Chapter 4

ANALYSIS OF [FeFe] HYDROGENASE SEQUENCES FROM THE HYDROGEN RICH GUTS OF HIGHER TERMITES REVEALS CORRELATION BETWEEN GUT ECOSYSTEM PARAMETERS AND SEQUENCE COMMUNITY COMPOSITION

Abstract

Hydrogen is the central free intermediate in the degradation of wood by termite gut microbes and can reach concentrations exceeding those measured for any other biological system. Degenerate primers targeting the largest family of [FeFe] hydrogenases observed in a termite gut metagenome (Warnecke, F., *et al.* 2007. Nature 450: 560-569) have been used to explore the evolution and representation of these enzymes in termites. Sequences were cloned from the guts of the higher termites *Amitermes* sp. Cost010, *Amitermes* sp. JT2, *Gnathamitermes* sp. JT5, *Microcerotermes* sp. Cost008, *Nasutitermes* sp. Cost003, and *Rhyncotermes* sp. Cost004. Each gut sample harbored a more rich and evenly distributed population of hydrogenase sequences than observed previously in the guts of lower termites and *C. punctulatus* (see Chapter 3). This accentuates the physiological importance of hydrogen to higher termite gut ecosystems and may reflect an increased metabolic burden imposed by a lack of gut protozoa. The sequences were phylogenetically distinct from previously sequenced [FeFe] hydrogenases. Phylogenetic and Unifrac comparisons revealed congruence between host phylogeny and hydrogenase sequence library clustering patterns. This may reflect the combined influences of the stable intimate relationship of gut microbes with their host and environmental alterations in the gut that have occurred over the course of termite evolution. Interestingly, host feeding habits were similarly observed to correlate with sequence library clustering in Unifrac. These results accentuate the physiological importance of hydrogen to termite

gut ecosystems and imply that gut microbes of wood feeding insects may have "co-evolved" with their hosts.

Introduction

Hydrogen plays a pivotal role in the digestion of wood by termites (3, 8, 16, 38, 39, 42).

Concentrations in the guts of some species can reach concentrations exceeding those

measured for any other biological system (16, 20, 42, 43, 46, 48, 49). The turnover of the

gas in the gut has been measured in some species at rates as high as 33 m^3 H$_2$ per m^3 gut

volume per day (42). The environment is also spatially complex, comprising a matrix of

microenvironments characterized by different hydrogen concentrations (10, 11, 16, 27,

28, 42).

This hydrogen is produced during the fermentation of lignocellulosic polysaccharides by

the symbiotic microbial community residing in the termite gut (19, 21, 22, 38, 39, 52, 56,

57). The termites are dependent upon this complex symbiosis for the degradation of

wood (4-6, 9, 14, 15, 38). The primary product of this symbiosis is acetate, which the

termites use as their primary carbon and energy source (40). The majority of the

hydrogen in the gut is used by bacteria in reductive acetogenesis to produce up to 1/3 of

this acetate (3, 8, 29, 40, 42). A small portion of the hydrogen in the gut is used by

methanogenic archaea (3, 27, 42).

Termites can be classified as belonging to one of two phylogenetic groups, higher

termites or lower termites (25). Higher termites characteristically lack protozoa in their

guts, which are abundant in the guts of lower termites, and have more highly segmented

gut structures than lower termites (14, 36, 37). Of the over 2600 known species of

termites, over 70% are higher termites (25, 55). They represent the largest and most

diverse group of termites (24, 55). Yet, most of what we know about termite gut

microbes comes from work done with lower termites and comparatively little work has

been done with the communities of higher termites (4, 5, 7, 9). The primary reason for this is that it was believed until recently that the gut microbes of higher termites played only a minor role in wood digestion (47, 51, 53). This changed with the recent publication of the gut metagenome of a higher termite where it was found that the gut community encodes genes for reductive acetogenesis, polysaccharide degradation, and an abundance of [FeFe] hydrogenases, all pointing in the direction of a more active role in wood degradation (53). This previously under-acknowledged role for the gut microbes has also found support in the findings of Toduda and Watanabe (51).

Wood feeding insects have shared a stable and intimate mutualism with their respective gut microbial communities over the course of their evolution (54). The composition of these communities has been shown to vary substantially with host feeding habits (35, 45, 50). Interestingly, a study on the distribution of formyltetrahydrofolate synthetase (FTHFS) genes in the guts of higher termites has provided evidence that feeding habits have an important influence on community composition (41). Moreover, it has been proposed that the gut microbes of lower termites and *Cryptocercus* may "co-evolve" with their respective hosts (1, 12, 13, 17).

Here we report a phylogenetic analysis of [FeFe] hydrogenase genes cloned from the guts of higher termites. The objective was to better understand the diversity, adaptation, and evolution of the genes in these hydrogen-metabolizing ecosystems. Moreover, the influence of host ecosystem variations on the hydrogenase sequence composition of their associated microbial communities was investigated through cross-comparisons with sequence libraries reported previously for lower termite and wood-roach samples (see Chapter 3).

Methods

Termites. *Nasutitermes* sp. Cost003 and *Rhyncotermtes* sp. Cost004 were collected in the INBIO forest preserve in Guápiles, Costa Rica. Cost003 was collected at a height of 1.2 m from a *Psidium guajaba* tree and was believed to be feeding on deadwood. Cost004 was collected from a nest located under a Bromeliad. Feeding trails leading from this nest to a pile of decaying wood and plant material suggested litter feeding. *Microcerotermes* sp. Cost008 was collected from the base of a palm tree about 100 m from the beach at Cahuita National Park in Costa Rica, and appeared to be feeding on the palm tree. *Amitermes* sp. Cost010 was collected from the roots of dead sugar cane plants at a plantation in Costa Rica. *Amitermes* sp. JT2 and *Gnathamitermes* sp. JT5 were collected from subterranean nests at Joshua Tree National Park (Permit#: JOTR:2008-SCI-002).

Termites were identified in a previous study (41) using insect mitochondrial cytochrome oxidase subunit II (COXII) gene sequences and morphology. The COXII genes were amplified directly from the DNA samples that hydrogenases were cloned from. COXII was amplified using the primers CI-J-1773 and B-tLys and cycling conditions described by Miura *et al.* (34) FailSafe PremixD (Epicentre) and Expand High Fidelity Taq (Roche) were substituted for the polymerase and buffers, respectively. Sequences were edited and analyzed in a manner analogous to that described below for cloned [FeFe] hydrogenase sequences. Samples were identified as belonging to the genus of the termite harboring harboring the COXII sequence to which they were found most near in phylogenetic analyses.

DNA Extraction and Cloning. DNA was extracted from whole dissected guts and quantitated as described previously (33) and in Chapter 3. Degenerate primers designed

in Chapter 3 for the specific amplification of Family 3 [FeFe] hydrogenases were used for the cloning of gut sequences as described there. Family 3 [FeFe] hydrogenases, first defined by Warnecke *et al.*, were the most highly represented group of enzymatic hydrogenases observed in the *Nasutitermes* hindgut metagenome sequence (53). Family 3 [FeFe] hydrogenases were the only group of hydrogenases observed in the *Nasutitermes* hindgut metagenome whose *in situ* translation was verified by MS (53). The degenerate primer sequences, which were ordered from IDT DNA, were WSICCICARCARATGATGG and CCIIKRCAIGCCATIACYTC for the forward and reverse primers, respectively, where "I" represents inositol.

RFLP Analysis and Sequencing. For each termite gut, 96 clones were randomly selected for RFLP analysis as described previously in Chapter 3. Sequences representing each unique RFLP pattern observed were arbitrarily selected and submitted for sequencing, as described previously in Chapter 3. The sequences obtained were manually trimmed in SeqMan, available from DNA* as part of the Lasergene software suite, to remove the plasmid and degenerate primer sequences.

The identity of each sequence as a hydrogenase was verified using by BLASTing it against GeneBank. Sequences that did not have hydrogenases as the top hits were not included in further analyses. Also, sequences that in subsequent analyses aligned poorly with other cloned hydrogenase sequences were re-sequenced and analyzed manually for frame-shift mutations if they continued to align poorly or contain internal stop codons. Frame-shift mutations were identified and manually corrected at the DNA level for three clones based upon sequence alignments and careful inspection of sequencer trace files, see footnotes to Table 4-S1 in this chapter's appendix.

Phylogenetic Analysis. An operational taxonomic unit (OTU) was defined as those peptide sequences sharing a minimum of 97% sequence identity. Sequences were grouped into OTUs using the furthest-neighbor algorithm in DOTUR (44).

The ARB software environment (32) was used for phylogenetic analysis of hydrogenase sequences, which was completed as described previously in Chapter 3. Cloned sequences and their OTUs used in these analyses are listed in Table 4-S1 in the appendix to this chapter. Trees were constructed using 173 unambiguously aligned amino acid positions with distance matrix (Fitch), maximum parsimony (Phylip PROTPARS), and maximum likelihood (PhylipPROML) treeing methods. The following sequences comprised the outgroup used to construct Figures 4-2 and 4-3: *Pseudotrichonympha grassii* (AB331668); uncultured parabasilid (AB331670); *Holomastigotoides mirabile* (AB331669), *Pseudotrichonympha grassii* (AB331667), *Treponema primitia* ZAS-1 (HndA1, accession), *T. primitia* ZAS-2 (HndA2, accession), *T. primitia* ZAS-2 (HndA3, accession), *T. primitia* ZAS-1 (HydA1, accession). The following Family 3 [FeFe] hydrogenase sequences reported in Chapter 2, were also used to construct Figures 4-2 and 4-3: *Treponema primitia* strain ZAS-2 (HndA1, Chapter 2); *Treponema azotonutricium* strain ZAS-9 (HndA, Chapter 2).

Diversity and Sequence Richness Calculations. Chao1 sequence richness and Shannon diversity indices for each clone set were calculated using EstimateS version 8.0.0 for Macintosh computers, written and made freely available by Robert K. Colwell (http://viceroy.eeb.uconn.edu/EstimateS). OTUs and their respective sequence abundances were used as inputs to the program.

Community Comparisons. Unifrac (31) was used for quantitative comparisons of the higher termite [FeFe] hydrogenase sequences with each other or with those from lower termites and *C. punctulatus* reported previously in Chapter 3. Maximum likelihood trees were constructed according to the methods described above and subsequently used as the input for Unifrac analyses. 173 unambiguously alighted amino acids were used in treeing calculations. Each termite or *C. punctulatus* sequence library was designated as a unique environment. The number of cloned sequences represented by each OTU was input to Unifrac to be used for calculating abundance weights. The environments were compared using the Unifrac jackknife and principle component analyses. Normalized abundance weights were used in all calculations. The jackknife calculation was completed with 1000 samplings and using 75% of the OTUs contained in the smallest environment sample as the minimum number of sequences to keep.

Results

Sequences cloned. Hydrogenase sequences representing as many as 44 sequence OTUs were cloned from each of the higher termites, see Table 4-1. Table 4-S1 in the chapter's appendix lists all clones, and their corresponding OTUs, analyzed in this study. The collector's curves for each sequence library are provided as Figure 4-1. Microcerotermes was the only sample having 75% of all cloned sequences distributed among less than 7 OTUs. The Shannon diversity index and the Chao1 species richness index for each sequence library are listed in Table 4-1.

Phylogenetic analysis. In phylogenetic analyses comparing the cloned sequences to publically available [FeFe] hydrogenase sequences in our database, all but one (see

Table 4-1. Quantifying hydrogenase clone library diversity.

	RFLPs[a]	OTUs[b]	Chao1 Mean[c]	Chao1 95% CI Lower Bound[c]	Chao1 95% CI Upper Bound[c]	Shannon Mean[d]
Amitermes sp. Cost010	60	31	45	35.02	79.77	2.97
Amitermes sp. JT5	33	22	32.67	24.18	74.18	2.65
Gnathamitermes sp. JT5	44	30[e]	40.29	32.75	68.49	3.05
Microcerotermes sp. Cost008	36	21	29.1	22.84	56.57	2.33
Nasutitermes sp. Cost003	38	25	43	29.54	96.38	2.69
Rhyncotermes sp. Cost004	54	44	68.05	52.73	110.24	3.53

[a]Number of unique restriction fragment polymorphism patterns (RFLPs) observed.

[b]Number of operational taxonomic units (OTUs); calculated using the furthest-neighbor method and a 97% amino-acid sequence similarity cut-off.

[c]Chao1 species-richness index calculated using the classic method in EstimateS. OTUs representing Family 3 [FeFe] hydrogenases with their respective abundances were used as program inputs.

[d]Shannon diversity index calculated using EstimateS. OTUs representing Family 3 [FeFe] hydrogenases with their respective abundances were used as program inputs.

[e]One of these OTUs represented Family 7 [FeFe] hydrogenase sequences (see Table 4-S1 in the chapter's appendix) and was not used in the calculation of the diversity indices.

Figure 4-1.

Figure 4-1. Collector's curves. The horizontal brackets in each figure indicate the number of OTUs comprising 75% of all sequences cloned. Each bin represents an OTU calculated with the furthest-neighbor method a 97% amino-acid similarity cut-off using DOTUR (44). A) *Amitermes* sp. JT5, B) *Gnathamitermes* sp. JT5, C) *Microcerotermes* sp. Cost008, D) *Amitermes* sp. Cost010, E) *Rhyncotermes* sp. Cost004, F) *Nasutitermes* sp. Cost003.

footnote to Table 4-S1 in the chapter's appendix) formed a single large clade to the exclusion of all non-termite bacterial sequences, data not shown. Within this clade were included Family 3 [FeFe] hydrogenase sequences from a *Nasutitermes* gut metagenome (53) and from the genome sequences of two treponemes isolated from *Zootermopsis angusticolis*, *T. primitia* ZAS-2 and *T. azotonutricum* ZAS-9 (see Chapter 2), data not shown. A maximum likelihood tree for all of the cloned [FeFe] hydrogenase sequences is provided as Figure 4-2.

Upon a cursory inspection of phylogenetic groupings, the hydrogenase sequences appeared to cluster in a manner roughly congruent with the phylogeny of their hosts. For example, both *Amitermes* samples tended to group with each other in phylogenetic analyses. *Gnathamitermes* and *Amitermes* were the only higher termite samples analyzed in this study whose COII sequences formed a tight, coherent clade with each other in phylogenetic analyses (41). Interestingly, hydrogenase sequences from these samples tended to group with one another as well. Moreover, hydrogenase sequences from a given termite sample tended to cluster with each other.

Sequence library cross-comparisons. A maximum likelihood tree comparing all of the Family 3 hydrogenases cloned from the higher termite samples to those cloned previously from *C. punctulatus* and lower termite gut samples (see Chapter 3) is provided as Figure 4-3. There is a clear separation between the higher termite sequences and those from *C. punctulatus* and lower termites. The latter two groups of sequences appear to intermingle with each other in the tree. This apparent congruence between the phylogenetic clustering of the cloned hydrogenases and that of their respective hosts was much more striking in the Unifrac jackknife clustering of the samples, see Figure 4-4. In this

Figure 4-2.

ZAS-9 (HndA)

ZAS-2 (HndA1)

Outgroup

0.10 (changes / position)
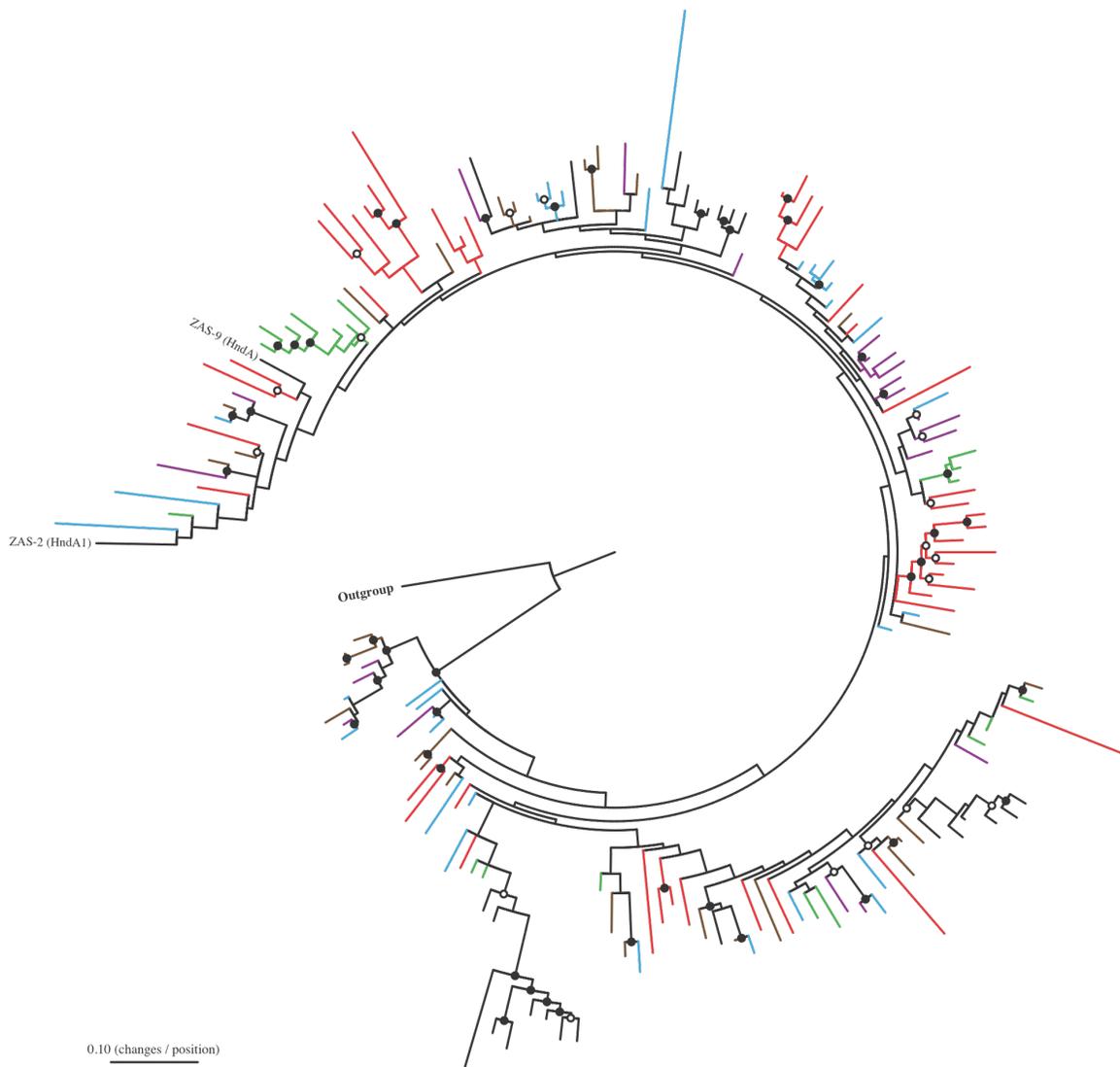
Figure 4-2. Phylogram for Family 3 [FeFe] hydrogenases cloned from the guts of higher termites. The tree was calculated using a maximum likelihood (Phylip ProML) method with 173 unambiguously aligned amino acid positions. Open circles designate groupings also supported by either parsimony (Phylip PROTPARS, 1000 bootstraps) or distance matrix (Fitch) methods. Closed circles designate groupings supported by all three methods. Each leaf represents an OTU. Leaves and branches representing OTUs cloned from *Amitermes* sp. Cost010 = blue; *Amitermes* sp. JT2 = purple; *Gnathamitermes* sp. JT5 = brown; *Microcerotermes* sp. JT5 = green; *Nasutitermes* sp. Cost003 = black; *Rhyncotermes* sp. Cost004 = red. Hydrogenase sequences taken from *T. primitia* ZAS-2 and *T. primitia* ZAS-9 are labeled as ZAS-2 (HndA1) and ZAS-9 (HndA), respectively. Tree drawn using Phylip drawgram (18).

Figure 4-3.

0.10 (changes / position)
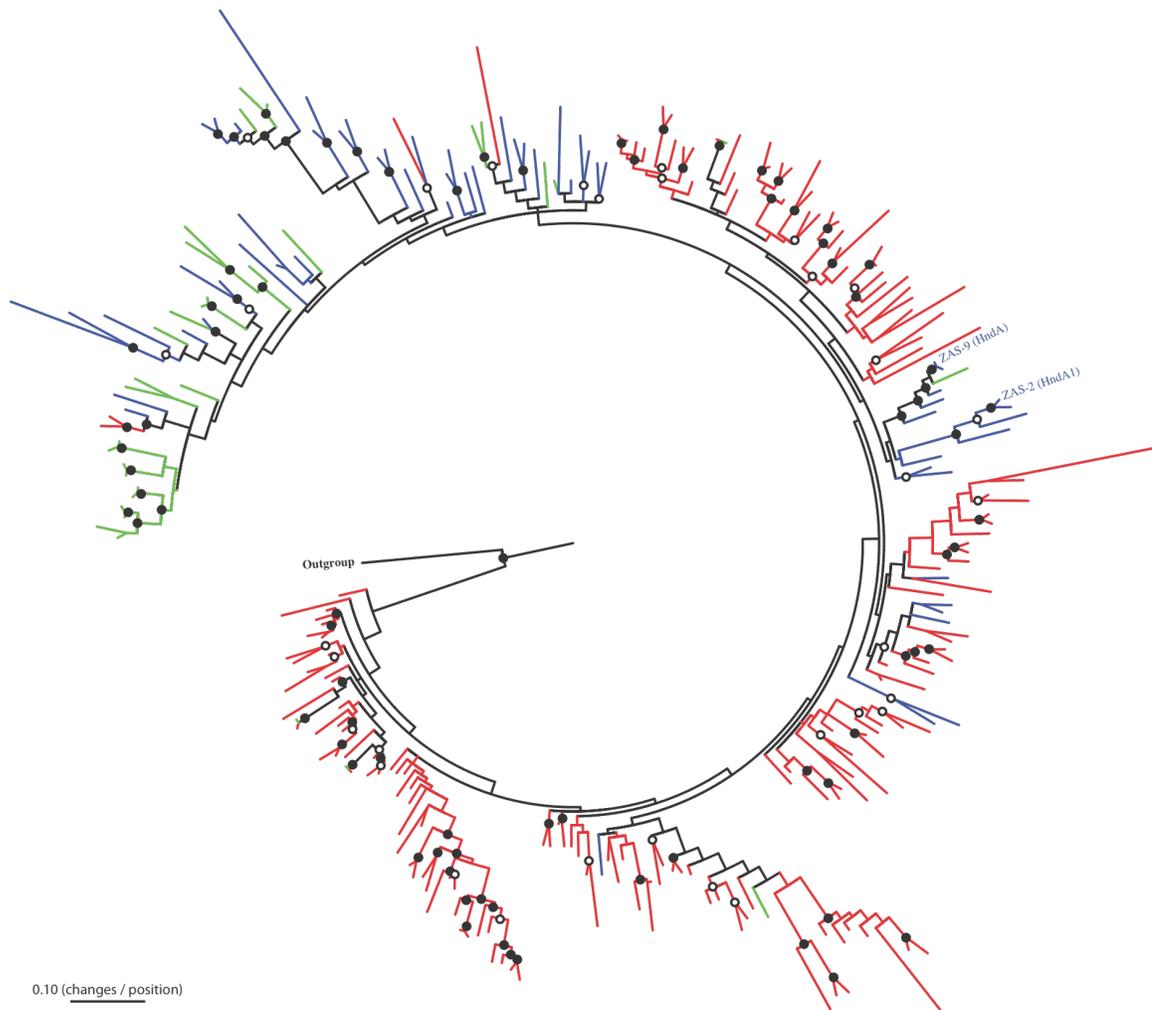
Figure 4-3. Phylogram comparing Family 3 [FeFe] hydrogenases cloned from higher termites to sequences cloned previously from *C. punctulatus* and lower termites. See Figure 4-2 caption for description of open and closed black circles and tree construction methods. Each leaf represents an OTU. Leaves and branches representing OTUs cloned from lower termites are in blue, from *C. punctulatus* are in green, and from higher termites are in red. Hydrogenase sequences taken from *T. primitia* ZAS-2 and *T. primitia* ZAS-9 are labeled as ZAS-2 (HndA1) and ZAS-9 (HndA), respectively. Tree drawn using Phylip drawgram (18).

Figure 4-4.
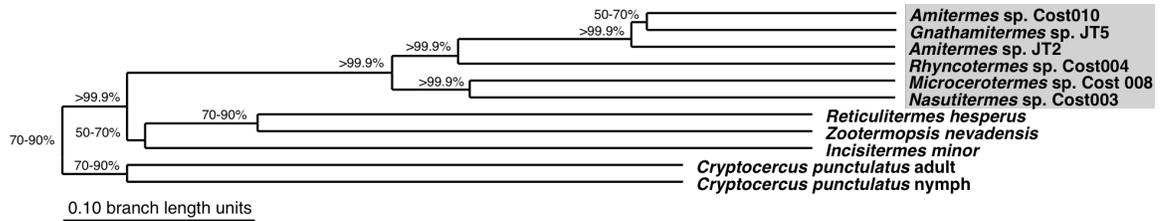
Figure 4-4. Unifrac jackknife analysis of Family 3 [FeFe] hydrogenase sequences cloned from higher termites, lower termites, and *C. punctulatus*. The maximum-likelihood tree shown in Figure 4-3 and the OTUs with their respective abundance weights listed in Table 4-S1 of the appendix to this chapter and taken from Table 3-S1 in the appendix to Chapter 3 were used as inputs to Unifrac. The analysis was completed using normalized abundance weights, 1000 samplings, and keeping a number of sequences equal to 75% of the number of OTUs represented by the smallest sample analyzed. Each insect sample was designated as a unique environment. The grey box highlights all higher termite environments. The numbers designate the percentage of samplings supporting a particular cluster.

analysis, the clustering of the hydrogenase sequences was congruent with the phylogeny of their respective hosts reported by Legendre *et al*. and Inward *et al*. (23, 24, 30). This clustering was further supported by the Unifrac PCA analysis of the sequences, see Figure 4-5. In the PCA analysis, there is a distinguishable separation between sequences from each of the three groups representing higher termites, lower termites, and *C. punctulatus*. Principle component 1, which accounted for the separation of higher termites from lower termites and *C. punctulatus*, explained 34.87% of the variation.

A Unifrac principle component analysis of the [FeFe] hydrogenase sequences cloned from higher termites is provided as Figure 4-6. Sequences from *Amitermes* sp. Cost010, *Amitermes* sp. JT5, and *Gnathamitermes* sp. JT5 clustered together. These termite samples are unique from the others because of their close phylogenetic relationship to one another, as discussed above, and because they were collected from sub-terranean nests. These samples could be distinguished from the others according to principle component 1, which explained 30.68% of the variation.

Discussion

High [FeFe] hydrogenase sequence diversity in higher termites. The abundance of [FeFe] hydrogenases cloned from the guts of higher termites, representing as many as 45 OTUs in the case of *Rhyncotermes* sp. Cost004, emphasizes the physiological importance of these enzymes to these complex ecosystems. Moreover, these cloned sequences, with the exception of one, belong to the largest family of [FeFe] hydrogenase sequences observed in a higher termite gut metagenome. There is good reason to believe that this is only a sampling of a much larger diversity because only one of a total of 9 families reported in the *Nasutitermes* gut metagenome sequence was targeted in this analysis. Interestingly, all of the higher termite sequences grouped with one another to the

Figure 4-5.
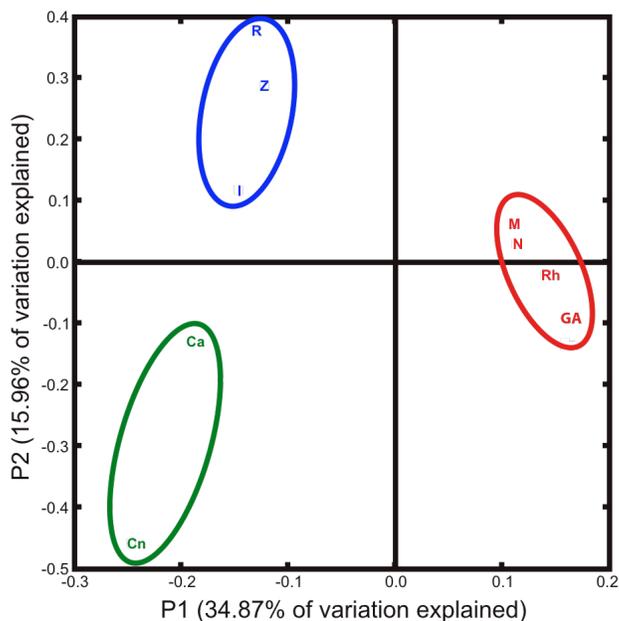
Figure 4-5. Unifrac principle component analysis of Family 3 [FeFe] hydrogenase sequences cloned from the guts of higher termites, lower termites, and *C. punctulatus*. The maximum-likelihood tree shown in Figure 4-3 and the OTUs with their respective abundance weights given in Table 4-S1 of the appendix to this chapter and taken from Table 3-S1 in the appendix to Chapter 3 were used as inputs to Unifrac. Principle components were calculated using normalized abundance weights. Each termite or *C. punctulatus* sample was designated as a unique environment. Higher termite environments are in red, lower termite environments are in blue, and *C. punctulatus* environments are in green. P1 = principle component 1, P2 = principle component 2. Ca = *C. punctulatus* adult, Cn = *C. punctulatus* nymph, GA = a cluster of samples comprising *Amitermes* sp. Cost010, *Amitermes* sp. Cost003, and *Gnathamitermes* sp. JT5, I = *Incisitermes minor* isolate collection Pas1, M = *Microcerotermes* sp. Cost008, N = *Nasutitermes* sp. Cost003, R = *Reticulitermes Hesperus* collection ChiA2, Rh = *Rhyncotermes* sp. Cost004, Z = *Zootermopsis nevadensis* collection ChiA1.

Figure 4-6.
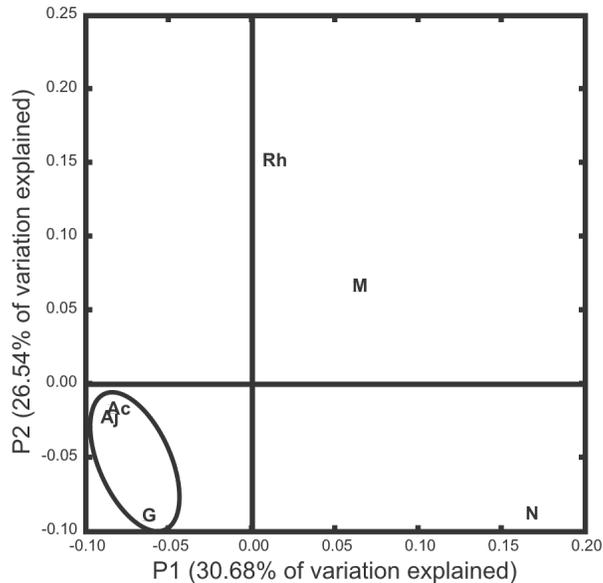
Figure 4-6. Unifrac principle component analysis of Family 3 [FeFe] hydrogenase sequences cloned from higher termites in this study. The maximum-likelihood tree shown in Figure 4-2 and the OTUs with their respective abundance weights given in Table 4-S1 of the appendix to this chapter were used as inputs to Unifrac. Principle components were calculated using normalized abundance weights. Each termite sample was designated as a unique environment. Environments representing sub-terranean termites are in purple, those representing all other higher termites are in red. P1 = principle component 1, P2 = principle component 2. Ac = *Amitermes* sp. Cost010, Aj = *Amitermes* sp. JT2, G = *Gnathamitermes* sp. JT5, M = *Microcerotermes* sp. Cost008, N = *Nasutitermes* sp. Cost003, Rh = *Rhyncotermes* sp. Cost004.

exclusion of all other non-termite [FeFe] hydrogenase sequences in our database. This may imply unique adaptations of these sequences to the termite gut ecosystem. Similar community-wide adaptations of [FeFe] hydrogenase sequences from unique ecosystems has been reported previously as reported in Chapter 3 and elsewhere (2).

Higher termites characteristically lack protozoa in the gut (14). Lower termites and *C. punctulatus* have an abundance of protozoa in their guts that are largely responsible for the fermentation of lignocellulosic polysaccharides and the concomitant production of most of the hydrogen in the termite gut (4, 7, 9, 15, 26, 42, 52). The abscence of protozoa in higher termite guts may introduce important selective forces on bacteria unique to these ecosystems including a greater burden to produce and consume hydrogen. As one might expect then, the hydrogenases cloned from the higher termites tended to have a more even distribution and broader sequence diversity than sequences cloned from *C. punctulatus* or lower termites, compare Table 4-1 and Figure 4-1 from this study to Table 3-1 and Figures 3-1 and 3-2 from Chapter 3.

Congruence of [FeFe] hydrogenase and host phylogeny. [FeFe] hydrogenases cloned from closely related termites had a tendency to group with one another in phylogenetic analyses, see Figure 4-2. For example, sequences from both *Amitermes* gut samples tended to group together despite their being collected from locations separated by a great distance – California and Costa Rica. Sequence OTUs from a particular termite tended to group with one another rather than with sequences from other termites. In a phylogenetic analysis of the COII sequences used for molecular characterization of the termite samples, *Gnathamitermes* sp. JT5 and *Amitermes* sp. JT2 were found to be the most closely related of any of the higher termites used in this study (41). Correspondingly,

there was a tendency for sequences from *Gnathamitermes* sp. JT5 to group with those from the *Amitermes* sp. samples. As one would expect, sequences taken from the genomes *T. primitia* ZAS-2 and *T. azotonutricium* ZAS-9, each isolated from the gut of a lower termite, did not group strongly with any of the sequences cloned from the higher termites, see Figure 4-2.

This congruence was further supported by phylogenetic comparisons of the higher termite sequences to lower termite and *Cryptocercus* sequences cloned previously. In the maximum likelihood tree shown in Figure 4-3, there is a clear segregation of the higher termite hydrogenase sequences from those of *Cryptocercus* and lower termites. The lack of clear segregation of the lower termite sequences from those of *C. punctulatus* is in agreement with the close evolutionary relatedness of these insects (23, 24, 30). A Unifrac principle component analysis using the maximum likelihood tree shown in Figure 4-5 further supported these qualitative observations. The 1st principle component, explaining 34.87% of the variation, separated the higher termites from *Cryptocercus* and lower termites. The jackknife clustering of the [FeFe] hydrogenease communities mimicked previously proposed termite phylogenies remarkably (23, 24, 30).

The observed congruence between [FeFe] hydrogenase phylogeny and that of the host may imply that hydrogenases, and by extension their respective gut communities, have co-evolved in an intimate relationship with their host termites. This is in agreement with previous proposals of termite or *Cryptocercus* gut microbes having co-evolved with their host (1, 12, 13, 17). Perhaps more accurately, this observation may be explained as a consequence of the influence of environmental alterations in the gut, such as the presence

or lack of protozoa or various anatomical alterations, that have developed over the course of termite evolution.

Influence of host feeding habits. Unifrac principle component and jackknife clustering analyses of a maximum likelihood tree of all higher termite sequences, see Figures 4-6, revealed a close clustering of the *Amitermes* sp. samples and *Gnathamitermes* sp. JT5 samples. This clustering was apparent when the 1st and 2nd principle components, collectively explaining 57.22% of variation, were plotted against each other. In addition to sharing the close phylogenetic relationship discussed above, the *Amitermes* sp. and the *Gnathamitermes* sp. JT5 termite samples were all collected from sub-terranean gallies implying a grass- or soil-feeding diet and increased exposure to humics. Elizabeth Ottesen has reported a similar distinguishability between sub-terranean higher termites and other higher termites in her work on the FTHFS gene using the same higher termite samples used in this study (41). Previous studies have also shown that higher termites with different feeding habits have markedly different compositions of symbiotic bacteria in their guts (35, 45, 50). Feeding habits may be an important parameter, intimately associated with host phylogeny, influencing the [FeFe] hydrogenase sequence representation in the termite gut.

Conclusions. Termites are a rich reservoir of uniquely adapted [FeFe] hydrogenase gene diversity. The high representation of [FeFe] hydrogenases observed in the guts of higher termites accentuates the physiological importance of these ecosystems. The enzymes had a higher representation and more even population distributions than was observed previously in lower termites and a woodroach, see Chapter 3. This may be the

consequence of an increased metabolic burden on the gut bacteria in higher termites to metabolize hydrogen as a consequence of a lack of gut protozoa.

The congruence of [FeFe] hydrogenase sequence phylogeny with host phylogeny provides experimental support for the hypothesis that the gut microbial communities of termites and *Cryptocersus* have "co-evolved" with their host. This may reflect the combined influences of the stable, intimate relationship of gut microbes with their host and the environmental alterations in the gut, such as the presence or lack of protozoa or various anatomical or host nutritional alterations, that have occurred over the course of termite evolution. Unifrac analyses further revealed that long standing host-feeding preferences, a variable perhaps closely correlated with termite evolution, may have an important influence on the hydrogenase sequence population in the termite gut.

Surveying the representation of Family 3 [FeFe] hydrogenases has begun to shed light on the physiology and evolution of the gut microbial communities of termites.

References

1. **Berlanga, M., B. J. Paster, and R. Guerrero.** 2007. Coevolution of symbiotic spirochete diversity in lower termites. Int. Microbiol. **10:**133-139.

2. **Boyd, E. S., J. R. Spear, and J. W. Peters.** 2009. [FeFe] Hydrogenase Genetic Diversity Provides Insight into Molecular Adaptation in a Saline Microbial Mat Community. Appl. Environ. Microbiol. **75:**4620-4623.

3. **Brauman, A., M. D. Kane, M. Labat, and J. A. Breznak.** 1992. Genesis of Acetate and Methane by Gut Bacteria of Nutritionally Diverse Termites. Science **257:**1384-1387.

4. **Breznak, J. A.** 2000. Ecology of prokaryotic microbes in the guts of wood- and litter-feeding termites, p. 209-231. *In* T. Abe, D. E. Bignell, and M. Higashi (ed.), Termites: evolution, sociality, symbioses, ecology. Kluwer Academic Publishers, Boston.

5. **Breznak, J. A.** 1982. Intestinal microbiota of termites and other xylophagous insects. Annu. Rev. Microbiol. **36:**323-343.

6. **Breznak, J. A.** 1975. Symbiotic Relationships Between Termites and Their Intestinal Microbiota. Symp. Soc. Exp. Biol. **:**559-580.

7. **Breznak, J. A., and B. A.** 1994. Role of microorganisms in the digestion of lignocellulose by termites. Annu. Rev. Entymology **39:**453-487.

8. **Breznak, J. A., and J. M. Switzer.** 1986. Acetate Synthesis from H2 plus CO2 by Termite Gut Microbes. Appl. Environ. Microbiol. **52:**623-630.

9. **Brune, A.** 2006. Symbiotic Associations Between Termites and Prokaryotes. Prokaryotes **1:**439-474.

10. **Brune, A.** 1998. Termite guts: the world's smallest bioreactors. Trends in Biotechnol. **16:**16-21.

11. **Brune, A., and M. Friedrich.** 2000. Microecology of the termite gut: structure and function on a microscale. Curr. Opin. Microbiol. **3:**263-269.

12. **Clark, J. W., S. Hossain, C. A. Burnside, and S. Kambhampati.** 2009. Coevolution between a cockroach and its bacterial endosymbiont: a biogeographical perspective. Proc. R. Soc. Lond., B., Biol. Sci. **268:**393-398.

13. **Clark, J. W., and S. Kambhampati.** 2003. Phylogenetic analysis of *Blattabacterium*, endosymbiotic bacteria from the wood roach, *Cryptocercus* (Blattodea: Cryptocercidae), including a description of three new species. Mol. Phylogenet. Evol. **26:**82-88.

14. **Cleveland, L. R.** 1923. Correlation Between the Food and Morphology of Termites and the Presence of Intestinal Protozoa. Am. J. Hyg. **3:**444-461.

15. **Cleveland, L. R.** 1923. Symbiosis between Termites and Their Intestinal Protozoa. Proc. Natl. Acad. Sci. U.S.A. **9:**424-428.

16. **Ebert, A., and A. Brune.** 1997. Hydrogen Concentration Profiles at the Oxic-Anoxic Interface: a Microsensor Study of the Hindgut of the Wood-Feeding Lower Termite *Reticulitermes flavipes* (Kollar). Appl. Environ. Microbiol. **63:**4039-4046.

17. **Eggleton, P.** 2006. The Termite Gut Habitat: Its Evolution and Co-Evolution, p. 373-404. *In* H. König and A. Varma (ed.), Intestinal Microorganisms of Soil Invertebrates, vol. 6. Springer-Verlag, Berlin Heidelberg.

18. **Felsenstein, J.** 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics **5:**164-166.

19. **Graber, J. R., J. R. Leadbetter, and J. A. Breznak.** 2004. Description of *Treponema azotonutricium* sp. nov. and *Treponema primitia* sp. nov., the First Spirochetes Isolated from Termite Guts. Appl. Environ. Microbiol. **70:**1315-1320.

20. **Hoehler, T. M., B. M. Bebout, and D. J. Des Marais.** 2001. The role of microbial mats in the production of reduced gases on the early Earth. Nature **412:**324-327.

21. **Hungate, R. E.** 1939. Experiments on the Nutrition of Zootermopsis. III. The anaerobic carbohydrate dissimilation by the intestinal protozoa. Ecology **20:**230-245.

22. **Hungate, R. E.** 1943. Quantitative Analysis on the Cellulose Fermentation by Termite Protozoa. Ann. Entomol. Soc. Am. **36:**730-9.

23. **Inward, D., G. Beccaloni, and P. Eggleton.** 2007. Death of an order: a comprehensive molecular phylogenetic study confirms that termites are eusocial cockroaches. Biol. Lett. **3:**331-335.

24. **Inward, D. J. G., A. P. Volger, and P. Eggleton.** 2007. A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspects of their evolutionary biology. Mol. Phylogenet. Evol. **44:**953-967.

25. **Kambhampati, S., and P. Eggleton.** 2000. Taxonomy and Phylogeny of Termites, p. 1-23. *In* T. Abe, D. E. Bignell, and M. Higashi (ed.), Termites: evolution, sociality, symbioses, ecology. Kluwer Academic Publishers, Boston.

26. **Klass, K.-D., C. Nalepa, and N. Lo.** 2008. Wood-feeding cockroaches as models for termite evolution (Insecta: Dictyoptera): *Cryptocercus* vs. *Parasphaeria boleiriana*. Mol. Phylogenet. Evol. **46:**809-817.

27. **Leadbetter, J. R.** 1996. Physiological ecology of *Methanobrevibacter cuticularis* sp. nov. and *Methanobrevibacter curvatus* sp. nov., isolated from the hindgut of the termite *Reticulitermes flavipes*. Appl. Environ. Microbiol. **62:**3620-31.

28. **Leadbetter, J. R., L. D. Crosby, and J. A. Breznak.** 1998. *Methanobrevibacter filiformis* sp. nov., a filamentous methanogen from termite hindguts. Arch. Microbiol. **169:**287-92.

29. **Leadbetter, J. R., T. M. Schmidt, J. R. Graber, and J. A. Breznak.** 1999. Acetogenesis from H2 Plus CO2 by Spirochetes from Termite Guts. Science **283:**686-689.

30. **Legendre, F., M. F. Whiting, C. Bordereau, E. M. Cancello, T. A. Evans, and P. Grandcolas.** 2008. The phylogeny of termites (Dictyoptera: Isoptera) based on mitochondrial and nuclear markers: Implications for evolution of the worker and pseudergate castes, and foraging behaviors. Mol. Phylogenet. Evol. **48:**615-627.

31. **Lozupone, C., M. Hamady, and R. Knight.** 2006. UniFrac - An online tool for comparing microbial community diversity in a phylogenetic context. BMC Bioinformatics **7**.

32. **Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Förster, I. Brettske, S. Gerber, A. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. König, T. Liss, R. Lüssmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N.**

Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, and K. Schleifer. 2004. ARB: a software environment for sequence data. Nucleic Acids Res. **32:**1363-1371.

33.   **Matson, E. G., E. A. Ottesen, and J. R. Leadbetter.** 2007. Extracting DNA from the gut microbes of the termite (*Zootermopsis nevadensis*), J. Vis. Exp.

34.   **Miura, T., K. Maekawa, O. Kitade, T. Abe, and T. Matsumoto.** 1998. Phylogenetic relationships among subfamilies in higher termites (Isoptera: Termitidae) based on mitochondrial COII gene sequences. Ann. Entomol. Soc. Am. **91:**515-523.

35.   **Miyata, R. N., N. Noda, H. Tamaki, T. Kinjyo, H. Aoyagi, H. Uchiyama, and H. Tanaka.** 2007. Influence of feed components on symbiotic bacterial community structure in the gut of the wood-feeding higher termite *Nasutitermes takasagoensis*. Biosci. Biotechnol. Biochem. **71:**1244-1251.

36.   **Noirot, C.** 2001. The Gut of Termites (Isoptera) Comparative Anatomy, Systematics, Phylogeny. II. -- Higher Termites (Termitidae). Ann. Soc. Entomol. Fr. **37:**431-471.

37.   **Noirot, C.** 1995. The Gut of Termites (Isoptera). Comparative Anatomy, Systematics, Phylogeny. I. Lower Termites. Ann. Soc. Entomol. Fr. **31:**197-226.

38.   **Odelson, D. A., and J. A. Breznak.** 1985. Cellulase and Other Polymer-Hydrolyzing Activities of *Trichomitopsis termopsidis*, a Symbiotic Protozoan from Termites. Appl. Environ. Microbiol. **49:**622-626.

39.   **Odelson, D. A., and J. A. Breznak.** 1985. Nutrition and Growth Characteristics of *Trichomitopsis termopsidis*, a Cellulolytic Protozoan from Termites. Appl. Environ. Microbiol. **49:**614-621.

40.   **Odelson, D. A., and J. A. Breznak.** 1983. Volatile Fatty Acid Production by the Hindgut Microbiota of Xylophagous Termites. Appl. Environ. Microbiol. **45:**1602-1613.

41.   **Ottesen, E. A.** 2009. The Biology and Community Structure of CO_2-Reducing Acetogens in the Termite Hindgut. California Institute of Technology, Pasadena.

42.   **Pester, M., and A. Brune.** 2007. Hydrogen is the central free intermediate during lignocellulose degradation by termite gut symbionts. ISME J. **1:**551-565.

43.   **Schink, B., F. S. Lupton, and J. G. Zeikus.** 1983. Radioassay for Hydrogenase Activity in Viable Cells and Documentation of Aerobic Hydrogen-Consuming Bacteria Living in Extreme Environments. Appl. Environ. Microbiol. **45:**1491-1500.

44.   **Schloss, P. D., and J. Handelsman.** 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol. **71:**1501-1506.

45.   **Schmitt-Wagner, D., M. W. Friedrich, B. Wagner, and A. Brune.** 2003. Phylogenetic diversity, abundance, and axial distribution of bacteria in the intestinal tract of two soil-feeding termites (*Cubitermes* spp.) Appl. Environ. Microbiol. **69:**6007-6017.

46. **Scranton, M. I., P. C. Novelli, and P. A. Loud.** 1984. The distribution and cycling of hydrogen gas in the waters of two anoxic marine environments. Limnol. Oceanogr. **29:**993-1003.

47. **Slaytor, M.** 1992. Cellulose Digestion in Termites and Cockroaches: What Role do Symbionts Play. Comp. Biochem. Physiol. **103B:**775-784.

48. **Smolenski, W. J., and J. A. Robinson.** 1988. In situ rumen hydrogen concentrations in steers fed eight times daily, measured using a mercury reduction detector. FEMS Microbiol. Ecol. **53:**95-100.

49. **Sugimoto, A., and N. Fujita.** 2006. Hydrogen concentrations and stable isotopic composition of methane in bubble gas observed in a natural wetland. Biogeochemistry **81:**33-44.

50. **Tanaka, H., H. Aoyagi, S. Shina, Y. Dodo, T. Yoshimura, R. Nakamura, and H. Uchiyama.** 2006. Influence of the diet components on the symbiotic microorganisms community in hindgut of *Coptotermes formosanus* Shiraki. Appl. Microbiol. Biotechnol. **71:**907-917.

51. **Tokuda, G., and H. Watanabe.** 2007. Hidden cellulases in termites: revision of an old hypothesis. Biol. Lett. **3:**336-339.

52. **Trager, W.** 1934. The Cultivation of a Cellulose-Digesting Flagellate, *Trichomonas termopsidis*, and of Certain Other Termite Protozoa. Biol. Bull. **66:**182-190.

53. **Warnecke, F., P. Luginbühl, N. Ivanova, M. Ghassemian, T. H. Richardson, J. T. Stege, M. Cayouette, A. C. McHardy, G. Djordjevic, N. Aboushadi, R. Sorek, S. G. Tringe, M. Podar, H. G. Martin, V. Kunin, D. Dalevi, J.**

Madejska, E. Kirton, D. Platt, E. Szeto, A. Salamov, K. Barry, N. Mikhailova, N. C. Kyrpides, E. G. Matson, E. A. Ottesen, X. Zhang, M. Hernández, C. Murillo, L. G. Acosta, I. Rigoutsos, G. Tamayo, B. D. Green, C. Chang, E. M. Rubin, E. J. Mathur, D. E. Robertson, P. Hugenholtz, and J. R. Leadbetter. 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature **450:**560-569.

54. **Wier, A., M. Dolan, D. Grimaldi, R. Guerrero, J. Wagensberg, and M. Margulis.** 2002. Spirochete and protist smbionts of a termite (*Mastotermes electrodominicus*) in Miocene amber. Proc. Natl. Acad. Sci. **99:**1410-1413.

55. **Wood, T., and R. Johnson.** 1986. The biology, physiology, and ecology of termites, p. 1-68. *In* S. B. Vinson (ed.), Economic impact and control of insects. Praeger, New York.

56. **Yamin, M. A.** 1979. Cellulolytic activity of an axenically-cultivated termite flagellate, Trichomitopsis termopsidis. J. Gen. Microbiol. **113:**417-20.

57. **Yamin, M. A.** 1981. Cellulose Metabolism by the Flagellate *Trichonympha* from a Termite Is Independent of Endosymbiotic Bacteria. Science **211:**58-59.

Appendix

Table 4-S1. Sequences cloned.

Table 4-S1. Sequences cloned.

Amitermes sp. Cost010

Phylotype[a]	Genotype[a]	Number[b]	%[c]
A1	A1, C7, D12, H11, A3, B7, D7, F9, E6	18	19
A10	A10, A4, H5, C8, G2, G3	10	11
A8	A8, C1	4	4
A9	A9, G1	2	2
B11	B11	1	1
B12	B12, C6, F8	4	4
B3	B3	1	1
B5	B5	1	1
B6	B6, G9	2	2
B8	B8	1	1
B9	B9, C9, E11, F7	6	6
C10	C10	4	4
C12	C12, C4, D4, D10, F2	11	12
C2	C2	2	2
C3	C3	1	1
C5	C5	2	2
D1	D1	1	1
D11	D11	3	3
D3	D3	1	1
D5	D5	1	1
D6	D6	1	1
E4	E4, G7	4	4
E7	E7	1	1
E9	E9, H12	2	2
F1	F1	3	3
F10	F10	1	1
F12	F12	2	2
F3	F3	2	2
G4	G4	1	1
G8	G8	1	1
H10	H10	1	1
	Total:	95	

Amitermes sp. JT2

Phylotype[a]	Genotype[a]	Number[b]	%[c]
A1	A1, E1	5	5
A2	A2	1	1
A7	A7, C2, E3	8	8
A8	A8	2	2
B11	B11, G1	3	3
B2	B2, E9, H12	10	10
B5	B5, D10, E2	5	5
B6	B6	1	1
B7	B7	1	1
B8	B8	1	1
C11	C11	2	2
D2	D2	6	6
D6	D6	5	5
E10	E10	3	3
F12	F12	1	1
F2	F2	8	8
F3	F3	1	1
F4	F4	22	23
F5	F5, H1, G8	7	7
G3	G3	1	1
G7	G7	2	2
H5	H5	1	1
	Total:	96	

Continuing Table 4-S1.

Phylotype[a]	Genotype[a]	Number[b]	%[c]
Gnathamitermes sp. JT5			
A10	A10, A3, G4	8	8
A11	A11	2	2
A2	A2, H11	2	2
A8	A8	4	4
B1	B1	1	1
B11	B11, D12, G6, E9	5	5
B12	B12, G12, B9, F8	13	14
B2	B2, H2	2	2
B5	B5	4	4
C1	C1, F11	10	11
C12	C12	4	4
C3	C3	1	1
C6	C6	1	1
C8	C8	1	1
D1[e]	D1	2	2
D10	D10	1	1
D11	D11	7	7
D3	D3	1	1
E12	E12	1	1
E2	E2	5	5
E3	E3	1	1
E7	E7	1	1
F2	F2	1	1
F4	F4, H8	2	2
F5	F5	3	3
F6	F6, F7	6	6
G11	G11	1	1
G8	G8	2	2
H4	H4	2	2
H5	H5	1	1
	Total:	95	
Microcerotermes sp. Cost008			
A12	A12, C3, E12, H7	5	5
A3	A3, B1, F4, H12, C12, C4	25	26
A4	A4	2	2
A8	A8	4	4
A9	A9	2	2
B7	B7	25	26
B8	B8	2	2
C1	C1	1	1
C10	C10[d]	1	1
C5	C5, D11	2	2
C6	C6, D2	5	5
C9	C9	1	1
D5	D5, G1	5	5
D8	D8	1	1
E3	E3	1	1
E5	E5	2	2
F10	F10[d]	1	1
F5	F5, G11	8	8
G3	G3	1	1
H11	H11	1	1
H8	H8	1	1
	Total:	96	

Continuing Table 4-S1.

Phylotype[a]	Genotype[a]	Number[b]	%[c]
Nasutitermes sp. Cost003			
A11	A11, D5, G1	4	4
A12	A12	1	1
A6	A6	4	4
A9	A9, B5, D12, E1	10	11
B10	B10	2	2
B2	B2, C4, D4, D1, G12	19	21
B3	B3	1	1
B9	B9	1	1
C1	C1, C5	10	11
C2	C2	12	13
C8	C8, D9	5	6
D2	D2	1	1
E2	E2	1	1
E3	E3, D7	2	2
E4	E4	2	2
F1	F1	1	1
F2	F2	1	1
F5	F5	1	1
F8	F8	1	1
G10	G10	2	2
H1	H1	1	1
H11	H11	3	3
H2	H2	1	1
H5	H5	1	1
H6	H6	3	3
	Total:	90	
Rhynchotermes sp. Cost004			
A1	A1	1	1
A12	A12	2	2
A2	A2	2	2
A3	A3	1	1
A4	A4	2	2
A5	A5	1	1
A6	A6	1	1
A7	A7, E8	4	4
A9	A9	5	5
B1	B1	4	4
B11	B11	3	3
B2	B2, C8, H6	6	6
B3	B3	1	1
B6	B6	1	1
B7	B7, H7	2	2
C11	C11	2	2
C12	C12	1	1
C4	C4, E12, H1	3	3
C5	C5, G5, H12	7	7
D10	D10, G11	7	7
D11	D11	1	1
D5	D5	2	2
D6	D6	2	2
E2	E2	2	2
E3	E3	1	1
E4	E4	3	3
E6	E6	2	2
F1	F1	1	1
F10	F10	1	1
F11	F11	1	1
F2	F2[d]	6	6
F6	F6	1	1
F8	F8	1	1
F9	F9	1	1
G1	G1	1	1
G2	G2	1	1
G3	G3	1	1
G4	G4	1	1
G9	G9	1	1
H11	H11	1	1
H2	H2	2	2
H3	H3	1	1
H5	H5	1	1
H8	H8	2	2
	Total:	93	

[a]OTUs calculated using the furthest-neighbor method in DOTUR with a 97% amino-acid similarity cut-off.

[b]Number of cloned sequences grouped within each OTU.

[c]Percent of cloned sequences represented by each OTU.

[d]Sequences containing frame-shift mutations that were "corrected" manually by the addition or subtraction of nucleotides at the DNA sequence level to allow for phylogenetic analyses using amino acid sequences.

[e]Sequence OUT grouping phylogenetically with sequences previously classified as Family 7 [FeFe] hydrogenases (53).