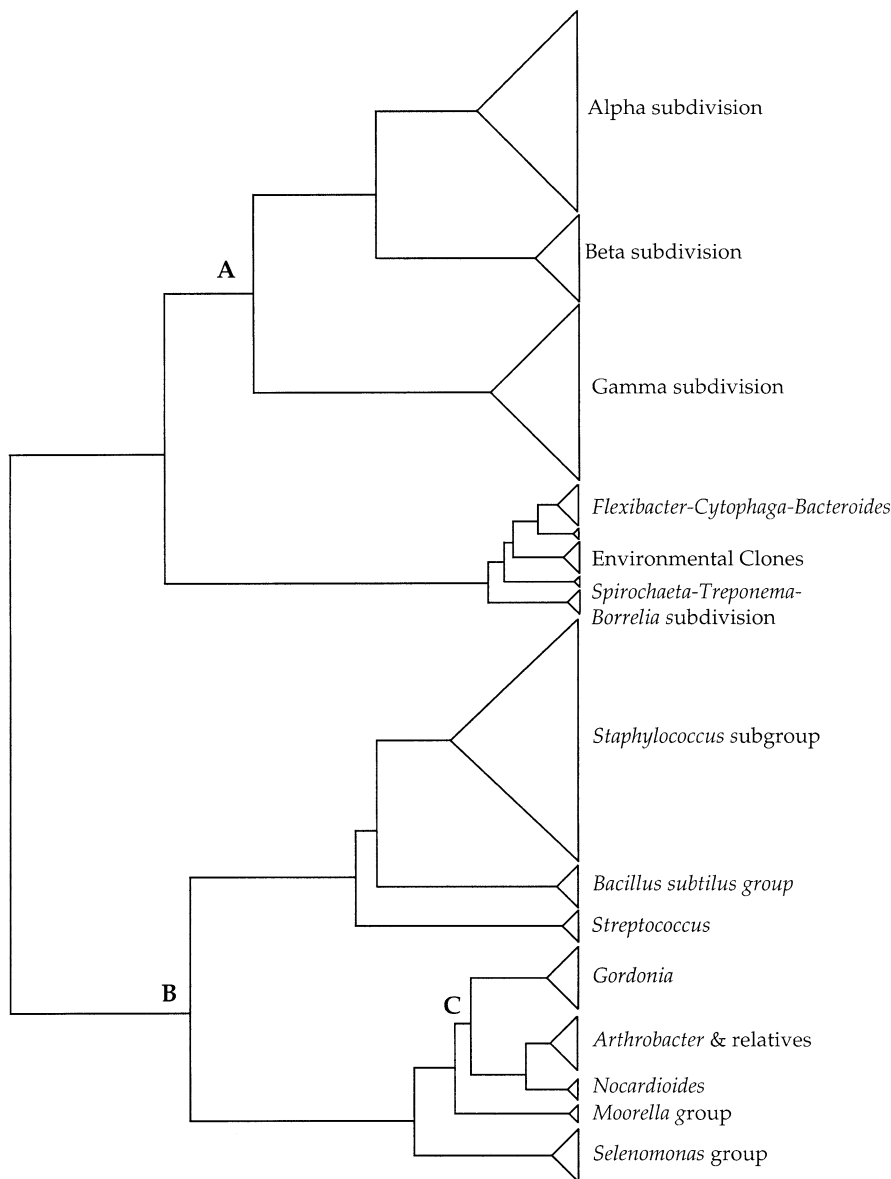


Fig. 1. Phylogenetic tree based on a parsimony analysis of partial 16S rDNA performed in the ARB computer package. Similar taxa have been grouped to show the major clades (*Proteobacteria* labeled A, *Firmicutes* labeled B, and *Actinomycetales* labeled C) of the 16S rDNA bacteria clones obtained from chewing louse hosts. The sizes of triangles are proportional to the number of species in each group (both clones and reference sequences from the ARE database).

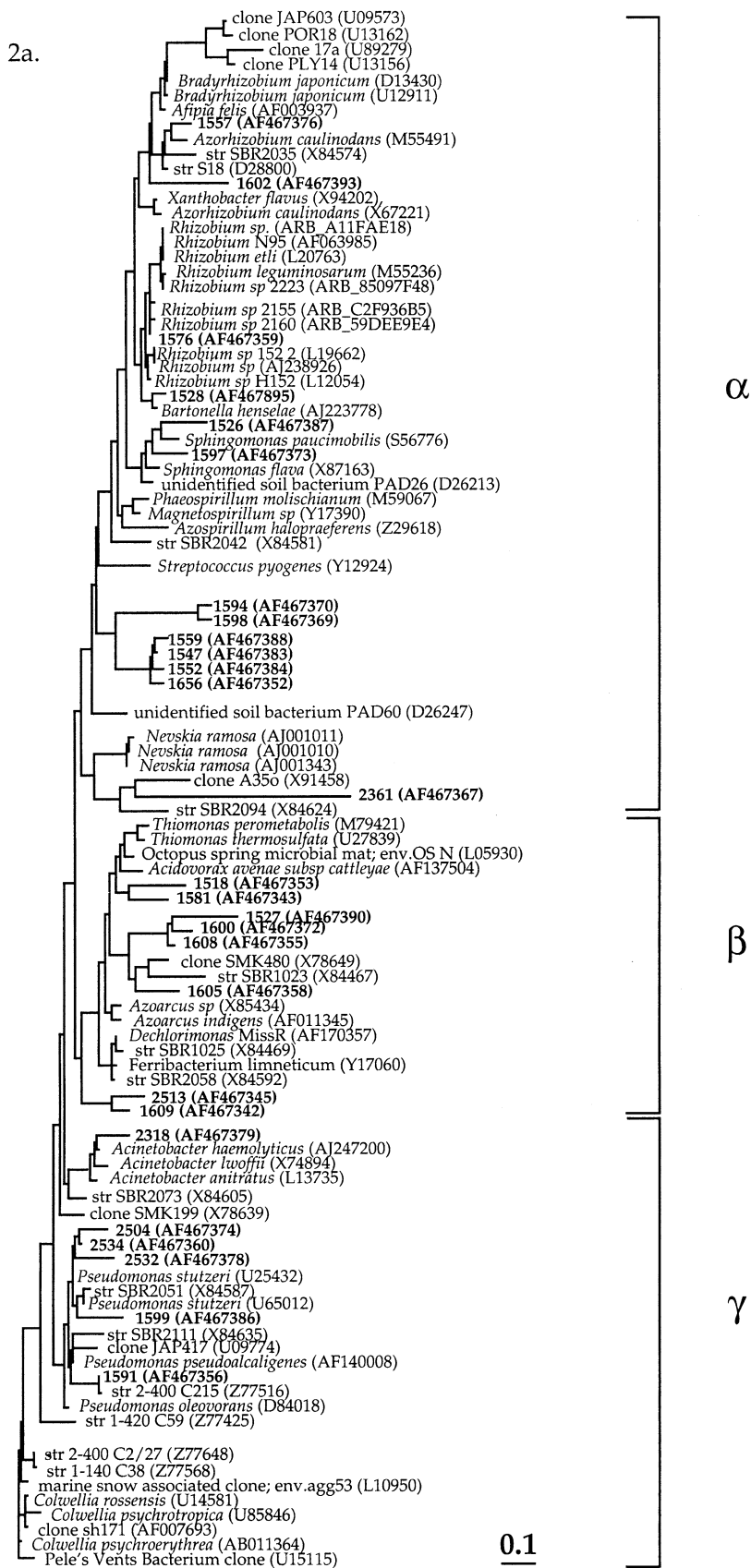


Fig. 2. Phylogenetic trees from maximum likelihood analyses (fastDNAmI in ARB program) showing affiliations of partial 16S rDNA clones (in bold) obtained from chewing lice with reference sequences (accession numbers in parentheses). (a)  $\alpha$ -,  $\beta$ -, and  $\gamma$ -proteobacteria, (b) environmental clones and allies, (c) species of *Staphylococcus*, *Bacillus*, and *Streptococcus*, and (d) Actinomycetales and the *Selenomonas* subgroup. Phylogenetic analyses were performed using multiple outgroup taxa from adjacent clades in the ARB database.

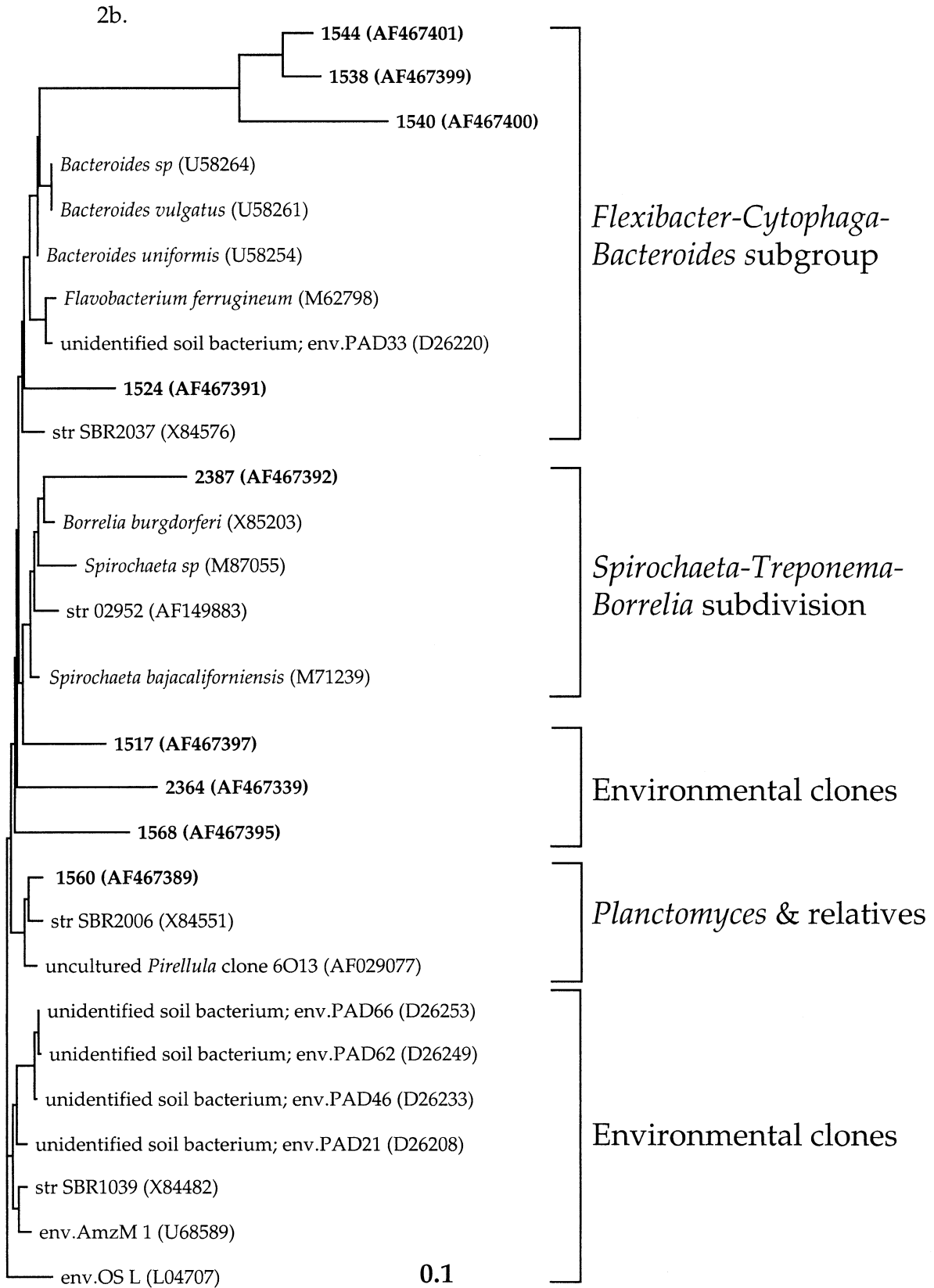


Fig. 2. Continued

2c.

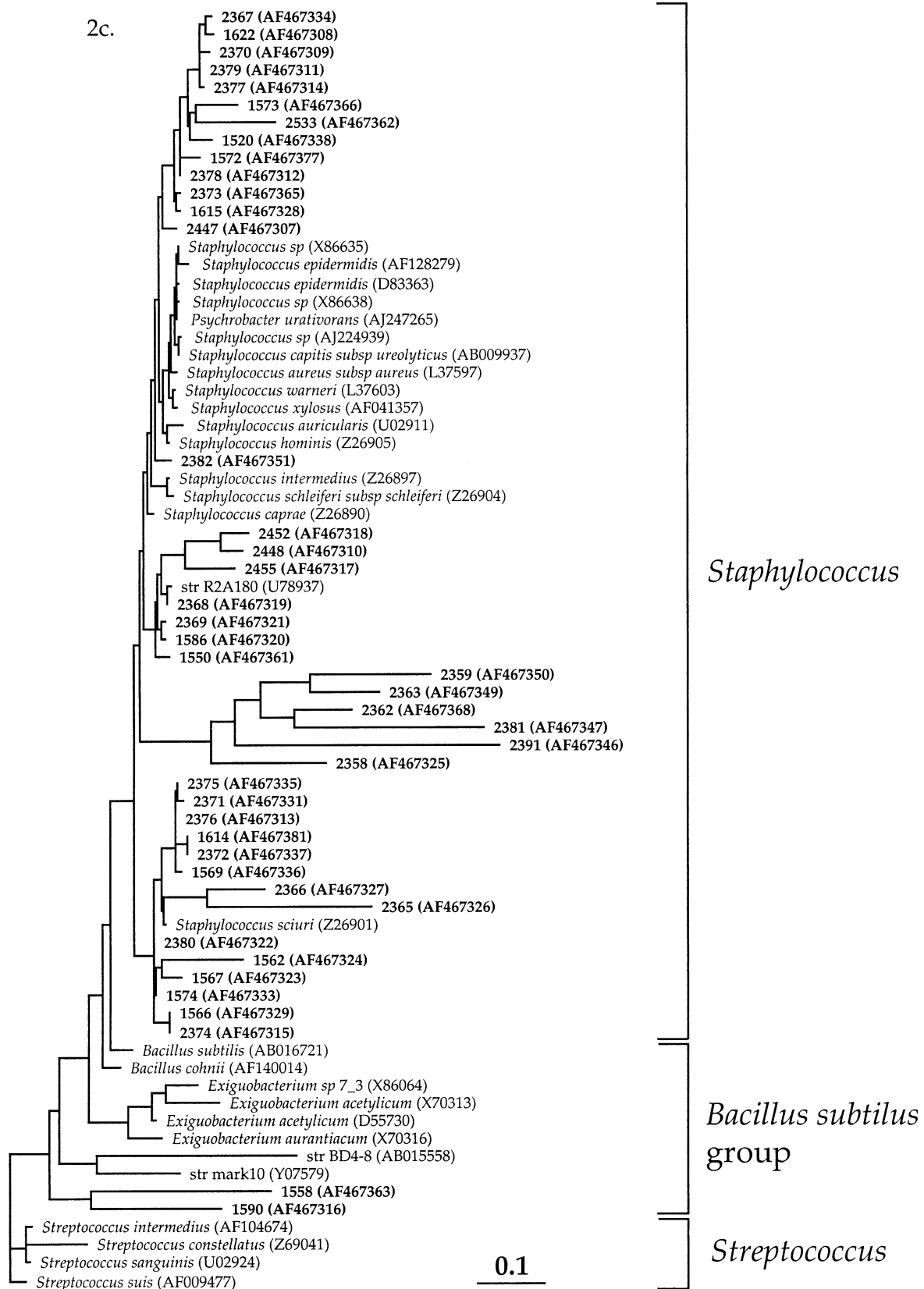


Fig. 2. Continued

2d.

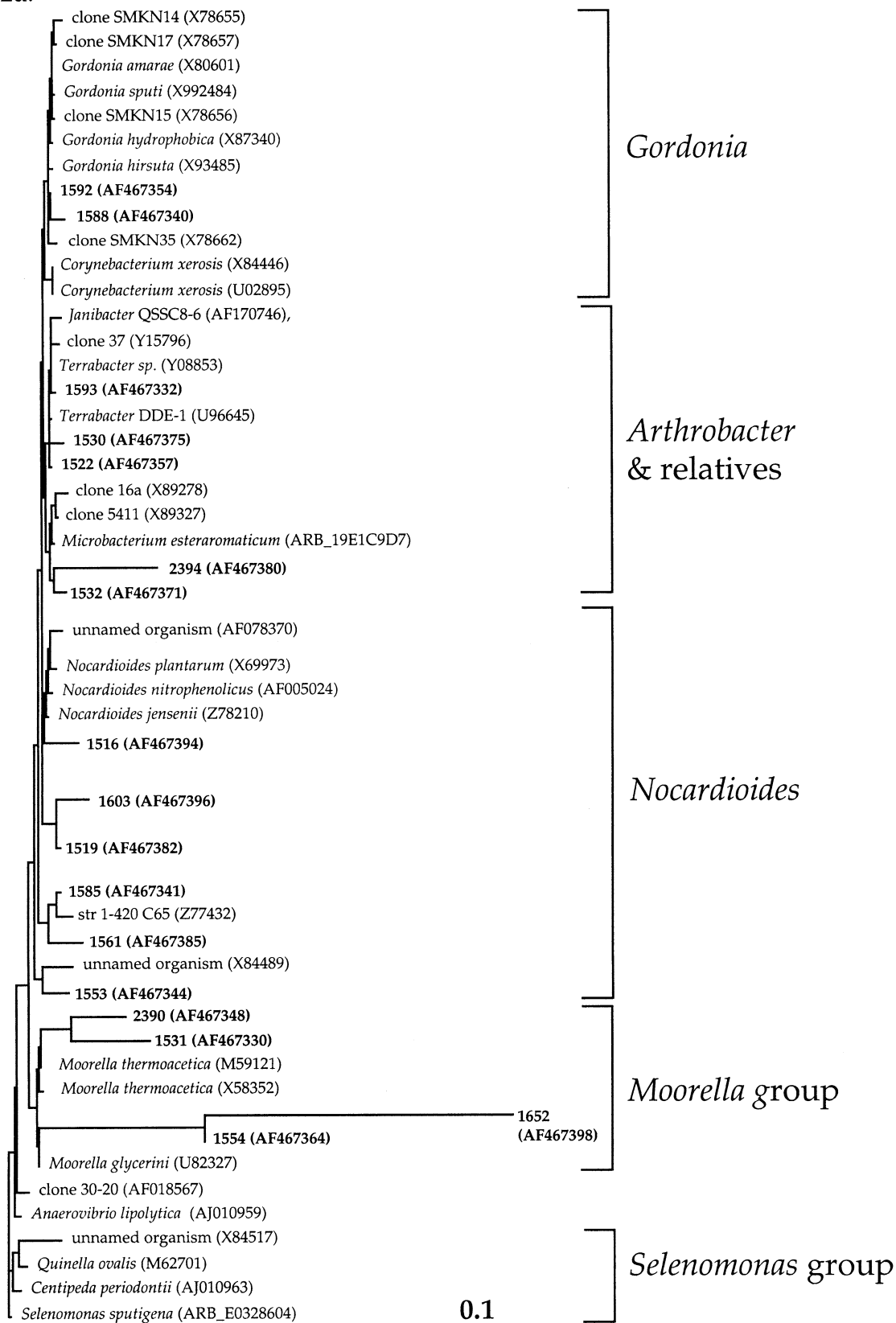


Fig. 2. Continued

taxonomically diverse (Figs. 1 and 2). From this analysis we determined that the majority of 16S rDNA clones grouped among Firmicutes (62.2%) and the *Proteobacteria* (31.6%). Three subgroups of the *Proteobacteria* were represented (alpha [13.7%], beta [8.4%], and gamma [6.3%]; Fig. 2a). Some of our 16S rDNA clones clustered with relatives of the Cytophaga–Flexibacter–Bacteroides (CFB) phylum (4.2%), Spirochaetes (1.1%), clades of environmental clones (3.2%), and Planctomycetales (1.1%; Fig. 2b). The *Bacillus subtilis* subgroup accounts for 1.1% of our clones (Fig. 2c). The remainder of our bacterial clones clustered with other Firmicutes taxa such as *Gordonia* (2.1%), *Arthrobacter* (5.3%), *Nocardioidea* (6.3%), and *Moorella* (4.2%; Fig. 2d). This phylogenetic method resulted in the identification of ca. 31 bacterial lineages.

BLAST searches of a larger dataset ( $n = 234$  partial 16S rDNA sequences) identified 35 lineages of bacteria in eight divergent groups of the domain Bacteria: CFB phylum, *Cyanobacteria*, *Fibrobacter/Acidobacteria* group, Firmicutes, *Planctomycetales*, *Proteobacteria*, *Thermus/Deinococcus* group, and *Verrucomicrobia* (Table 1). Most sequences belong to the *Proteobacteria* (59%) and the Firmicutes (34%) groups. Of all the subgroups, the  $\beta$ -*Proteobacteria* was represented most frequently (35% of clones).

Table 1 shows pocket gopher and chewing louse hosts, their general collection locality, and the number and type of bacteria found from BLAST searches. Many bacteria ( $n = 28$ ) were found infrequently, sometimes in only a single host (e.g., *Acidovorax*, *Brevibacterium*, and *Ensifer*). These may represent transient bacteria that are found locally in the gopher/louse habitat and were perhaps acquired through the louse's diet. It is also possible that these bacteria are more widespread in this system but have yet to be sampled in other taxa of lice. Twelve bacterial lineages (Table 1) were found in three or more species of chewing lice and may represent either geographically widespread bacteria or bacteria that are specifically associated with the gopher/louse system. Six lineages (e.g., *Alcaligenes*, *Azoarcus*, *Burkholderia*, *Methylobacterium*, *Propionibacterium*, and *Staphylococcus*) were found in four or more DNA extracts from chewing lice. *Burkholderia* and *Staphylococcus* clones had the highest prevalence and were found in six and seven of eight louse species, respectively (Table 1). Representatives of two bacterial genera, *Methylobacterium* and *Propionibacterium*, were found in five species of chewing lice. Two species of chewing lice (*Geomydoecus expansus* and *G. geomydis*)

hosted the greatest diversity of bacteria (19 and 20 lineages, respectively; Table 1). *Geomydoecus centralis* and *Thomomydoecus minor* hosted a combined total of only six lineages of bacteria, as did *G. costaricensis*. The predominance of only a few bacterial lineages in these hosts could be caused by PCR or DNA extraction bias, local extinction of other bacteria, or incomplete taxon sampling.

### Phylogenetic Analysis

16S clones identified as *Acinetobacter* and *Staphylococcus* based on BLAST searches of partial sequences ( $n = 37$ ) were chosen for further phylogenetic analysis. *Acinetobacter* was chosen because it is a member of the  $\gamma$ -*Proteobacteria*, which has given rise to several insect endosymbiont lineages [13]. *Staphylococcus* was chosen because it was found in seven of the eight species of chewing lice. The ModelTest program [18] identified the Tamura–Nei model of nucleotide evolution as the best-fit model for both the *Acinetobacter* and *Staphylococcus* datasets. The Tamura–Nei model allows for two rates of transitions ( $A \leftrightarrow G$  and  $C \leftrightarrow T$ ) and one rate for transversions, and it allows for unequal base frequencies. ModelTest determined that the addition of both an invariant sites parameter and a variable rate parameter (according to a gamma distribution) significantly increased the fit of the model.

The best-fit ML tree (Fig. 3) for the *Acinetobacter* dataset showed that the clones extracted from chewing louse samples were nested both within and outside the *Acinetobacter* sequences downloaded from GenBank. The uncorrected sequence divergence between pairs of taxa ranged from 0.0 to 18.9%. When complete 16S rDNA sequences from *Acinetobacter*-like clones were compared to sequences in GenBank, several clones were most closely related to the genera *Xanthomonas* and *Pseudomonas*, indicating a misidentification from earlier BLAST searches based on 500 bp sequence fragments. Species from these two genera were added to the phylogenetic analysis to determine the true placement of these clones. The best-fit ML tree (Fig. 4) for the *Staphylococcus* dataset showed that the clones extracted from chewing louse samples were nested both within and outside the *Staphylococcus* sequences downloaded from GenBank. The uncorrected sequence divergence between pairs of taxa ranged from 0.0 to 12.4% for the *Staphylococcus* taxa examined. Some of the *Staphylococcus* clones were closely related to known taxa whereas others seem to be phylogenetically distinct.

**Table 1.** Number of unique 16S rDNA bacterial sequences identified from 8 taxa of chewing lice by means of environmental extraction, PCR, sequencing, and searching GenBank databases via BLAST nucleotide recognition search protocol

Bacteria	Chewing Louse Host								
	<i>Geomydoecus expansus</i> <sup>1</sup>	<i>Geomydoecus geomydis</i> <sup>2</sup>	<i>Geomydoecus scleritus</i> <sup>3</sup>	<i>Geomydoecus panamensis</i> <sup>4</sup>	<i>Geomydoecus seizeri</i> <sup>5</sup>	<i>Geomydoecus costaricensis</i> <sup>6</sup>	<i>Thomomydoecus minor</i> and <i>Geomydoecus centralis</i> <sup>7</sup>	Number of clones	Number of hosts
CFB Group									
Unknown CFB				1				3	2
Cyanobacteria		2							
Oscillatoriales									
<i>Lyngbya</i>	1							1	1
Fibrobacter Group									
Unknown	1	5	1		1			8	4
Firmicutes									
Unknown			1					1	1
Firmicutes									
Actinobacteria									
Actinomycetales									
<i>Brevibacterium</i>	1							1	1
<i>Gordonia</i>		2						2	1
<i>Janibacter</i>		3				1		4	2
<i>Kitasatosporia</i>	1							1	1
<i>Micrococcus</i>	1	1						1	1
<i>Nocardioides</i>	1	1	2					4	3
<i>Propionibacterium</i>	4			5	2	2	3	16	5
Bacillus/Staph. Group									
<i>Facklamia</i>		1						1	1
<i>Leuconostic</i>			1					1	1
<i>Staphylococcus</i>	16	4		2	2	2	17	43	6
Clostridiaceae									
<i>Peptostreptococcus</i>	1				1			2	2
Streptococcaceae									
<i>Streptococcus</i>	2	1			2			5	3
Proteobacteria									
$\alpha$ -Proteobacteria									
Unknown $\alpha$ -Proteo. Rhizobiaceae		2	4					6	2
<i>Ensifer</i>							1	1	1
<i>Methylobacterium</i>	1	3		1	4	6		15	5
<i>Ochrobactum</i>		1				1		2	2
Rhodospirillaceae									
<i>Phaeospirillum</i>			1					2	2
<i>Rhodospirillum</i>			1					1	1



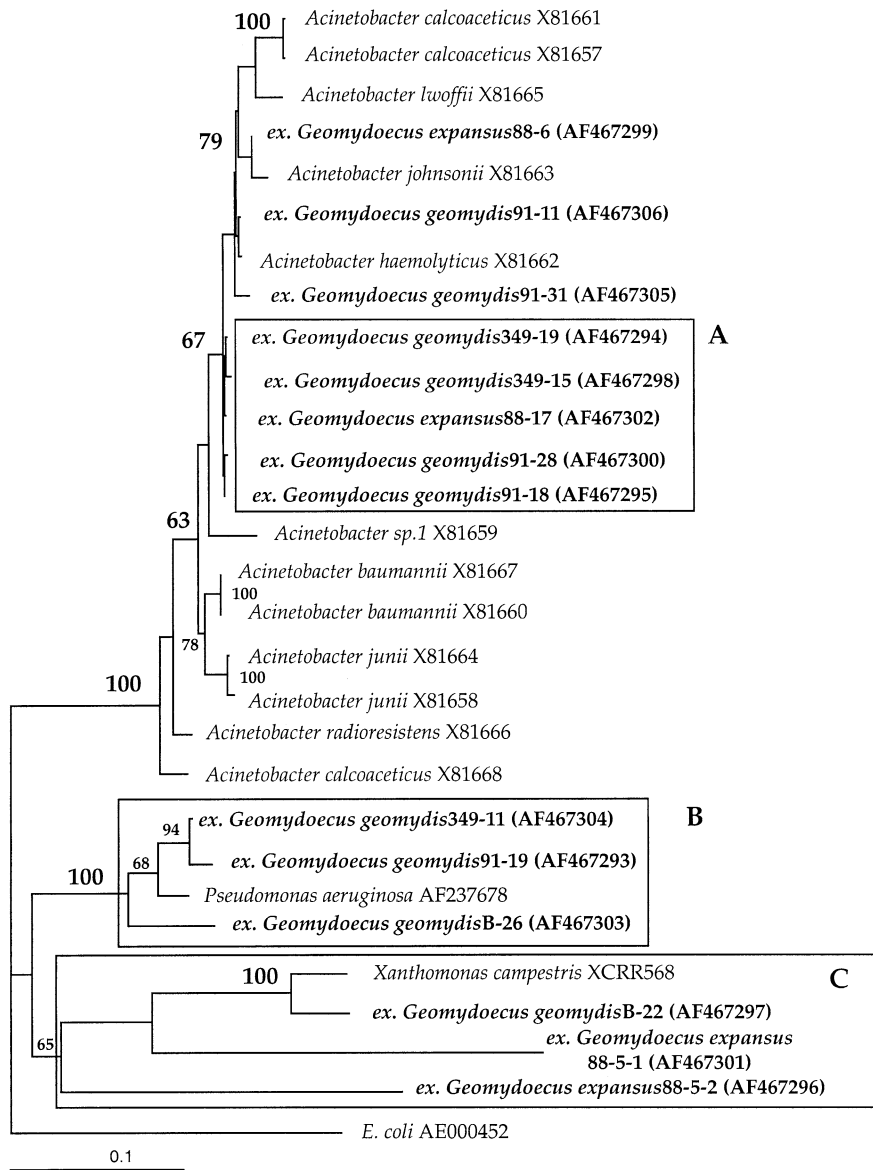


Fig. 3. Best-fit maximum likelihood phylogeny (based on complete 16S rDNA sequences) for  $\gamma$ -Proteobacteria using the Tamura-Nei + I + G model of nucleotide evolution. Bacterial taxa from chewing louse hosts are shown in boldface; other taxa were retrieved from GenBank (accession numbers in parentheses). *E. coli* was used to root the tree. Bootstrap values (1000 replicates) are based on the parsimony optimality criterion. Values less than 50% are not shown.

Several *Staphylococcus* clones found in *Thomomydoecus minor* are basal to the known *Staphylococcus* species (Fig. 4). When complete 16S rDNA sequences from these clones were compared to sequences in the NCBI database using BLAST searches, the most similar taxa were those downloaded for this phylogenetic analysis. In other words, these clones have 16S rDNA sequences that appear to represent *Staphylococcus* lineages distinct from those surveyed elsewhere and deposited in GenBank.

## Discussion

The phylogenetic approach to bacteria identification (e.g., using the ARB database) yields more reliable information

about bacterial relationships because the identification is based on phylogenetic information that takes into account evolutionary history. In contrast, the BLAST method identifies the bacterium based solely on local sequence similarity. One benefit of the BLAST method is that it generates an expectation value (E) that measures the level of similarity between the query sequence and the most similar sequences in a particular database. In addition, BLAST searches do not require aligned sequences from a single region of an amplified gene, as is required for any phylogenetic analysis. This allows the investigator to identify all clones sequenced (e.g., 234 sequence fragments in our study), rather than only those that overlap sufficiently for phylogenetic analysis (e.g., 95 sequences used in the computer program ARB in our study).

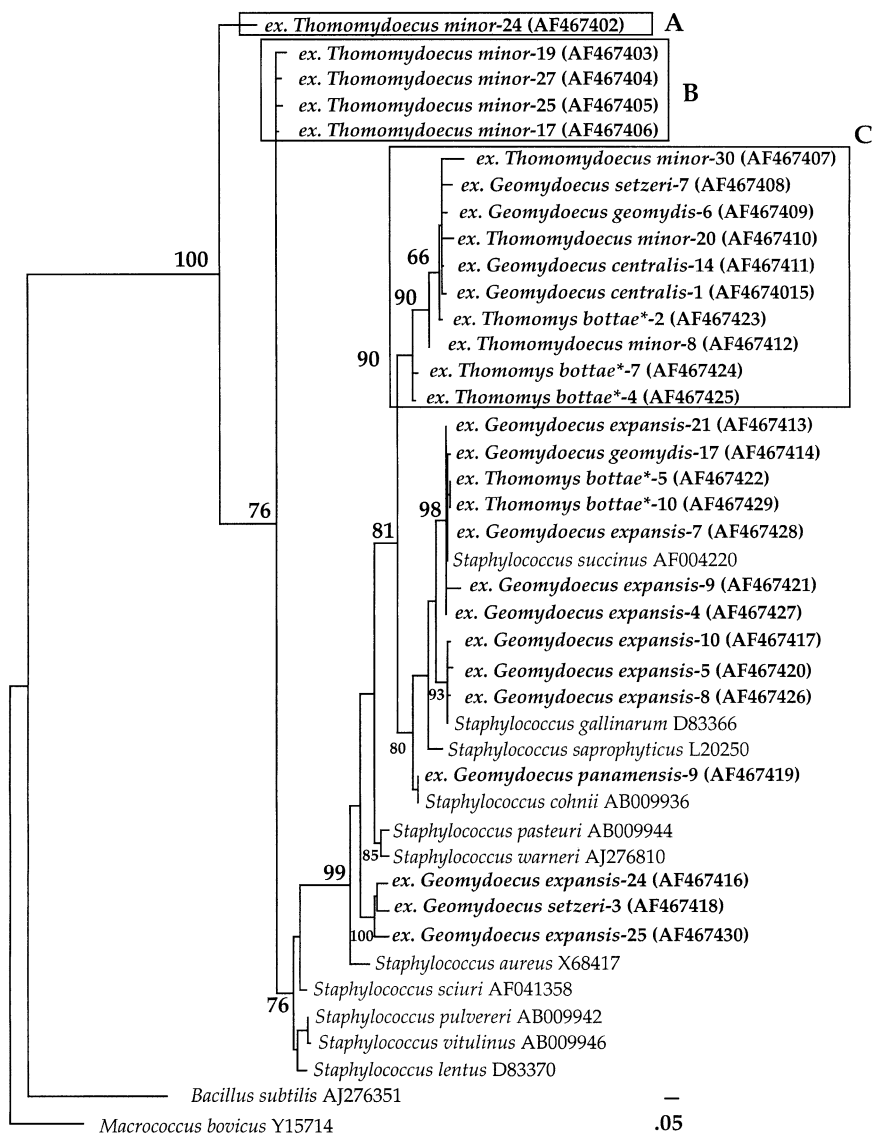


Fig. 4. Best-fit maximum likelihood phylogeny (based on complete 16S rDNA sequences) for *Staphylococcus* clones using the Tamura-Nei + I + G model. Bacterial taxa from chewing louse hosts are shown in boldface; other taxa were retrieved from GenBank (accession numbers in parentheses). Clones marked with an asterisk could be from either of two louse hosts, *Thomomydoecus minor* or *Geomydoecus centralis*, which occur on the same rodent host, *Thomomys bottae*. *Bacillus subtilis* and *Macroccoccus bovicus* were used as outgroup taxa. Bootstrap values (1000 replicates) are based on the parsimony optimality criterion. Values less than 50% are not shown.

Our survey of the bacteria associated with chewing lice identified a large and diverse assemblage of bacterial taxa (Fig. 1 and Table 1). Some of these taxa appeared to be confined to a single host taxon, whereas others were widespread among the hosts examined. However, these data (including both number of bacterial lineages per host and the host distribution of each bacterial taxon) should be interpreted with caution, considering that the survey was not exhaustive. Rarefaction curves (not shown) indicate that a much larger number of clones must be examined to ensure adequate sampling of the bacterial community associated with chewing lice. We propose that future surveys include a screening method, such as amplified rDNA restriction analysis (ARDRA), to increase the efficiency of finding novel 16S rDNA clones.

Some of the clones found in only one or two louse species may be local soil- or plant-associated bacteria incidentally ingested by chewing lice. In contrast, the more abundant and widespread clones may represent bacteria that are a part of the natural gut flora of the chewing louse and aid in digestion of the skin detritus consumed by the lice. It is also possible that endosymbiotic bacteria reside elsewhere in louse tissues and perform other functions. Thus far, electron microscopy has revealed no evidence of specialized bacterial cells, or mycetocytes; however, not all endosymbiotic bacteria of insects are membrane bound. Whereas we are reasonably certain that no bacteria were sampled from the exterior surface of the lice, we cannot determine where these bacteria reside within the lice. DNA sequencing from a wide variety of lice from birds and mammals have shown that several lineages of bacteria

seem to be widespread in lice, suggesting a long-term association (unpublished data). Future studies using *in situ* hybridization should provide information on localization and concentration of bacteria in chewing lice.

The wide range in number of clones recovered per bacterial taxon (1–43, Table 1) may represent real differences in bacterial abundance, but most likely is the result of either PCR amplification or DNA extraction bias. It is not possible to assess the relative abundance of bacterial species merely by counting the number of clones of a particular bacterium because it is known that bacterial DNA sequences occurring in low abundance in a DNA extract can amplify more rapidly than more abundant DNA sequences if the PCR conditions are better suited to the rarer sequence [24]. Apparent absence of a bacterium in a mixed-template PCR reaction may be an artifact of PCR bias as well. Although use of primers from conserved regions of the 16S rDNA gene permits study of bacterial diversity, conserved primers also introduce biases that must be considered when interpreting results of molecular surveys.

#### Phylogenetic Analyses of the $\gamma$ -Proteobacteria

Certain of our *Acinetobacter*-like clones grouped within the *Acinetobacter* species, although five of those clones (clade A, Fig. 3) appear to be somewhat distinct from species of *Acinetobacter* downloaded from GenBank. Three clones (clade B, Fig. 3) were closely related to *Pseudomonas aeruginosa* and three clones (clade C, Fig. 3) cluster with *Xanthomonas campestris*. Figure 3 was rooted with the outgroup taxon *E. coli*.

Bootstrap analysis (1000 replicates) using the parsimony optimality criterion shows some ambiguity in the relationships of these  $\gamma$ -Proteobacteria (Fig. 3). It appears that the low bootstrap support for some clades is a product of the nucleotide substitution patterns inherent in the ribosomal DNA sequences of these taxa. There are long regions of invariable sites punctuated by regions of high variability. Where nucleotide changes occur between taxa, the changes contain homoplasious information and provide limited phylogenetic signal. As a result, some clades remain unresolved. A genetic marker that evolves at a rate faster than the 16S rRNA gene may help to clarify these relationships.

#### Phylogenetic Analyses of *Staphylococcus*

Maximum likelihood analysis of *Staphylococcus*-like sequences showed that many of the clones extracted from

chewing lice were closely related to known *Staphylococcus* species, whereas other clones (clades A and B in Fig. 4) appeared as basal lineages relative to the known *Staphylococcus* species. It is interesting to note that the chewing louse host, *Thomomydoecus minor*, also is more deeply rooted with respect to other chewing lice and its mammal host, *Thomomys bottae*, is more deeply rooted with respect to other pocket gophers [4, 7]. Other *Staphylococcus* clones (clade C in Fig. 4) formed a monophyletic clade nested within the known *Staphylococcus* species. Without further taxon sampling, however, it would be premature to formally name these taxa based solely on their 16S rDNA sequences. Nevertheless, both of the basal lineages found on *Thomomydoecus minor* hosts (clades A and B) and this more-derived lineage (clade C) appear to be quite distinct from the taxa of *Staphylococcus* for which 16S rDNA sequences are known (Fig. 4).

Bootstrap analysis (1000 replicates) using the parsimony optimality criterion, shows moderate to strong support for most of the clades in the analysis of *Staphylococcus* clones (Fig. 4). It seems that the 16S rRNA gene is able to resolve the more divergent relationships in this clade, but it fails to resolve relationships among more closely related taxa. Again, a genetic marker that evolves at a rate faster than the 16S rRNA gene would be beneficial in future analyses of these bacteria.

---

## Conclusions

This study was the first to document the association of bacteria with chewing lice of the family Trichodectidae. We found at least 35 lineages of bacteria representing eight widely divergent groups within the domain Bacteria. Certain of these lineages were found in most chewing louse species examined, whereas others were found in only a single host species. Of the two methods used for taxonomic identification (phylogenetic analysis and BLAST searches) phylogenetic analysis resulted in a more complete identification of bacterial clones. However, fewer clones could be examined using the phylogenetic method, and as a result BLAST searches identified a greater diversity of bacterial types. Each method found approximately the same number of lineages; however, the lineages differed between analyses.

Phylogenetic analysis of complete 16S rDNA sequences for two genera (*Acinetobacter* and *Staphylococcus*) provides evidence that bacterial lineages may have diversified

within chewing lice. However, it is possible that lineages that appear to be unique to chewing lice are found in other habitats, but have no 16S rDNA sequences deposited in GenBank. Whereas the 16S rRNA gene is useful for identification of bacteria, it is likely not the best molecular marker for phylogenetic analyses, given that our phylogenetic trees were not well resolved for closely related taxa. Phylogenetic analyses of *Acinetobacter* (Fig. 3) and *Staphylococcus* (Fig. 4) revealed bacterial clades that appear distinct from known clades within each group.

This particular host–parasite system (gophers, lice, and bacteria) seems exceptionally promising for study of bacterial differentiation because of the high level of population isolation documented for the mammal and insect lineages. Future work should continue to characterize the microbial community associated with pocket gophers, chewing lice, and their subterranean burrow systems. Understanding the basic natural history of the bacteria associated with the gopher–louse system will enable future studies of cospeciation to focus on lineages of bacteria that have biologically meaningful interactions with chewing lice and pocket gophers.

---

## Acknowledgments

Fred Rainey and Naomi Ward assisted in the early development of this project. D. Lewis, M. Digby, J. Lyman, and R. Miller assisted in the PCR, cloning, and sequencing of the material presented. This work was supported, in part, by the National Science Foundation (DEB-9527583 and DEB-0075381 to MSH and DBI-0102112 to DLR).

---

## References

- Baumann P, Moran NA (1997) Non-cultivable microorganisms from symbiotic associations of insects and other hosts. *Antonie van Leeuwenhoek* 72:39–48
- Cunningham CW, Zhu H, Hillis DM (1998) Best-fit maximum likelihood models for phylogenetic inference: empirical tests with known phylogenies. *Evolution* 52:978–987
- Demastes JW, Hafner MS (1993) Cospeciation of pocket gophers and their chewing lice. *J Mamm* 74:521–530
- Douglas AE (1989) Mycetocyte symbiosis in insects. *Biol Rev Cambridge Phil Soc* 64:409–434
- Hafner MS, Nadler SA (1988) Phylogenetic trees support the coevolution of parasites and their hosts. *Nature* 332:258–259
- Hafner MS, Demastes JW, Spradling TA, Reed DL In press. Cophylogeny between pocket gophers and chewing lice. In: RDM Page (ed), *Tangled Trees: Phylogenies, Cospeciation, and Coevolution*. University of Chicago Press
- Hafner MS, Sudman PD, Villablanca FX, Spradling TA, Demastes JW, Nadler SA (1994) Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science* 265:1087–1090
- Hugenholtz P, Goebel BM, Pace NR (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* 180:4765–4774
- Ishikawa H (1989) Biochemical and molecular aspects of endosymbiosis in insects. *Int Rev Cytol* 116:1–45
- Jeyaprakash A, Hoy MA (2000) Long PCR improves *Wolbachia* DNA amplification: *wsp* sequences found in 76% of sixty-three arthropod species. *Insect Mol Biol* 9:393–405
- Komatsoulis GA, Waterman MS (1997) A new computational method for detection of chimeric 16S rRNA artifacts generated by PCR amplification from mixed bacterial populations. *Appl Environ Microbiol* 63:2338–2346
- Lambert LH, Cox T, Mitchell K, Rosselló-Mora RA, Del Cueto C, Dodge DE, Orkland P, Cano RJ (1998) *Staphylococcus succinus* sp. nov., isolated from Dominican amber. *Int J Syst Bacteriol* 48:511–518
- Moran NA, Telang A (1998) Bacteriocyte-associated symbiosis of insects. *Bioscience* 48:295–304
- Nardon P, Grenier AM (1989) CRC Handbook on endocytobiosis in Coleoptera: Biological, biochemical, and genetic aspects. In: W Schwemmler, G Gassner (eds) *Insect Endocytobiosis: Morphology, Physiology, Genetics, and Evolution*. Boca Raton, FL pp 175–216
- Pace NR (1997) A molecular view of microbial diversity and the biosphere. *Science* 276:734–740
- Page RDM (1990) Temporal congruence and cladistic analysis of biogeography and cospeciation. *Syst Zool* 39:205–226
- Patton JL, Smith MF, Price RD, Hellenthal RA (1984) Genetics of hybridization between the pocket gophers *Thomomys bottae* and *Thomomys townsendii* in northeastern California. *Great Basin Naturalist* 44:431–440
- Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817–818
- Rainey FA, Lang E, Stackebrandt E (1994) The phylogenetic structure of the genus *Acinetobacter*. *FEMS Microbiol Lett* 124:349–354
- Ries E (1931) Die symbiose der Läuse und Federlinge. *Zeitschr Morph Ökol Tiere* 20:233–367
- Robison-Cox JF, Bateson MM, Ward DM (1995) Evaluation of nearest-neighbor methods for detection of chimeric small-subunit rRNA sequences. *Appl Environ Microbiol* 61:1240–1245
- Saffo MB (1992) Invertebrates in endosymbiotic associations. *Am Zoologist* 32:57–565
- Saxena AK, Agarwal GP (1985) Mycetocytes in *Aegypocercus perspicuus* (Phthiraptera). *Curr Sci* 54(15):763–764
- Suzuki MT, Giovannoni SJ (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl Environ Microbiol* 62:625–630

25. Swofford DL (2001) PAUP\*. Phylogenetic Analysis Using Parsimony (and Other Methods). Version 4. Sinauer Associates, Sunderland, MA
26. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 24:4876–4882
27. Timm RM (1983) Farenholz's Rule and resource tracking: A study of host–parasite coevolution. In: M Nitecki (ed) *Coevolution*. University of Chicago Press, Chicago, pp 225–266
28. Wang GC-Y, Wang Y (1997) Frequency of formation of chimeric molecules as a consequence of PCR coamplification of 16S rRNA genes from mixed bacterial genomes. *Appl Environ Microbiol* 63:4645–4650
29. Ward DM, Ferris MJ, Nold SC, Bateson MM (1998) Natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbial Molec Biol Rev* 62:1–18
30. Woese CR (1994) There must be prokaryote somewhere: microbiology's search for itself. *Microbiol Rev* 58:1–9