Chapter 5

MICROFLUIDIC DIGITAL PCR REVEALS THAT TREPONEMES MAY BE A PREDOMINANT GENUS OF EUBACTERIA ENCODING AN IMPORTANT FAMILY OF [FeFe] HYDROGNEASES IN THE GUT OF *RETICULITERMES TIBIALIS*

Abstract

Hydrogen is an important free intermediate in the degradation of wood by termites and is turned over at high fluxes and maintained at concentrations exceeding those measured for any other biological system. We have employed microfluidic digital PCR to identify bacteria encoding [FeFe] hydrogenase genes and, therefore, potentially participating in hydrogen metabolism in the gut of *Reticulitermes tibialis*. We successfully designed degenerate primers specifically targeting the largest group of [FeFe] hydrogenases observed in a termite hindgut metagenome. Nucleotide sequences gathered in previous molecular inventories were utilized in probe design. 27 16S rRNA – Family 3 [FeFe] hydrogenase gene pairs from putative single cell genomes were successfully co-amplified by multiplex PCR in microfluidic chambers and subsequently sequenced. 22 of the 16S rRNA phylotypes were treponemal, and of these 16 fell within the termite cluster of treponemes. All instances of the same 16S rRNA – hydrogenase gene pairings observed in multiple independent chambers, referred to as *"Reticulitermes* environmental genomovars, corresponded to treponeme phylotypes. The non-treponemal phylotypes fell within the β- and ε-Proteobacteria and Bacteroidetes phyla. All but 3 of the phylotypes grouped closely with phylotypes previously sequenced from the guts of *Reticulitermes* termites. A number of the hydrogenase gene peptide sequences grouped closely with sequences cloned from *Reticulitermes hesperus* in a previous study. Our

results provide compelling evidence that treponemes, particularly of the termite cluster, represent an important genus encoding Family 3 [FeFe] hydrogenases in the gut of *R. tibialis* that perhaps is comprised of members making an important contribution to bacterially mediated hydrogen metabolism.

Introduction

Hydrogen plays a pivotal role in the breakdown of lignocellulosic biomass by termites (3, 7, 15, 39, 40, 49). It can reach concentrations in some species exceeding those measured for any other biological system (15, 18, 49, 52, 55-57) and fluxes have been measured as high as 33 m^3/m^3 gut volume (49). Moreover, the gut ecosystem is spatially complex being comprised of numerous microenvironments formed as a consequence of the numerous chemical gradients extant in the gut, hydrogen being among the most prominent (9, 10, 15, 24, 25, 49).

This hydrogen is produced and turned-over as an important metabolic free intermediate in the breakdown of lignocellulosic biomass by the complex microbial community residing in the termite gut (17, 21, 22, 40, 58, 65, 66). Termites are entirely dependent upon this symbiosis, which can include representatives of all three domains of life, to derive energy and carbon from wood (4-6, 8, 11, 12, 39). Acetate, which serves as a termite's primary carbon and energy source, hydrogen and carbon dioxide are the primary products produced in this symbiosis (41). Only a small fraction of the hydrogen and carbon dioxide are emitted from the gut; indeed, these potential waste products are used by bacteria in reductive acetogenesis to produce up to 33% of the total acetate pool in the gut (3, 7, 26, 41, 49). This contributes substantially to the metabolic efficiency of the system and has led some to postulate that it is among the smallest and most efficient bioreactors found in nature (49). Only a small fraction of the hydrogen and carbon dioxide is lost from the system as methane produced by methanogens (3, 24).

The sequencing of a termite hindgut metagenome has revealed the presence of an abundance and striking diversity of hydrogenase-like proteins in the termite gut (61).

The vast majority of these were classified as [FeFe] hydrogenase-like proteins (61). Through for some of these sequences predictions as to the identity of the host of origin were made based upon nucleotide composition using Phylopythia (34, 61), it remains nebulous precisely what bacterial species encode these sequences or play an important role in hydrogen metabolism in the termite gut.

Traditionally, efforts to identify microbes filling a particular physiological niche in an environment have relied upon molecular inventories of a structural gene essential for a phenotype of interest. Numerous studies have taken this approach, for instance, to gain a better understanding of nitrogen metabolism (64, 68, 69), methanogenesis (14), or hydrogen metabolism (2, 16, 62, 63) in environmental samples. Several notable studies have taken this approach to analyze the diversity of microbes participating in reductive acetogenesis (48, 51) or nitrogen metabolism (38, 43, 64) in the termite gut. To infer a bacterium's phylogeny, these studies, by necessity, assume a reliable correlation between host phylogeny and the structural gene of interest. In some instances this may indeed be a reliable assumption, but it is notably not the case for [FeFe] hydrogenases (54, 59). Schmidt *et al.* noted in their 2010 study, "the presence of homologous multiple hydrogenases per organism and inconsistencies between 16S rRNA gene and [FeFe]-hydrogenase based phylogenies" (54). One might propose addressing this challenge by isolating bacteria from the termite gut and subsequently annotating hydrogenases encoded in their genome sequences, as has been done in the cases of *Treponema primitia* ZAS-1 and ZAS-2 and *Treponema azotonutricium* ZAS-9 in Chapter 2. But, on account of the resistance of termite gut bacteria to cultivation (23), one can not feasibly

investigate global community characteristics of the termite gut by isolating a large sampling of its component species.

Microfluidic digital PCR (dPCR) was developed by Ottesen *et al*. in 2006 and successfully used to co-amplify formyl-tetrahydrofolate synthetase (FTHFS) gene sequences and 16S rRNA gene sequences from single bacterial cells from a termite gut (45). This method allowed for the unambiguous pylogenetic identification of bacteria encoding FTHFS (45), an important marker for reductive acetogenesis (27, 28). Moreover, since a single microbial genome is used as a template in microfluidic digital PCR, the challenges of primer bias or chimera formation that plague the traditional molecular profiling techniques discussed above (1) were largely obviated.

Here we report the use of microfluidic digital PCR to unambiguously identify bacterial species encoding [FeFe] hydrogenases in the gut microbial community of *Reticulitermes tibialis.* Our objective was to use this metabolic marker as a means to identify bacteria that may contribute to hydrogen metabolism in the termite gut.

Methods

Sample collection and classification. *Reticulitermes tibialis* collection JT2 termites were collected from a single colony found in a fallen tree along the side of a road in Joshua Tree National Park (Permit#: JOTR:2008-SCI-002) in Southern California on December 15, 2009 at 3:00 PM. The GPS coordinates of the site were 34° 1’ 7.2” N and 116° 10’ 6.9” W. Termites were immediately used in experiments. During the 5 days duration of experiments, they were stored in a room-temperature glass aquarium with moist, sterile sand, and wood cut from the tree from which they were collected. The atmosphere of the chamber was maintained at 95% humidity.

The termites were classified using insect mitochondrial cytrochome oxidase subunit II (COII) gene sequences. The COII genes were amplified from the head of one of the termites used in a digital PCR experiment. The head was macerated with a sterile glass rod in a 2 ml tube with 50 μl TRIS-EDTA (Sigma). The liquid fraction was transferred to another tube and incubated at 95°C for 10 min and subsequently diluted 10-fold in nuclease-free water. 3 μl of this solution was used as a template in a 50 μl PCR reaction. The template was amplified using primers COII-R and COII-F (see Table 5-1), each at a final concentration of 1 μM, Expand High Fidelity Taq Polymerase (Roche), and FailSafe Premix D (Epicentre). The thermal cycling regimen was 94°C 3 min, (94°C 30s, 50°C 30s, 72°C 1.5 min) x 30, 72°C 10 min. Amplicons were cloned into One Shot Top10 chemical competent *Escherichia coli* cells (Invitrogen) using a QIAGEN PCR Cloning Kit following manufacturers' protocols. Colonies were submitted to GENEWIZ for sequencing. Sequences were edited manually using SeqMan, available from DNA* as part of the Lasergene software suite, and analyzed phylogenetically using the ARB software environment (32) in a manner analogous to that described below. The termites were classified as *Reticultermes tibialis* (see Figure 5-1) and given the identifier "collection JT2."

Microbial Strains. Microbial isolate *Treponema primitia* str. ZAS-2 was grown in anaerobic YACo medium under a headspace of 80% H_2 + 20% CO_2, as described previously (26, 29). *Treponema azotonutricium* str. ZAS-9 was grown in a similar medium (26, 29).

Primer and Probe Design. Degenerate primers and probes for the detectable amplification of "Family 3" (61) [FeFe] hydrogenases in quantitative PCR (qPCR) and

Table 5-1. Primers and probes used.

Primer	Sequence[a]	Target[b]	Reference
386 F'	5' - CIC GIA TGA THA ARC ARG CIG G - 3'	Family 3 [FeFe] hydrogenase	this study
467 R'	5' - CCA TYT GRT GIG CIA YIG C - 3'	Family 3 [FeFe] hydrogenase	this study
1100R	5' - AGG GTT GCG CTC GTT G - 3'	16S	-
1492RL2D	5' - TAC GGY TAC CTT GTT ACG ACT T - 3'	Gen. Bac. 16S rRNA	(45)
357F	5' - CTC CTA CGG GAG GCA GCA G - 3'	Gen. Bac. 16S rRNA	-
COII-R	5' - GTT TAA GAG ACC AGT ACT TG - 3'	cytochrome oxidase II gene	(31)
COII-F	5' - ATG GCA GAT TAG TGC AAT GG - 3'	cytochrome oxidase II gene	(31)
LNA H2-1a[c]	5' - 56-FAM-CT**T CC**A **TGA** C-3BHQ_1 - 3'	Family 3 [FeFe] hydrogenase (probe)	this study
LNA H2-1b[c]	5' - 56-FAM-C**TT CC**A T**A**A **C**-3BHQ_1 - 3'	Family 3 [FeFe] hydrogenase (probe)	this study
Hex-1389Prb	5' - 5HEX-CTT GTA CAC ACC GCC CGT C-3BHQ_1 -3'	Gen. Bac. 16S rRNA (probe)	(45)

[a]"I" represents inositol.

[b]Gen. Bac., general bacterial.

[c]Boldface residues are locked nucleic acids.

Figure 5-1.

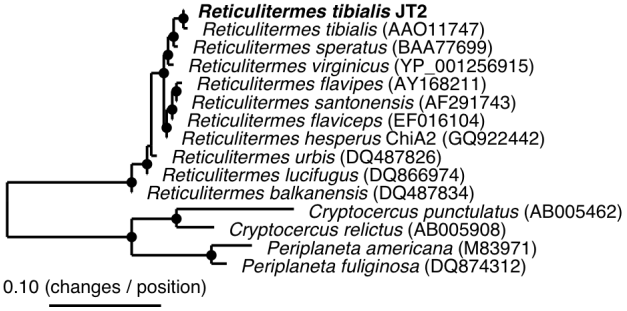

Figure 5-1. Mitochondrial cytochrome oxidase II (COII) phylogeny of the termite sample used in this study. Sequence accession numbers are listed in parentheses. The COII sequence cloned from the termite sample used in this study is in bold. The tree was calculated using a maximum likelihood (Phylip ProML) method with 225 unambiguously aligned amino acid positions. Closed circles designate groupings also supported by parsimony (Phylip PROTPARS, 1000 bootstraps) and distance matrix (Fitch) methods.

dPCR were designed manually from multiple sequence alignments, see Figures 5-S1 and 5-S2. Family 3 [FeFe] hydrogenases were the most highly represented group of enzymatic hydrogenases observed in the *Nasutitermes* gut metagenome sequence (61) and have also been observed, see Chapter 2, in the genome sequences of treponemes isolated from the gut of *Zootermopsis angusticolis*. They are the only group of hydrogenases whose *in situ* translation has been verified (61).

An alignment prepared using DIALIGN (36), available on the Mobyle Portal (37), of [FeFe] hydrogenase gene sequences cloned previously from *Reticulitermes hesperus* collection ChiA2 revealed a highly conserved nucleotide region, see Figure 5-S2 in the chapter appendix. Due to the short length of the region, 10 base pairs, it was necessary to use locked nucleic acids (LNA) (37) to design probes targeting it with sufficiently high melting temperatures to be use in qPCR or dPCR. The probes designed, see Table 5-1, target 89% of the hydrogenase sequences cloned in a gut microbe molecular inventory prepared from *R. hesperus*, see Chapter 3. Probes were manufactured by Integrated DNA Technologies.

For degenerate primer design, the peptide sequences for all bacterial [FeFe] hydrogenase genes previously cloned from termite guts and a wood roach (see Warnecke *et al*. (61) and Chapters 2, 3 and 4) were aligned using DIALIGN on the Mobyle Portal. A portion of the alignment is shown in Figure 5-S2 in the chapter appendix. Highly conserved regions identified in this alignment, and flanking the sequence targeted by the probes designed as described above, were used in the design of degenerate primer combinations for use in dPCR. A functional primer set and optimal conditions for gene amplification were determined empirically. All oligomers were ordered from Integrated DNA

Technologies. Initial screens for functionality of candidate primer sets and PCR conditions were done using qPCR. Genomic DNA purified from *Treponema primitia* ZAS-2 or *T. azotonutricium* ZAS-9 was used as template in these screens. The degenerate primers targeting Family 3 [FeFe] hydrogenases reported in Chapter 3 did not function well in qPCR. Successful primer sets and PCR conditions were defined as those facilitating successful quantitative discrimination between different template concentrations in a dilution series in a multiplex qPCR reaction with the primer and probe set targeting 16S rRNA gene sequences, see Table 5-1. The best conditions identified using qPCR, with minor modifications introduced following control dPCR experiments, are described below as those used for all dPCR experiments. The best degenerate primer set identified is listed in Table 5-1. The primers target the H domain known to be highly conserved among all [FeFe] hydrogenases (35, 60). The amplified region corresponds approximately to the regions spanning I393-V470 and I272-V349 in the [FeFe] hydrogenases from *Clostridium pasteurianum* (P29166) and *Desulfovibrio vulgaris* (YP_010987), respectively.

Template Preparation for Digital PCR. For each digital PCR run, a single *Reticulitermes tibialis* collection JT2 gut was dissected out of a worker termite and resuspended in 250 μl of a "synthetic gut fluid" (SGF) solution. The SGF comprised 500 ng/ml bovine pancreas RNase (Roche), 10 mM Tris pH 8, 1 mM EDTA, 30 mM NaCl, and 60 mM KCl in water. The contents of the whole gut was suspended in this solution by pipetting up and down and crushing it against the sides of a centrifuge tube using a 1 ml tip. This suspension was used as the template for use in digital PCR. In control experiments, aliquots of *Treponema azotonutricium* ZAS-9 or *Treponema primitia* ZAS-

2 cells in mid-log phase growth were added to the SGF solution instead of a termite gut suspension. The crushed gut suspensions were allowed to stand for 30s to allow large particles to settle. The templates were diluted to working concentrations in SGF. These template solutions were then mixed with the PCR reaction mixture described below and immediately loaded onto a microfluidic digital array.

Digital PCR Protocol. Equipment for microfluidic digital PCR was purchased from Fluidigm Corporation and included a BioMark with its controlling software, a Nanoflex IFC controller with its controlling software, and custom microfluidic digital array 12.765P digital PCR devices. On-chip PCR reaction solutions contained iQ Multiplex Powermix (BioRad), 0.1% Tween-20 (Sigma), and 400 nM Rox standard (Quanta Biosciences). The primers and probes used are all listed in Table 5-1 and were used at the following concentrations: 154 nM each of 357F and 1492RL2D; 304 nM each of 386 F' and 467 R'; 450 nM each of LNA H2-1a and LNA H2-1b; 300 nM Hex-1389Prb.

Arrays were loaded and PCR was performed as recommended by Fluidigm. The thermal cycling protocol was 95°C 5 min, (95°C 15 s, 60°C 45 s) x 45, 60°C 10 min, 20°C 10 min. Data was analyzed using the Fluidigm Digital PCR Analysis program version 2.1.1, build 20090521.1140. The program data analysis parameters were set to the following to detect wells containing putative positive co-amplifications: target Ct range of 23-45, linear base correction, user data (global) Ct threshold method, a quality threshold of 0.65, and thresholds of 0.03 for FAM-MGB and 0.05 for VIC-MGB signals. FAM-MGB and VIC-MGB thresholds were selected such that the number of putative amplifications detected in a no template control panel run on a digital array were routinely on the order of 1.5% or less.

Sample Retrieval and Analysis. The thermal conducting silicon wafer was removed from each digital PCR array device using a device purchased for this purpose from Fluidigm. Pressure was released from each device after this process. Products were only retrieved from panels with less than 33% of the chambers containing putative 16S rRNA gene amplicons. Assuming a Poisson distribution of particles, this precaution should have assured that only 6% of the chambers contained multiple particles. Chambers containing putative successful co-amplifications were located with the aid of a dissecting microscope and pierced with the tip of a 26-gage needle. The needle was then mixed briefly in 10 μl TRIS-EDTA (Sigma).

Retrieved samples were initially screened for the presence of an [FeFe] hydrogenase gene product by PCR. 2 μl of each suspension were added as template to PCR reaction mixes comprised of 1.4 U Expand High Fidelity Taq (Roche) and FailSafe PCR Premix D (Epicentre), and 1 μM each of primers 386F' and 467R', see Table 5-1. The thermal cycling program used was 95°C 5 min, (95°C 15 s, 60°C 45 s, 72°C 60 s) x 40. Reactions yielding a product detectable in gel electrophoresis were PCR purified using a QIAquick PCR Purification Kit (QIAGEN) following the manufacturer's protocol. 2 μl of the resulting solutions were then used as template in PCR reaction mixes comprised of FailSafe PCR Premix D (Epicentre), 1 U Taq polymerase (Roche), and 500 nM each of primers 386 F' and 467 R', see Table 5-1. The thermal cycling regimen was 95°C 5 min, (95°C 15 s, 60°C 45 s, 72°C 60 s) x 15. The products of these reactions were cloned into One Shot® Top10 chemical competent *Escherichia coli* cells (Invitrogen) using the QIAGEN PCR Cloning Kit following manufacturer protocols.

All samples that yielded a detectable product in the above screen were further screened for the presence of a 16S rRNA gene product. 2 μl of each resuspension were added to PCR mixes containing iQMultiplex Powermix (BioRad) and 1 μM each of primers 1492RL2D and 357F, see Table 5-1. The thermal cycling regimen was 95°C 5 min, (95°C 15 s, 60°C 45 s, 72°C 60 s) x 30. Samples yielding a product detectable in gel electrophoresis were purified using a QIAquick PCR Purification kit (QIAGEN) following the manufacturer's protocol. 2 μl of the resulting suspensions were used as template in PCR reactions mixes comprised of FailSafe PCR Premix D (Epicentre), 1 U Taq polymerase (Roche), and 500 nM each of primers 1492RL2D and 357F, see Table 5-1. The thermal cycling regimen was 95°C 5 min, (95°C 15 s, 60°C 45 s, 72°C 60 s) x 15. The products of these reactions were cloned into One Shot Top10 chemical competent *Escherichia coli* cells (Invitrogen) using a TOPO TA Cloning Kit (Invitrogen) following manufacturer protocols. Clones corresponding to array chambers A1, A8, A11, A12, B5, B2, B6, B9, C7, and D9 were prepared using a QIAGEN PCR Cloning Kit (QIAGEN) following manufacturer protocols.

For each dPCR amplicon, 10 individual clones were selected at random for restriction fragment length polymorphism (RFLP) analysis. Each clone was suspended in TRIS-EDTA (Sigma) and used as a template for PCR. Sequences were amplified in PCR using T7 and T3 primers, NEB Taq Polymerase (New England Biolabs) and FailSafe PCR Premix H (Epicentre). Sequences cloned using the QIAquick PCR purification kit were amplified suing SP6 and T7 primers. The temperature cycling program was 95°C 5 min, (95°C 30 s, 55°C 30s, 72°C 5 min) x 25, 72°C 10 min.

The products of each reaction were subjected to digestion with HinPI1 and the resulting RFLP patterns were analyzed by agarose gel electrophoresis. Cloned sequences representing unique RFLP patterns were arbitrarily selected for submission to GENEWIZ for colony sequencing.

Sequences were manually trimmed in SeqMan, available from DNA* as part of the Lasergene software suite, to remove plasmid and degenerate primer sequences. The 16S rRNA sequences were checked for chimeras using GreenGenes/Bellepheron (13) available on the GreenGenes website (13). A more rigorous screening for chimeras was not deemed necessary because digital PCR, by its very nature, dramatically reduces the likelihood of their formation because a single cell is used as the template in PCR reactions. The [FeFe] hydrogenase sequences were each BLASTed against the *Nasutitermes* hindgut metagenome sequence on the IMG/M server (33) to verify, by means of the top hits, their identity as Family 3 [FeFe] hydrogenase sequences.

The ARB software environment (32) was used for phylogenetic analysis of the 16S-rRNA gene and [FeFe] hydrogenase peptide sequences. Sequences cloned in this study and used for tree construction are listed in Table 5-S1 in the chapter appendix. Peptide sequences were aligned using DIALIGN (36), available on the Mobyle server (37). 16S rRNA gene sequences were aligned using the Silva aligner (53). If 16S rRNA gene nucleotide sequences or [FeFe] hydrogenase peptide sequences corresponding to amplicons from a dPCR array chamber shared less than 97% sequence identity, all sequences from the chamber were discarded as a precaution assuming that this may be an indication of either contamination or co-localization of multiple cells.

Nomenclature. All cloned dPCR products were given the prefix "Rt" designating the termite host of origin, *Reticulitermes tibialis* collection JT2, followed by an “H2” for all hydrogenase sequences or an “R” for all 16S rRNA gene sequences. A dash followed by a letter-number combination uniquely identifying the dPCR array chamber from which a given sequence was collected was then added to this prefix. This chamber designator was then appended on the right by a period and a number designating the identity of the specific clone out of the 10 prepared for each hydrogenase or 16S rRNA gene amplicon retrieved from a dPCR array chamber.

Reticulitermes environmental genomovars (REG), using terminology proposed by Ottesen *et al*. (45), were defined as sets of sequences co-amplified in dPCR for which the 16S rRNA gene nucleotide sequences and [FeFe] hydrogenase peptide sequences each share at least 97% identity. REGs were arbitrarily assigned a number identifying the REG followed by a period and another number identifying a unique dPCR chamber that sequences co-amplified in. The final “r” or “h” in each REG name serves to identify a sequence comprising the REG as a 16S rRNA gene sequence or an [FeFe] hydrogenase peptide sequence, respectively.

Results

Sequence amplification and retrieval. Efforts to employ the method of Ottesen *et al*. (44) used in the past in the design of degenerate primers with an appended probe binding site (70) for use in digital PCR failed. This method is typically employed by necessity when little is known about an environmental gene sequence of interest. This obstacle was overcome in this study by utilizing nucleotide sequences of Family 3 [FeFe] hydrogenases cloned previously from the gut of *R. hesperus* collection ChiA2, see

Chapter 3. These sequences allowed for the design of probes targeting termite gut Family 3 [FeFe] hydrogenases, as described in the methods section.

In control experiments, the Family 3 [FeFe] hydrogenase and 16S rRNA gene sequences of *T. azotonutricium* ZAS-9 and *Treponema primitia* ZAS-2 co-amplified in dPCR. In these control experiments, there were, on average (2 panels analyzed in each of two dPCR experiments), 9.5 16S rRNA gene amplification false positives and less than 1 [FeFe] hydrogenase false positive observed in array panels loaded with PCR mix containing no template.

A heat map for a representative panel from which PCR products amplified from *Reticulitermes tibialis* collection JT2 gut microbes were retrieved is shown in Figure 5-2. The digital array devices, panels, and chambers from which samples putatively amplified from a single cell were successfully retrieved and subsequently sequenced are listed in Table 5-2. Samples were retrieved from 16 panels distributed across four digital PCR array devices. The panels from which product was retrieved had an average of 180 ± 51 positive 16S rRNA gene signals, 11 ± 5 [FeFe] hydrogenase gene signals, and 4 ± 3 putative co-amplifications.

Sequence Analysis. Twenty-seven chambers yielded co-amplification products successfully sequenced and putatively corresponding to a single cell genome according to the criterion put forward in the methods section, see Table 5-S1 in the chapter appendix and Table 5-2. Of these, 22, or 81%, were classified as treponemal, see Figure 5-3. All of the treponemal 16S rRNA gene sequences fell within one of two groupings. Either sub-group 2 of known treponemes defined previously by Paster *et al*. (47), also referred to as "sub-cluster II" by Hongoh *et al*. (20), and the group of phylotypes referred to as

Figure 5-2.

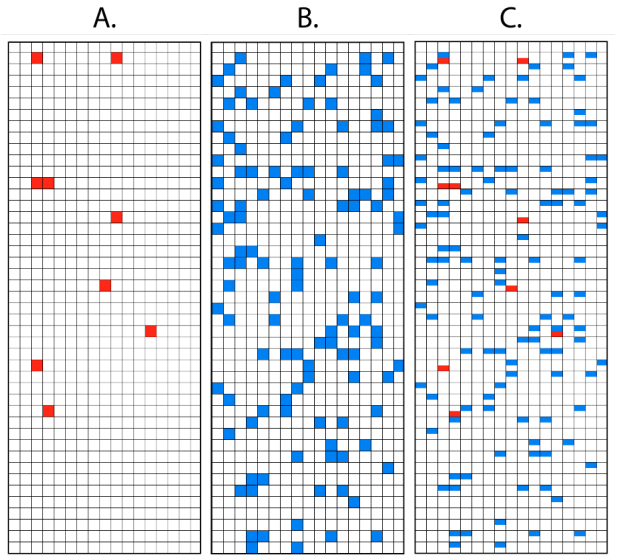

Figure 5-2. Representative heat map for a digital PCR array panel from which amplicons were retrieved for analysis. The heat maps were constructed using the Fluidigm Digital PCR Analysis program version 2.1.1, build 20090521.1140. The images correspond to panel 6 of the device with serial number 1151065028. A.) Wells with positive FAM-MGB signals corresponding to putative Family 3 FeFe hydrogenase gene amplicons, B.) Wells with positive VIC-MGB signals corresponding to 16S rRNA gene amplicons, C.) An overlay of the heat maps represented in A. and B. used to identify wells corresponding to positive co-amplifications.

Table 5-2. Digital PCR array wells from which amplicons were retrieved.

Device Serial No.	Panel[a]	H_2ase signal[b]	16S signal[c]	Co-amp.[d]	Well[e]
1151065021	2	3	129	1	A1
1151065021	7	6	129	3	A8
1151065021	8	12	152	5	A11, A12
1151065021	9	8	253	4	B2, B5
1151065021	10	6	127	1	B6
1151065021	11	14	212	9	B9
1151065030	9	22	244	10	C7
1151065030	11	12	248	4	D9
1151065028	1	9	195	4	D11, E4
1151065028	3	12	231	4	E5
1151065028	6	9	119	2	E9
1151065028	8	7	106	2	E12
1151065028	12	9	205	3	F4, F2
1151065037	7	14	211	3	G6, G7
1151065037	8	21	185	7	G12, G10, H2
1151065037	10	16	135	9	H6, H11, H8, H10, H7

[a]Panel (1 of 12) of the microfluidic digital PCR array.

[b]Number of wells having a positive FAM-MGB signal (see Methods), corresponding to a putative [FeFe] hydrogenase gene amplification.

[c]Number of wells having a positive VIC-MGB signal (see Methods), corresponding to a putative 16S rRNA gene amplification.

[d]Number of wells predicted to have successful 16S rRNA gene and [FeFe] hydrogenase gene amplifications.

[e]The names chosen for the wells from which amplicons were retrieved and subsequently yielded products putatively originating from a single cell genome (see Methods) and used in sequence analyses.

Figure 5-3.

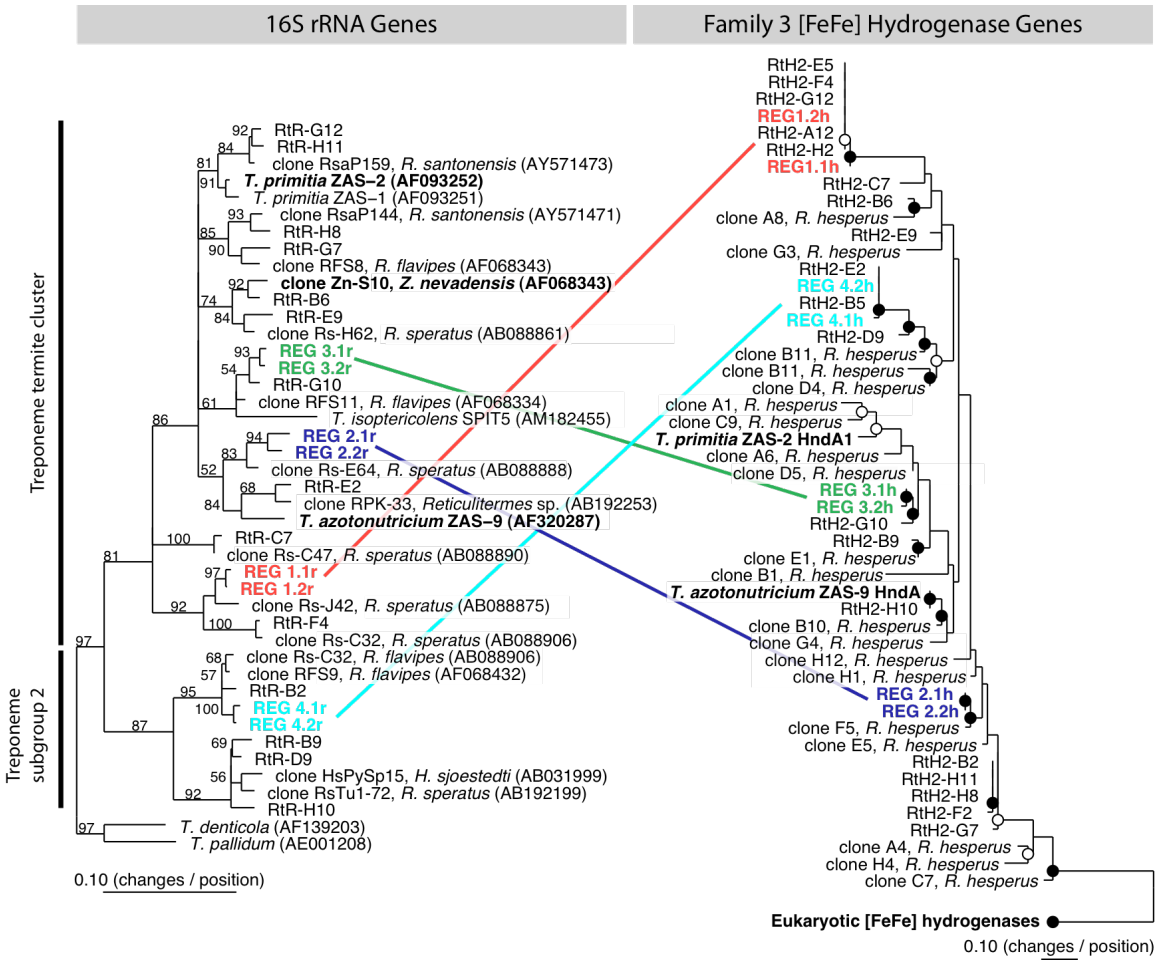
16S rRNA Genes
Family 3 [FeFe] Hydrogenase Genes
Treponeme termite cluster
Treponeme subgroup 2
RtR-G12
RtR-H11
clone RsaP159, R. santonensis (AY571473)
T. primitia ZAS–2 (AF093252)
T. primitia ZAS–1 (AF093251)
clone RsaP144, R. santonensis (AY571471)
RtR-H8
RtR-G7
clone RFS8, R. flavipes (AF068343)
clone Zn-S10, Z. nevadensis (AF068343)
RtR-B6
RtR-E9
clone Rs-H62, R. speratus (AB088861)
REG 3.1r
REG 3.2r
RtR-G10
clone RFS11, R. flavipes (AF068334)
T. isoptericolens SPIT5 (AM182455)
REG 2.1r
REG 2.2r
clone Rs-E64, R. speratus (AB088888)
RtR-E2
clone RPK-33, Reticulitermes sp. (AB192253)
T. azotonutricium ZAS–9 (AF320287)
RtR-C7
clone Rs-C47, R. speratus (AB088890)
REG 1.1r
REG 1.2r
clone Rs-J42, R. speratus (AB088875)
RtR-F4
clone Rs-C32, R. speratus (AB088906)
clone Rs-C32, R. flavipes (AB088906)
clone RFS9, R. flavipes (AF068432)
RtR-B2
REG 4.1r
REG 4.2r
RtR-B9
RtR-D9
clone HsPySp15, H. sjoestedti (AB031999)
clone RsTu1-72, R. speratus (AB192199)
RtR-H10
T. denticola (AF139203)
T. pallidum (AE001208)
0.10 (changes / position)
RtH2-E5
RtH2-F4
RtH2-G12
REG1.2h
RtH2-A12
RtH2-H2
REG1.1h
RtH2-C7
RtH2-B6
clone A8, R. hesperus
RtH2-E9
clone G3, R. hesperus
RtH2-E2
REG 4.2h
RtH2-B5
REG 4.1h
RtH2-D9
clone B11, R. hesperus
clone B11, R. hesperus
clone D4, R. hesperus
clone A1, R. hesperus
clone C9, R. hesperus
T. primitia ZAS-2 HndA1
clone A6, R. hesperus
clone D5, R. hesperus
REG 3.1h
REG 3.2h
RtH2-G10
RtH2-B9
clone E1, R. hesperus
clone B1, R. hesperus
T. azotonutricium ZAS-9 HndA
RtH2-H10
clone B10, R. hesperus
clone G4, R. hesperus
clone H12, R. hesperus
clone H1, R. hesperus
REG 2.1h
REG 2.2h
clone F5, R. hesperus
clone E5, R. hesperus
RtH2-B2
RtH2-H11
RtH2-H8
RtH2-F2
RtH2-G7
clone A4, R. hesperus
clone H4, R. hesperus
clone C7, R. hesperus
Eukaryotic [FeFe] hydrogenases
0.10 (changes / position)

Figure 5-3. Phylograms of 16S rRNA gene sequences (left panel) and putative Family 3 [FeFe] hydrogenase peptide sequences (right panel) corresponding to products co-amplified in dPCR. The 16S rRNA gene sequence tree was calculated using the TreePuzzle algorithm (37) with 1,000 puzzling steps and 963 unambiguously aligned nucleotides. The numbers adjacent to branches are bootstrap values. Sequences from the genomes of termite gut treponeme isolates known to encode Family 3 [FeFe] hydrogenases in their genomes are in boldface, as is a 16S rRNA gene (clone Zn-S10) shown previously to co-amplify with a gene for FTHFS in dPCR (45). The hydrogenase peptide sequence tree was calculated using a maximum likelihood (Phylip ProML) method with 70 unambiguously aligned amino acids. On the hydrogenase tree, open circles designate groupings also supported by either parsimony (Phylip PROTPARS, 1000 bootstraps) or distance matrix (Fitch) methods. Closed circles on the hydrogenase tree designate groupings supported by all three methods. For each REG (see text for definition), its associated sequences share the same color and a line has been connecting its hydrogenase and 16S sequences across the two trees. Sequence accession numbers are listed in parentheses. Clone names of sequences taken from public databases that are derived from a termite gut are each listed followed by a comma and the name of the termite host of origin. *Reticulitermes hesperus* hydrogenase sequences are taken from Chapter 3.

"the termite cluster" by Lilburn *et al*. (30), or as "subcluster I" by Hongoh *et al*. (20). Most, 16 of 22 total, of the treponemal 16S rRNA gene phylotypes fell within the termite cluster. All but one, RtR-B6, of the treponemal phylotypes grouped with sequences cloned in previous studies from termites belonging to the genus *Reticulitermes*. These sequences shared an average of 98.2% ± 0.86% (minimum of 96.5%), across 917 aligned nucleotides, with the *Reticulitermes* gut microbe sequences available in public databases to which they associated most closely in phylogenetic analyses. One 16S rRNA gene sequence grouped closely (98.1% sequence identity across 917 aligned nucleotides) in phylogenetic analyses with clone ZN-S10 from the gut of *Z. nevadensis* proposed previously results to originate from a bacterium also encoding a gene for FTHFS (45). The 4 REGs observed in this study, see Table 5-3, have been predicted to correspond to treponemes, see Figure 5-3.

Five of the 16S rRNA gene sequences fell outside the phylum Spirochaetes in phylogenetic analyses, see Figure 5-4. Three grouped with Beta-Proteobacteria, one with Bacteroidetes, and one with Epsilon-Proteobacteria. Three of these sequences shared an average of 98.2% ± 0.93% (minimum of 97.4%), across 917 aligned nucleotides, with the *Reticulitermes* gut microbe sequences available in public databases to which they associated mostclosely in phylogenetic analyses. The other two sequences did not group with any sequences cloned previously from termites belonging to the genus *Reticulitermes*. There was not an unambiguous correlation between 16S rRNA gene and [FeFe] hydrogenase peptide sequence phylogenies, see Figures 5-3 and 5-4. In 10 instances, Family 3 [FeFe] hydrogenase peptide sequences grouped closely in phylogenetic analyses with sequences cloned previously, see Chapter 3, from the gut

Table 5-3. *Reticulitermes* environmental genomovars proposed in this study.

Clone Name	Gene	REG	Min. REG Seq. Id. (%)[a]
RsH2-A1	[FeFe] hydrogenase	REG 1.1h	98.5%
RsH2-E12	[FeFe] hydrogenase	REG 1.2h	98.5%
RsH2-A11	[FeFe] hydrogenase	REG 2.1h	98.5%
RsH2-A8	[FeFe] hydrogenase	REG 2.2h	98.5%
RsH2-H6	[FeFe] hydrogenase	REG 3.1h	98.5%
RsH2-H7	[FeFe] hydrogenase	REG 3.2h	98.5%
RsH2-D11	[FeFe] hydrogenase	REG 4.1h	98.5%
RsH2-G6	[FeFe] hydrogenase	REG 4.2h	98.5%
RsR-A1	16S	REG 1.1r	99.0%
RsR-E12	16S	REG 1.2r	99.0%
RsR-A11	16S	REG 2.1r	98.5%
RsR-A8	16S	REG 2.2r	98.5%
RsR-H6	16S	REG 3.1r	98.9%
RsR-H7	16S	REG 3.2r	98.9%
RsR-D11	16S	REG 4.1r	98.5%
RsR-G6	16S	REG 4.2r	98.5%

[a]The minimum percent identity shared by all sequences corresponding to a given set of hydrogenase peptide sequences or 16S gene sequences comprising a REG (see Table 5-S1 in the chapter appendix for a list of all sequences)

Figure 5-4.

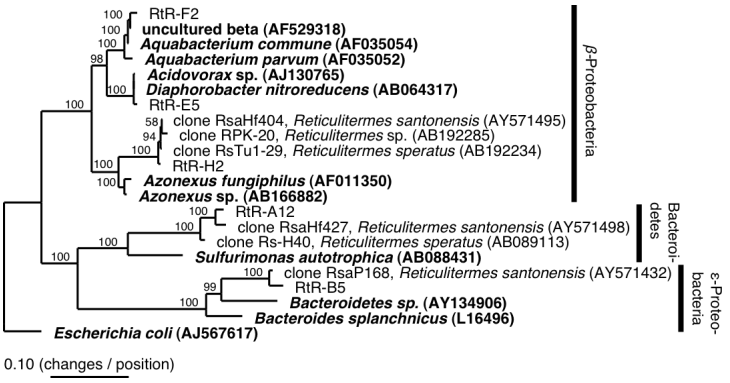

Figure 5-4. Phylogram of non-treponemal 16S rRNA gene sequences co-amplified with Family 3 [FeFe] hydrogenase genes in dPCR. The tree was calculated using the TreePuzzle algorithm (37) with 1,000 puzzling steps and 931 unambiguously aligned nucleotides. The numbers adjacent to branches are bootstrap values. 16S gene sequences from pure culture isolates are in boldface. Sequence accession numbers are listed in parentheses. Clone names of sequences taken from public databases that are derived from a termite gut are listed followed by a comma and the name of the termite host of origin.

microbial community of *Reticulitermes hesperus* collection ChiA2. In one instance, a hydrogenase peptide sequence, predicted to be treponemal in origin, was identical across 70 aligned amino acid residues with the Family 3 FeFe hydrogenase from *T. azotonutricium* ZAS-9.

Discussion

We have used microfluidic digital PCR to directly associate 16S rRNA gene sequences with Family 3 [FeFe] hydrogenase structural genes from the genomes of individual members of the gut microbial community of a lower termite. Hydrogen metabolism plays a prominent role in the degradation of wood in the termite gut (3, 7, 15, 39, 40, 49). By employing this microfluidic approach, we overcame many of the limitations inherent in traditional gene-inventory techniques for the characterization of bacterial genre or phyla filling a physiological niche in an environment.

Traditionally, to identify important microbial contributors to a particular physiological function in the environment, degenerate primers are designed targeting a gene important to a physiology of interest and environmental gene inventories are prepared using them (2, 14, 16, 62-64, 68, 69). The identity of the microbes from which the genes originate is then inferred based upon the phylogeny of the functional gene assuming a close correlation between the evolution of the gene and its cognate genome. This is not always a valid assumption, particularly in the case of [FeFe] hydrogenases whose phylogeny has been shown to be an unreliable predictor of the genus of its host of origin (54, 59). In fact, in this study we observed instances of very closely related [FeFe] hydrogenase gene sequences associating with 16S rRNA gene phylogtypes from different bacterial phyla, see Figures 5-3 and 5-4. This necessitates the use of methods such as microfluidic digital

PCR, a technique recently developed by Ottesen *et al*. (45), that allow for multiplex PCR using a single cell genome as template.

Employing the technique of microfluidic digital PCR, we have shown that treponemes may be an important, perhaps dominant, bacterial group encoding Family 3 [FeFe] hydrogenases in the gut of *Reticulitermes tibialis* collection JT2. This is important because Family 3 [FeFe] hydrogenase were the predominant group of hydrogenases observed in a termite hindgut metagenome and the only family for which translated products were detected *in situ* by mass spectroscopy (61). This family of [FeFe] hydrogenase was also observed in the genome sequences of two treponemes isolated from the gut of *Zootermopsis angusticolis,* see Chapter 2. Our results provide evidence consistent with an important role for treponemes in the metabolism of hydrogen, the "central free intermediate" in the degradation of wood by termites (49), in the gut microbial communities of termites. This is in agreement with findings supporting a similar contribution of treponemes to acetogenesis (48, 51), an important hydrogen sink in the termite gut ecosystem (3, 7, 26, 41, 49). In fact, one of the 16S rRNA gene phylotypes grouped closely with sequence Zn-S10 cloned from the gut of *Z. nevadensis*, shown previously in microfluidic digital PCR to co-amplify with an FTHFS gene, an important marker for acetogenesis (27, 28). Moreover, the majority of the 16S rRNA gene phylotypes fall within the termite cluster of treponemes defined by Lilburn *et al*. (30), which is the cluster into which most termite gut treponemal phylotypes have associated in previous 16S rRNA gene inventories (20, 30, 42, 67). The remaining treponemal 16S rRNA gene phylotypes associated with treponeme subgroup 2, as defined

by Paster *et al*. (47), and not subgroup 1, which is also consistent with the results of previous molecular inventories (20, 30, 42, 67).

In microfluidic digital PCR, there is always the possibility that false associations may be observed as a consequence of multiple cells or genomic fragments localizing to a single chamber. For this reason, we are suspicious of a 16S rRNA gene – [FeFe] hydrogenase gene co-localization in a dPCR chamber until the same co-localization is observed in another, independent, microfluidic chamber. Importantly, the only instances of such observed multiple co-localizations, referred to as REGs, corresponded to treponemal 16S rRNA gene phylotypes.

All but 3 of the 27 16S rRNA gene phylotypes observed grouped closely (98.2 ± 0.85% sequence identity across 917 aligned nucleotides) with phylotypes cloned previously from *Reticulitermes* termites, rather than those from any other genus. This provides further evidence for previous proposals of the "co-evolution" of the symbiotic microbes in the termite gut with their host proposed in Chapter 4 and elsewhere (19, 30, 67).

16S rRNA gene phylotypes were observed that fell outside the phylum Spirochaetes. These other phyla included Epsilon-Proteobacteria, Bacteroidetes, and Beta-Proteobacteria. It would not be unexpected to observe bacteria encoding hydrogenases within these phyla (50, 59), but we view these co-localizations with caution because no REGs were observed falling within any of these phyla in this study.

On account of the short length of the [FeFe] hydrogenase sequences cloned, corresponding to 70 unambiguously aligned amino acid residues, it is difficult to draw precise conclusions with regard to their phylogeny, though two important observations are noteworthy. First, clone RtH2-H10, which co-localized in dPCR with a treponemal

16S rRNA gene sequence, shared 100% sequence identity across 70 aligned amino acid residues with the Family 3 [FeFe] hydrogenase from *T. azotonutricium* ZAS-9. This provides support for the relevance of the gene of this strain, isolated from the gut of *Zootermopsis angusticolis* (26), to termite gut ecosystems and their hydrogen metabolism. Second, several of the sequences grouped with those cloned from *R. hesperus* collection ChiA2 in a previous study, see Chapter 3. This provides support broader relevance of the cloned hydrogenase sequences cloned in this study to *Reticulitermes* termite gut ecosystems.

Conclusions. We have used the technique of microfluidic digital PCR to demonstrate directly that treponemes may be an important, perhaps dominant, bacterial group encoding Family 3 [FeFe] hydrogenases in the gut of *Reticulitermes tibialis* collection JT2. This is in agreement with studies supporting an important role for termite gut treponemes in acetogenesis (48, 51) and the high fraction that they comprise of all gut eubacteria (8, 46). It is now of interest to identify what bacterial phyla other families of [FeFe] hydrogenases may associate with, or to employ microfluidic techniques to sequence termite gut microbe genomes to elucidate the precise physiological context of hydrogenase genes in the termite gut. Work should also be done to further understand the *in situ* expression of hydrogenases in the termite gut.

References

1. **Acinas, S. G., R. Sarma-Rupavtarm, V. Klepac-Ceraj, and M. F. Polz.** 2005. PCR-induced sequence artifacts and bias: Insights from comparison of two 16S rRNA clone libraries custructed from teh same sample. Appl. Environ. Microbiol. **71:**8966-8969.
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 J. M. Switzer.** 1986. Acetate Synthesis from H2 plus CO2 by Termite Gut Microbes. Appl. Environ. Microbiol. **52:**623-630.
8. **Brune, A.** 2006. Symbiotic Associations Between Termites and Prokaryotes. Prokaryotes **1:**439-474.

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

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

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

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

13. **DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen.** 2006. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl. Environ. Microbiol. **72:**5069-5072.

14. **Dumont, M. G., and J. C. Murrell.** 2005. Community-level analysis: key genes of aerobic methane oxidation. Methods. Enzymol. **397:**413-427.

15. **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.

16. **Fang, H. H. P., T. Zhang, and C. Li.** 2006. Characterization of Fe-hydrogenase genes diversity and hydrogen-production population in an acidophilic sludge. Journal of Bacteriology **126:**357-364.

17. **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.

18. **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.

19. **Hongoh, Y., P. Deevong, T. Inoue, S. Moriya, S. Trakulnaleamsai, M. Ohkuma, C. Vongkaluang, N. Noparatnaraporn, and T. Kudo.** 2005. Intra- and Interspecific Comparisons of Bacterial Diversity and Community Structure Support Coevolution of Gut Microbiota and Termite Host. Appl. Environ. Microbiol. **71:**6590-6599.

20. **Hongoh, Y., M. Ohkuma, and T. Kudo.** 2003. Molecular analysis of bacterial microbiota in the gut of the termite *Reticulitermes speratus* (Isoptera; Rhinotermitidae). FEMS Microbiol. Ecol. **44:**231-242.

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. **Kudo, T.** 2009. Termite-Microbe Symbiotic System and Its Efficient Degradation of Lignocellulose. Biosci. Biotechnol. Biochem. **73:**2561-2567.

24. **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.

25. **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.

26. **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.

27. **Leaphart, A. B., M. J. Friez, and C. R. Lovell.** 2003. Formyltetrahydrofolate synthetase sequences from salt marsh plant roots reveal a diversity of acetogenic bacteria and other bacterial functional groups. Appl. Environ. Microbiol. **69:**693-696.

28. **Leaphart, A. B., and C. R. Lovell.** 2001. Recovery and analysis of formyltetrahydrofolate synthetase gene sequences from natual populations of acetogenic bacteria. Appl. Environ. Microbiol. **67:**1392-1395.

29. **Lilburn, T. G., K. S. Kim, N. E. Ostrom, K. R. Byzek, J. R. Leadbetter, and J. A. Breznak.** 2001. Nitrogen Fixation by Symbiotic and Free-Living Spirochetes. Science **292:**2495-2498.

30. **Lilburn, T. G., T. M. Schmidt, and J. A. Breznak.** 1999. Phylogenetic diversity of termite gut spirochaetes. Environ. Microbiol. **1:**331-345.

31. **Lo, N., G. Tokuda, H. Watanabe, H. Rose, M. Slaytor, K. Maekawa, C. Bandi, and H. Noda.** 2000. Evidence from multiple gene sequences indicates that termites evolved from wood-feeding cockroaches. Curr. Biol. **23:**515-538.

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. **Markowitz, V. M., N. N. Ivanova, E. Szeto, K. Palaniappan, K. Chu, D. Dalevi, I.-M. A. Chen, Y. Grechkin, I. Dubchak, I. Anderson, A. Lykidis, K. Mavromatis, P. Hugenholtz, and N. C. Kyrpides.** 2008. IMG/M: a data management and analysis system for metagenomes. Nucleic Acids Res. **36:**D534-D538.

34. **McHardy, A. C., H. G. Martin, A. Tsirigos, P. Hugenholtz, and I. Rigoutsos.** 2007. Accurate phylogenetic classification of variable-length DNA fragments. Nat. Meth. **4:**63.

35. **Meyer, J.** 2007. [FeFe] hydrogenases and their evolution: a genomic perspective. Cell. Mol. Life Sci. **64:**1063-1084.

36. **Morgenstern, B.** 1999. DIALIGN 2: improvement of the segment-to-segment approach to multiple sequence alignment. Bioinformatics **15:**211-218.

37. **Néron, B., H. Ménager, C. Maufrais, N. Joly, T. Pierre, and C. Letondal.** 2008. Presented at the Bio Open Source Conference, Toronto.

38. **Noda, S., M. Ohkuma, R. Usami, K. KHorikoshi, and T. Kudo.** 1999. Culture-Independent Characterization of a Gene Responsible for Nitrogen Fixation in the Symbiotic Microbial Community in the Gut of the Termite *Neotermes koshunensis*. Appl. Environ. Microbiol. **65:**4935-4942.

39. **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.

40. **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.

41. **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.

42. **Ohkuma, M., T. Iida, and T. Kudo.** 1999. Phylogenetic relationships of symbiotic spirochetes in the gut of diverse termites. FEMS Microbiol. Lett. **181:**123-129.

43. **Ohkuma, M., S. Noda, and T. Kudo.** 1999. Phylogenetic Diversity of Nitrogen Fixation Genes in the Symbiotic Microbial Community in the Gut of Diverse Termites. Appl. Environ. Microbiol. **65:**4926-4934.

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

45. **Ottesen, E. A., J. W. Hong, S. R. Quake, and J. R. Leadbetter.** 2006. Microfluidic Digital PCR Enables Multigene Analysis of Individual Environmental Bacteria. Science **314:**1464-1467.

46. **Paster, B. J., F. E. Dewhirst, S. M. Cooke, V. Fussing, L. K. Poulsen, and J. A. Breznak.** 1996. Phylogeny of Not-Yet-Cultured Spirochetes from Termite Guts. Appl. Environ. Microbiol. **62:**347-352.

47. **Paster, B. J., F. E. Dewhirst, W. G. Weisburg, L. A. Tordoff, G. J. Fraser, R. B. Hespell, T. B. Stanton, L. Zablen, L. Mandelco, and C. R. Woese.** 1991. Phylogenetic Analysis of Spirochetes. Journal of Bacteriology **173:**6101-6109.

48. **Pester, M., and A. Brune.** 2006. Expression profiles of *fhs* (FTHFS) genes support the hypothesis that spirochaetes dominate reductive acetogenesis in the hindgut of lower termites. Environ. Microbiol. **8:**1261-1270.

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

50. **Robson, R.** 2001. Biodiversity of hydrogenases, p. 9-32. *In* R. Cammack, M. Frey, and R. Robson (ed.), Hydrogen as a Fuel: Learning from Nature. Taylor & Francis Inc., London.

51. **Salmassi, T. M., and J. R. Leadbetter.** 2003. Analysis of genes of tetrahydrofolate-dependent metabolism from cultivated spirochaetes and the gut community of the termite *Zootermopsis angusticollis*. Microbiology **149:**2529-2537.

52. **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.

53. **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.

54. **Schmidt, O., H. L. Drake, and M. A. Horn.** 2010. Hitherto unknown [Fe-Fe]-Hydrogenase Gene Diversity in Anaerobes and Anoxic Enrichments from a Moderately Acidic Fen. Appl. Environ. Microbiol. **76:**2027-2031.

55. **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.

56. **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.

57. **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.

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

59. **Vignais, P. M., and B. Billoud.** 2007. Occurence, Classification, and Biological Function of Hydrogenases: An Overview. Chem. Rev. **107:**4206-4272.

60. **Vignais, P. M., B. Billoud, and J. Meyer.** 2001. Classification and phylogeny of hydrogenases. FEMS Microbiol. Rev. **25:**455-501.

61. **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.

62. **Wawer, C., M. S. M. Jetten, and G. Muyzer.** 1997. Genetic Diversity and Expression of the [NiFe] Hydrogenase Large-Subunit Gene of *Desulfovibrio* spp. in Environmental Samples. Appl. Environ. Microbiol. **63:**4360-4369.

63. **Xing, D., N. Ren, and B. E. Rittmann.** 2008. Genetic Diversity of Hydrogen-Producing Bacteria in an Acidophilic Ethanol-H_2-Coproducing System, Analyzed Using the [Fe]-Hydrogenase Gene. Appl. Environ. Microbiol. **74:**1232-1239.

64. **Yamada, A., T. Inoue, S. Noda, Y. Hongoh, and M. Ohkhuma.** 2007. Evolutionary trend of phylogenetic diversity of nitrogen fixation genes in the gut community ofwood-feeding termites. Mol. Ecol. **16:**3768-3777.

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

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

67. **Yang, H., D. Schmitt-Wagner, U. Stingl, and A. Brune.** 2005. Niche heterogeneity determites bacterial community structure in the termite gut (*Reticulitermes santonensis*). Environ. Microbiol. **7:**916-932.

68. **Zehr, J. P., B. D. Jenkins, S. M. Short, and G. F. Steward.** 2003. Nitrogenase gene diversity and microbial community structure: a cross-comparison. Environ. Microbiol. **5**.

69. **Zehr, J. P., and L. A. McReynolds.** 1989. Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine cyanobacterium *Trichodesmium thiebautii*. Appl. Environ. Microbiol. **55:**2522-2526.

70. **Zhang, Y., D. Zhang, W. Li, J. Chen, Y. Peng, and W. Cao.** 2003. A novel real-time quantative PCR method using attached universal template probe. Nuc. Acids Res. **31:**e123.

Appendix

Table S1. Sequences cloned.

Figure S1. Alignment used in primer design.

Figure S2. Alignment used in probe design.

Table 5-S1. Sequences cloned and proposed *Reticulitermes* environmental genomovars (REG).

Clone Name[a]	RFLP Abundance[b]	Min. RFLP Seq. Id. (%)[c]	Gene	REG[d]	Min. REG Seq. Id. (%)[e]	Accession
RsH2-A1.2	1	97.1%	[FeFe] hydrogenase			local ARB database
RsH2-A1.3	1	97.1%	[FeFe] hydrogenase			local ARB database
RsH2-A1.6	2	97.1%	[FeFe] hydrogenase	REG 1.1h	98.5%	local ARB database
RsH2-A1.7	2	97.1%	[FeFe] hydrogenase			local ARB database
RsH2-A11.1	5	98.5%	[FeFe] hydrogenase	REG 2.1h	98.5%	local ARB database
RsH2-A11.10	1	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-A11.2	1	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-A11.3	1	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-A12.3	9	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-A12.5	1	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-A8.2	2	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-A8.4	4	100.0%	[FeFe] hydrogenase	REG 2.2h	98.5%	local ARB database
RsH2-A8.9	1	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-B2.1	8	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-B2.5	1	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-B5.1	9	-	[FeFe] hydrogenase			local ARB database
RsH2-B6.2	9	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-B6.4	1	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-B9.1	9	-	[FeFe] hydrogenase			local ARB database
RsH2-C7.10	10	-	[FeFe] hydrogenase			local ARB database
RsH2-D11.6	10	-	[FeFe] hydrogenase	REG 4.1h	98.5%	local ARB database
RsH2-D9.1	1	97.1%	[FeFe] hydrogenase			local ARB database
RsH2-D9.2	2	97.1%	[FeFe] hydrogenase			local ARB database
RsH2-D9.4	7	97.1%	[FeFe] hydrogenase			local ARB database
RsH2-E12.1	10	-	[FeFe] hydrogenase	REG 1.2h	98.5%	local ARB database
RsH2-E2.5	5	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-E2.7	1	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-E5.2	10	-	[FeFe] hydrogenase			local ARB database
RsH2-E9.1	10	-	[FeFe] hydrogenase			local ARB database
RsH2-F2.1	10	-	[FeFe] hydrogenase			local ARB database
RsH2-F4.1	10	-	[FeFe] hydrogenase			local ARB database
RsH2-G10.1	6	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-G10.5	1	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-G10.7	1	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-G10.9	1	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-G6.3	10	-	[FeFe] hydrogenase	REG 4.2h	98.5%	local ARB database
RsH2-G7.2	3	98.3%	[FeFe] hydrogenase			local ARB database
RsH2-G7.3	1	98.3%	[FeFe] hydrogenase			local ARB database
RsH2-H10.1	3	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-H10.3	1	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-H10.6	3	100.0%	[FeFe] hydrogenase			local ARB database
RsH2-H2.1	9	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-H2.6	1	98.5%	[FeFe] hydrogenase			local ARB database
RsH2-H6.1	10	-	[FeFe] hydrogenase	REG 3.1h	98.5%	local ARB database
RsH2-H7.1	9	-	[FeFe] hydrogenase	REG 3.2h	98.5%	local ARB database
RsH2-H8.1	10	-	[FeFe] hydrogenase			local ARB database
RsH2-G12.1	10	-	[FeFe] hydrogenase			local ARB database
RsH2-H11.1	10	-	[FeFe] hydrogenase			local ARB database
RsR-A1.1	10	-	16S	REG 1.1r	99.0%	local ARB database

Continuing Table 5-S1.

Clone Name[a]	RFLP Abundance[b]	Min. RFLP Seq. Id. (%)[c]	Gene	REG[d]	Min. REG Seq. Id. (%)[e]	Accession
RsR-A11.2	5	97.8%	16S	REG 2.1r	98.5%	local ARB database
RsR-A11.4	4	97.8%	16S			local ARB database
RsR-A11.1	1	97.8%	16S			local ARB database
RsR-A12.1	8	97.8%	16S			local ARB database
RsR-A12.2	1	97.8%	16S			local ARB database
RsR-A12.3	1	97.8%	16S			local ARB database
RsR-A8.1	5	98.0%	16S	REG 2.2r	98.5%	local ARB database
RsR-A8.3	2	98.0%	16S			local ARB database
RsR-A8.8	1	98.0%	16S			local ARB database
RsR-B2.1	5	-	16S			local ARB database
RsR-B5.1	7	98.3%	16S			local ARB database
RsR-B5.9	1	98.3%	16S			local ARB database
RsR-B5.5	1	98.3%	16S			local ARB database
RsR-B5.8	1	98.3%	16S			local ARB database
RsR-B6.2	8	98.8%	16S			local ARB database
RsR-B6.9	1	98.8%	16S			local ARB database
RsR-B9.1	6	99.1%	16S			local ARB database
RsR-B9.5	3	99.1%	16S			local ARB database
RsR-B9.10	1	99.1%	16S			local ARB database
RsR-C7.4	10	-	16S			local ARB database
RsR-D11.1	6	98.3%	16S	REG 4.1r	98.5%	local ARB database
RsR-D11.3	1	98.3%	16S			local ARB database
RsR-D9.1	9	99.1%	16S			local ARB database
RsR-D9.2	1	99.1%	16S			local ARB database
RsR-E12.1	9	99.3%	16S	REG 1.2r	99.0%	local ARB database
RsR-E12.4	1	99.3%	16S			local ARB database
RsR-E2.2	9	-	16S			local ARB database
RsR-E5.1	9	-	16S			local ARB database
RsR-E9.1	8	98.8%	16S			local ARB database
RsR-E9.2	1	98.8%	16S			local ARB database
RsR-E9.10	1	98.8%	16S			local ARB database
RsR-F2.3	7	98.3%	16S			local ARB database
RsR-F2.2	1	98.3%	16S			local ARB database
RsR-F2.6	1	98.3%	16S			local ARB database
RsR-F4.1	6	98.7%	16S			local ARB database
RsR-F4.5	1	98.7%	16S			local ARB database
RsR-F4.8	1	98.7%	16S			local ARB database
RsR-G10.1	5	99.4%	16S			local ARB database
RsR-G10.2	1	99.4%	16S			local ARB database
RsR-G10.7	1	99.4%	16S			local ARB database
RsR-G10.8	1	99.4%	16S			local ARB database
RsR-G12.1	6	99.2%	16S			local ARB database
RsR-G12.9	1	99.2%	16S			local ARB database
RsR-G12.10	1	99.2%	16S			local ARB database
RsR-G12.6	1	99.2%	16S			local ARB database
RsR-G6.1	9	99.3%	16S	REG 4.2r	98.5%	local ARB database
RsR-G6.5	1	99.3%	16S			local ARB database
RsR-G7.1	8	98.8%	16S			local ARB database
RsR-G7.3	1	98.8%	16S			local ARB database

Continuing Table 5-S1.

Clone Name[a]	RFLP Abundance[b]	Min. RFLP Seq. Id. (%)[c]	Gene	REG[d]	Min. REG Seq. Id. (%)[e]	Accession
RsR-H10.2	8	98.8%	16S			local ARB database
RsR-H10.8	1	98.8%	16S			local ARB database
RsR-H11.1	5	98.7%	16S			local ARB database
RsR-H11.2	1	98.7%	16S			local ARB database
RsR-H11.5	1	98.7%	16S			local ARB database
RsR-H11.7	1	98.7%	16S			local ARB database
RsR-H11.9	1	98.7%	16S			local ARB database
RsR-H11.10	1	98.7%	16S			local ARB database
RsR-H2.1	9	-	16S			local ARB database
RsR-H6.1	8	99.0%	16S	REG 3.1r	98.9%	local ARB database
RsR-H6.5	1	99.0%	16S			local ARB database
RsR-H7.1	9	-	16S	REG 3.2r	98.9%	local ARB database
RsR-H8.1	10	-	16S			local ARB database

[a]In boldface are sequences selected as representative of each set of sequences.

[b]Number of RFLP patterns corresponding to each sequence. For each set of clones, this is out of a total of 10 clones selected at random from a clone library prepared from a digital PCR well. Some sequences have less than 10 total RFLP patterns reported because several of the clones were false positives.

[c]The minimum percent sequence (peptide for hydrogenases and nucleotide for 16S rRNA genes) identity shared among each clone set.

[d]*Reticulitermes* environmental genomovar names.

[e]The minimum percent sequence identity shared by *all* hydrogenease gene (peptide sequence) or 16S gene (nucleotide sequence) clones corresponding to a particular REG.

Figure 5-S1.

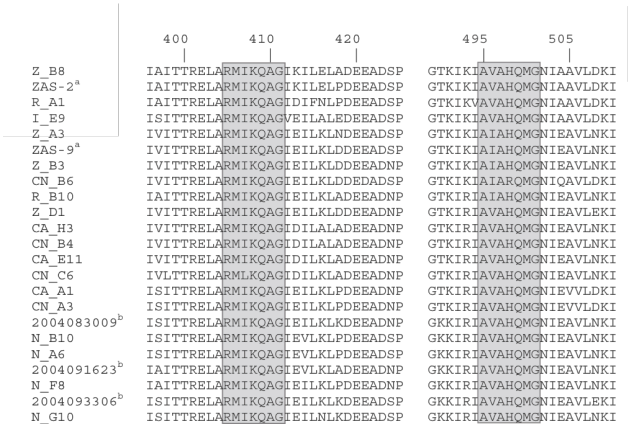

Figure 5-S1. Alignment used to design degenerate primers for use in digital PCR. The highlighted highly conserved regions were used in the design of degenerate primers. The following host designators were used: CA = *Cryptocercus punctulatus* adult, CN = *Cryptocercus punctulatus* nymph, N = *Nasutitermes* sp. Cost003, R = *Reticulitermes Hesperus* collection ChiA2, Z = *Zootermopsis nevadensis* collection ChiA1. The numbers-letter names following each host designator correspond to sample names taken from Chapters 3 and 4**.** [a]Family 3 [FeFe] hydrogenases identified in the genomes of *Treponema primitia* ZAS-2 (HndA1), and *Treponema azotonutricium* ZAS-9 (HndA), [b]IMG Gene Object Identifier.

Figure 5-S2.

Figure 5-S2. Alignment used in the design of locked nucleic acid (LNA) probes for use in digital PCR. See Figure 5-S1 for a description of sequence names used in the alignment. The highlighted, highly conserved region was used in the design of LNA probes.