

**The Biology and Community Structure of CO₂-Reducing
Acetogens in the Termite Hindgut**

Thesis by
Elizabeth Ann Ottesen

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended September 25, 2008)

© 2009

Elizabeth Ottesen

All Rights Reserved

Acknowledgements

Much of the scientist I have become, I owe to the fantastic biology program at Grinnell College, and my mentor Leslie Gregg-Jolly. It was in her molecular biology class that I was introduced to microbiology, and made my first attempt at designing degenerate PCR primers. The year I spent working in her laboratory taught me a lot about science, and about persistence in the face of experimental challenges.

At Caltech, I have been surrounded by wonderful mentors and colleagues. The greatest debt of gratitude, of course, goes to my advisor Jared Leadbetter. His guidance has shaped much of how I think about microbes and how they affect the world around us. And through all the ups and downs of these past six years, Jared's enthusiasm for microbiology—up to and including the occasional microscope session spent exploring a particularly interesting puddle—has always reminded me why I became a scientist in the first place.

The Leadbetter Lab has been a fantastic group of people. In the early days, Amy Wu taught me how much about anaerobic culture work and working with termites. These last few years, Eric Matson has been a wonderful mentor, endlessly patient about reading drafts and discussing experiments. Xinning Zhang also read and helped edit much of this work. As for the rest of the crew: Jean Huang, Yajuan Wang, Jong-In Han, Paul Orwin, Suvi Flagan, Andrew Hawkins, Abbie Green, Nick Ballor, and recently Adam Rosenthal; thanks so much for being there to bounce ideas off of and just hang around with!

Another amazing group to work with has been my collaborators in microfluidics. The last two chapters of this thesis wouldn't exist without the support and advice of Dr. Steve Quake. Both Steve and my coauthor Jong Wook Hong stuck with me and with the work all the way through the long early days before we finally managed to make it work. In later days, Luigi Warren helped a great deal with on-chip molecular biology and data analysis. More recently, Paul Blainey, Yann Marcy, Christina Fan, and Rick White have been great to work with, and endlessly helpful about negotiating the logistics of long-distance research.

Finally, I need to thank my family. There's no question that I wouldn't be where I am without my parent's support and the sacrifices they've made to give each and every one their children the best opportunities life can grant. Particular thanks are also owed to my sister Jen, who is yet another valuable member of my proofreading crew. The rest of my siblings—well, they're more likely than not to roll their eyes and ignore me when I start to talk about science (except for Eric, whom Jen and I have finally managed to lure to the scientist side of the force) but I wouldn't be the person I am without them. Last but certainly not least, this is for my grandfather, Alfred Hieronymus, and his insistence that if I wanted to claim there were three domains of life then I'd better get used to calling my work *Triology*.

Thank you, to everyone who has made this work possible.

Abstract

In the guts of wood-feeding termites, CO₂-reductive acetogenesis serves as the dominant sink for H₂ generated during the fermentation of wood polysaccharides. This activity can generate up to 1/3 of the acetate that powers the energy metabolism of the host insect. The gene for formyl-tetrahydrofolate synthetase (FTHFS), a key gene in the acetyl-CoA pathway, can be used as a genetic marker of acetogenic capability. The dominant FTHFS types in the guts of wood-feeding termites are known to cluster phylogenetically with those from acetogenic *Treponemes*. In this work, we present the discovery that the guts of wood-feeding roaches are also dominated by *Treponeme*-like sequences. Phylogenetic analysis of roach-derived FTHFS sequences reveals a cluster that forms a basal radiation of the termite *Treponeme* cluster. This suggests that the *Treponemes* found in roach guts represent an ancient divergence, present in the last common ancestor of these insects, rather than a modern lineage acquired by cross-species symbiont transfer. The FTHFS sequences present in the guts of higher termites were also examined. Wood-, palm-, and litter-feeding termites were found to be dominated by acetogenic *Treponemes*, while subterranean soil/grass feeders were found to be dominated by a novel cluster of *Firmicute*-like FTHFS types. Also presented herein is the development of microfluidic digital PCR for molecular characterization of individual bacteria from environmental samples. We used this technique to retrieve FTHFS and 16S rRNA gene sequences from single bacterial cells, thereby discovering the 16S rRNA sequences of uncultured acetogens in the termite gut. This technique should provide a valuable tool for molecular analyses of termite gut acetogens, and can potentially be adapted for the characterization of uncultured bacteria that carry any metabolic gene of interest.

Table of Contents

ACKNOWLEDGEMENTS	iii
ABSTRACT	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
LIST OF TABLES	x
<hr/>	
GENERAL INTRODUCTION	1-1
TERMITE PHYLOGENY AND BIOLOGY	1-2
TERMITE GUT MICROORGANISMS	1-4
ROLES FOR MICROBES IN TERMITE NUTRITION	1-7
INTRODUCTION TO ACETOGENESIS	1-9
ACETOGENESIS IN THE TERMITE GUT	1-12
ACETOGENIC BACTERIA ISOLATED FROM THE TERMITE HINDGUT	1-13
MOLECULAR COMMUNITY ANALYSIS OF TERMITE GUT ACETOGENS	1-16
ON THE ORGANIZATION OF THIS THESIS	1-19
REFERENCES	1-22
<hr/>	
<u>MOLECULAR COMMUNITY ANALYSIS OF ACETOGENIC BACTERIA IN ROACHES AND LOWER TERMITES: EVOLUTION OF THE SYMBIOSIS BETWEEN TERMITES AND ACETOGENIC <i>SPIROCHETES</i></u>	2-1
ABSTRACT	2-1
INTRODUCTION	2-1
MATERIALS AND METHODS	2-4
RESULTS	2-7
DISCUSSION	2-11
CHAPTER TWO APPENDIX	2-14
REFERENCES	2-18
<hr/>	
<u>MOLECULAR COMMUNITY ANALYSIS OF ACETOGENIC BACTERIA IN THE GUTS OF HIGHER TERMITES: COMMUNITY STRUCTURE IN TERMITES WITH DIVERSE FEEDING STRATEGIES</u>	3-1
ABSTRACT	3-1
INTRODUCTION	3-2
MATERIALS AND METHODS	3-4
RESULTS	3-7
DISCUSSION	3-18
CHAPTER THREE APPENDIX	3-22
REFERENCES	3-29
<hr/>	
<u>MICROFLUIDIC DIGITAL PCR FOR MULTIGENE ANALYSIS OF INDIVIDUAL ENVIRONMENTAL BACTERIA</u>	4-1

ABSTRACT	4-1
ARTICLE TEXT	4-2
MATERIALS AND METHODS	4-13
CHAPTER FOUR APPENDIX	4-19
REFERENCES	4-27

**MICROFLUIDIC DIGITAL PCR WITH DEGENERATE PRIMERS: MULTIPLEX
MOLECULAR COMMUNITY ANALYSIS OF ACETOGENIC BACTERIA IN THE
TERMITE HINDGUT** **5-1**

ABSTRACT	5-1
INTRODUCTION	5-1
MATERIALS AND METHODS	5-5
RESULTS	5-10
DISCUSSION	5-17
CHAPTER FIVE APPENDIX	5-20
REFERENCE	5-22

CONCLUSIONS **6-1**

List of Figures

Figure 1.1. Wood-Ljungdahl Pathway for CO ₂ -reductive acetogenesis.....	1-10
Figure 1.2. Phylogenetic analysis of FTHFS genes from acetogenic bacteria.	1-18
Figure 2.1. Mitochondrial cytochrome oxidase II phylogeny of termites and roaches. .	2-6
Figure 2.2. Phylogenetic analysis of termite <i>Treponeme</i> FTHFS sequences.....	2-9
Figure 2.3. Phylogenetic analysis of FTHFS sequences from roaches and lower termites.	2-10
Figure 3.1. Mitochondrial cytochrome oxidase II phylogeny of termites and roaches.	3-8
Figure 3.2. Phylogeny of major FTHFS clades found in termites and relatives.....	3-9
Figure 3.3. FTHFS sequences from potential acetogens.	3-12
Figure 3.4. Higher termite clade of termite <i>Treponemes</i>	3-13
Figure 3.5. Amitermes clade of probable <i>Firmicute</i> acetogens.	3-14
Figure 3.6. Phylogeny of <i>Moorella</i> / <i>Sporomusa</i> FTHFS clade.	3-15
Figure 3.7. Putative amino acid or purine-degrading FTHFS clades.....	3-16
Figure 3.8. Nonacetogenic FTHFS sequences.	3-17
Figure 4.1. Microfluidic Digital PCR Chip Architecture.....	4-6
Figure 4.2. Multiplex microfluidic digital PCR of single cells in environmental samples.	4-7
Figure 4.3. Phylogenetic Analysis of <i>Treponemal</i> 16S rRNA sequences retrieved from microfluidic chips.	4-8
Figure 4.4. Phylogenetic Analysis of 16S rRNA sequences retrieved from microfluidic chips and close relatives.....	4-9

Figure 4.5. “Clone H” and “Clone P Group” FTHFS genes are encoded by not-yet-cultivated termite gut treponemes.....	4-11
Figure 4.6. FTHFS primer specificity and demonstration of single copy sensitivity...	4-22
Figure 5.1. Microfluidic Digital PCR for “Lovell cluster” FTHFS and all-bacterial 16S rRNA genes.....	5-10
Figure 5.2. Phylogenetic analysis of FTHFS and 16S rRNA gene sequences amplified using microfluidic digital PCR.	5-12
Figure 5.3. Phylogenetic analysis of ClpX and 16S rRNA gene sequences amplified using microfluidic digital PCR.	5-14
Figure 5.4. Phylogenetic analysis of FTHFS and ClpX gene sequences amplified using microfluidic digital PCR.....	5-16

List of Tables

Table 1.1. Abundance of key bacterial phyla in termites of different feeding groups ^a ..	1-6
Table 1.2. Rates of acetogenesis and methanogenesis in the guts of selected termites	1-13
Table 1.3. CO ₂ -reducing acetogens isolated from termite guts.....	1-14
Table 2.1. Composition of FTHFS libraries constructed from <i>C. punctulatus</i> and <i>Incisitermes</i> sp. Pas1 ^a	2-8
Table 2.2. Operational Taxonomic Unit Grouping of FTHFS sequences identified in this study	2-15
Table 2.3. Sequences used in FTHFS phylogenetic analysis.....	2-16
Table 2.4. Sequences used in COII phylogenetic analysis.....	2-17
Table 3.1. FTHFS libraries constructed in this study.....	3-8
Table 3.2. Composition of FTHFS libraries from the hindgut microbiota of termites and relatives ^a	3-10
Table 3.3. Operational taxonomic unit grouping of FTHFS sequences identified in this study	3-23
Table 3.4. Sequences used in FTHFS phylogenetic analysis.....	3-26
Table 3.5. Sequences used in COII phylogenetic analysis.....	3-28
Table 4.1. Sequences Used for Phylogenetic Analysis.....	4-24
Table 5.1. Primers Used in Microfluidic Digital PCR	5-6
Table 5.2. FTHFS/16S rRNA gene pairs ^a	5-13
Table 5.3. ClpX/16S rRNA gene pairs.....	5-14
Table 5.4. FTHFS/ClpX gene pairs ^a	5-15
Table 5.5. Proposed Environmental Genomovars.....	5-17

Table 5.6. Sequences used in phylogenetic analysis.....5-20

General Introduction

The study of acetogenesis in the termite hindgut began with two key papers on methane production. The first was the 1982 proposal by Zimmerman et al. that the world termite population (2.4×10^{17}) consumed 28% of the total biomass produced each year, and could be responsible for 15%–56% of global yearly methane production (97). The second was Odelson and Breznak's 1983 study of fatty acid production in termite guts, and their observation that the ratio of CH₄ to CO₂ emitted by termites was far lower than that expected based on the current understanding of acetate fermentation in that system (66).

Later studies challenged Zimmerman's estimate; it is currently accepted that termites are responsible for up to 2% of global CO₂ and 2%–4% of global CH₄ production (81). However, Odelson and Breznak's observation has withstood the test of time; as they hypothesized, the dominant H₂ sink in wood-feeding termites is not methanogenesis but CO₂-reductive acetogenesis. As a result of acetogenesis, wood-feeding termites emit only trace quantities of methane (68), in stark contrast to the superficially similar cellulose-fermenting ecosystem of the cow rumen, the source of 8% of global methane production (48) (ruminants in general are responsible for 15%–19% (30)).

My work focuses on the bacteria responsible for acetogenesis in the termite gut: their evolutionary history, the effect of termite lifestyle on acetogen population structure, and

the development of molecular techniques for improved enumeration and identification of uncultured bacteria affiliated with this group.

In the introductory section of this thesis, I will briefly summarize key elements of termite phylogeny and nutritional ecology, the microbes present in the termite gut, and the roles played by gut microbes in termite nutrition. The focus will then shift to acetogenesis, its relationship to termite nutrition, the biology of acetogenic isolates from termites, and the results of previous molecular characterizations of the termite gut acetogens.

Termite Phylogeny and Biology

Termites are insects of the order *Isoptera*. *Isoptera* encompasses over 281 genera and 2,600 species (50). There are 7 generally accepted termite families, 6 of lower termites and the “higher termite” family *Termitidae*. Termites associate phylogenetically with the roach and mantid insect orders (46).

Given the focus of this work on the gut microbiota, it is important to discuss briefly the gut morphology of termites and its relationship to termite diet and microbial composition. The termite gut is divided into the foregut, which contains crop and gizzard, the midgut and the hindgut, which is the major site of microbial activity. The hindgut is divided into P1–P5 sections: the P1, a chamber of greater or lesser size, the P2, a valve between P1 and P3, the P3 paunch, the largest chamber of the hindgut, the P4 colon, and the P5 rectum.

Lower termites feed exclusively on dead plant material, primarily wood (some species eat grass) (50). They have a relatively simple gut structure, with a minimal P1 and the bulk of the symbiotic microbial community housed in a single, large chamber that encompasses both the expanded paunch (P3) and a tapering colon (P4) (63). All lower termites have symbiotic protists in their guts, which are thought to aid cellulose digestion (discussed in detail in later sections).

Higher termites are divided into 4–6 subfamilies. The *Macrotermitinae*, thought to be the most basal group of higher termites (47), are the fungus-cultivating termites. These termites harvest plant material and build it into combs for fungal growth; the fungus-degraded material is then digested by the termite (70). *Macrotermitinae*, likely due to the externalization of many symbiotic functions, have expanded midguts and reduced, relatively simple hindguts (64).

The remaining 3–5 subfamilies of higher termite make use of a diversity of feeding strategies, including wood-, grass-, litter-, and soil-feeding. These higher termites have complex hindguts, with well-defined P1 and P3 segments and frequently at least one additional segmentation in the P4/P5 region (64). Each of these chambers is relatively independent, with distinctive pH (4, 17, 18) and microbial communities (34, 73, 74, 85). In wood-feeding termites, the P1 segment has a pH 10-11 and a circumneutral P3 segment (17). Soil-feeders have P1 segments with pH 11-12.5, P3 with pH > 10, and neutral P4b (18).

Termite Gut Microorganisms

Termite guts contain complex microbial communities that span all three domains of life (13, 16). This assemblage represents a stable association; termite gut microorganisms are distinct from those present in the food supply and immediate environment of the host, and organisms found in one termite species are generally most closely related to microbes associated with other termites. In this work, I refer to this association as a *symbiosis* according to the original definition of that term, a close association of two or more organisms; this does not necessarily imply a beneficial or mutualistic relationship.

Termite Gut Protists

Termite gut protists are among the most visually striking and longest-studied termite gut symbionts. All lower termites harbor from 1–11 species of protists, which are key to the ability of these termites to digest wood (44). Protist species composition is generally host specific (53). These protists fall into three orders: *Hypermastigida*, *Trichomonadida*, and *Oxymonadida*. Hypermastigotes and Trichomonads have been shown to digest cellulose in axenic culture (93, 94). The only evidence for cellulose digestion by Oxymonads is the differential survival of some species in xylan fed vs. cellulose fed *Reticulitermes speratus* (79).

Many protist species within the termite have further symbioses with bacteria. Several termite gut flagellates have endosymbionts that may provide amino acid and cofactor synthetic capabilities (see comments on *Endomicrobia* in the gut bacteria section). Other prokaryote-protist symbioses include endosymbiotic methanogens (presumably involved

in H₂ transfer) (58), the use of ectosymbiotic bacteria to provide motility (22, 82), and the use of ectosymbiotic bacteria for osmotic regulation and sensory functions by *Streblomastix strix* (29).

Archaea in the Termite Gut

Archaea appear to represent a minor but constant population within the termite gut. In a dot-blot analysis, Archaea represented 0.83%–1.78% of the prokaryotic SSU rRNA in lower termites, 0.13%–1.68% in wood-feeding higher termites, and 1.42%–3.22% in soil-feeding higher termites (8). The best-studied archaeal group in the guts of termites are the methanogens. Wood-feeding termites produce little methane, but a few (presumably specialized) methanogens are present. They have been observed as symbionts of certain protist species (58) and colonizing the gut wall of *R. flavipes* (55). Soil-feeding termites, on the other hand, produce on average more methane (9); this is most likely reflected in the increased abundance of Archaea listed above. In these termites, methanogens are specifically associated with P4 and P5 gut compartments (85). Nonmethanogenic archaea are also abundant in the guts of some termites (8, 25, 34, 76), where their function remains ambiguous.

Bacteria of the Termite Hindgut

The guts of termites, like most animals, host a large diversity of bacteria. A summary of 16S rRNA analyses of gut bacterial diversity in representatives of each of the major feeding classes (wood-feeding lower termite, wood-feeding higher termite, fungus cultivating, soil-feeding) is presented in Table 1.1.

Table 1.1. Abundance of key bacterial phyla in termites of different feeding groups^a

Species	Food	<i>Spirochetes</i>	TG1	<i>Fibrobacter</i>	<i>Firmicutes</i>	<i>Proteobacteria</i>	CFB	Other
<i>Reticulitermes speratus</i> ^b	Wood	42–63	4–11	–	3–19	–	6–16	–
<i>Nasutitermes takasagoensis</i> ^c	Wood	62	–	10	10	3	8	–
<i>Odontotermes formosanus</i> ^d	Fungus	–	–	–	54	14	31	–
<i>Cubitermes</i> sp. ^e	Cultivating	–	–	–	–	–	–	–
	Soil	–	–	–	–	–	–	–
P1	–	–	–	–	96	–	4	–
P3	–	8	–	–	72	4	12	4
P4	–	10	–	–	50	20	10	10
P5	–	–	–	–	48	21	28	3

^a Abundance given as percent of total bacterial 16S rRNA sequences, N.D. not detected

^b From Hongoh, Ohkuma, and Kudo (40)

^c From Miyata et al. (61)

^d From Shinzato et al. (77)

^e From Schmitt-Wagner et al. (74)

The most abundant bacterial group in wood-feeding termites are the *Spirochetes*. Termite gut spirochetes largely affiliate with the genus *Treponema*; termite gut *Treponemes* have been implicated in acetogenesis (56), nitrogen fixation (59), and lignocellulose degradation (91). Bacteria from TG1 (Termite Group 1, sometimes referred to as *Endomicrobia*) are largely present as endosymbionts of gut protists (43, 80), and may be involved in amino acid and cofactor synthesis to supplement host nutrition (41). *Fibrobacter*-like bacteria (including the TG3 group) make up approximately 10% of the bacterial complement of higher termites (39), and may be involved in cellulose degradation (91). *Firmicutes* are abundant in the guts of many animals; termite-relevant physiological capabilities include acetogenesis (6, 15, 51, 52) and cellulose degradation (38). CFB group bacteria (mainly *Bacteroides*) and *Proteobacteria* are present in many termite species, but little is known about their physiologies in this environment.

Roles for Microbes in Termite Nutrition

Gut microbes play several important roles in termite nutrition. Given the emphasis of this work on CO₂-reductive acetogenesis, I will focus on cellulose fermentation, the major source of available reducing power (H₂). The nature of the soil components utilized by soil-feeding termites is still poorly understood as is the role of gut bacteria in substrate transformations (reviewed in (7)). Gut microbes have also been implicated in termite nitrogen balance, specifically nitrogen acquisition through nitrogen fixation and reclamation of nitrogenous waste through uric acid degradation (reviewed in (11)).

Cellulose Fermentation in Lower Termites

Early studies of cellulose fermentation focused on lower termites. In 1924, Cleveland demonstrated that termites could not survive on wood or cellulose when their symbiotic protozoa were removed (20, 21). Trager (89) and Hungate (42) extended Cleveland's work with demonstrations of cellulose decomposition by termite gut protozoa in mixed cultures. In 1978 and 1981, Yamin reported the first axenic cultures of *Trichomitopsis termopsidis* (93) and *Trichonympha sphaerica* (94) from *Z. angusticollis*, which allowed the unambiguous demonstration of cellulose degradation by these protists.

Cellulose Fermentation in Higher Termites

Wood-feeding higher termites harbor few gut protists. As a result, it has been proposed that these termites have gained the capability of digesting cellulose without the aid of digestive symbionts (78). Several endoglucanases have indeed been isolated from termite tissues (35, 45, 86, 88). However, transcriptional studies show that in wood-feeding

higher termites these enzymes are largely expressed in the salivary glands and midgut, with only low levels of expression in the hindgut (35).

Alternatively, symbiotic bacteria may have replaced protists as the primary cellulose fermenters in higher termites. This hypothesis was initially dismissed due to a lack of cultivated cellulose-degrading bacteria from termite guts and low rates of cellulase activity in the hindgut (78, 88). In 2007, Tokuda and Watanabe demonstrated bacteria-associated cellulase activity in the hindguts of two *Nasutitermes* species; the methodological change associated with this discovery was the performance of cellulase assays utilizing bacterial cell pellets rather than crude lysates (87). The recent metagenomic analysis by Warnecke et al. (91) demonstrated that in another wood-feeding *Nasutitermes* species: 1, the genomes of gut bacteria encode numerous putative endoglucanases and xylanases, 2, several of these genes have demonstrable activity when expressed in *E. coli*) and 3, proteins corresponding to these genes can be detected in P3 fluid (host-derived enzymes were not detected). This suggests that termite gut bacteria indeed play a significant role in degradation of wood polysaccharides by higher termites.

Hydrogen Production and Cellulose Degradation

The major products of cellulose degradation by termite gut protists are H₂, CO₂, and acetate according to the equation below (42, 65, 95).



The stoichiometry of cellulose fermentation within the guts of wood-feeding higher termites is unknown, but presumed to follow a similar pattern.

Acetate has been shown to accumulate in the guts of both higher and lower termites at concentrations of up to 80 mM, and oxidation of this acetate can account for up to 100% of the respiratory activity of the termite host (66). However, in 1983 Odelson and Breznak (66) observed that H₂ emissions from live termites were not sufficient to balance the observed rates of acetate production according to the equation above. While low rates of methane emission (a key H₂ sink in anaerobic ecosystems) were present, they too were insufficient to account for the missing electrons. As a result, they proposed that the H₂ generated during cellulose fermentation was being utilized for the reduction of CO₂ to acetate by acetogenic bacteria.

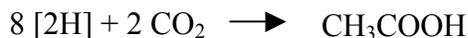
Introduction to Acetogenesis

H₂-mediated reduction of CO₂ is an important electron sink in many anaerobic ecosystems. In most environments, this niche is dominated by methanogenic archaea. CO₂-reductive acetogenesis is less energetically favorable ($\Delta G^{\circ} = -94.9$ kJ/mol for acetogenesis vs. -131.0 for methanogenesis) (72). However, for unknown reasons, acetogens can coexist with and even outcompete methanogenic archaea in some environments, including the termite gut (9, 49, 69).

Acetogenesis by bacteria from H₂ and CO₂ was first reported in 1932 by Fischer et al. (reported in German (31), reviewed in (26)). The model acetogen, *Moorella thermoacetica*, was at first characterized as a glucose-fermenting organism that produced acetate as the sole end product with stoichiometry 3 mol/mol glucose (32).



The discrepancy was proposed, and later demonstrated (2, 3), to be due to the pairing of glucose fermentation to acetate and CO_2 with reduction of that CO_2 to acetate, wherein:



The fermentation in the termite gut was projected to follow a similar pattern, but with protists carrying out the glucose fermentation and transferring the 8 reducing equivalents to acetogenic bacteria for the reduction of CO_2 to acetate (66).

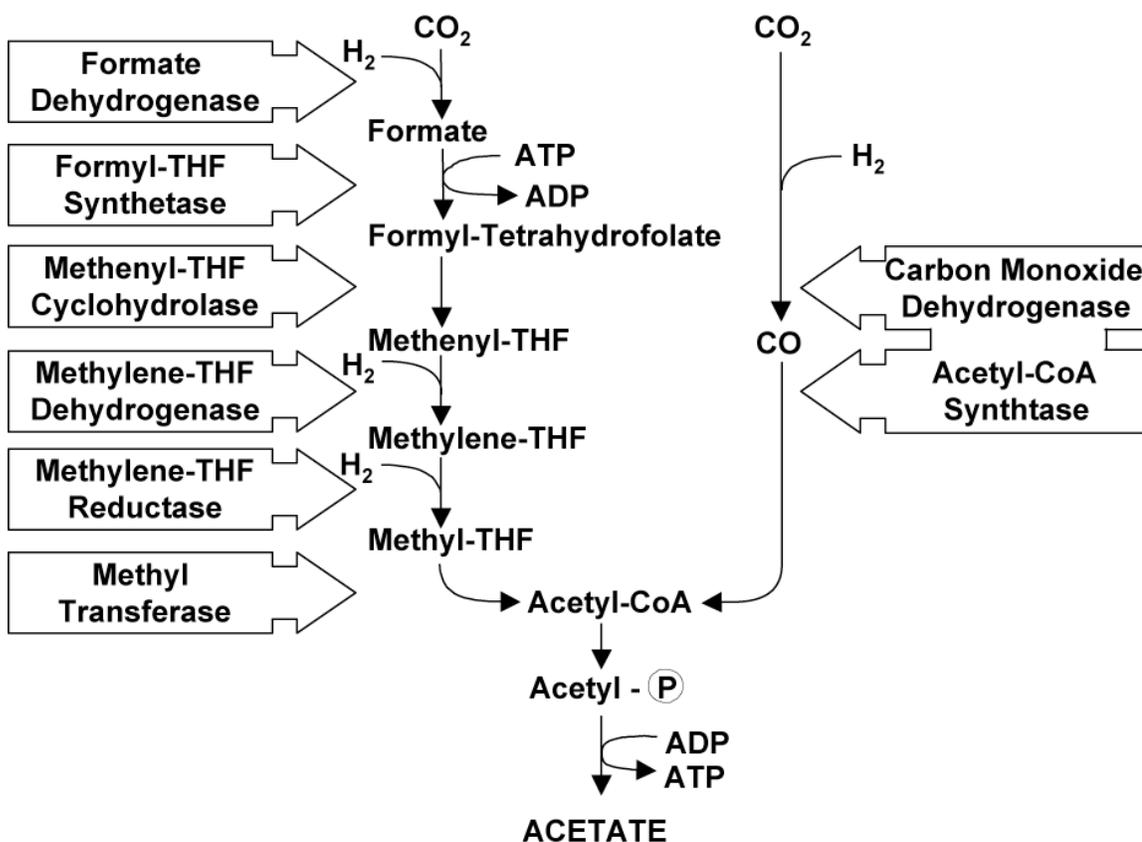
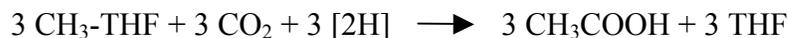


Figure 1.1. Wood-Ljungdahl Pathway for CO_2 -reductive acetogenesis. Reducing equivalents depicted as H_2 .

CO₂-reductive acetogenesis occurs via the Wood-Ljungdahl or acetyl-CoA cycle (60) (Figure 1.1). *Acetogens* are generally defined as bacteria utilizing this pathway as a major source of energy for growth and for CO₂ fixation into organic carbon (26).

One of the remarkable features of acetogenic bacteria is their metabolic flexibility. Methanogenic archaea are highly specialized, using H₂/CO₂, acetate, and a few other C1 compounds (92). The acetogens, however, make use of a diversity of substrates (26). Most acetogenic bacteria ferment a variety of carbohydrates and funnel the resultant reducing equivalents into the reduction of CO₂ (extrinsic or intrinsically generated during pyruvate conversion to acetate) (26). Additionally, many acetogens can directly feed reduced C1 units such as carbon monoxide, formate, and methanol into the acetyl-CoA pathway according to their redox potentials (24, 26). These reactions generally proceed as a disproportionation, where a subset of substrate molecules are oxidized in order to generate the required reducing equivalents for reduction of CO₂ to the carbonyl group of acetate. An example discussed in later sections is acetogenic o-demethylation of methoxylated aromatics. The methyl groups from these compounds enter the pathway at the level of methyl-THF; one methyl unit is oxidized to CO₂, (generating 3 reducing equivalents), for every 3 methyl units condensed with CO₂ to form acetate (33).



Finally, several acetogens have been reported to utilize alternative electron acceptors, such as nitrate (75) and the C=C double bonds in phenylacrylate derivatives (90).

While patterns of carbon flow during acetogenesis are fairly well understood, the energetics of acetogenesis are a bit harder to pin down. As can be seen in Figure 1.1, no net ATP is generated via substrate-level phosphorylation during acetogenesis from H₂ and CO₂. As a result, ATP must be generated via chemiosmotic phosphorylation. The reactions most likely to yield sufficient energy to translocate ions are the two final methyl transformations, catalyzed by methylene-THF reductase and methyl transferase (23). Acetogens can be grouped into those that depend on a proton or a sodium motive force; methylene-THF reductase has been proposed to drive proton translocation, while methyl transferase is considered a more likely driver of sodium translocation (62). However, the exact patterns of electron flow in these organisms remain unclear.

Acetogenesis in the Termite Gut

In 1986, H₂-dependent ¹⁴CO₂ reduction to acetate was demonstrated in termite gut homogenates, where it was found to occur at rates that were 2- to 33-fold higher than rates of methanogenesis (14). Table 1.2 presents measured rates of acetogenesis and methanogenesis from selected termites examined in this and an expanded study carried out in 1992 (9). A general pattern was observed in which wood-feeding lower and higher termites (represented here by *R. flavipes* and *N. nigriceps*) and the wood-feeding roach *C. punctulatus* had acetogenesis rates that outpaced methanogenesis. However, the reverse was observed in the guts of soil-feeding termites (*C. speciosus*) and the common cockroach (*P. americana*).

Table 1.2. Rates of acetogenesis and methanogenesis in the guts of selected termites

Species	Rate of acetogenesis from CO ₂ in gut homogenates ($\mu\text{mol acetate/g/hr}$)		Rate of CH ₄ emission from live animals ($\mu\text{mol CH}_4/\text{g/hr}$)
	Under N ₂	Under H ₂	
<i>Reticulitermes flavipes</i> ^a	0.09	0.93	0.10
<i>Nasutitermes nigriceps</i> ^a	0.89	3.68	0.24
<i>Cubitermes speciosus</i> ^a	0.01	0.02	0.85
<i>Cryptocercus punctulatus</i> ^b	0.04	0.14	<0.01
<i>Periplaneta americana</i> ^b	0.02	0.04	2.02
Beef Cow Rumen	0.00 ^c	0.05 ^c	0.9-1.1 ^d

a. From Brauman et al. (9)

b. From Breznak and Switzer (14) CH₄ production measured as ¹⁴CH₄ production in gut homogenates in presence of ¹⁴CO₂ and N₂ headspace (rather than emission)

c. From Le Van et al. (54)

d. Calculated based on 60-71 kg/cow/yr (48), assumes 450kg animal.

In 2007, Pester and Brune measured acetogenesis rates in three species of wood-feeding lower termites by microinjection of ¹⁴C-bicarbonate into guts that had been extracted, intact, from living termites (68). They observed rates of CO₂ fixation to acetate that corresponded to 22%–26% of the respiratory carbon turnover, confirming a major role for acetogenic bacteria in fueling host metabolism.

Acetogenic Bacteria Isolated from the Termite Hindgut

Over 100 species of acetogenic bacteria have been described (26). Of these, the overwhelming majority are *Firmicutes*. However, acetogenic capability is not monophyletic; several different lineages of acetogenic bacteria have been described, and many acetogens are closely related to nonacetogenic strains (26, 83).

Six species of acetogenic bacteria have been isolated from the guts of termites (Table 1.3). Four are acetogenic *Firmicutes*: *A. longum*, *C. mayombei*, *S. aerovorans*, and *S. termitida*. *A. longum* was isolated from the gut of a wood-feeding lower termite, and was

isolated from the highest dilution of the six strains. *S. termitida* was isolated from an enrichment using a single whole gut from a wood-feeding higher termite. *C. mayombei* and *S. aerovorans* were both isolated from soil feeders. The remaining two isolates, *T. primitia* strains ZAS-1 and ZAS-2, are acetogenic spirochetes isolated from the guts of the wood-feeding lower termite *Z. angusticollis*. This was the first report of acetogenesis, or chemolithoautotrophy in general, in a spirochete (56).

Table 1.3. CO₂-reducing acetogens isolated from termite guts

Species	Termite	Dilution	Reference
<i>Acetonema longum</i>	<i>Pterotermes occidentis</i>	10 ⁻⁶ dilution	(52)
<i>Clostridium mayombei</i>	<i>Cubitermes speciosus</i>	Not reported	(51)
<i>Sporomusa aerivorans</i>	<i>Thoracotermes macrothorax</i>	10 ⁻³ dilution	(5, 6)
<i>Sporomusa termitida</i>	<i>Nasutitermes nigriceps</i>	1 gut/tube	(12, 15)
<i>Treponema primitia</i> ZAS-1	<i>Zootermopsis angusticollis</i>	1 gut/tube	(36, 37, 56)
<i>Treponema primitia</i> ZAS-2	<i>Zootermopsis angusticollis</i>	1 gut/tube	(36, 37, 56)

Nutritional Characteristics of Termite Gut Acetogens

Acetogenesis from H₂ and CO₂ is the form most discussed in the context of the termite gut. This is due in part to evidence that it does play a major role in carbon cycling in the termite; as discussed above, H₂ is a major product of cellulose fermentation by termite gut protists, and observed rates of ¹⁴CO₂ reduction to acetate are sufficient to account for 22%–26% of the respiratory activity of the termite. However, termite gut acetogens are capable of utilizing a wide range of carbon sources, including mono- and disaccharides such as glucose (*A. longum*, *C. mayombei*, ZAS-1, ZAS-2), xylose (*C. mayombei*, ZAS-1, ZAS-2), and cellobiose (*C. mayombei*, ZAS-1) (see references in Table 1.3). Lactate and formate were identified by Tholen and Brune (84) as intermediates generated when ¹⁴C-glucose was injected into *R. flavipes*; the *Sporomusa* strains used both compounds (*S.*

termitida is noted as growing only weakly on formate), and *C. mayombei* utilized formate but not lactate.

Furthermore, *S. termitida* and *T. primitia* ZAS-2 are both capable of mixotrophic growth, simultaneously utilizing $H_2 + CO_2$ and organic substrates for carbon and energy (12, 36). This could allow these organisms to increase both the amount of energy per unit time generated by the cell and the amount of energy generated per mol H_2 (12). This ability has been invoked as a possible cause of the ability of acetogenic bacteria to outcompete methanogens in the termite gut (10). While the remaining termite gut acetogens are also capable of utilizing organic compounds, their ability to benefit from mixotrophy has not been investigated.

O-Demethylation of Aromatic Side Chains by Termite Gut Acetogens

In 1981, *Acetobacterium woodii* was shown to be capable of O-demethylation of methoxylated aromatic acids (1), and this activity has since been identified in many acetogens (33). While there is limited evidence for degradation of core lignin compounds in the guts of wood-feeding termites, lignin monomers can be utilized (13, 19). Ring cleavage appears to be minimal in the absence of oxygen, but side chain modifications are carried out under anaerobic conditions (19). Some of this activity might be attributable to acetogens; *S. aerivorans*, *S. termitida*, and *T. primitia* ZAS-2 are capable of growth by demethylation of methoxylated aromatics (see references in Table 1.3). *A. longum* and *C. mayombei* were each listed as growing consistently but weakly on a single modified aromatic (2,3,5-trimethoxybenzoate and syringate, respectively), while *T.*

primitia ZAS-1 did not utilize any of the four methoxylated aromatics tried (syringate, vanillate, ferulate, 2,3,5-trimethoxybenzoate).

Oxygen Reduction by Termite Gut Acetogens

The traditional view of the termite gut is that of a strictly anaerobic fermentation. This view is based on the oxygen sensitivity of termite gut protists (20) and the importance of anaerobic metabolic activities such as acetogenesis. However, Brune et al. have demonstrated conclusively that the gut epithelium does not serve as a barrier to oxygen diffusion, and that as a result the periphery of the gut may be microoxic (17).

This finding has stimulated research into oxygen tolerance and utilization by gut microbes. Termite gut acetogens *A. longum*, *S. aerivorans*, and *S. termitida* have been shown capable of growth when inoculated into media with up to 1.5% (*S. aerivorans*) O₂ in the headspace (5). These strains did not grow in the presence of oxygen; rather, the bacteria were able to use H₂ in the headspace to reduce oxygen, and resumed growth once the medium was anoxic. *T. primitia* (both ZAS-1 and ZAS-2) are described as “tolerating” O₂ concentrations of up to 0.5%, but it was not specified whether they grew in the presence of this oxygen or responded by reducing it prior to resuming growth (36).

Molecular Community Analysis of Termite Gut Acetogens

Although four of the six acetogenic isolates are *Firmicutes*, the numerical dominance of spirochetes in the termite gut makes it tempting to suggest *Treponemes* as key acetogens

in this environment. However, the fact that *T. primitia* ZAS-1 and ZAS-2 were not isolated from high dilutions made this difficult to prove.

Cultivation-based techniques are inherently limited to identification and enumeration of bacteria that will grow in the media chosen. Given the cryptic nutritive requirements of environmental bacteria, it is impossible to say with any certainty that the most abundant organisms cultured are indeed key players in the environment. Indeed, despite the abundance (more than half of the bacterial population in some termites) and importance of spirochetes in the termite gut, only 3 additional species have been successfully cultivated in the over 100 years of termite research; *Treponema azotonutricium* (37), *Spirochaeta coccoides* (27), and *Treponema isoptericolans* (28).

In the section on the bacteria of the termite hindgut, I discussed the results of 16S rRNA molecular community profiling studies carried out on gut bacteria from different termite species (Table 1.1). Acetogenic capability is not restricted to a single bacterial grouping (83), so 16S rRNA is not a suitable tool for characterization of acetogens. In 2001, Leaphart and Lovell discovered that the gene for formyl-tetrahydrofolate synthetase (FTHFS) from acetogenic *Firmicutes* was distinct from that of nonacetogens, and designed primers that would specifically amplify genes that fall within this cluster (57).

Leaphart and Lovell did not design their primer set to target spirochetal acetogens, as the FTHFS sequences from *T. primitia* ZAS-1 and ZAS-2 were not yet known. In 2003, Salmassi and Leadbetter demonstrated that the Lovell primers could amplify FTHFS

genes from ZAS-1 and ZAS-2 (71). Phylogenetic analysis of the recovered sequences found that they were not closely related to the *Treponema denticola* FTHFS sequence, but rather were closely related to FTHFS genes from “Lovell cluster” acetogenic *Firmicutes* (Figure 1.2). This suggests that *T. primitia* acquired its acetogenic capability by lateral gene transfer from a *Firmicute* acetogen.

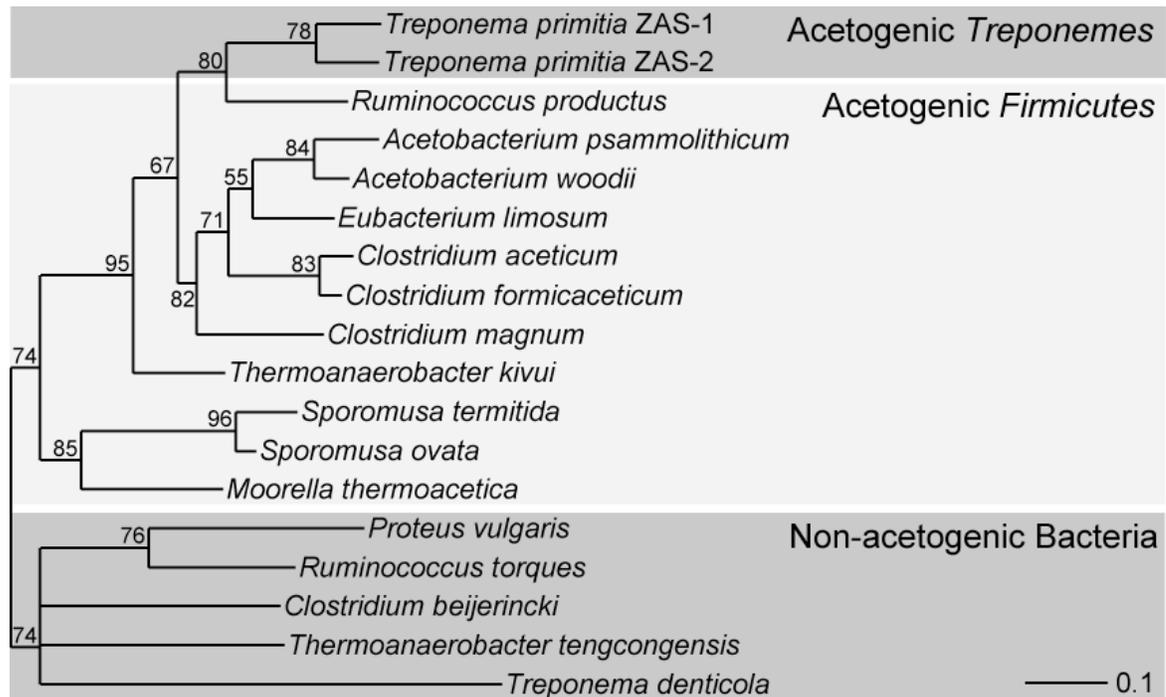


Figure 1.2. Phylogenetic analysis of FTHFS genes from acetogenic bacteria. Tree built using the TreePuzzle algorithm, with 1,000 puzzling steps and 345 unambiguously aligned amino acid positions. Scale bar represents 0.1 amino acid changes per alignment position.

Salmassi and Leadbetter also used the Lovell primers to build a community gene inventory of FTHFS sequences present in DNA extracted from the hindguts of *Zootermopsis nevadensis* workers (71). The majority of FTHFS types amplified from this environment grouped phylogenetically with ZAS-1 and ZAS-2 sequences, suggesting

that spirochetes are indeed the most abundant acetogens in this termite. Pester and Brune (67) reported similar results for two more wood-feeding lower termites, *Reticulitermes santonensis* and *Cryptotermes secundus*. Furthermore, they demonstrated that termite gut *Treponeme* FTHFS sequences were the most abundant FTHFS types in the community mRNA pool, showing that these organisms were actively utilizing the acetyl-CoA cycle *in situ* (67). Taken together, this evidence suggests a major role for spirochetes in acetate formation within the guts of wood-feeding lower termites.

On the Organization of This Thesis

My work has focused on furthering our understanding of the roles and community structure of acetogenic bacteria in the termite hindgut. In the second chapter of this thesis, I utilize the FTHFS-based community analysis method to examine the diversity of acetogenic bacteria present in the guts of wood-feeding roaches (*C. punctulatus*). These roaches, as shown in Table 1.2, have high rates of acetogenesis, but the nature of the acetogenic bacteria present in their guts was unknown. We demonstrated that wood-feeding roaches, like lower termites, host a diversity of acetogenic spirochetes. This, in addition to phylogenetic evidence placing roach-hosted spirochetes as basal to at least two key radiations of termite-derived sequences, suggests that acetogenic spirochetes arose prior to the roach-termite divergence. Additionally, it suggests that a diversity of sequence types were present in the last common ancestor, and that these bacteria gave rise to the complex species assemblage seen in lower termites today.

The third chapter of this thesis presents work in which I utilize the same techniques to explore the diversity of FTHFS-bearing organisms in higher termites. Higher termites, as mentioned above, have adapted to a variety of lifestyles. Some, like lower termites, feed exclusively on wood, but other termite species have adapted to using food sources at different stages of decomposition, up to and including soil. I explored the acetogenic community of 6 species of higher termite, 4 tropical species collected in Costa Rica and 2 desert-adapted species from California. A striking bifurcation was noted, as wood-, palm-, and litter-feeding species were dominated by *Treponeme*-like FTHFS types, while soil-exposed (and potentially soil-feeding) subterranean termite species were dominated by novel *Firmicute*-like FTHFS types. This suggests that the environmental conditions that allow high rates of acetogenesis in the guts of wood-feeding termites may correspond with those that favor *Treponemes* over other acetogenic bacteria.

In the fourth and fifth chapters of this thesis, I discuss the development of microfluidics-based tools for molecular characterization of uncultured microorganisms. In the previous section, I presented evidence that *Treponemes* are the dominant CO₂-reductive acetogens in the guts of wood-feeding termites. However, this hypothesis is based on the phylogenetic affiliation of a large cluster of FTHFS genes with those from *T. primitia* strains ZAS-1 and ZAS-2. Given that these *Treponemes* are believed to have acquired their FTHFS gene by lateral gene transfer, this affiliation should not be taken as definitive proof of identity. The fourth chapter of this thesis describes the development of technique for highly parallel, multiplex PCR interrogation of single bacterial cells from environmental samples. We used a microfluidic device to separate individual

microbes from the guts of *Z. nevadensis* and perform multiplex PCR reactions for simultaneous amplification and detection of bacterial 16S rRNA genes and a key FTHFS sequence type. Retrieval and analysis of PCR products from successful reactions allowed the rRNA-based species characterization of bacteria that hosted the targeted FTHFS gene, confirming the phylogenetics-based hypothesis that it was derived from a spirochete.

The final chapter of this thesis presents an expansion of the microfluidic technique described in chapter 4. The approach described in chapter 4 utilized sequence-specific Taqman probes to detect on-chip amplification of FTHFS and 16S rRNA genes. However, conventional Taqman probes are limited to detection of simple target populations. The presence of highly conserved sequence regions in bacterial rRNA genes allows the design of “*all-bacterial*” probes, but FTHFS probes were constrained to small clusters of highly similar sequences. In chapter 5, I present a modified “universal template probe” (96) strategy that allows multiplex detection of amplicons generated using degenerate primers. Using this system, we have developed a novel “Lovell cluster” FTHFS assay for detection and characterization of acetogenic bacteria using multiplex microfluidic PCR.

References

1. **Bache, R., and N. Pfennig.** 1981. Selective isolation of *Acetobacterium woodii* on methoxylated aromatic acids and determination of growth yields. Arch Microbiol **130**:255-261.
2. **Barker, H. A.** 1944. On the role of carbon dioxide in the metabolism of *Clostridium Thermoaceticum*. Proc Natl Acad Sci USA **30**:88-90.
3. **Barker, H. A., and M. D. Kamen.** 1945. Carbon dioxide utilization in the synthesis of acetic acid by *Clostridium thermoaceticum*. Proc Natl Acad Sci USA **31**:219-225.
4. **Bignell, D. E., and P. Eggleton.** 1995. On the elevated intestinal pH of higher termites (*Isoptera, Termitidae*). Insectes Sociaux **42**:57-69.
5. **Boga, H. I., and A. Brune.** 2003. Hydrogen-dependent oxygen reduction by homoacetogenic bacteria isolated from termite guts. Appl Environ Microbiol **69**:779-786.
6. **Boga, H. I., W. Ludwig, and A. Brune.** 2003. *Sporomusa aerivorans* sp. nov., an oxygen-reducing homoacetogenic bacterium from the gut of a soil-feeding termite. Int J Syst Evol Microbiol **53**:1397-1404.
7. **Brauman, A., D. E. Bignell, and I. Tayasu.** 2000. Soil-feeding termites: biology, microbial associations and digestive mechanisms, p. 233-259. In T. Abe, D. E. Bignell, and M. Higashi (ed.), Termites: Evolution, Sociality, Symbioses, Ecology. Kluwer Academic Publishers, Boston, MA.

8. **Brauman, A., J. Dore, P. Eggleton, D. Bignell, J. A. Breznak, and M. D. Kane.** 2001. Molecular phylogenetic profiling of prokaryotic communities in guts of termites with different feeding habits. *FEMS Microbiol Ecol* **35**:27-36.
9. **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.
10. **Breznak, J. A.** 1994. Acetogenesis from carbon dioxide in termite guts, p. 303-330. *In* H. A. Drake (ed.), *Acetogenesis*. Chapman and Hall, New York.
11. **Breznak, J. A.** 2000. Ecology of prokaryotic microbes in the guts of wood- and litter-feeding termites, p. 209-233. *In* T. Abe, D. E. Bignell, and M. Higashi (ed.), *Termites: Evolution, Sociality, Symbioses, Ecology*. Kluwer Academic Publishers, Boston, MA.
12. **Breznak, J. A., and J. S. Blum.** 1991. Mixotrophy in the termite gut acetogen, *Sporomusa termitida*. *Arch Microbiol* **156**:105-110.
13. **Breznak, J. A., and A. Brune.** 1994. Role of microorganisms in the digestion of lignocellulose by termites. *Ann Rev Entomol* **39**:453-487.
14. **Breznak, J. A., and J. M. Switzer.** 1986. Acetate synthesis from H₂ plus CO₂ by termite gut microbes. *Appl Environ Microbiol* **52**:623-630.
15. **Breznak, J. A., J. M. Switzer, and H. J. Seitz.** 1988. *Sporomusa termitida* sp. nov., an H₂/CO₂-utilizing acetogen isolated from termites. *Arch Microbiol* **150**:282-288.

16. **Brune, A.** 2006. Symbiotic associations between termites and prokaryotes, p. 439-474. *In* M. Dworkin, S. Falkow, E. Rosenber, K. H. Schleifer, and E. Stackebrandt (ed.), *The Prokaryotes*, Third ed, vol. I. Springer.
17. **Brune, A., D. Emerson, and J. A. Breznak.** 1995. The termite gut microflora as an oxygen sink: microelectrode determination of oxygen and pH gradients in guts of lower and higher termites. *Appl Environ Microbiol* **61**:2681-2687.
18. **Brune, A., and M. Kuhl.** 1996. pH profiles of the extremely alkaline hindguts of soil-feeding termites (Isoptera: Termitidae) determined with microelectrodes. *J Insect Physiol* **42**:1121-1127.
19. **Brune, A., E. Miambi, and J. A. Breznak.** 1995. Roles of oxygen and the intestinal microflora in the metabolism of lignin-derived phenylpropanoids and other monoaromatic compounds by termites. *Appl Environ Microbiol* **61**:2688-2695.
20. **Cleveland, L. R.** 1925. The effects of oxygenation and starvation on the symbiosis between the termite, *Termopsis*, and its intestinal flagellates. *Biol Bull* **48**:309-312.
21. **Cleveland, L. R.** 1924. The physiological and symbiotic relationships between the intestinal protozoa of termites and their host, with special reference to *Reticulitermes flavipes* Kollar. *Biol Bull* **46**:178-201.
22. **Cleveland, L. R., and A. V. Grimstone.** 1964. Fine structure of flagellate *Mixotricha paradoxa* and its associated microorganisms. *Proc Royal Soc London B Biol Sci* **159**:668-686.

23. **Diekert, G., and G. Wohlfarth.** 1994. Energetics of acetogenesis from C1 units, p. 157-179. *In* H. Drake (ed.), Acetogenesis. Chapman and Hall, New York, NY.
24. **Diekert, G., and G. Wohlfarth.** 1994. Metabolism of homocetogens. *Antonie Van Leeuwenhoek* **66**:209-221.
25. **Donovan, S. E., K. J. Purdy, M. D. Kane, and P. Eggleton.** 2004. Comparison of *Euryarchaea* strains in the guts and food-soil of the soil-feeding termite *Cubitermes fungifaber* across different soil types. *Appl Environ Microbiol* **70**:3884-3892.
26. **Drake, H., K. Küsel, and C. Matthies.** 2006. Acetogenic prokaryotes, p. 354-420. *In* M. Dworkin, S. Falkow, E. Rosenber, K. H. Schleifer, and E. Stackebrandt (ed.), *The Prokaryotes*, Third ed. Springer.
27. **Dröge, S., J. Fröhlich, R. Radek, and H. König.** 2006. *Spirochaeta coccoides* sp. nov., a novel coccoid spirochete from the hindgut of the termite *Neotermes castaneus*. *Appl Environ Microbiol* **72**:392-397.
28. **Dröge, S., R. Rachel, R. Radek, and H. König.** 2008. *Treponema isoptericolens* sp. nov., a novel spirochaete from the hindgut of the termite *Incisitermes tabogae*. *Int J Syst Evol Microbiol* **58**:1079-1083.
29. **Dyer, B. D., and O. Khalsa.** 1993. Surface bacteria of *Streblomastix strix* are sensory symbionts. *Biosystems* **31**:169-180.
30. **Ehhalt, D., M. Parther, F. Dentener, R. Derwent, E. Dlugokencky, E. Holland, I. Isaksen, J. Katima, V. Kirchhoff, P. Matson, P. Midgley, and M. Wang.** 2001. Atmospheric chemistry and greenhous gases, p. 241-287. *In* J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. Van der Linden, X. Dai, K.

Maskell, and C. A. Johnson (ed.), *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, New York, NY.

31. **Fischer, F., R. Lieske, and K. Winzer.** 1932. Biological gas reactions II Concerning the formation of acetic acid in the biological conversion of carbon oxide and carbonic acid with hydrogen to methane. *Biochem Z* **245**:2-12.
32. **Fontaine, F. E., W. H. Peterson, E. McCoy, M. J. Johnson, and G. J. Ritter.** 1942. A new type of glucose fermentation by *Clostridium thermoaceticum* n sp. *J Bacteriol* **43**:701-715.
33. **Frazer, A. C.** 1994. O-Demethylation and other transformations of aromatic compounds by acetogenic bacteria, p. 445-483. *In* H. L. Drake (ed.), *Acetogenesis*. Chapman and Hall, New York, NY.
34. **Friedrich, M. W., D. Schmitt-Wagner, T. Lueders, and A. Brune.** 2001. Axial differences in community structure of *Crenarchaeota* and *Euryarchaeota* in the highly compartmentalized gut of the soil-feeding termite *Cubitermes orthognathus*. *Appl Environ Microbiol* **67**:4880-4890.
35. **Gaku Tokuda, N. L. H. W. G. A. T. M. H. N.** 2004. Major alteration of the expression site of endogenous cellulases in members of an apical termite lineage. *Mol Ecol* **13**:3219-3228.
36. **Graber, J. R., and J. A. Breznak.** 2004. Physiology and nutrition of *Treponema primitia*, an H₂/CO₂-acetogenic spirochete from termite hindguts. *Appl Environ Microbiol* **70**:1307-1314.

37. **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.
38. **Hethener, P., A. Brauman, and J. L. Garcia.** 1992. *Clostridium termitidis* sp. nov., a cellulolytic bacterium from the gut of the wood-feeding termite, *Nasutitermes lujae*. *Syst Appl Microbiol* **15**:52-58.
39. **Hongoh, Y., P. Deevong, S. Hattori, T. Inoue, S. Noda, N. Noparatnaraporn, T. Kudo, and M. Ohkuma.** 2006. Phylogenetic diversity, localization, and cell morphologies of members of the candidate phylum TG3 and a subphylum in the phylum *Fibrobacteres*, recently discovered bacterial groups dominant in termite guts. *Appl Environ Microbiol* **72**:6780-6788.
40. **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.
41. **Hongoh, Y., V. K. Sharma, T. Prakash, S. Noda, T. D. Taylor, T. Kudo, Y. Sakaki, A. Toyoda, M. Hattori, and M. Ohkuma.** 2008. Complete genome of the uncultured Termite Group 1 bacteria in a single host protist cell. *Proc Natl Acad Sci USA* **105**:5555-5560.
42. **Hungate, R. E.** 1943. Quantitative analyses on the cellulose fermentation by termite protozoa. *Ann Entomol Soc of Amer* **36**:730-739.
43. **Ikeda-Ohtsubo, W., M. Desai, U. Stingl, and A. Brune.** 2007. Phylogenetic diversity of "Endomicrobia" and their specific affiliation with termite gut flagellates. *Microbiol* **153**:3458-3465.

44. **Inoue, T., O. Kitade, T. Yoshimura, and I. Yamaoka.** 2000. Symbiotic association with protists, p. 275-288. *In* T. Abe, D. E. Bignell, and M. Higashi (ed.), *Termites: Evolution, Sociality, Symbioses, Ecology*. Kluwer Academic Publishers, Dordrecht.
45. **Inoue, T., S. Moriya, M. Ohkuma, and T. Kudo.** 2005. Molecular cloning and characterization of a cellulase gene from a symbiotic protist of the lower termite, *Coptotermes formosanus*. *Gene* **349**:67-75.
46. **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.
47. **Inward, D. J., A. P. Vogler, and P. Eggleton.** 2007. A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspects of their evolutionary biology. *Mol Phylogenet Evol* **44**:953-967.
48. **Johnson, K. A., and D. E. Johnson.** 1995. Methane emissions from cattle. *J Anim Sci* **73**:2483-2492.
49. **Jones, J. G., and B. M. Simon.** 1985. Interaction of acetogens and methanogens in anaerobic freshwater sediments. *Appl Environ Microbiol* **49**:944-948.
50. **Kambhampati, S., and P. Eggleton.** 2000. Taxonomy and phylogeny of termites, p. 1-24. *In* T. Abe, D. E. Bignell, and M. Higashi (ed.), *Termites: Evolution, Sociality, Symbioses, Ecology*. Kluwer Academic Publishers, Dordrecht.

51. **Kane, M. D., A. Brauman, and J. A. Breznak.** 1991. *Clostridium mayombei* sp. nov., an H₂/CO₂ acetogenic bacterium from the gut of the African soil-feeding termite, *Cubitermes speciosus*. Arch Microbiol **156**:99-104.
52. **Kane, M. D., and J. A. Breznak.** 1991. *Acetonema longum* gen. nov. sp. nov., an H₂/CO₂ acetogenic bacterium from the termite, *Pterotermes occidentis*. Arch Microbiol **156**:91-98.
53. **Kitade, O., and T. Matsumoto.** 1993. Symbiotic protistan faunas of *Reticulitermes* (Isoptera, Rhinotermitidae) in the Japan Archipelago. Sociobiol **23**:135-153.
54. **Le Van, T. D., J. A. Robinson, J. Ralph, R. C. Greening, W. J. Smolenski, J. A. Leedle, and D. M. Schaefer.** 1998. Assessment of reductive acetogenesis with indigenous ruminal bacterium populations and *Acetitomaculum ruminis*. Appl Environ Microbiol **64**:3429-3436.
55. **Leadbetter, J. R., and J. A. Breznak.** 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**:3620-3631.
56. **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.
57. **Leaphart, A. B., and C. R. Lovell.** 2001. Recovery and analysis of formyltetrahydrofolate synthetase gene sequences from natural populations of acetogenic bacteria. Appl Environ Microbiol **67**:1392-1395.

58. **Lee, M. J., P. J. Schreurs, A. C. Messer, and S. H. Zinder.** 1987. Association of methanogenic bacteria with flagellated protozoa from a termite hindgut. *Curr Microbiol* **15**:337-341.
59. **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.
60. **Ljungdahl, L. G.** 1986. The autotrophic pathway of acetate synthesis in acetogenic bacteria. *Ann Rev Microbiol* **40**:415-450.
61. **Miyata, R., N. Noda, H. Tamaki, K. 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.
62. **Müller, V.** 2003. Energy conservation in acetogenic bacteria. *Appl Environ Microbiol* **69**:6345-6353.
63. **Noirot, C.** 1995. The gut of termites (*Isoptera*) comparative anatomy, systematics, phylogeny. I. Lower termites. *Ann Soc Entomol Fr* **31**:197-226.
64. **Noirot, C.** 2001. The gut of termites (*Isoptera*) comparative anatomy, systematics, phylogeny. II. - Higher termites (*Termitidae*). *Ann Soc Entomol Fr* **37**:431-471.
65. **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.

66. **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.
67. **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.
68. **Pester, M., and A. Brune.** 2007. Hydrogen is the central free intermediate during lignocellulose degradation by termite gut symbionts. *Isme J* **1**:551-565.
69. **Phelps, T. J., and J. G. Zeikus.** 1984. Influence of pH on terminal carbon metabolism in anoxic sediments from a mildly acidic lake. *Appl Environ Microbiol* **48**:1088-1095.
70. **Rouland-Lefèvre, C.** 2000. Symbiosis with Fungi, p. 289-306. *In* T. Abe, D. E. Bignell, and M. Higashi (ed.), *Termites: Evolution, Sociality, Symbioses, Ecology*. Kluwer Academic Publishers, Boston, MA.
71. **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*. *Microbiol* **149**:2529-2537.
72. **Schink, B.** 1997. Energetics of syntrophic cooperation in methanogenic degradation. *Microbiol Mol Biol Rev* **61**:262-280.
73. **Schmitt-Wagner, D., M. W. Friedrich, B. Wagner, and A. Brune.** 2003. Axial dynamics, stability, and interspecies similarity of bacterial community structure in the highly compartmentalized gut of soil-feeding termites (*Cubitermes* spp.). *Appl Environ Microbiol* **69**:6018-6024.

74. **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.
75. **Seifritz, C., S. L. Daniel, A. Gössner, and H. L. Drake.** 1993. Nitrate as a preferred electron sink for the acetogen *Clostridium thermoaceticum*. *J Bacteriol* **175**:8008-8013.
76. **Shinzato, N., T. Matsumoto, I. Yamaoka, T. Oshima, and A. Yamagishi.** 1999. Phylogenetic diversity of symbiotic methanogens living in the hindgut of the lower termite *Reticulitermes speratus* analyzed by PCR and in situ hybridization. *Appl Environ Microbiol* **65**:837-840.
77. **Shinzato, N., M. Muramatsu, T. Matsui, and Y. Watanabe.** 2007. Phylogenetic analysis of the gut bacterial microflora of the fungus-growing termite *Odontotermes formosanus*. *Biosci Biotechnol Biochem* **71**:906-915.
78. **Slaytor, M.** 1992. Cellulose digestion in termites and cockroaches - What role do symbionts play? *Comp Biochem Physiol B Biochem Mol Biol* **103**:775-784.
79. **Slaytor, M., A. Sugimoto, J. Azuma, K. Murashima, and T. Inoue.** 1997. Cellulose and xylan utilisation in the lower termite *Reticulitermes speratus*. *J Insect Physiol* **43**:235-242.
80. **Stingl, U., R. Radek, H. Yang, and A. Brune.** 2005. "Endomicrobia": cytoplasmic symbionts of termite gut protozoa form a separate phylum of prokaryotes. *Appl Environ Microbiol* **71**:1473-1479.

81. **Sugimoto, A., D. E. Bignell, and J. A. MacDonald.** 2000. Global impact of termites on the carbon cycle and atmospheric trace gases, p. 409-435. *In* T. Abe, D. E. Bignell, and M. Higashi (ed.), *Termites: Evolution, Sociality, Symbioses, Ecology*. Kluwer Academic Publishers, Norwell, MA.
82. **Tamm, S. L.** 1982. Flagellated ecto-symbiotic bacteria propel a eukaryotic cell. *J Cell Biol* **94**:697-709.
83. **Tanner, R. S., and C. R. Woese.** 1994. A phylogenetic assessment of the acetogens, p. 254-269. *In* H. L. Drake (ed.), *Acetogenesis*. Chapman and Hall, New York, NY.
84. **Tholen, A., and A. Brune.** 2000. Impact of oxygen on metabolic fluxes and *in situ* rates of reductive acetogenesis in the hindgut of the wood-feeding termite *Reticulitermes flavipes*. *Environ Microbiol* **2**:436-449.
85. **Tholen, A., and A. Brune.** 1999. Localization and *in situ* activities of homoacetogenic bacteria in the highly compartmentalized hindgut of soil-feeding higher termites (*Cubitermes* spp.). *Appl Environ Microbiol* **65**:4497-4505.
86. **Tokuda, G., N. Lo, H. Watanabe, M. Slaytor, T. Matsumoto, and H. Noda.** 1999. Metazoan cellulase genes from termites: intron/exon structures and sites of expression. *Biochem Biophys Acta* **1447**:146-159.
87. **Tokuda, G., and H. Watanabe.** 2007. Hidden cellulases in termites: revision of an old hypothesis. *Biology Letters* **3**:336-339.
88. **Tokuda, G., H. Watanabe, T. Matsumoto, and H. Noda.** 1997. Cellulose digestion in the wood-eating higher termite, *Nasutitermes takasagoensis*

(Shiraki): Distribution of cellulases and properties of endo- β -1,4-glucanase.
Zoolog Sci **14**:83-93.

89. **Trager, W.** 1934. The cultivation of a cellulose-digesting flagellate, *Trichomonas termopsidis*, and of certain other termite protozoa. Biol Bull **66**:182-190.
90. **Tschech, A., and N. Pfennig.** 1984. Growth yield increase linked to caffeate reduction in *Acetobacterium woodii*. Arch Microbiol **137**:163-167.
91. **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**:560-565.
92. **Whitman, W., T. Bowen, and D. Boone.** 2006. The methanogenic bacteria, p. 165-207. In M. Dworkin, S. Falkow, E. Rosenber, K. H. Schleifer, and E. Stackebrandt (ed.), The Prokaryotes, Third ed. Springer.
93. **Yamin, M. A.** 1978. Axenic cultivation of the cellulolytic flagellate *Trichomitopsis termopsidis* (Cleveland) from the termite *Zootermopsis*. J Protozool **25**:535-538.
94. **Yamin, M. A.** 1981. Cellulose metabolism by the flagellate *Trichonympha* from a termite is independent of endosymbiotic bacteria. Science **211**:58-59.

95. **Yamin, M. A.** 1980. Cellulose metabolism by the termite flagellate *Trichomitopsis termopsidis*. *Appl Environ Microbiol* **39**:859-863.
96. **Zhang, Y., D. Zhang, W. Li, J. Chen, Y. Peng, and W. Cao.** 2003. A novel real-time quantitative PCR method using attached universal template probe. *Nucleic Acids Res* **31**:e123.
97. **Zimmerman, P. R., J. P. Greenberg, S. O. Wandiga, and P. J. Crutzen.** 1982. Termites: a potentially large source of atmospheric methane, carbon dioxide, and molecular hydrogen. *Science* **218**:563-565.

Molecular Community Analysis of Acetogenic Bacteria in Roaches and Lower Termites: Evolution of the Symbiosis between Termites and Acetogenic *Spirochetes*

Abstract

The termite gut is host to a highly active population of acetogenic bacteria, which can fuel up to 1/3 of the energy metabolism of the host insect. In order to shed light on the roots of this symbiosis, we carried out molecular community analysis of acetogens present in the guts of the wood-feeding roach *Cryptocercus punctulatus* and lower termites of the genus *Incisitermes*. Acetogenesis in the termite gut is carried out primarily by spirochetes from the genus *Treponema*. Termite *Treponemes* appear to have acquired the ability to carry out acetogenesis by lateral gene-transfer from acetogenic *Firmicutes*, and this capacity has not yet been identified in free-living spirochetes. Phylogenetic analysis of the gene for formyl-tetrahydrofolate synthetase (FTHFS) suggests that spirochetes acquired this component of the acetogenic pathway prior to the roach-termite divergence, and that at least three species of FTHFS-bearing spirochetes were present in the last common ancestor of *Cryptocercus* and *Isoptera*.

Introduction

Molecular phylogeny of *Isoptera* and related insect species confirms that eusocial termites diverged from wood-feeding roaches whose modern representatives exist as the roach family *Cryptocercidae* (9, 16). Wood-feeding *Cryptocercus* are known to possess

complex microbial communities that share many characteristics with the microbiota of termites, including the presence of cellulolytic flagellates (5) and high rates of reductive acetogenesis (3). However, a long-standing question remains as to whether this similarity is due to vertical transmission from a wood-feeding common ancestor or whether the roach hindgut community was acquired from true termites via a lateral community-transfer event (23, 36, 37). Phylogenetic examination of ribosomal RNA genes from cellulolytic flagellates and gut bacteria have yet to yield clear evidence of either vertical or horizontal transmission of symbionts (i.e., branch patterns that robustly demonstrate congruent (or incongruent) host-symbiont evolutionary histories) (4, 8, 25, 33).

CO₂-reductive acetogenesis is a major electron sink in the guts of wood-feeding termites, accounting for 18%–26% of respiratory electron flow (2, 3, 29) and generating 10%–30% of the acetate produced in this environment (3, 34). Microbially generated acetate is the principal source of carbon for oxidation and biosynthesis for the termite host (24). The diversity of acetogens present in an environmental sample can be investigated using the gene for formyl-tetrahydrofolate synthetase (FTHFS), a key enzyme in the Wood-Ljungdahl pathway of reductive acetogenesis (14), as a marker of acetogenic capability (12, 17).

FTHFS diversity has been examined in three lower termites: *Zootermopsis nevadensis*, *Cryptotermes secundus*, and *Reticulitermes santonensis* (28, 31). The hindgut communities of all three termites were dominated by FTHFS sequences that cluster phylogenetically with the FTHFS genes of acetogenic spirochetes isolated from the

termite *Zootermopsis angusticolis* (11, 31). Fragmentary FTHFS gene sequences present in the metagenome of *Nasutitermes* termites also fell within this “termite *Treponeme* cluster” (38). Termite *Treponeme* FTHFS sequences cluster more broadly with FTHFS sequences from acetogenic *Firmicutes*, rather than an FTHFS from *Treponema denticola*, suggesting that this gene may have been acquired via lateral gene-transfer (31).

To date, the guts of termites are the only environment in which acetogenic spirochetes have been identified. This suggests that the acquisition of acetogenic capability by members of the genus *Treponema* occurred within the context of this symbiosis. Additionally, the presence of acetogenic spirochetes in members of four major lineages of termites on four continents suggests that this lateral gene-transfer event took place early in the evolutionary history of this insect. FTHFS and 16S rRNA phylogenies both suggest some degree of coevolution between lower termites and their symbiotic spirochetes (1, 28). However, within the termite gut *Treponeme* clade, each lower termite carries multiple, polyphyletic FTHFS sequence groups. This diversity might have been generated by either repeated horizontal symbiont-transfer or by vertical transmission from a common ancestor with multiple FTHFS-bearing spirochetes.

In this study, we examine the acetogenic bacteria present in roaches of the family *Cryptocercidae*. By examining the diversity and phylogeny of FTHFS genes present in wood-feeding roaches and lower termites, we hope to trace the evolutionary history of this symbiosis, shedding light on the mechanisms that have generated and maintained this remarkable association.

Materials and Methods

Insect Collection

Incisitermes minor termites were collected from a wood pile in Pasadena. *Cryptocercus punctulatus* were collected by Christina Nalepa (NC State University). The adult sample was from a roach collected at Mt. Collins, the nymphs from the South Mountains. The insects were shipped priority mail; upon receipt, they were maintained in glass jars in the dark.

DNA Extraction

DNA was extracted from whole dissected guts. The *C. punctulatus* adult sample contained a single gut, the *C. punctulatus* nymph sample contained three pooled guts, and the *Incisitermes* sample contained the guts of 7 workers. *C. punctulatus* guts were prepared within a week of receipt, *Incisitermes* within 24 hours of collection. DNA was purified as described by Matson, Ottesen and Leadbetter (20). The purified DNA was quantified using the Hoefer DyNAQuant 200 fluorometer and DNA quantification system (amersham pharmacia biotech) using reagents and procedures directed in the manual (DQ200-IM, Rev C1, 5-98).

FTHFS Amplification, Cloning, and RFLP Analysis

FTHFS genes were amplified from insect guts as described in Leaphart and Lovell (12). Primers with 5' phosphate groups were purchased from Integrated DNA Technologies. Amplification reactions for cloning contained 1 μ M each primer, 1X Failsafe Premix D

(Epicentre), 0.0525 U/ μ L Expand High Fidelity Taq polymerase (Roche) and 1 ng/ μ L template, using the recommended step-down protocol (12) and 25 cycles at 55 °C. PCR reactions were purified using QIAquick PCR purification kits (Qiagen), and cloned using a GC Cloning and amplification kit with LC-Kan vector (Lucigen).

Cloned PCR products were screened by RFLP analysis. Isolated colonies were picked into 10 μ L 1X TE, then incubated at 95 °C for 5 min. This lysate was used to provide template for amplification reactions generating both RFLP analyses and sequencing template. Inserts were amplified using vector primers SL1 and SR2 (from GC Cloning Kit manual), FailSafe Premix A (Epicentre), and 0.05 U/ μ L Taq polymerase (New England Biolabs). The thermocycling protocol was as follows: 3 min at 95 °C, 30 cycles of (95 °C 30 s, 55 °C 30 s, 72 °C 1 min 30 s), then 10 min at 72 °C. PCR product was analyzed by gel electrophoresis to verify the presence of full-length insert. RFLP typing used the enzyme HinP1I (New England Biolabs): 6 μ L of the PCR product was added to 0.4 μ L 10X NEB Buffer 2, 0.3 μ L HinP1I (New England Biolabs), and 3.3 μ L H₂O, then digested at 37 °C for 4 hr. Digested product was analyzed by gel electrophoresis using a 2.5% agarose gel.

A single representative of each RFLP type was selected for sequencing. Samples were amplified using vector primers SL1 and SR2 as described above, but with the substitution of EXPAND high fidelity polymerase. Cycle sequencing reactions were carried out by Laragen (Los Angeles, CA).

COII Identification of Termites and Cockroaches

Roach and termite identifications were confirmed using insect mitochondrial cytochrome oxidase subunit II (COII) gene sequences (Figure 2.1). COII genes were amplified directly from the DNA samples used for FTHFS analysis. *Cryptocercus punctulatus* COII was amplified using primers and cycling conditions described in Park et al. (27). *Incisitermes* COII was amplified using the primers CI-J-1773 and B-tLys and cycling conditions described in Miura et al. (21). FailSafe Premix D (Epicentre) and Expand high fidelity Taq (Roche) were substituted for the polymerase and buffers described.

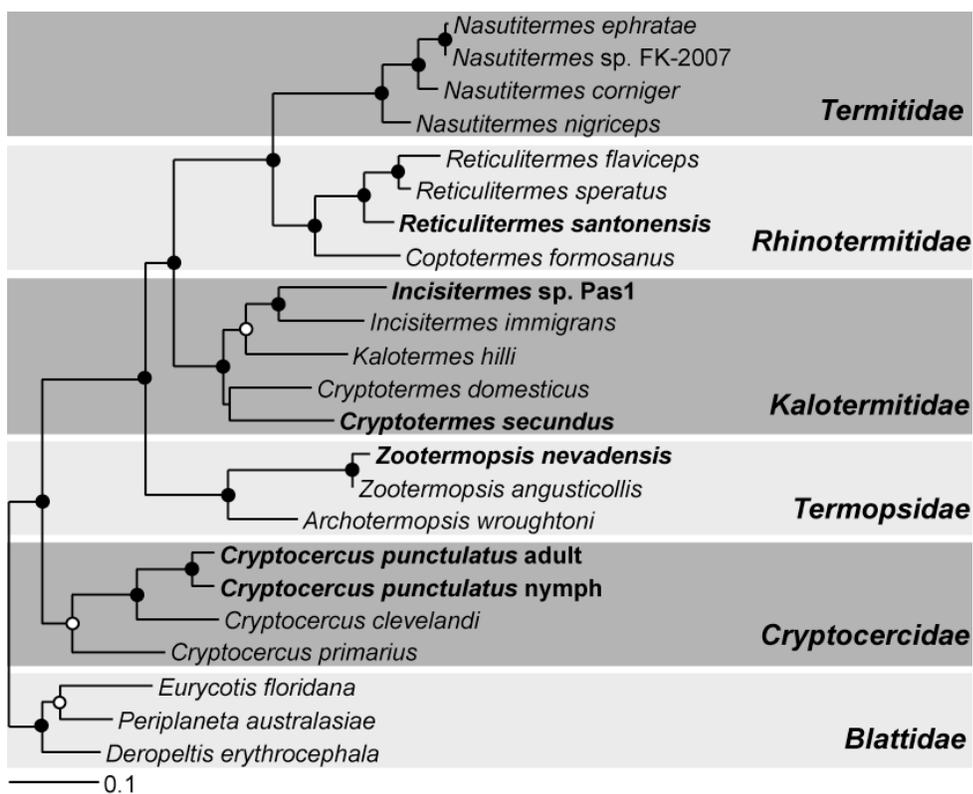


Figure 2.1. Mitochondrial cytochrome oxidase II phylogeny of termites and roaches. Species from which gut FTHFS diversity has been examined marked in bold. Tree calculated using AxML and 396 unambiguously aligned DNA bases. Open circles mark groupings also supported by either Phylip DNAPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 base pair changes per alignment position. Alternate *R. santonensis* and *C. secundus* COII sequences used to represent those from Pester and Brune (28), which were truncated.

Sequence Analysis

Sequence reads were assembled and edited using the Lasergene software package (version 7.2.1, DNASTAR). FTHFS protein sequences were aligned using MUSCLE (6), and phylogenetic analyses were carried out using the ARB software package (18). A single chimeric sequence was identified in the *Incisitermes* library using the Bellerophon program (7), and eliminated from further analysis. Sequence similarities given in the text of this paper represent amino acid similarities calculated using an ARB neighbor joining matrix and 352 alignment positions.

Results

DNA was extracted from the pooled guts of 7 *Incisitermes* workers and from 2 *C. punctulatus* whole gut samples, one containing a single adult gut and the other 3 pooled nymph guts. FTHFS sequences from each sample were amplified and cloned to generate libraries that represent a snapshot of the acetogenic diversity in these environments. The libraries were sorted by restriction fragment length polymorphism (RFLP), and a single representative of each RFLP type sequenced, generating 26, 29, and 16 nonchimeric phylotypes corresponding to 60, 88, and 90 total clones, respectively. These phylotypes were characterized by phylogenetic analysis (Table 2.1), and were further binned into operational taxonomic units with a cutoff of 98% amino acid sequence similarity (Table 2.2).

Table 2.1. Composition of FTHFS libraries constructed from *C. punctulatus* and *Incisitermes* sp. Pas1^a

Species	<i>Treponemes</i>			Acetogenic Firmicutes	Clone E / <i>Streptococcus</i>	Non-acetogenic
	I	II	III			
<i>C. punctulatus</i> adult	49	3	27	6	2	6
<i>C. punctulatus</i> nymph	22	6	8	6	36	2
<i>Incisitermes</i> sp. Pas1	37	63	-	-	-	-

^a Abundance given as percentage of full-length clones

All of the FTHFS sequences from *Incisitermes* termites, 79% of those from the adult roach, and 36% of those amplified from the roach nymph fell within the termite *Treponeme* cluster (Figure 2.2). Many of the FTHFS sequences recovered from *Incisitermes* grouped closely with sequences identified in *C. secundus*, which also falls within the *Kalotermitidae* family of lower termites. The *C. punctulatus* individuals contained three major groups of termite *Treponeme*-like FTHFS sequences. Roach group III formed a cluster that was identified as basal to the termite *Treponeme* radiation by two of the three treeing methods used (the Fitch distance algorithm clustered Roach group II and several associated lower termite sequences with Roach group III). Sequences of Roach group III lacked the hexapeptide insert characteristic of the termite *Treponeme* clade. Roach group II was the least abundant of the three groups, and affiliated with sequences basal to most of the termite *Treponeme* cluster. Roach group I lies near the middle of the termite *Treponeme* radiation, but is basal to a radiation that includes the cultured acetogen *Treponema primitia* strain ZAS-1 and sequences from representatives of all three families of lower termites examined.

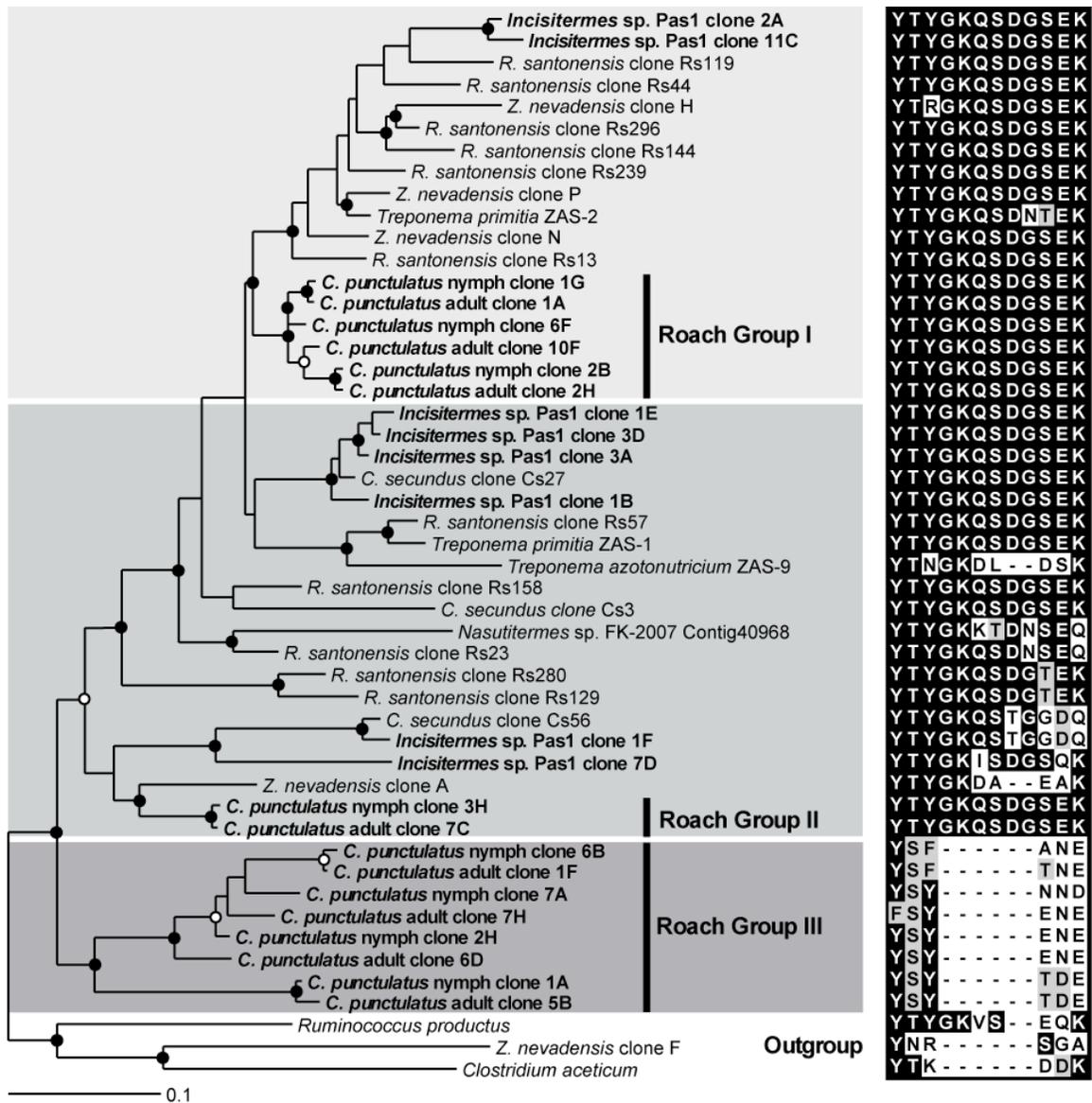


Figure 2.2. Phylogenetic analysis of termite *Treponeme* FTHFS sequences. **Left.** Tree constructed using Phylip PROML algorithm and 351 unambiguously aligned amino acid positions. Open circles mark groupings also supported by either Phylip PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position. **Right.** A highly variable region of the protein sequence, corresponding to residues 229–234 in *M. thermoacetica*. Each line of the alignment corresponds to a sequence in the tree at left.

Six percent of the FTHFS sequences in each *C. punctulatus* individual group phylogenetically with FTHFS sequences from acetogenic *Firmicutes* (Figure 2.3). Two

of these sequence types contained clones from both adult and nymph, and may represent small but stable populations within the roach gut. While the phylogenetic analysis presented in Figure 2.3 suggests that *Z. nevadensis* clone F and affiliated sequences may be basal to FTHFS sequences present in *C. punctulatus*, this relationship is not supported when the analysis is carried out with a larger sequence library.

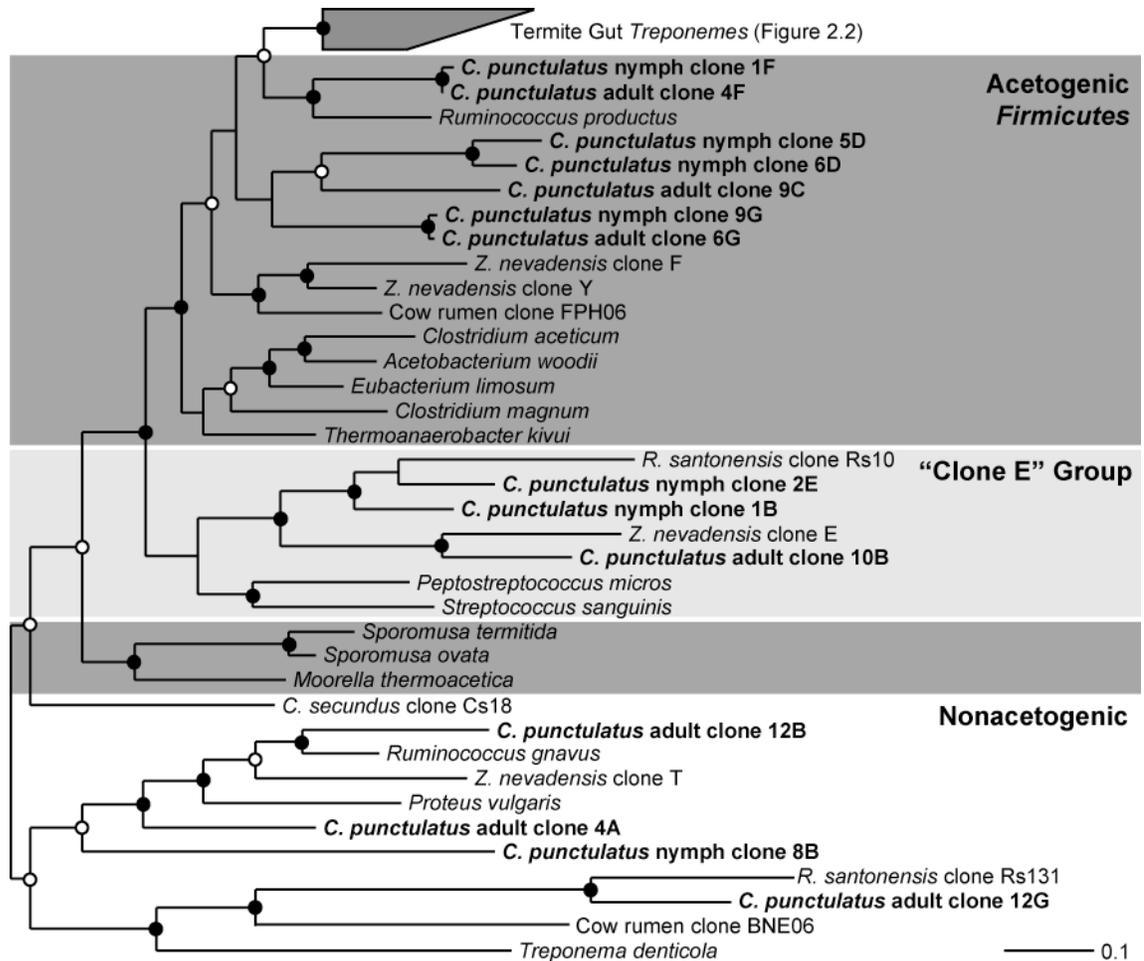


Figure 2.3. Phylogenetic analysis of FTHFS sequences from roaches and lower termites. Tree constructed using Phylip PROML algorithm and 317 unambiguously aligned amino acid positions. Open circles mark groupings also supported by either Phylip PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position. Fifteen sequences from *C. punctulatus* and *Z. nevadensis* were chosen to represent the termite *Treponeme* group in this analysis.

Thirty-six percent of sequences from the roach nymph, and two percent of those from the adult, clustered with *Z. nevadensis* clone E and *R. santonensis* clone Rs10. This group was clustered with sequences from *Peptostreptococcus micros* and *Streptococcus pyogenes*, and is most likely nonacetogenic. Remaining FTHFS types were most likely nonacetogenic but were not closely related to cultured organisms.

Discussion

The presence of acetogenic spirochetes in the hindgut appears to be a common characteristic of termites and wood-feeding roaches. In the three termite species previously examined, *Treponeme*-like FTHFS sequences were more abundant than those from acetogenic *Firmicutes* (28, 31). In this study, we confirmed that this pattern is repeated in the lower termite *Incisitermes* sp. Pas1, and extends to the gut of the wood-feeding roach *C. punctulatus*.

All wood-feeding insects examined hosted multiple sequence types within the termite *Treponeme* clade, with the least diverse, *C. secundus*, containing three distinct groups of FTHFS sequence, and the most diverse, *R. santonensis*, containing 11 groups. The most closely related termites, *Incisitermes* sp. and *C. secundus*, contained FTHFS types that were 97% similar, despite extensive geographic separation (*Incisitermes* were collected in Pasadena, CA, while *C. secundus* was collected in Darwin, Australia). This tendency of sequences from a single species of termite to group more closely with each other than with those from other termite species, and the tendency of closely related termites to host phylotypes more similar to each other than to distantly related termites, has been

observed in spirochete rRNA analyses (1, 13), and suggests a high degree of host-symbiont coevolution.

The *C. punctulatus* individuals examined in this study contained three novel lineages of termite *Treponeme*-like FTHFS sequences. One, referred to in this study as roach group III, is basal to the termite *Treponeme* clade, and lacked a hexapeptide insert characteristic of other termite gut *Treponeme* FTHFS types. The basal position of this clade and its absence from lower termites suggests that it represents an evolutionarily ancient lineage, present in the last common ancestor of termites and roaches, that was lost prior to the radiation of the lower termites examined. Because roach group III clustered consistently with termite gut *Treponemes* to the exclusion of all other FTHFS types, we propose that it represents a line of *Treponemes* that diverged following the acquisition of acetogenic capability by spirochetes but prior to the acquisition of the hexapeptide insert. Alternatively, this group may represent descendants of the *Firmicute* from which termite *Treponemes* acquired their FTHFS genes.

C. punctulatus hosts two additional groups of *Treponeme*-like FTHFS sequences (roach groups I and II). These were again distinct from termite-derived FTHFS types, with less than 93% amino acid similarity to the most closely related termite-derived sequences. Phylogenetic analysis shows that roach group I is basal to a termite *Treponeme* subclade that encompasses sequences present in all three lower termite families examined. The remaining FTHFS sequences (associated with roach group II) do not appear to be

monophyletic, and may be descendants of a more extensive radiation of acetogenic spirochetes present in the last common ancestor but lost from *C. punctulatus*.

In conclusion, we posit that the acquisition of acetogenic capability by gut spirochetes occurred prior to the divergence of *Cryptocercidae* and *Isoptera*. Furthermore, the three lineages of *Treponeme*-like FTHFS types identified in *C. punctulatus* are proposed to represent an ancestral radiation of acetogenic spirochetes, whose further divergence gave rise to the rich diversity of FTHFS types observed in wood-feeding lower termites.

Chapter Two Appendix

- 1. Table 2.2. Operational taxonomic unit grouping of FTHFS sequences identified in this study**
- 2. Table 2.3. Sequences used in FTHFS phylogenetic analysis**
- 3. Table 2.4. Sequences used in COII phylogenetic analysis**

Table 2.2. Operational Taxonomic Unit Grouping of FTHFS sequences identified in this study

Group	Phylotype	Abundance (%) ^a	Genotypes ^b
<i>C. punctulatus</i> adult			
<i>Treponeme</i> Group I	1A	40	1A , 1B, 1C, 3B, 3D, 4B, 5D, 10E
	2H	7	2H
	10F	2	10F
<i>Treponeme</i> Group II	7C	3	7C
<i>Treponeme</i> Group III	1F	20	1F , 1G, 6E
	7H	3	7H
	5B	2	5B
Acetogenic <i>Firmicutes</i>	6D	2	6D
	4F	2	4F
	6G	2	6G
Clone E Group	9C	2	9C
	10B	2	10B
	Nonacetogenic	4A	2
12B		2	12B
12G		2	12G
<i>C. punctulatus</i> nymph			
<i>Treponeme</i> Group I	1G	16	1G , 1E, 11G
	2B	15	2B , 1C, 3G, 10H
	6F	2	6F , 6E
<i>Treponeme</i> Group II	3H	6	3H , 7F
<i>Treponeme</i> Group III	2H	3	2H , 1D, 9D
	1A	2	1A , 12G
	9C	2	7A , 9C
Acetogenic <i>Firmicutes</i>	6B	1	6B
	1F	3	1F , 6A
	5D	1	5D
Clone E Group	6D	1	6D
	9G	1	9G
	1B	35	1B , 3A, 3C
Nonacetogenic	2E	1	2E
	8B	2	8B
<i>Incisitermes</i> sp. Pas1			
<i>Treponeme</i> Group I	2A	34	2A , 3C, 3F, 3G, 8B
	11C	2	11C , 11G
<i>Treponeme</i> Group II	1B	40	1B , 11F
	3D	12	3D , 4B, 11B
	1F	4	1F
	1E	2	1E
	3A	2	3A
	7D	2	7D

^aDefined as percent of full-length clones^bSequenced RFLP type clones. Group representative marked in bold.

Table 2.3. Sequences used in FTHFS phylogenetic analysis

Source / Sequence Type	Designation	Accession	Reference
<i>T. primitia</i> ZAS-1	ZAS-1a	AY162313	(31)
<i>T. primitia</i> ZAS-2	ZAS-2	AY162315	(31)
<i>T. azotonutricium</i> ZAS-9	ZAS-9	AY162316	(31)
<i>Z. angusticollis</i> Gut Clone	A	AY162294	(31)
<i>Z. angusticollis</i> Gut Clone	E	AY162296	(31)
<i>Z. angusticollis</i> Gut Clone	F	AY162298	(31)
<i>Z. angusticollis</i> Gut Clone	H	AY162302	(31)
<i>Z. angusticollis</i> Gut Clone	N	AY162306	(31)
<i>Z. angusticollis</i> Gut Clone	P	AY162307	(31)
<i>Z. angusticollis</i> Gut Clone	T	AY162309	(31)
<i>Z. angusticollis</i> Gut Clone	Y	AY162311	(31)
<i>C. secundus</i> Gut Clone	Cs3	DQ278251	(28)
<i>C. secundus</i> Gut Clone	Cs18	DQ278253	(28)
<i>C. secundus</i> Gut Clone	Cs27	DQ278254	(28)
<i>C. secundus</i> Gut Clone	Cs56	DQ278258	(28)
<i>R. santonensis</i> Gut Clone	Rs10	DQ278259	(28)
<i>R. santonensis</i> Gut Clone	Rs13	DQ278232	(28)
<i>R. santonensis</i> Gut Clone	Rs23	DQ278210	(28)
<i>R. santonensis</i> Gut Clone	Rs44	DQ278211	(28)
<i>R. santonensis</i> Gut Clone	Rs57	DQ278215	(28)
<i>R. santonensis</i> Gut Clone	Rs119	DQ278226	(28)
<i>R. santonensis</i> Gut Clone	Rs129	DQ278222	(28)
<i>R. santonensis</i> Gut Clone	Rs131	DQ278221	(28)
<i>R. santonensis</i> Gut Clone	Rs144	DQ278223	(28)
<i>R. santonensis</i> Gut Clone	Rs158	DQ278226	(28)
<i>R. santonensis</i> Gut Clone	Rs239	DQ278201	(28)
<i>R. santonensis</i> Gut Clone	Rs280	DQ278207	(28)
<i>R. santonensis</i> Gut Clone	Rs296	DQ278208	(28)
<i>Nasutitermes</i> sp. FK-2007	Contig40968	JGI GOI: 2004144560	(38)
Cow Rumen Clone	BNE06	AB085284	Database only
Cow Rumen Clone	FPH06	AB085574	Database only
<i>Acetobacterium woodii</i>		AF295701	(12)
<i>Clostridium aceticum</i>		AF295705	(12)
<i>Clostridium magnum</i>		AF295703	(12)
<i>Eubacterium limosum</i>		AF295706	(12)
<i>Moorella thermoacetica</i>		NC_007644	(30)
<i>Peptostreptococcus micros</i>		NZ_ABEE02000017	Database only
<i>Proteus vulgaris</i>		AF295710	(12)
<i>Ruminococcus gnavus</i>		NZ_AAYG02000005	Database only
<i>Ruminococcus productus</i>		AF295707	(12)
<i>Sporomusa ovata</i>		AF295708	(12)
<i>Sporomusa termitida</i>		AF295709	(12)
<i>Streptococcus sanguinis</i>		NC_009009	(39)
<i>Thermoanaerobacter kivui</i>		AF295704	(12)
<i>Treponema denticola</i>		NC_002967	(32)

Table 2.4. Sequences used in COII phylogenetic analysis

Source	Accession	Reference
<i>Archotermopsis wroughtoni</i>	DQ442080	(10)
<i>Deropeltis erythrocephala</i>	DQ874271	(9)
<i>Coptotermes formosanus</i>	AB109529	(26)
<i>Cryptocercus clevelandi</i>	DQ007626	(15)
<i>Cryptocercus primarius</i>	DQ007644	(15)
<i>Cryptotermes domesticus</i>	AF189086	(35)
<i>Cryptotermes secundus</i>	AF189093	(35)
<i>Incisitermes immigrans</i>	AB109542	(26)
<i>Kaloterme hilli</i>	AF189101	(35)
<i>Nasutitermes corniger</i>	AB037327	(22)
<i>Nasutitermes ephratae</i>	AB037328	(22)
<i>Nasutitermes nigriceps</i>	AB037329	(22)
<i>Nasutitermes sp. FK-2007</i>	EU236539	(38)
<i>Eurycotis floridana</i>	DQ874283	(9)
<i>Periplaneta australasiae</i>	DQ874310	(9)
<i>Reticulitermes flaviceps</i>	AB109532	(26)
<i>Reticulitermes santonensis</i>	AF291743	(19)
<i>Reticulitermes speratus</i>	AB109530	(26)
<i>Zootermopsis angusticollis</i>	DQ442267	(10)

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. **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.
3. **Breznak, J. A., and J. M. Switzer.** 1986. Acetate synthesis from H₂ plus CO₂ by termite gut microbes. *Appl Environ Microbiol* **52**:623-630.
4. **Carpenter, K. J., and P. J. Keeling.** 2007. Morphology and phylogenetic position of *Eucomonympha imla* (Parabasalia: Hypermastigida). *J Eukaryot Microbiol* **54**:325-332.
5. **Cleveland, L. R.** 1934. The wood-feeding roach *Cryptocercus*, its protozoa, and the symbiosis between protozoa and roach. *Mem Amer Acad Arts Sci* **17**:185-342.
6. **Edgar, R. C.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**:113.
7. **Huber, T., G. Faulkner, and P. Hugenholtz.** 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**:2317-2319.
8. **Ikeda-Ohtsubo, W., M. Desai, U. Stingl, and A. Brune.** 2007. Phylogenetic diversity of "Endomicrobia" and their specific affiliation with termite gut flagellates. *Microbiol* **153**:3458-3465.

9. **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.
10. **Inward, D. J., A. P. Vogler, and P. Eggleton.** 2007. A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspects of their evolutionary biology. *Mol Phylogenet Evol* **44**:953-967.
11. **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.
12. **Leaphart, A. B., and C. R. Lovell.** 2001. Recovery and analysis of formyltetrahydrofolate synthetase gene sequences from natural populations of acetogenic bacteria. *Appl Environ Microbiol* **67**:1392-1395.
13. **Lilburn, T. G., T. M. Schmidt, and J. A. Breznak.** 1999. Phylogenetic diversity of termite gut spirochaetes. *Environ Microbiol* **1**:331-345.
14. **Ljungdahl, L. G.** 1986. The autotrophic pathway of acetate synthesis in acetogenic bacteria. *Ann Rev Microbiol* **40**:415-450.
15. **Lo, N., P. Luykx, R. Santoni, T. Beninati, C. Bandi, M. Casiraghi, W. H. Lu, E. V. Zakharov, and C. A. Nalepa.** 2006. Molecular phylogeny of *Cryptocercus* wood-roaches based on mitochondrial COII and 16S sequences, and chromosome numbers in Palearctic representatives. *Zoolog Sci* **23**:393-398.
16. **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* **10**:801-804.

17. **Lovell, C. R., and A. B. Leaphart.** 2005. Community-level analysis: key genes of CO₂-reductive acetogenesis. *Methods Enzymol* **397**:454-469.
18. **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. W. 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. H. Schleifer.** 2004. ARB: a software environment for sequence data. *Nucl Acids Res* **32**:1363-1371.
19. **Marini, M., and B. Mantovani.** 2002. Molecular relationships among European samples of *Reticulitermes* (Isoptera, Rhinotermitidae). *Mol Phylogenet Evol* **22**:454-459.
20. **Matson, E., E. A. Ottesen, and J. R. Leadbetter.** 2007. Extracting DNA from the gut microbes of the termite (*Zootermopsis nevadensis*). *J Visualized Exper* **4**.
21. **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 of Amer* **91**:515-523.
22. **Miura, T., Y. Roisin, and T. Matsumoto.** 2000. Molecular phylogeny and biogeography of the nasute termite genus *Nasutitermes* (Isoptera: Termitidae) in the pacific tropics. *Mol Phylogenet Evol* **17**:1-10.
23. **Nalepa, C. A.** 1991. Ancestral transfer of symbionts between cockroaches and termites: an unlikely scenario. *Proc Biol Sci* **246**:185-189.

24. **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.
25. **Ohkuma, M., K. Ohtoko, T. Iida, M. Tokura, S. Moriya, R. Usami, K. Horikoshi, and T. Kudo.** 2000. Phylogenetic identification of hypermastigotes, *Pseudotrichonympha*, *Spirotrichonympha*, *Holomastigotoides*, and parabasalium symbionts in the hindgut of termites. *J Eukaryot Microbiol* **47**:249-259.
26. **Ohkuma, M., H. Yuzawa, W. Amornsak, Y. Sornnuwat, Y. Takematsu, A. Yamada, C. Vongkaluang, O. Sarnthoy, N. Kirtibutr, N. Noparatnaraporn, T. Kudo, and T. Inoue.** 2004. Molecular phylogeny of Asian termites (Isoptera) of the families Termitidae and Rhinotermitidae based on mitochondrial COII sequences. *Mol Phylogenet Evol* **31**:701-710.
27. **Park, Y. C., K. Maekawa, T. Matsumoto, R. Santoni, and J. C. Choe.** 2004. Molecular phylogeny and biogeography of the Korean woodroaches *Cryptocercus* spp. *Mol Phylogenet Evol* **30**:450-464.
28. **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.
29. **Pester, M., and A. Brune.** 2007. Hydrogen is the central free intermediate during lignocellulose degradation by termite gut symbionts. *Isme J* **1**:551-565.
30. **Pierce, E., G. Xie, R. D. Barabote, E. Saunders, C. S. Han, J. C. Detter, P. Richardson, T. S. Brettin, A. Das, L. G. Ljungdahl, and S. W. Ragsdale.**

2008. The complete genome sequence of *Moorella thermoacetica* (f. *Clostridium thermoaceticum*). *Environ Microbiol* **10**:2550-2573.
31. **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*. *Microbiol* **149**:2529-2537.
32. **Seshadri, R., G. S. Myers, H. Tettelin, J. A. Eisen, J. F. Heidelberg, R. J. Dodson, T. M. Davidsen, R. T. DeBoy, D. E. Fouts, D. H. Haft, J. Selengut, Q. Ren, L. M. Brinkac, R. Madupu, J. Kolonay, S. A. Durkin, S. C. Daugherty, J. Shetty, A. Shvartsbeyn, E. Gebregeorgis, K. Geer, G. Tsegaye, J. Malek, B. Ayodeji, S. Shatsman, M. P. McLeod, D. Smajs, J. K. Howell, S. Pal, A. Amin, P. Vashisth, T. Z. McNeill, Q. Xiang, E. Sodergren, E. Baca, G. M. Weinstock, S. J. Norris, C. M. Fraser, and I. T. Paulsen.** 2004. Comparison of the genome of the oral pathogen *Treponema denticola* with other spirochete genomes. *Proc Natl Acad Sci USA* **101**:5646-5651.
33. **Stingl, U., R. Radek, H. Yang, and A. Brune.** 2005. "Endomicrobia": cytoplasmic symbionts of termite gut protozoa form a separate phylum of prokaryotes. *Appl Environ Microbiol* **71**:1473-1479.
34. **Tholen, A., and A. Brune.** 2000. Impact of oxygen on metabolic fluxes and *in situ* rates of reductive acetogenesis in the hindgut of the wood-feeding termite *Reticulitermes flavipes*. *Environ Microbiol* **2**:436-449.
35. **Thompson, G. J., L. R. Miller, M. Lenz, and R. H. Crozier.** 2000. Phylogenetic analysis and trait evolution in Australian lineages of drywood termites (Isoptera, Kalotermitidae). *Mol Phylogenet Evol* **17**:419-429.

36. **Thorne, B. L.** 1990. A case for ancestral transfer of symbionts between cockroaches and termites. *Proc Biol Sci* **241**:37-41.
37. **Thorne, B. L.** 1991. Ancestral transfer of symbionts between cockroaches and termites: an alternative hypothesis. *Proc Biol Sci* **246**:191-195.
38. **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**:560-565.
39. **Xu, P., J. M. Alves, T. Kitten, A. Brown, Z. Chen, L. S. Ozaki, P. Manque, X. Ge, M. G. Serrano, D. Puiu, S. Hendricks, Y. Wang, M. D. Chaplin, D. Akan, S. Paik, D. L. Peterson, F. L. Macrina, and G. A. Buck.** 2007. Genome of the opportunistic pathogen *Streptococcus sanguinis*. *J Bacteriol* **189**:3166-3175.

Molecular Community Analysis of Acetogenic Bacteria in the Guts of Higher Termites: Community Structure in Termites with Diverse Feeding Strategies

Abstract

CO₂-reductive acetogenesis is a key bacterial activity in the termite hindgut, capable of fueling up to 30% of the metabolism of wood- and grass-feeding termites. In wood-feeding lower termites, acetogenesis is known to be carried out by acetogenic spirochetes. However, the acetogens of higher termites have not been extensively characterized. In this study, we examine the acetogenic bacteria hosted by 6 higher termites species through preparation and phylogenetic analysis of functional gene inventories for formyl-tetrahydrofolate synthetase (FTHFS), a key enzyme in the acetyl-CoA pathway. In wood-, palm-, and litter-feeding higher termites, the dominant acetogens appear to be termite gut *Treponemes* similar to those found in wood-feeding lower termites. However, in subterranean termites, whose diet likely includes some degree of soil-feeding, the dominant acetogens were represented by a novel clade of *Firmicute*-like FTHFS sequences. *Firmicute* acetogens are widespread in the environment, whereas acetogenic *Treponemes*, to date, have only been identified in the guts of termites and wood-feeding roaches. The relative dominance of acetogenic *Firmicutes* in the guts of termites utilizing alternate substrates suggests that the fermentation of wood polysaccharides (and similar substrates) in the termite hindgut establishes a uniquely favorable environment of acetogenesis by spirochetes.

Introduction

The symbiosis between termites and their gut microbes is a highly complex, obligate mutualism. The hindgut community acts as a highly efficient bio-reactor, converting complex substrates to acetate, the principle source of energy for the termite (27). In wood-feeding termites, H_2 is the central free intermediate in the degradation of lignocellulose, representing 22%–26% of the respiratory activity of the termites (31). Microtracer experiments suggest that rates of CO_2 -reductive acetogenesis represent 83–100% of hydrogen turnover in these experiments, corresponding to 18%–26% of the termite's respiratory activity (31).

Isoptera is divided into 7 major families. 6 of these families are comprised of “lower termites,” which are exclusively wood and/or grass feeders. The “higher termites” are a single family (*Termitidae*), which nonetheless contains about 85% of known genera (13). Higher termites are able to utilize a much broader range of substrates than lower termites; in addition to wood- and grass-feeding, higher termite species have evolved fungus-cultivating, litter- and soil-feeding lifestyles. In lower termites, microbial fermentation of cellulose takes place in a single hindgut paunch. Many higher termites have a more complex gut structure; the five gut compartments of soil-feeding *Cubitermes* sp. have been shown to have distinct physical conditions (pH, metabolite concentrations) (6) and bacterial communities (39). Higher termites with different feeding habits have been found to have vastly different complements of symbiotic bacteria (24, 39, 41, 46).

This diversification of feeding habits and prokaryotic community structure corresponds with altered patterns of acetogenesis and methanogenesis. While wood- and grass-feeding higher termites were found to have high rates of acetogenesis and low rates of methanogenesis, this relationship was reversed in soil-feeding and fungus-cultivating termites (4). Additionally, domain-level phylogenetic profiling found that soil-feeding higher termites have a lower ratio of *Bacteria* to *Archaea* than wood-feeding termites, suggesting a larger methanogenic population (3). However, it has been demonstrated that soil-feeding termites with low rates of *in situ* CO₂ reduction to acetate nonetheless have substantial populations of acetogenic bacteria (>10⁶ cells/mL) (42). As a result, it has been hypothesized that acetogens in this environment subsist on alternative substrates or within microniches.

While most termite gut acetogens remain uncultured, the diversity of organisms capable of carrying out this activity can be assessed using molecular ecology-based techniques. Leaphart and Lovell (14, 15) have designed primers that target the gene for formyl-tetrahydrofolate synthetase (FTHFS), a key enzyme in the Wood-Ljungdahl pathway of reductive acetogenesis (16). In lower termites, the dominant FTHFS types group phylogenetically with FTHFS genes from acetogenic spirochetes of the genus *Treponema* (30, 37).

The recent metagenome of microbes inhabiting the gut of the wood-feeding higher termite *Nasutitermes* revealed the presence of termite *Treponema*-like FTHFS genes (46). However, the fragmentary nature of that data precludes detailed phylogenetic analysis of

these sequences. An exhaustive survey of acetogenesis genes in other species of higher termites has not yet been presented. Here, we will explore the diversity of acetogenic organisms present in 6 species of higher termites with diverse feeding regimes.

Materials and Methods

Insect Collection

Nasutitermes sp. Cost003 and *Rhynchotermes* sp. Cost004 were collected in the INBio forest preserve in Guápiles, Costa Rica. Cost003 was collected at a height of 1.2 m in a *Psidium guajaba* tree and appeared to be feeding on deadwood. Cost004 was collected in the same area, from a nest located under an unidentified *Bromeliad*. Extensive feeding trails led from this nest to a large pile of decaying wood and plant material, suggesting a litter-feeding lifestyle. *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 dead portions of the same plant. *Amitermes* sp. Cost010 was collected from the roots of dead sugar cane plants at a Costa Rican plantation. Costa Rican termite derived materials were collected, processed, exported, and imported under existing permits between INBio (Costa Rica) and Diversa Corporation (Verenium). Work with these samples at Caltech was subject to guidelines established within a material transfer agreement between the three parties. *Amitermes* sp. JT2 and *Gnathamitermes* sp. JT5 were collected from subterranean nests at Joshua Tree National Park.

DNA Extraction

Guts were extracted from termites within 48 hours of collection. Whole guts were collected from 20 workers of each species. Extracted whole guts were suspended in 500 μ L 1X TE (10mM Tris, 1mM EDTA, pH 7.4) and stored at -20 °C until DNA purification. DNA was purified from gut samples as described by Matson, Ottesen and Leadbetter (20). The purified DNA was quantified using the Hoefer DyNAQuant 200 fluorometer and DNA quantification system (Amersham Pharmacia Biotech) using reagents and procedures directed in the manual (DQ200-IM, Rev C1, 5-98).

FTHFS Amplification, Cloning, and RFLP Analysis

FTHFS genes were amplified from insect guts as described in Leaphart and Lovell (15). Primers with 5' phosphate groups were purchased from Integrated DNA Technologies. Amplification reactions for cloning contained 1 μ M each primer, 1X Failsafe Premix D (Epicentre), and 0.0525 U/ μ L Expand High Fidelity Taq polymerase (Roche). FTHFS was amplified from Cost003 in reactions containing 1 ng/ μ L template and following the recommended step-down protocol (15) followed by 25 cycles at 55 °C. All other termite samples contained low levels of PCR-inhibiting compounds and required further dilution; these reactions contained 0.1 ng/ μ L template and were amplified for an additional 5 cycles at 55 °C to generate a similar final concentration of product. PCR reactions were purified using QIAquick PCR purification kits (Qiagen), and cloned using a GC Cloning and amplification kit with LC-Kan vector (Lucigen).

Cloned PCR products were screened by RFLP analysis. Isolated colonies were picked into 10 μL 1X TE, then incubated at 95 $^{\circ}\text{C}$ for 5 min. This lysate was used to provide template for amplification reactions generating both RFLP analyses and sequencing template. Inserts were amplified using vector primers SL1 and SR2 (from GC Cloning Kit manual), FailSafe Premix A (Epicentre), and 0.05 U/ μL Taq polymerase (New England Biolabs). The thermocycling protocol was as follows: 3 min at 95 $^{\circ}\text{C}$, 30 cycles of (95 $^{\circ}\text{C}$ 30 s, 55 $^{\circ}\text{C}$ 30 s, 72 $^{\circ}\text{C}$ 1 min 30 s), then 10 min at 72 $^{\circ}\text{C}$. RFLP typing used the enzyme HinP1I (New England Biolabs): 6 μL of the PCR product was added to 0.4 μL 10X NEB buffer 2, 0.3 μL HinP1I (New England Biolabs), and 3.3 μL H₂O, then digested at 37 $^{\circ}\text{C}$ for 4 hr. Digested product was analyzed by gel electrophoresis using a 2.5% agarose gel. A single representative clone of each RFLP type was amplified for sequencing using the protocol above and substituting Expand high fidelity polymerase (Roche) for Taq DNA polymerase.

COII Identification of Termites

Termites identifications were confirmed using insect mitochondrial cytochrome oxidase subunit II (COII) gene sequences. COII genes were amplified directly from the DNA samples used for FTHFS analysis for the Costa Rican termites. JT2 and JT5 COII sequences were amplified from single termites. Single termites were placed in 2 mL microcentrifuge tubes with 50 μL , then crushed using a sterile glass rod. The supernatant from this disruption was transferred to a 200 μL PCR tube, then incubated at 95 $^{\circ}\text{C}$ for 10 min to lyse suspended cells and inactivate cellular protein. The resultant solution was clarified by centrifugation 1 min at 13,000 x g, and the resultant supernatant used directly

for COII amplification. Termite COII was amplified using the primers CI-J-1773 and B-tLys and cycling conditions described in Miura et al. (22). Reactions included FailSafe Premix D (Epicentre) and Expand high fidelity Taq (Roche).

Sequence Analysis

Cycle sequencing was carried out by Laragen (Los Angeles, CA). Sequence reads were assembled and edited using the Lasergene software package (version 7.2.1, DNASTAR). FTHFS protein sequences were aligned using MUSCLE (8), and phylogenetic analyses were carried out using the ARB software package (18). Libraries were screened for chimeric sequences using the Bellerophon program (11); single RFLP types from Cost003 and Cost004 were eliminated from further analysis.

Results

FTHFS libraries were constructed from 4 species of higher termite from Costa Rica and 2 desert-adapted species from California (Table 3.1). *Nasutitermes* sp. Cost003, collected in the mountains of central Costa Rica, was clearly wood-feeding. *Rhynchotermes* sp. Cost004 was collected in the same area, and appears to be a litter feeder. *Microcerotermes* sp. Cost008, collected on the eastern coast of Costa Rica, was found feeding on dead portions of a palm tree. *Amitermes* sp. Cost010 was collected from the roots of a decaying sugarcane plant in Costa Rica. *Amitermes* sp. JT2 and *Gnathamitermes* sp. JT5 were collected from subterranean nests in the Mohave Desert. It has not been determined whether these three termites were feeding on soil or on nearby plant material.

Table 3.1. FTHFS libraries constructed in this study

Species	Full-Length Clones ^a	RFLP Types ^a	OTU (98% AA)
<i>Nasutitermes</i> sp. Cost003	52	19	14
<i>Rhynchotermes</i> sp. Cost004	63	42	30
<i>Microcerotermes</i> sp. Cost008	27	16	12
<i>Amitermes</i> sp. Cost010	27	18	17
<i>Amitermes</i> sp. JT2	90	24	20
<i>Gnathamitermes</i> sp. JT5	60	24	22

^a Excludes RFLP types and clones determined to be chimeric

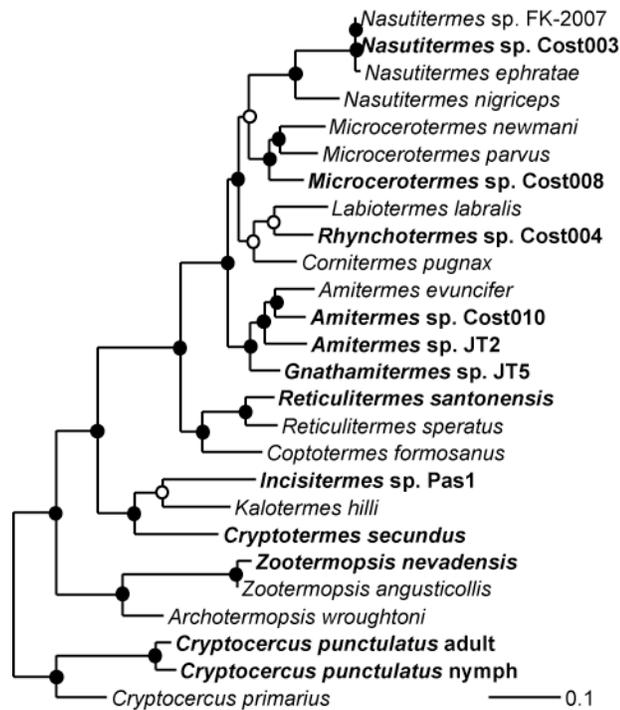


Figure 3.1. Mitochondrial cytochrome oxidase II phylogeny of termites and roaches. Species from which gut FTHFS diversity has been examined are marked in bold. Tree calculated using AxML and 396 unambiguously aligned DNA bases. Open circles mark groupings also supported by either Phylip DNAPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 base pair changes per alignment position.

Mitochondrial cytochrome oxidase II (COII) phylogeny was used to help identify collected termites (Figure 3.1). Cost003 was collected within 30 ft of *Nasutitermes* sp.

FK-2007 (the source of the 2007 metagenome by Warneke et al. (46)), and had an identical COII gene sequence. The identification of Cost008, Cost010, and JT2 termites could be confirmed to genus level with molecular phylogeny. No COII genes were available for *Rhynchotermes* or *Gnathamitermes*, so identification of Cost004 and JT5 relied on morphological characteristics. The COII gene from *Gnathamitermes* genus groups closely with sequences from *Amitermes* termites. The genus *Rhynchotermes* is typically classified as a member of the *Nasutiterminae* subfamily. However, this family is paraphyletic (2, 12), and Cost004 groups phylogenetically with termites from proposed subfamily *Syntermitinae* (9).

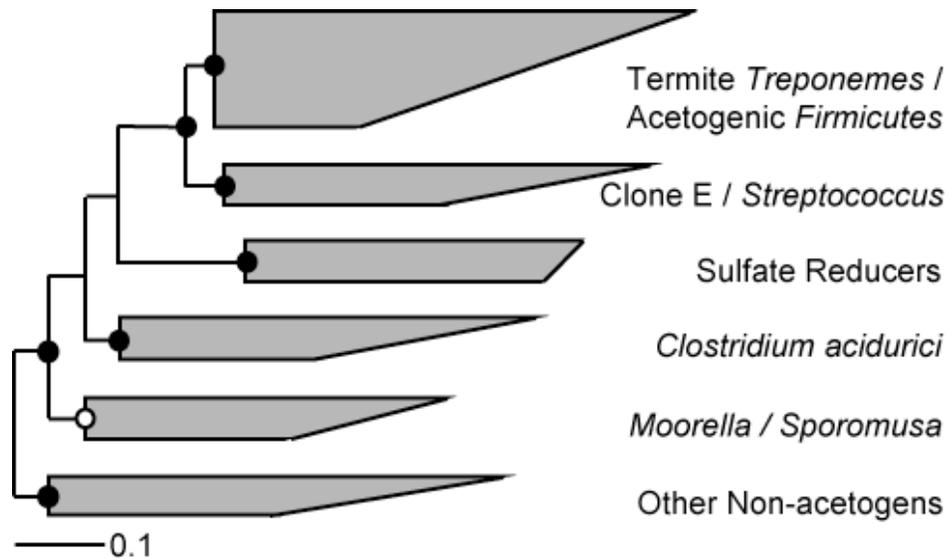


Figure 3.2. Phylogeny of major FTHFS clades found in termites and relatives. Tree calculated using PhyML PROML and 337 unambiguously aligned amino acids. Open circles mark groupings also supported by either PhyML PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position.

Table 3.2. Composition of FTHFS libraries from the hindgut microbiota of termites and relatives^a

Species	Food Source	Termite <i>Treponemes</i>	<i>Firmicute</i> Acetogens	<i>Moorella</i> / <i>Sporomusa</i>	Clone E / <i>Streptococcus</i>	<i>Clostridium acidurici</i>	Other Non-acetogenic
<i>C. punctulatus</i> adult	Wood	78	5		2		5
<i>C. punctulatus</i> nymph	Wood	50	7		41		2
<i>Z. nevadensis</i> ^c	Wood	77	10		4		9
<i>C. secundus</i> ^d	Wood	97					2
<i>Incisitermes</i> sp. Pas1	Wood	100					
<i>R. santonensis</i> ^d	Wood	98			1		1
<i>Nasutitermes</i> sp. Cost003	Wood ^b	98	2				
<i>Rhynchotermes</i> sp. Cost004	Litter ^b	37	6			45	10
<i>Microcerotermes</i> sp. Cost008	Palm ^b	89	11				
<i>Amitermes</i> sp. Cost010	Sugarcane / Soil ^b	12	85	4			
<i>Amitermes</i> sp. JT2	Grass / Soil ^b	1	87	6			3
<i>Gnathamitermes</i> sp. JT5	Grass / Soil ^b	2	28	2	37	10	17

^a Sequence abundance for each major FTHFS clade is given as percentage of total clones examined

^b Food source unknown, probable sources based on nest location and/or feeding trails.

^c From Salmassi and Leadbetter, (37)

^d From Pester and Brune, (30)

A diversity of FTHFS sequences were identified in higher termites. These FTHFS types were classed into 6 broad categories (Figure 3.2). Sequences from the termite *Treponeme* and acetogenic *Firmicute* groups were considered probable acetogens. The *Sporomusa/Moorella* group was considered indeterminate, and all other groups are considered probable nonacetogens. Phylogenetic analysis of FTHFS sequences from higher termites show striking variability in community composition (Table 3.2). Wood-feeding Cost003 and palm-feeding Cost008, similar to lower termites and *C. punctulatus*, are dominated by termite *Treponeme*-like FTHFS sequences. The library generated from Cost004, a litter feeder, was dominated by nonacetogenic FTHFS types, but the majority of acetogenic FTHFS sequences present were *Treponeme*-like. The remaining species of termite were subterranean and appeared to feed on soil and/or plant material. Cost010

and JT2 were both dominated by FTHFS types that grouped with acetogenic *Firmicutes*. Like Cost004, the JT5 library was dominated by nonacetogens. However, the majority of acetogenic FTHFS sequences from JT5 grouped with those from Cost010 and JT2.

FTHFS types from probable acetogens were identified in all higher termite species. Figure 3.3 summarizes the phylogenetic relationships amongst the “Lovell cluster” of probable acetogens (marked as node A). Groups A, B, and C have been termed *Firmicute* acetogens, as those represent the most closely affiliated characterized organisms. However, it should be noted that the distances between termite gut FTHFS types and those from *Firmicute* acetogens are relatively large, and that at least one incidence of horizontal gene-transfer (to generate the termite *Treponeme* clade) has been postulated within this cluster.

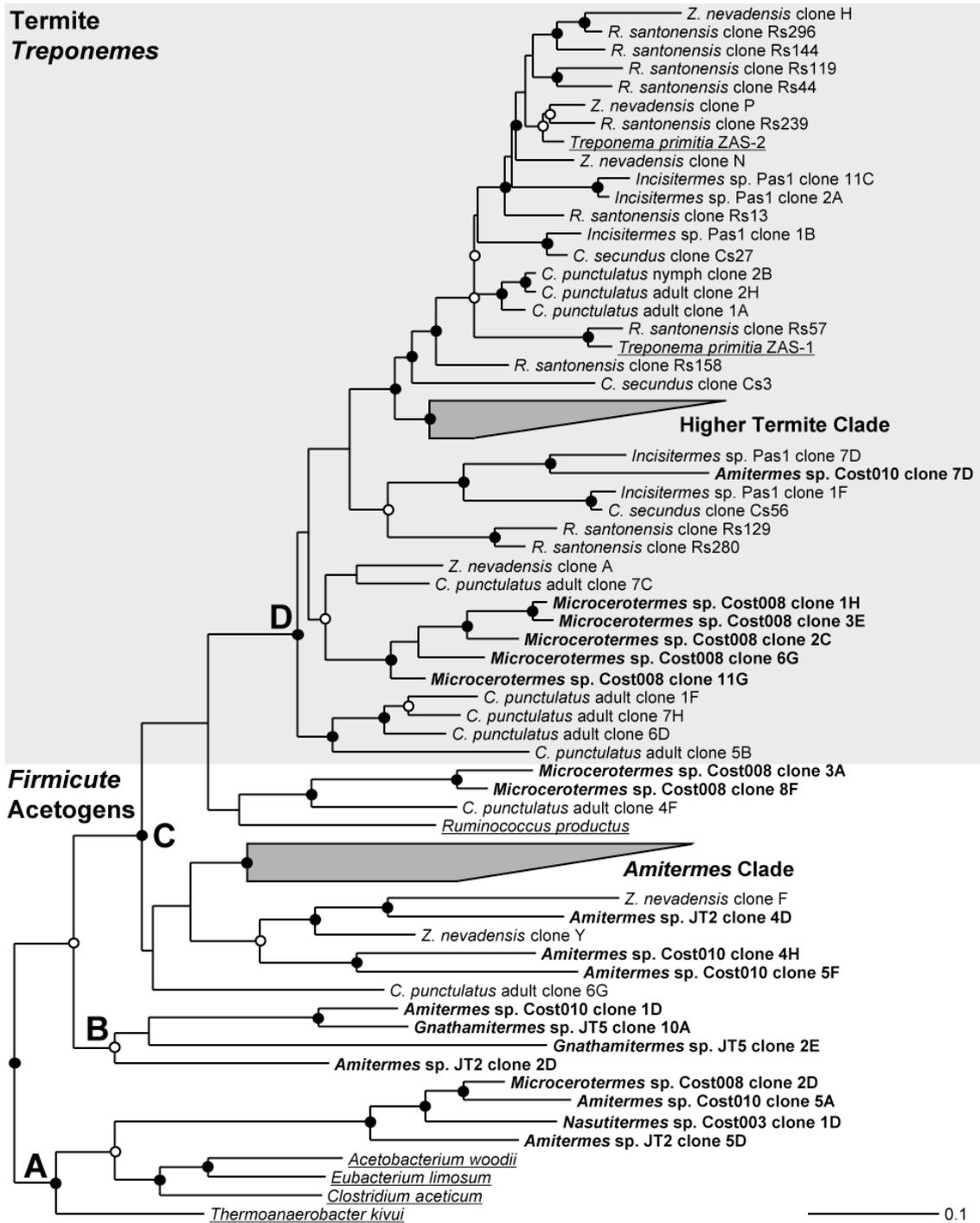


Figure 3.3. FTHFS sequences from potential acetogens. Sequences from this study marked in bold, known acetogens underlined. Tree calculated using Phylip PROML and 339 unambiguously aligned amino acids. Open circles mark groupings also supported by either Phylip PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position. Tree was rooted using 7 members of the *Moorella* / *Sporomusa* group of potential acetogens.

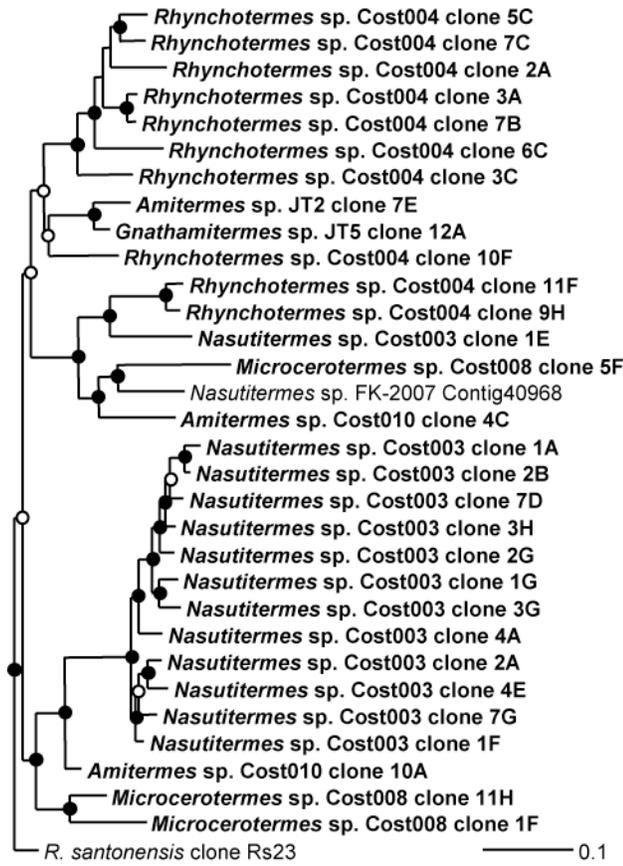


Figure 3.4. Higher termite clade of termite *Treponemes*. Tree calculated using Phylip PROML and 354 unambiguously aligned amino acids. Open circles mark groupings also supported by either Phylip PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position. An outgroup consisting 3 termite *Treponeme* isolates was used to root the cluster.

Node D was chosen as the boundary for the termite *Treponeme* clade of FTHFS sequences. Most of the sequences included in this group (*C. punctulatus* adult clone 7C and above) share a hexapeptide insert absent from other acetogens; the basal *Microcerotermes* and *C. punctulatus* clusters were included based on the strength of their phylogenetic association with these sequences. While Cost008 was dominated by a distinct group of termite *Treponeme* FTHFS sequences, most of the *Treponeme*-like sequences amplified from higher termites formed a single cluster (Figure 3.4). These sequences grouped to the exclusion of FTHFS types from lower termites. All higher

termites examined hosted sequences that fell within the “higher termite clade”; clones affiliated with this group represented 98% of acetogenic FTHFS sequences retrieved from Cost003, 85% in Cost004, 26% in Cost008, 8% in Cost010, 1% in JT2, and 6% in JT5.

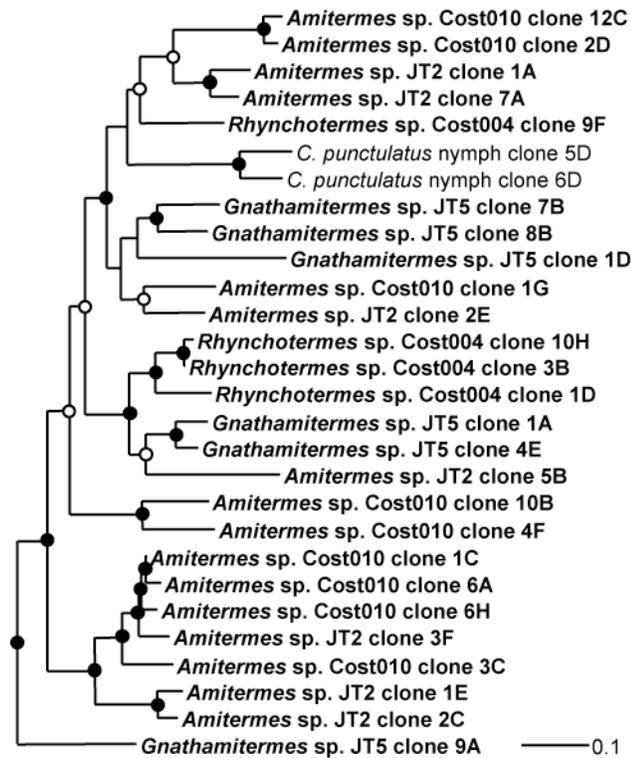


Figure 3.5. Amitermes clade of probable *Firmicute* acetogens. Tree calculated using Phylip PROML and 340 unambiguously aligned amino acids. Open circles mark groupings also supported by either Phylip PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position. An outgroup consisting of 6 cultured *Firmicute* acetogens was used to root the cluster.

The acetogenic community of the three subterranean termites (Cost010, JT2, and JT5) was dominated by a novel clade of *Firmicute*-like FTHFS sequences (the *Amitermes* clade in Figure 3.3, phylogenetic detail in Figure 3.5). This clade also included sequences from Cost004 and *C. punctulatus*. Sequences affiliated with this cluster

represented 89% of acetogenic FTHFS sequences found in JT2, 83% in JT5, 72% in Cost010, 15% in Cost004, and 4% of those found in the *C. punctulatus* nymph.

Cost010, JT2 and JT5 termites also contained sequences that fell within the *Moorella/Sporomusa* FTHFS clade (Figure 3.6). This clade contained FTHFS sequences both from true acetogens such as *Sporomusa termitida* (5) and *Moorella thermoacetica* (32) and from organisms that carry all or some of the machinery for the acetyl-CoA cycle but do not grow as CO₂-reductive acetogens, such as *Desulfitobacterium hafniense* (26), *Carboxydotherrmus hydrogenoformans* (48), and *Syntrophomonas wolfei* (21). As a result, while these sequences may belong to acetogenic organisms, we chose not to define them as such.

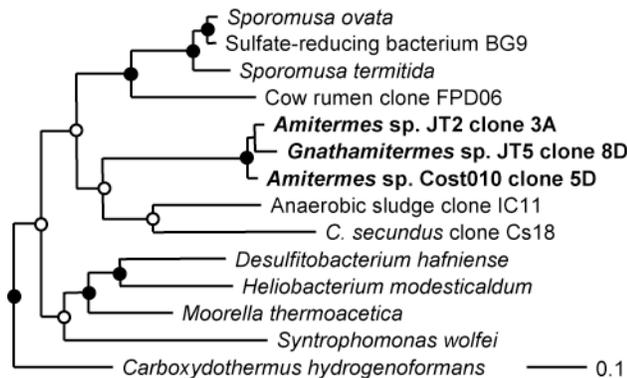


Figure 3.6. Phylogeny of *Moorella / Sporomusa* FTHFS clade. Tree calculated using Phylip PROML and 350 unambiguously aligned amino acids. Open circles mark groupings also supported by either Phylip PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position. An outgroup consisting of 7 cultured *Firmicute* acetogens was used to root the cluster.

The remaining FTHFS types were identified as probable nonacetogens (Figures 3-7 and 3-8). Two of these sequence groups can be assigned a probable role in amino acid or purine degradation. The first, described as the *Clostridium acidurici* group, included

clones that represent 45% of the FTHFS library from Cost004, and 10% of FTHFS sequences from JT5 (Figure 3.7a). This sequence cluster was closely related to FTHFS sequences from purinolytic *Firmicutes* *C. acidurici* (1), *Clostridium cylindrosporium* (1), and *Eubacterium acidaminophilum* (50). In these organisms, anaerobic degradation of purines results in the transfer of a formimino group to tetrahydrofolate (THF). Formimino-THF is converted to formyl-THF, and FTHFS is used to couple the release of formate and THF to generation of ATP via substrate-level phosphorylation. Uric acid degradation, which has been hypothesized as a role for gut bacteria in termite nitrogen conservation (33), can proceed via this pathway (45).

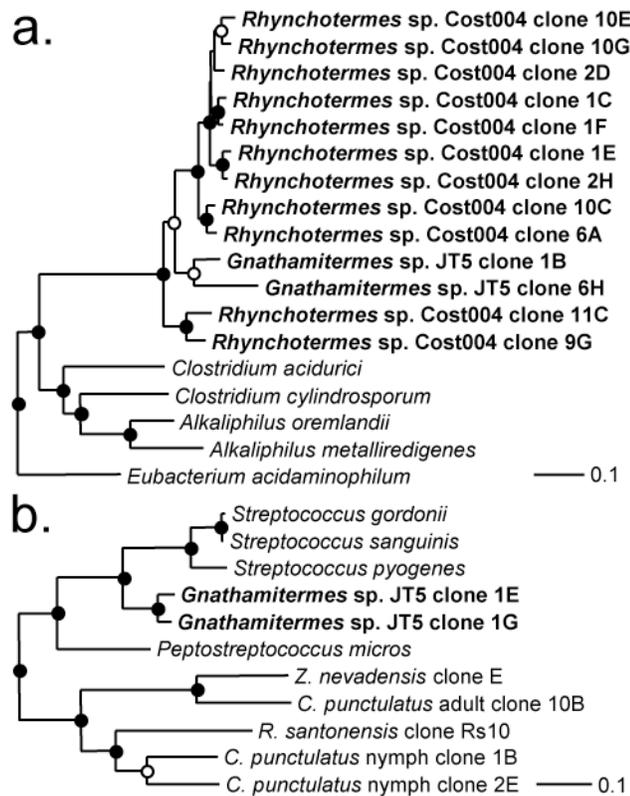


Figure 3.7. Putative amino acid or purine-degrading FTHFS clades. **a.** *C. acidurici* clade. **b.** Clone E/*Streptococcus* clade. Trees calculated using Phylip PROML and 351 unambiguously aligned amino acids. Open circles mark groupings also supported by either Phylip PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position. Trees rooted using 7 cultured *Firmicute* acetogens.

The second group of FTHFS types linked to purine or amino acid degradation is the Clone E/*Streptococcus* group (Figure 3.7b). Clones that represented 37% of the total FTHFS library from JT5 clustered with FTHFS sequences from *Peptostreptococcus micros* and three species of *Streptococcus* (*S. pyogenes*, *S. gordonii*, and *S. sanguinis*). While the genomic context of the FTHFS gene from *Peptostreptococcus micros* does not contain obvious functional clues, the Clone E-like FTHFS sequences in the three *Streptococci* are part of a conserved histidine degradation operon. In this context, FTHFS is again being used to generate ATP from the release of formate following the breakdown of a formimino group attached to THF. While the use of FTHFS to generate ATP from the release of formate during histidine degradation has not been formally reported in bacteria, the presence of glutamate formimidoyltransferase in certain bacterial histidine degradation operons has been observed via bioinformatics techniques (29).

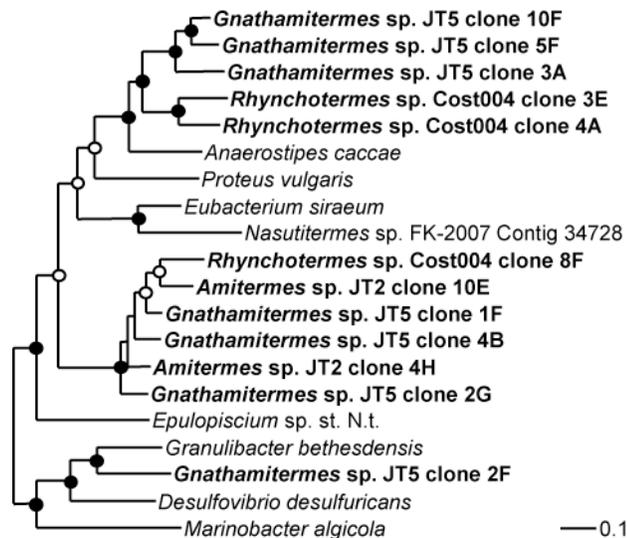


Figure 3.8. Nonacetogenic FTHFS sequences. Tree calculated using Phylip PROML and 350 unambiguously aligned amino acids. Open circles mark groupings also supported by either Phylip PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position.

Remaining nonacetogenic FTHFS sequences fall into three clusters (Figure 8). JT5 clone 2F appears to belong to a sulfate-reducing *Proteobacterium*. The cluster that includes JT5 clone 10F clone 4A appears to belong to nonacetogenic *Firmicutes*. Finally, a third cluster, including Cost004 clone 8F, does not cluster reliably with currently available FTHFS sequences.

Discussion

The diverse feeding habits of higher termites seem to have dramatic effects on the population of acetogenic bacteria in their guts. FTHFS libraries from lower termites have, without exception, proven to be dominated by sequences from the termite gut *Treponeme* clade. In higher termites, there appear to be two distinct scenarios for acetogenic bacteria. In Cost003, Cost008, and Cost004, the most abundant acetogenic bacteria appear to be termite *Treponemes* (comprising 98%, 89%, and 37% of total FTHFS). These termites appeared to primarily feed on wood, palm, and litter, respectively. In Cost010, JT2, and JT5, the most abundant acetogenic bacteria appear to be *Firmicutes* (85%, 87%, and 28% of FTHFS sequences), most of which fall within the novel “*Amitermes* clade.” These termites had subterranean lifestyles that are consistent with increased exposure to humics and a grass- or soil-feeding diet.

Termite *Treponeme* FTHFS sequences from higher termites largely fell within a single “higher termite clade.” This finding is a striking contrast to the diversity of FTHFS types found in lower termites, and may indicate an evolutionary bottleneck during which most lines of acetogenic *Treponemes* were lost. The higher termite clade may also represent a

symbiotic innovation that allowed this particular line of acetogens to outcompete other bacteria. The only termite gut *Treponeme* sequences that fell outside this cluster were a group of sequences found in Cost008 and a single sequence identified in Cost010. These may represent either FTHFS types lost from other lines of higher termites but retained in these insects or a reacquisition from lower termites of FTHFS types lost early in the higher termite radiation.

Cost010, JT2, and JT5 represent the first examples of termite gut communities that are not dominated by FTHFS sequences from the termite *Treponeme* clade, but rather by *Firmicute*-associated sequences. Given that *Treponeme*-associated FTHFS types are present in these termites, it seems likely that this shift in community structure is due to the presence of conditions that favor this group over acetogenic *Treponemes*. These termites have subterranean lifestyles and diets that potentially include soil-feeding. Soil-feeding *Cubitermes* spp. have been shown to have low rates of *in situ* CO₂ fixation to acetate. However, a robust population of CO₂-reductive acetogens can be detected in gut homogenates when incubated with inhibitors of methanogenesis (42). While studies in the termite gut have focused on acetogenesis from H₂ and CO₂, acetogens are capable of utilizing a wide variety of substrates (7). The acetogens present in soil-feeding termites may principally subsist on alternative sources of reducing equivalents and/or carbon, such as carbohydrates or methoxylated aromatics. The *Amitermes* clade of FTHFS types may represent organisms better adapted to this lifestyle than termite gut *Treponemes*.

Finally, although the primers utilized in this study were designed for specific detection of acetogenic FTHFS types, several intriguing nonacetogenic FTHFS types were identified in higher termite libraries. A group of FTHFS types was identified in Cost004 and JT5 that cluster with FTHFS sequences from purinolytic *Firmicutes*. Uric acid recycling by gut bacteria has been hypothesized to play a role in termite nitrogen conservation (33), and the presence of this FTHFS clade suggests that *Firmicutes* may be carrying out this activity within the guts of Cost004 and JT5. Litter-feeding *Rhynchotermes* termites have been shown to have lower rates of nitrogen fixation than wood-feeding *Nasutitermes* (35). While this was initially attributed to higher nitrogen content in their food source, uric acid recycling may also play a role. Additionally, these bacteria may aid in release of nitrogen from food material.

The Clone E / *Streptococcus* FTHFS clade also likely represents an alternative use of the FTHFS enzyme. The *Streptococci* associated with this clade appear to be utilizing FTHFS in the context of histidine degradation. The entire clade may represent FTHFS types employed in this manner or, alternatively, it may represent FTHFS adapted for formyl-THF metabolism rather than synthesis. One of the uricolytic strains isolated from *Reticulitermes flavipes* by Potrikus and Breznak (34) was a *Streptococcus* species; the termite-derived sequences may represent FTHFS genes from similar organisms.

In conclusion, the diversity of lifestyles and feeding strategies employed by higher termites coincides with a diversity of population structures among symbiotic acetogens. FTHFS sequences amplified from wood-, palm-, and litter-feeding higher termites

affiliate with the acetogenic *Treponemes* that dominate the guts of wood-feeding lower termites. However, subterranean termites, whose diets may include some level of soil-feeding (and who certainly experience greater exposure to soil), yielded a diversity of sequences that affiliate with acetogenic *Firmicutes* but few *Treponeme*-like FTHFS sequences. It has been broadly observed that wood-feeding termites (both higher and lower) have higher rates of acetogenesis than soil feeders; this may correlate with a uniquely favorable environment for acetogenic *Treponemes*.

Chapter Three Appendix

- 1. Table 3.3. Operational taxonomic unit grouping of FTHFS sequences identified in this study**
- 2. Table 3.4. Sequences used in FTHFS phylogenetic analysis**
- 3. Table 3.5. Sequences used in COII phylogenetic analysis**

Table 3.3. Operational taxonomic unit grouping of FTHFS sequences identified in this study

Group	Phylotype	Abundance (%) ^a	Genotypes ^b	
<i>Nasutitermes</i> sp. Cost003				
Termite <i>Treponemes</i>	1F	23	1F	
	2B	15	2B, 2F	
	1A	12	1A, 2D, 7B	
	2A	10	2A, 7A	
	1E	6	1E	
	2G	6	2G	
	3H	6	3H	
	1G	4	1G	
	4A	4	4A, 4B	
	4E	4	4E	
	7D	4	7D	
	3G	2	3G	
	7G	2	7G	
	Acetogenic <i>Firmicutes</i>	1D	2	1D
<i>Microcerotermes</i> sp. Cost008				
Termite <i>Treponemes</i>	1H	30	1H, 5E, 8H	
	1F	15	1F	
	2C	15	2C, 11A	
	6G	11	6G, 4H	
	11G	4	11G	
	11H	4	11H	
	3E	4	3E	
	5F	4	5F	
	9E	4	9E	
	Acetogenic <i>Firmicutes</i>	2D	4	2D
	3A	4	3A	
	8F	4	8F	
<i>Rynchotermes</i> sp. Cost004				
Termite <i>Treponemes</i>	7C	10	7C, 2E, 4E, 5A	
	3A	8	3A, 8C	
	6C	5	6C	
	3C	3	3C	
	5C	3	5C, 11A	
	10F	2	10F	
	11F	2	11F	
	2A	2	2A	
	7B	2	7B	
	9H	2	9H	
	Acetogenic <i>Firmicutes</i>	10H	2	10H
		1D	2	1D
		3B	2	3B
		9F	2	9F
<i>Clostridium acidiurici</i>	1C	13	1C, 2B	
	1E	10	1E, 7A, 8D, 4C, 6B	
	10C	3	10C	
	10E	3	10E, 12E	
	1F	3	1F	
	2H	3	2H, 9E	
	9G	3	9G, 12H	

Group	Phylotype	Abundance (%) ^a	Genotypes ^b
	10G	2	10G
	11C	2	11C
	2D	2	2D
	6A	2	6A
Nonacetogenic FTHFS	4A	6	4A
	3E	2	3E
	8F	2	8F
<i>Amitermes</i> sp. Cost010			
Termite <i>Treponemes</i>	10A	4	10A
	4C	4	4C
	7D	4	7D
Acetogenic <i>Firmicutes</i>	2D	23	2D, 8H
	4F	15	4F
	1G	8	1G
	10B	4	10B
	12C	4	12C
	1C	4	1C
	1D	4	1D
	3C	4	3C
	4H	4	4H
	5A	4	5A
	5F	4	5F
	6A	4	6A
	6H	4	6H
<i>Moorella</i> / <i>Sporomusa</i>	5D	4	5D
<i>Amitermes</i> sp. JT2			
Termite <i>Treponemes</i>	7E	1	7E
Acetogenic <i>Firmicutes</i>	1E	48	1E, 1B, 8G, 12B
	1A	20	1A, 2G, 1F
	2E	6	2E
	2C	2	2C
	3F	2	3F
	5D	3	5D, 8A
	7A	2	7A
	2D	1	2D
	4D	1	4D
	5B	1	5B
<i>Moorella</i> / <i>Sporomusa</i>	2H	6	2H, 3A
Nonacetogenic FTHFS	4H	2	4H, 3D
	10E	1	10E
<i>Gnathamitermes</i> sp. JT5			
Termite <i>Treponemes</i>	12A	2	12A
Acetogenic <i>Firmicutes</i>	4E	7	4E
	1D	5	1D
	8B	5	8B
	1A	3	1A
	9A	3	9A
	10A	2	10A
	2E	2	2E
	7B	2	7B
<i>Moorella</i> / <i>Sporomusa</i>	8D	2	8D
<i>Clostridium acidurici</i>	1B	7	1B
	6H	3	6H
Clone E / <i>Streptococcus</i>	1G	27	1G, 8E

Group	Phylotype	Abundance (%)^a	Genotypes^b
Nonacetogenic FTHFS	1E	10	1E
	3A	7	3A, 9F
	10F	2	10F
	1F	2	1F
	2F	2	2F
	2G	2	2G
	4B	2	4B
5F	2	5F	

^aDefined as percent of full-length, nonchimeric clones

^bSequenced RFLP type clones. Group representative marked in bold.

Table 3.4. Sequences used in FTHFS phylogenetic analysis

Source / Sequence Type	Designation	Accession	Reference
<i>Cryptocercus punctulatus</i> adult gut clone	1A		
<i>Cryptocercus punctulatus</i> adult gut clone	1F		
<i>Cryptocercus punctulatus</i> adult gut clone	4F		
<i>Cryptocercus punctulatus</i> adult gut clone	5B		
<i>Cryptocercus punctulatus</i> adult gut clone	6D		
<i>Cryptocercus punctulatus</i> adult gut clone	7C		
<i>Cryptocercus punctulatus</i> adult gut clone	7H		
<i>Cryptocercus punctulatus</i> adult gut clone	10B		
<i>Cryptocercus punctulatus</i> nymph gut clone	1B		
<i>Cryptocercus punctulatus</i> nymph gut clone	2E		
<i>Cryptocercus punctulatus</i> nymph gut clone	5D		
<i>Cryptocercus punctulatus</i> nymph gut clone	6D		
<i>Cryptotermes secundus</i> gut clone	Cs3	DQ278251	(30)
<i>Cryptotermes secundus</i> gut clone	Cs18	DQ278253	(30)
<i>Cryptotermes secundus</i> gut clone	Cs27	DQ278254	(30)
<i>Cryptotermes secundus</i> gut clone	Cs56	DQ278258	(30)
<i>Incisitermes</i> sp. Pas-1 gut clone	1B		
<i>Incisitermes</i> sp. Pas-1 gut clone	1F		
<i>Incisitermes</i> sp. Pas-1 gut clone	2A		
<i>Incisitermes</i> sp. Pas-1 gut clone	7D		
<i>Reticulitermes santonensis</i> gut clone	Rs10	DQ278259	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs13	DQ278232	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs23	DQ278210	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs44	DQ278211	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs57	DQ278215	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs119	DQ278226	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs129	DQ278222	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs144	DQ278223	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs158	DQ278226	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs239	DQ278201	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs280	DQ278207	(30)
<i>Reticulitermes santonensis</i> gut clone	Rs296	DQ278208	(30)
<i>Zootermopsis angusticollis</i> gut clone	A	AY162294	(37)
<i>Zootermopsis angusticollis</i> gut clone	E	AY162296	(37)
<i>Zootermopsis angusticollis</i> gut clone	F	AY162298	(37)
<i>Zootermopsis angusticollis</i> gut clone	H	AY162302	(37)
<i>Zootermopsis angusticollis</i> gut clone	N	AY162306	(37)
<i>Zootermopsis angusticollis</i> gut clone	P	AY162307	(37)
<i>Nasutitermes</i> sp. FK-2007 metagenome	Contig34728	JGI GOI: 2004131907	(46)
<i>Nasutitermes</i> sp. FK-2007 metagenome	Contig40968	JGI GOI: 2004144560	(46)
Cow Rumen clone	FPDO6	AB085528	Database only
Anaerobic sludge clone	IC11	EU009529	Database only
<i>Aceotbacterium woodii</i>		AF295701	(15)
<i>Alkaliphilus metalliredigenes</i> QYMF		CP000724	Database only
<i>Alkaliphilus oremlandii</i>		NC_009922	Database only
<i>Anaerostipes caccae</i>		ABAX03000038	Database only
<i>Carboxydotherrmus hydrogenoformans</i> Z-2901		ABB16038	(48)
<i>Clostridium aceticum</i>		AF295705	(15)
<i>Clostridium acidurici</i>		M21507	(47)
<i>Clostridium cylindrosporum</i>		L12465	(36)
<i>Clostridium formicaceticum</i>		AF295702	(15)

Source / Sequence Type	Designation	Accession	Reference
<i>Clostridium magnum</i>		AF295703	(15)
<i>Desulfitobacterium hafniense</i> st. Y51		NC_007907	(26)
<i>Desulfovibrio desulfuricans</i>		AJ494753	(14)
<i>Epulopiscium</i> sp. st. N.t. morphotype B		NZ_ABEQ01000077	Database only
<i>Eubacterium acidaminophilum</i>		AY722711	Database only
<i>Eubacterium limosum</i>		AF295706	(15)
<i>Eubacterium siraeum</i>		ABCA03000037	Database only
<i>Granulibacter bethesdensis</i>		NC_008343	(10)
<i>Heliobacterium modesticaldum</i>		NC_010337	(38)
<i>Marinobacter algicola</i>		ZP_01892361	Database only
<i>Moorella thermoacetica</i>		NC_007644	(32)
<i>Peptostreptococcus micros</i>		NZ_ABEE02000017	Database only
<i>Proteus vulgaris</i>		AF295710	(15)
<i>Ruminococcus productus</i>		AF295707	(15)
<i>Sporomusa ovata</i>		AF295708	(15)
<i>Sporomusa termitida</i>		AF295709	(15)
<i>Streptococcus gordonii</i>		NC_009785	(44)
<i>Streptococcus pyogenes</i> st. SSI-1		BAC64868	(25)
<i>Streptococcus sanguinis</i>		NC_009009	(49)
Sulfate-reducing bacterium BG9		AJ494757	(14)
<i>Syntrophomonas wolfei</i> subsp. <i>wolfei</i> str. Goettingen		EAO23711	Database only
<i>Thermoanaerobacter kivui</i>		AF295704	(15)
<i>Treponema azotonutricium</i> st. ZAS-9		AY162316	(37)
<i>Treponema denticola</i>		NC_002967	(40)
<i>Treponema primitia</i> st. ZAS-1	ZAS-1a	AY162313	(37)
<i>Treponema primitia</i> st. ZAS-2		AY162315	(37)

Table 3.5. Sequences used in COII phylogenetic analysis

Source	Accession	Reference
<i>Amitermes evuncifer</i>	DQ442066	(12)
<i>Archotermopsis wroughtoni</i>	DQ442080	(12)
<i>Coptotermes formosanus</i>	AB109529	(28)
<i>Cornitermes pugnax</i>	DQ442106	(12)
<i>Cryptocercus primarius</i>	DQ007644	(17)
<i>Cryptocercus punctulatus</i> adult		
<i>Cryptocercus punctulatus</i> nymph		
<i>Cryptotermes secundus</i>	DQ442111	(12)
<i>Incisitermes</i> sp. Pas1		
<i>Kalotermes hilli</i>	AF189101	(43)
<i>Labiotermes labralis</i>	DQ442149	(12)
<i>Microcerotermes newmani</i>	DQ442166	(12)
<i>Microcerotermes parvus</i>	DQ442167	(12)
<i>Nasutitermes ephratae</i>	AB037328	(23)
<i>Nasutitermes nigriceps</i>	DQ442193	(12)
<i>Nasutitermes</i> sp. FK-2007	EU236539	(46)
<i>Reticulitermes santonensis</i>	AF291742	(19)
<i>Reticulitermes speratus</i>	AB109530	(28)
<i>Zootermopsis angusticollis</i>	DQ442267	(12)
<i>Zootermopsis nevadensis</i>		

References

1. **Barker, H. A., and J. V. Beck.** 1942. *Clostridium acidi-uridi* and *Clostridium cylindrosporum*, organisms fermenting uric acid and some other purines. J Bacteriol **43**:291-304.
2. **Bitsch, C., and C. Noirot.** 2002. Gut characters and phylogeny of the higher termites (Isoptera : Termitidae). A cladistic analysis. Ann Soc Entomol Fr **38**:201-210.
3. **Brauman, A., J. Dore, P. Eggleton, D. Bignell, J. A. Breznak, and M. D. Kane.** 2001. Molecular phylogenetic profiling of prokaryotic communities in guts of termites with different feeding habits. FEMS Microbiol Ecol **35**:27-36.
4. **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.
5. **Breznak, J. A., J. M. Switzer, and H. J. Seitz.** 1988. *Sporomusa termitida* sp. nov., an H₂/CO₂-utilizing acetogen isolated from termites. Arch Microbiol **150**:282-288.
6. **Brune, A.** 1998. Termite guts: the world's smallest bioreactors. Trends Biotechnol **16**:16-21.
7. **Drake, H., K. Küsel, and C. Matthies.** 2006. Acetogenic prokaryotes, p. 354-420. In M. Dworkin, S. Falkow, E. Rosenber, K. H. Schleifer, and E. Stackebrandt (ed.), *The Prokaryotes*, Third ed. Springer.
8. **Edgar, R. C.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5**:113.

9. **Engel, M. S., and K. Krishna.** 2004. Family-group names for termites (Isoptera). *Amer Museum novitates* **3432**:1-9.
10. **Greenberg, D. E., S. F. Porcella, A. M. Zelazny, K. Virtaneva, D. E. Sturdevant, J. J. Kupko, III, K. D. Barbian, A. Babar, D. W. Dorward, and S. M. Holland.** 2007. Genome sequence analysis of the emerging human pathogenic acetic acid bacterium *Granulibacter bethesdensis*. *J Bacteriol* **189**:8727-8736.
11. **Huber, T., G. Faulkner, and P. Hugenholtz.** 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**:2317-2319.
12. **Inward, D. J., A. P. Vogler, and P. Eggleton.** 2007. A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspects of their evolutionary biology. *Mol Phylogenet Evol* **44**:953-967.
13. **Kambhampati, S., and P. Eggleton.** 2000. Taxonomy and phylogeny of termites, p. 1-24. *In* T. Abe, D. E. Bignell, and M. Higashi (ed.), *Termites: Evolution, Sociality, Symbioses, Ecology*. Kluwer Academic Publishers, Dordrecht.
14. **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.

15. **Leaphart, A. B., and C. R. Lovell.** 2001. Recovery and analysis of formyltetrahydrofolate synthetase gene sequences from natural populations of acetogenic bacteria. *Appl Environ Microbiol* **67**:1392-1395.
16. **Ljungdahl, L. G.** 1986. The autotrophic pathway of acetate synthesis in acetogenic bacteria. *Ann Rev Microbiol* **40**:415-450.
17. **Lo, N., P. Luykx, R. Santoni, T. Beninati, C. Bandi, M. Casiraghi, W. H. Lu, E. V. Zakharov, and C. A. Nalepa.** 2006. Molecular phylogeny of *Cryptocercus* wood-roaches based on mitochondrial COII and 16S sequences, and chromosome numbers in Palearctic representatives. *Zoolog Sci* **23**:393-398.
18. **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. W. 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. H. Schleifer.** 2004. ARB: a software environment for sequence data. *Nucl Acids Res* **32**:1363-1371.
19. **Marini, M., and B. Mantovani.** 2002. Molecular relationships among European samples of *Reticulitermes* (Isoptera, Rhinotermitidae). *Mol Phylogenet Evol* **22**:454-459.
20. **Matson, E., E. A. Ottesen, and J. R. Leadbetter.** 2007. Extracting DNA from the gut microbes of the termite (*Zootermopsis nevadensis*). *J Visualized Exper* **4**.

21. **McInerney, M. J., M. P. Bryant, R. B. Hespell, and J. W. Costerton.** 1981. *Syntrophomonas wolfei* gen. nov. sp. nov., an anaerobic, syntrophic, fatty acid-oxidizing bacterium. *Appl Environ Microbiol* **41**:1029-1039.
22. **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 of Amer* **91**:515-523.
23. **Miura, T., Y. Roisin, and T. Matsumoto.** 2000. Molecular phylogeny and biogeography of the nasute termite genus *Nasutitermes* (Isoptera: Termitidae) in the pacific tropics. *Mol Phylogenet Evol* **17**:1-10.
24. **Miyata, R., N. Noda, H. Tamaki, K. 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.
25. **Nakagawa, I., K. Kurokawa, A. Yamashita, M. Nakata, Y. Tomiyasu, N. Okahashi, S. Kawabata, K. Yamazaki, T. Shiba, T. Yasunaga, H. Hayashi, M. Hattori, and S. Hamada.** 2003. Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. *Genome Res* **13**:1042-1055.
26. **Nonaka, H., G. Keresztes, Y. Shinoda, Y. Ikenaga, M. Abe, K. Naito, K. Inatomi, K. Furukawa, M. Inui, and H. Yukawa.** 2006. Complete genome sequence of the dehalorespiring bacterium *Desulfitobacterium hafniense* Y51 and comparison with *Dehalococcoides ethenogenes* 195. *J Bacteriol* **188**:2262-2274.

27. **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.
28. **Ohkuma, M., H. Yuzawa, W. Amornsak, Y. Sorannuwat, Y. Takematsu, A. Yamada, C. Vongkaluang, O. Sarnthoy, N. Kirtibutr, N. Noparatnaraporn, T. Kudo, and T. Inoue.** 2004. Molecular phylogeny of Asian termites (Isoptera) of the families Termitidae and Rhinotermitidae based on mitochondrial COII sequences. *Mol Phylogenet Evol* **31**:701-710.
29. **Overbeek, R., T. Begley, R. M. Butler, J. V. Choudhuri, H. Y. Chuang, M. Cohoon, V. de Crécy-Lagard, N. Diaz, T. Disz, R. Edwards, M. Fonstein, E. D. Frank, S. Gerdes, E. M. Glass, A. Goesmann, A. Hanson, D. Iwata-Reuyl, R. Jensen, N. Jamshidi, L. Krause, M. Kubal, N. Larsen, B. Linke, A. C. McHardy, F. Meyer, H. Neuweger, G. Olsen, R. Olson, A. Osterman, V. Portnoy, G. D. Pusch, D. A. Rodionov, C. Ruckert, J. Steiner, R. Stevens, I. Thiele, O. Vassieva, Y. Ye, O. Zagnitko, and V. Vonstein.** 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* **33**:5691-5702.
30. **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.
31. **Pester, M., and A. Brune.** 2007. Hydrogen is the central free intermediate during lignocellulose degradation by termite gut symbionts. *Isme J* **1**:551-565.

32. **Pierce, E., G. Xie, R. D. Barabote, E. Saunders, C. S. Han, J. C. Detter, P. Richardson, T. S. Brettin, A. Das, L. G. Ljungdahl, and S. W. Ragsdale.** 2008. The complete genome sequence of *Moorella thermoacetica* (f. *Clostridium thermoaceticum*). *Environ Microbiol* **10**:2550-2573.
33. **Potrikus, C. J., and J. A. Breznak.** 1981. Gut bacteria recycle uric acid nitrogen in termites: A strategy for nutrient conservation. *Proc Natl Acad Sci USA* **78**:4601-4605.
34. **Potrikus, C. J., and J. A. Breznak.** 1980. Uric acid-degrading bacteria in guts of termites [*Reticulitermes flavipes* (Kollar)]. *Appl Environ Microbiol* **40**:117-124.
35. **Prestwich, G. D., B. L. Bentley, and E. J. Carpenter.** 1980. Nitrogen sources for neotropical Nasute termites - Fixation and selective foraging. *Oecologia* **46**:397-401.
36. **Rankin, C. A., G. C. Haslam, and R. H. Himes.** 1993. Sequence and expression of the gene for N(10)-formyltetrahydrofolate synthetase from *Clostridium cylindrosporium*. *Protein Sci* **2**:197-205.
37. **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*. *Microbiol* **149**:2529-2537.
38. **Sattley, W. M., M. T. Madigan, W. D. Swingley, P. C. Cheung, K. M. Clocksin, A. L. Conrad, L. C. Dejesa, B. M. Honchak, D. O. Jung, L. E. Karbach, A. Kurdoglu, S. Lahiri, S. D. Mastrian, L. E. Page, H. L. Taylor, Z. T. Wang, J. Raymond, M. Chen, R. E. Blankenship, and J. W. Touchman.** 2008. The genome of *Heliobacterium modesticaldum*, a phototrophic

representative of the *Firmicutes* containing the simplest photosynthetic apparatus.

J Bacteriol **190**:4687-4696.

39. **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.
40. **Seshadri, R., G. S. Myers, H. Tettelin, J. A. Eisen, J. F. Heidelberg, R. J. Dodson, T. M. Davidsen, R. T. DeBoy, D. E. Fouts, D. H. Haft, J. Selengut, Q. Ren, L. M. Brinkac, R. Madupu, J. Kolonay, S. A. Durkin, S. C. Daugherty, J. Shetty, A. Shvartsbeyn, E. Gebregeorgis, K. Geer, G. Tsegaye, J. Malek, B. Ayodeji, S. Shatsman, M. P. McLeod, D. Smajs, J. K. Howell, S. Pal, A. Amin, P. Vashisth, T. Z. McNeill, Q. Xiang, E. Sodergren, E. Baca, G. M. Weinstock, S. J. Norris, C. M. Fraser, and I. T. Paulsen.** 2004. Comparison of the genome of the oral pathogen *Treponema denticola* with other spirochete genomes. Proc Natl Acad Sci USA **101**:5646-5651.
41. **Shinzato, N., M. Muramatsu, T. Matsui, and Y. Watanabe.** 2007. Phylogenetic analysis of the gut bacterial microflora of the fungus-growing termite *Odontotermes formosanus*. Biosci Biotechnol Biochem **71**:906-915.
42. **Tholen, A., and A. Brune.** 1999. Localization and in situ activities of homoacetogenic bacteria in the highly compartmentalized hindgut of soil-feeding higher termites (*Cubitermes* spp.). Appl Environ Microbiol **65**:4497-4505.

43. **Thompson, G. J., L. R. Miller, M. Lenz, and R. H. Crozier.** 2000. Phylogenetic analysis and trait evolution in Australian lineages of drywood termites (Isoptera, Kalotermitidae). *Mol Phylogenet Evol* **17**:419-429.
44. **Vickerman, M. M., S. Iobst, A. M. Jesionowski, and S. R. Gill.** 2007. Genome-wide transcriptional changes in *Streptococcus gordonii* in response to competence signaling peptide. *J Bacteriol* **189**:7799-7807.
45. **Vogels, G. D., and C. Van der Drift.** 1976. Degradation of purines and pyrimidines by microorganisms. *Bacteriol Rev* **40**:403-468.
46. **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**:560-565.
47. **Whitehead, T. R., and J. C. Rabinowitz.** 1988. Nucleotide sequence of the *Clostridium acidiurici* ("*Clostridium acidi-urici*") gene for 10-formyltetrahydrofolate synthetase shows extensive amino acid homology with the trifunctional enzyme C1-tetrahydrofolate synthase from *Saccharomyces cerevisiae*. *J Bacteriol* **170**:3255-3261.

48. **Wu, M., Q. H. Ren, A. S. Durkin, S. C. Daugherty, L. M. Brinkac, R. J. Dodson, R. Madupu, S. A. Sullivan, J. F. Kolonay, W. C. Nelson, L. J. Tallon, K. M. Jones, L. E. Ulrich, J. M. Gonzalez, I. B. Zhulin, F. T. Robb, and J. A. Eisen.** 2005. Life in hot carbon monoxide: The complete genome sequence of *Carboxydotherrmus hydrogenoformans* Z-2901. PLoS Gen **1**:563-574.
49. **Xu, P., J. M. Alves, T. Kitten, A. Brown, Z. Chen, L. S. Ozaki, P. Manque, X. Ge, M. G. Serrano, D. Puiu, S. Hendricks, Y. Wang, M. D. Chaplin, D. Akan, S. Paik, D. L. Peterson, F. L. Macrina, and G. A. Buck.** 2007. Genome of the opportunistic pathogen *Streptococcus sanguinis*. J Bacteriol **189**:3166-3175.
50. **Zindel, U., W. Freudenberg, M. Rieth, J. R. Andreasen, J. Schnell, and F. Widdel.** 1988. *Eubacterium acidaminophilum* sp nov, a versatile amino acid-degrading anaerobe producing or utilizing H₂ or formate - Description and enzymatic studies. Arch Microbiol **150**:254-266.

Microfluidic Digital PCR for Multigene Analysis of Individual Environmental Bacteria

Originally presented in: Elizabeth A Ottesen, Jong Wook Hong, Stephen R. Quake, Jared R. Leadbetter (2006). Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science* **314**(5804): 1464-1467. DOI: 10.1126/science.1131370.

Reprinted with Permission

Abstract

Gene inventory and metagenomic techniques have allowed rapid exploration of bacterial diversity and the potential physiologies present within microbial communities. However, it remains nontrivial to discover the identities of environmental bacteria carrying two or more genes of interest. We have employed microfluidic digital PCR to amplify and analyze multiple, different genes obtained from single bacterial cells harvested from nature. A gene encoding a key enzyme involved in the mutualistic symbiosis occurring between termites and their gut microbiota was used as an experimental hook to discover the previously unknown rRNA-based species identity of several symbionts. The ability to systematically identify bacteria carrying a particular gene and to link any two or more genes of interest to single species residing in complex ecosystems opens up new opportunities for research on the environment.

Article Text

A major challenge of environmental science is the identification of microbial species capable of catalyzing important activities *in situ* (12). PCR-based techniques that use single genes as proxies for organisms or key microbial activities continue to provide valuable insights into microbial community diversity (17, 44, 60). However, it has been difficult to interrelate PCR-derived gene inventories to derive correspondences between any two or more specific genes of interest, or to determine the phylogenetic species identity of organisms carrying particular genetic capabilities. Metagenomic (41) analyses of complex communities are dominated by genome “shrapnel”; unless the microbial community is dominated by one or a few species (45, 50) resident genomes are not reliably reconstructed via computation (49, 51). A gene of interest can be attributed to a specific organism only if it is linked to an unambiguous phylogenetic marker, i.e., on the same genome fragment (7, 41). Both PCR and metagenomic studies are typically carried out on homogenized, whole-community genomic DNA preparations. Thus the cell as a distinct informational entity is almost entirely lost.

Outside of traditional culture-based isolation, few approaches can attribute multiple genes to a single species or cell type. Microautoradiography (33) and stable isotope probing (31) allow detection of cells or retrieval of genetic material from organisms utilizing a substrate of interest, but require active cellular incorporation of that substrate. Microscopy-based *in situ* hybridization-based techniques (FISH and variants (5, 61)) allow colocalization of sequences through probe hybridization, but require that both genes be 1) actively transcribed and their sequences 2) be known in advance and 3) be of

sufficient difference from related, nontarget genes for effective probe design and implementation. Single cell whole genome amplification has recently been reported for a highly abundant, culturable marine microbial species, but has not yet been shown to be scalable to interrogating multitudes of diverse, coresident microbes (59). Here, we have applied microfluidic devices to perform a variant of “digital PCR” (52), separating and interrogating hundreds of individual environmental bacteria in parallel.

Microfluidic devices allow control and manipulation of small volumes of liquid (14, 48), in this case allowing for rapid separation and partitioning of single cells from a complex parent sample. Single, partitioned cells served as templates for individual multiplex PCR reactions using primers and probes for simultaneous amplification of both small-subunit ribosomal RNA and metabolic genes of interest. Primers and probes with broad target specificities were employed with subsequent resolution of exact gene sequences after successful amplification and retrieval. This technique operates independent of gene expression, position of the gene on the genome, and the physiological state of the cell at the time of harvest. This resulted in the rapid colocalization of two genes (encoding 16S rRNA and a key metabolic enzyme) to single genome templates, along with the determination of the fraction of cells within the community that encoded them. Subsequent retrieval of PCR products from individual chambers allowed sequence analysis of both genes; phylogenetic analysis of the ribosomal RNA gene allows classification of the host bacterium and the metabolic gene is sequenced to confirm the cell carried the genotype of interest. Additionally, since microfluidic digital PCR yields fluorescent signal upon amplification of a gene regardless of the number of copies

present the cell, this approach can yield estimates of the fraction given species represent within the general microbial community. The number of *rrn* operons present in a genome can vary widely, ranging from 1 (e.g., *Rickettsia prowazekii* (37)) to 15 (e.g., *C. paradoxum* (40)), confounding the interpretation of traditional environmental gene inventories. Moreover, the use of single cell PCR to prepare clone libraries will avoid complications and PCR artifacts such as amplification biases and unresolvable chimeric products (4).

We employed this technique to examine a complex, species-rich environment: the lignocellulose-decomposing microbial community resident in the hindguts of wood-feeding termites. Therein, the bacterial metabolism known as CO₂-reductive homoacetogenesis is one of the major sources of the bacterial fermentation product, acetate (10). Acetogenic bacteria must compete for hydrogen with *Archaea* that generate methane, a potent greenhouse gas for which termites are considered a small yet significant source. Because of their high rates of bacterially mediated homoacetogenesis, many termites contribute significantly much less to the global methane budget than they might otherwise (8). Additionally, acetate serves as the insect host's major carbon and energy source, literally fueling a large proportion of this mutualistic symbiosis (10, 35, 47). A key gene of the homoacetogenesis pathway encodes formyl-tetrahydrofolate synthetase (FTHFS) (27). Previously, a diversity of termite hindgut community FTHFS variants were inventoried (42), but the identities of the organisms dominating homoacetogenesis in termites had remained uncertain. Here using microfluidics, we

discovered the identities of a multitude of FTHFS-encoding organisms by determining their specific 16S rRNA gene sequences.

The “*Clone H Group*” of FTHFS genotypes corresponds to a large fraction of the sequences collected during an inventory of FTHFS genes present in the termite hindgut environment (42). We designed a specific primer set and a fluorescein-labeled probe capable of on-chip detection and amplification of the genotypes comprising this FTHFS group. We also redesigned broad-specificity “*all bacterial*” 16S rRNA gene primers and employed a previously published probe (46) to amplify and detect bacterial rRNA genes. Both the all bacterial 16S rRNA gene and Clone H Group FTHFS primer/probe sets showed single molecule sensitivity in multiplex on-chip reactions using purified plasmid or termite gut community DNA. The observed success rate for the amplification of individual genes from single molecule templates was 40% (see chapter appendix), thus the success rate for coamplification of two genes from single molecule templates is estimated to be ca. 1 in 7.

Freshly collected termite hindgut contents were suspended in a PCR reaction buffer and loaded into a microfluidic device. Each microfluidic panel uses micromechanical valves to randomly partition a single PCR mixture into 1,176 independent 6.25 nL reaction chambers (Figure 4.1). We considered single-cell separation to be achieved when fewer than one third of chambers showed rRNA gene amplification. Assuming a Poisson distribution of cells, under such conditions 6% of chambers should have contained multiple cells or cell aggregates (1). PCR was carried out on a conventional flat-block

thermocycler. Amplification was monitored using 5' nuclease probes to generate a fluorescent signal detected with a modified microarray scanner.

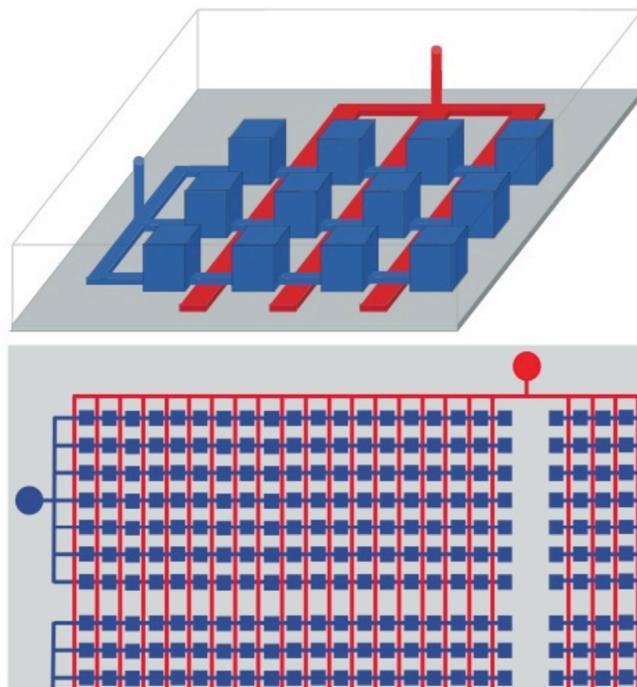


Figure 4.1. Microfluidic Digital PCR Chip Architecture. **Top**, schematic diagram showing many parallel chambers (blue) connected by channels to a single input. When pressure is applied to the control channel network (red), the membranes between the red and blue channels are deflected upward, creating micromechanical valves. When the valves are closed, the continuous blue network is partitioned into independent PCR reactors. **Bottom**, schematic showing how a single valve connection can be used to partition thousands of chambers. In the device used, each experimental sample could be partitioned into 1,176 chambers, and each device contained 12 such sample panels.

Multiplex PCR amplifications from single cells or cell aggregates were successfully performed using diluted gut contents that had been partitioned on-chip (Figure 4.2, left).

We found global averages of $1.2 \pm 0.8 \times 10^8$ total bacterial 16S rRNA gene encoding units and $1.5 \pm 1.0 \times 10^6$ total Clone H Group FTHFS gene encoding units per *Zootermopsis nevadensis* termite (2). This suggests that, in *Z. nevadensis*, these particular FTHFS genes are carried by a minority population representing ca. 1% of gut symbionts.

The observed variability of these measurements was not surprising as the *Z. nevadensis*

specimens examined were collected from different colonies and locations, and had been maintained in captivity for varying periods of time.

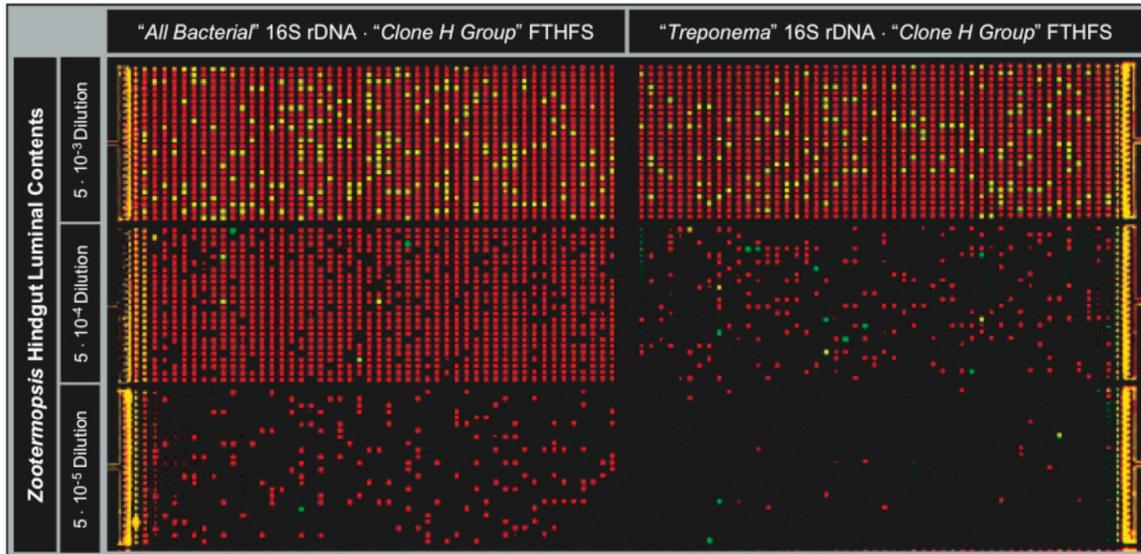


Figure 4.2. Multiplex microfluidic digital PCR of single cells in environmental samples. Six panels from a representative experiment showing microfluidic digital PCR on hindgut contents harvested from a single *Z. nevadensis* individual. Left, multiplex PCR using “all bacterial” 16S rRNA gene (red fluorescence) and “Clone H Group” (42) FTHFS gene (green fluorescence) primers and probes. Reaction chambers that contained both genes in 1/500,000 dilutions from this and other on-chip experiments were sampled and the PCR products were analyzed (see Figure 4.5). Right, the same, except that 16S rRNA primers specifically targeted members of the “termite cluster” (26) of the spirochetal genus *Treponema*.

Amplification products were retrieved from reaction chambers via syringe needle and were reamplified, cloned, sequenced, and analysed using standard methods. Twenty randomly selected chambers that had amplified only a 16S rRNA gene (and not FTHFS) yielded a diversity of *Endomicrobia*, *Firmicutes*, *Bacteroidetes*, *Proteobacteria* and *Spirocheates* ribotypes that was expected based upon prior 16S rRNA gene clone libraries (36) (Figure 4.3 & 4.4). Two thirds of chambers positive for FTHFS genes did not amplify 16S rRNA genes when either all bacterial or termite treponeme-specific rRNA gene primers were employed. This amplification success rate is comparable to that

observed when purified, single molecule templates were used and remains a target of refinement and improvement in the future.

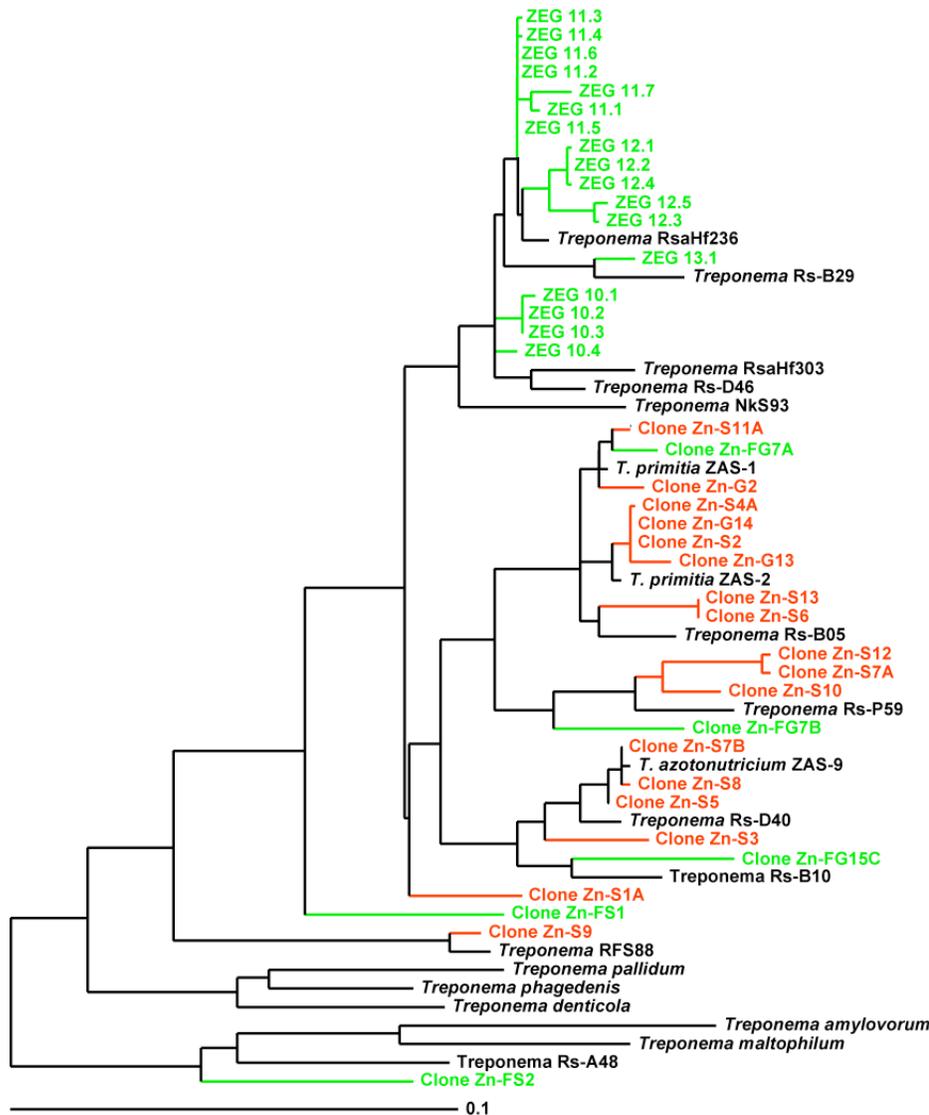


Figure 4.3. Phylogenetic Analysis of *Treponemal* 16S rRNA sequences retrieved from microfluidic chips. Sequences recovered from chambers in which only 16S rRNA genes were amplified are marked in red; a Zn-G moniker denotes that “all bacterial” primers were employed, Zn-S spirochete-specific primers. Sequences corecovered with FTHFS sequences are marked in green; those that fell outside the ZEG cluster were assigned a Zn-FG or Zn-FS moniker according to the 16S rRNA primer set employed. ZEG 11.5-11.7 and 12.5 were identified in experiments using spirochete-specific rRNA primers. Tree calculated using Phylip distance methods and 630 unambiguous, aligned residues. Scale bar represents 0.1 changes per alignment position.

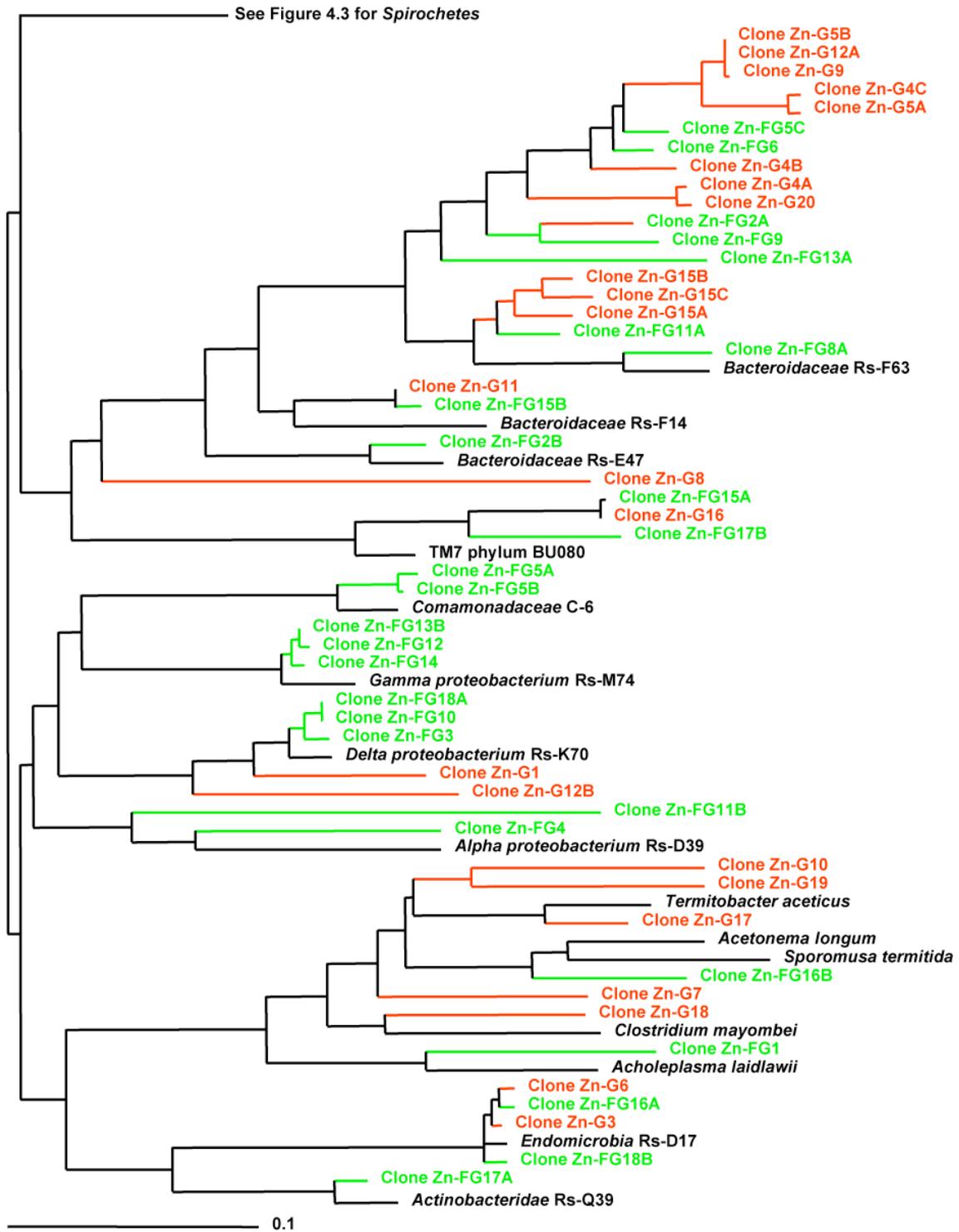


Figure 4.4. Phylogenetic Analysis of 16S rRNA sequences retrieved from microfluidic chips and close relatives. Sequence naming and color coding as described in Figure 4.3. Tree was calculated using Phylip distance methods and 630 unambiguous, unaligned residues. Scale bar represents 0.1 changes per alignment position.

PCR products were retrieved and analyzed from 28 reaction chambers that coamplified both FTHFS and 16S rRNA genes. In ten of those reactions, sequence analyses revealed that the FTHFS gene had coamplified with a clade of closely related 16S rRNA gene sequences affiliating with within the “termite spirochete cluster” (26) of the genus *Treponema*. Members of this novel clade were never observed in chambers that lacked FTHFS gene amplification. An additional three chambers contained a single FTHFS type and multiple 16S rRNA genotypes, one of which in each affiliated with the above mentioned group (ZEG 11.4, 10.2, 10.1). These latter reactions also contained: two additional other *Spirochaetes* (Zn-FG7A&B in Figure 4.3) in one chamber, a single γ -*Proteobacterium* sequence (Zn-FG12) in the second, and a *Firmicutes* sequence (ZN-FG1) in the third. The remaining fifteen chambers analyzed (that coamplified FTHFS and rRNA genes) yielded 16S rDNA sequences in proportions that corresponded well with the ribotype diversity encountered in the general non-FTHFS encoding population. Based on this evidence, we conclude that the unique cluster of termite gut treponeme rRNA gene sequences that were repeatedly identified in FTHFS-containing chambers represent the ribotype of the FTHFS-encoding cells. We attribute the instances of FTHFS colocalization with other rRNA gene sequences to cell-cell aggregations. The latter is not to be unexpected in a complex, wood-particle-filled and sticky environment such as the termite hindgut (9, 21). Such aggregations appear to be largely random, though there may be a slight enrichment of proteobacterial sequences in comparison to the general population (Figure 4.4). Our results show that FTHFS sequences present in ca. 1% of all bacterial cells were, in 13 out of 28 trials, found in association with a 16S rRNA sequence type not identified in 20 random samplings of the all bacterial population

(16S rRNA only chambers) at large. The probability of a 16S rRNA gene sequence type that is present at less than 5% of the population randomly colocalizing with FTHFS in 13 out of 28 trials is low, on the order of 10^{-10} (3).

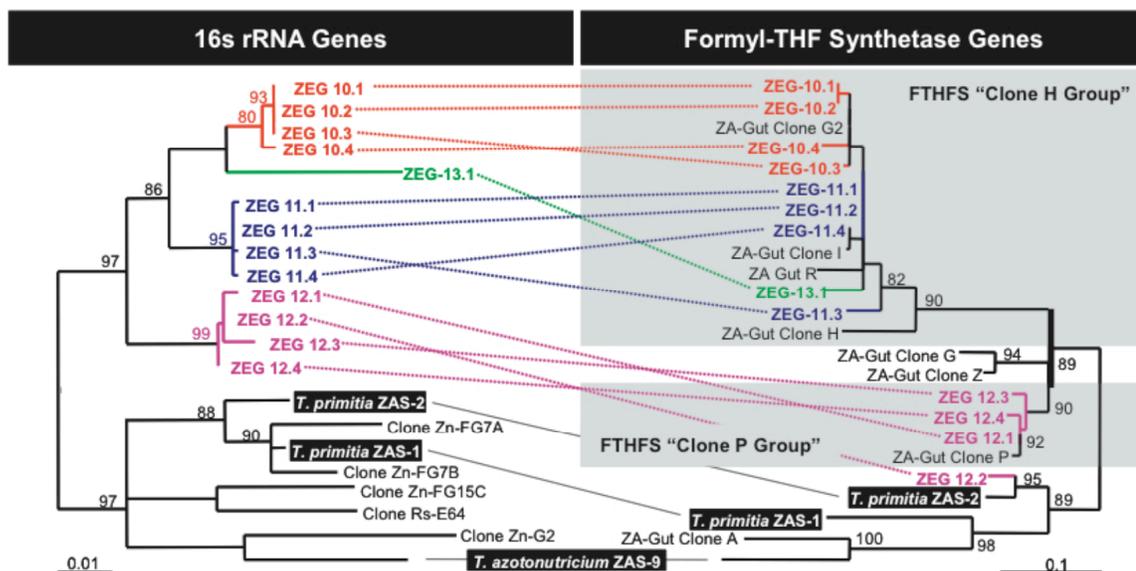


Figure 4.5. “Clone H” and “Clone P Group” FTHFS genes are encoded by not-yet-cultivated termite gut treponemes. **Left**, phylogenetic tree of 16S rRNA genes cloned from cultivated strain isolates (orange) and from hindgut community microbiota. **Right**, phylogenetic tree of FTHFS genes from the termite hindgut. Dotted lines connect genes believed to originate from the same genome. Incongruent gene phylogenies implicate acquisition of FTHFS genes via lateral gene-transfer and can be observed in both isolated species (*T. primitia* ZAS-1) and proposed “environmental genomovars” (ZEG 12.2). Scale bars represent substitutions per alignment position. The trees were constructed using TreePuzzle (43); 630 (16S rDNA) and 249 (FTHFS) nucleotide positions were used.

Refined phylogenetic analysis of 16S rRNA gene sequences that were repeatedly isolated from FTHFS-containing reaction chambers revealed that all such 16S rRNA gene sequences affiliated within the termite gut treponeme cluster of *Spirochaetes*. These 16S rRNA genes group into four distinct ribotype clusters (Figure 4.5). These four sequence types share >99% sequence identity within-group and between-group identities of 95%–99%. We propose the term “environmental genomovar” (genome variant) to describe

not-yet-cultivated organisms shown to encode two or more known genes of interest. Here, we use the epithets ZEG-10 (for *Zootermopsis* Environmental Genomovar) through ZEG-13 to describe the four 16S ribotypes identified (9 termite gut treponemes have been isolated and assigned the strain epithets ZAS-1 (for *Zootermopsis* Acetogenic Spirochete) through ZAS-9 (22, 25)). Genomovars ZEG-10, 11, and 13 encode Clone H Group FTHFS sequences, while one ZEG-12 genomovar encodes a Clone P Group FTHFS sequence.

To build additional support for a spirochetal origin of Clone H Group FTHFS genotypes, we designed and employed a termite treponeme-specific 16S rRNA gene primer set and gene probe, with the aim of reducing nonspirochetal background (Figure 4.2, right). The frequency with which Clone H Group FTHFS genes were recovered increased from 1 in 175 cells of the general bacterial population, to 1 in 16 treponemal cells (several termite gut treponemes are already known or suspected to encode FTHFS genotypic variants that would not amplify with the Clone H group FTHFS primer and probe set (42), see Figure 4.3). Similar to the amplification success rates observed in experiments using the “all bacterial” 16S rRNA gene primers (Figure 4.2, left) and those using the Clone H primers against purified single molecule templates ca. 1/3 of FTHFS-positive reaction chambers also amplified detectable levels of 16S rRNA gene. Treponemal cells were deduced to comprise 10%–12% of the bacterial community of *Z. nevadensis* (comparing amplification frequencies in the left and right panels of Figure 4.2). These results are in good agreement with the results of a traditional 16S rRNA clone inventory from *Z.*

nevadensis, which suggested that 15% of clones corresponded to treponemes (unpublished data).

In summary: specific not-yet-cultivated *Treponema* species encode variants of a key gene underlying the dominant bacterial metabolism known to impact the energy needs of their termite hosts. The microfluidic, multiplex digital PCR approach taken here can be extended to expand our understanding of the genetic capacities of not-yet-cultivated species, and to collect and collate genetic information in a manner that builds conceptual genomovars that directly represent the organisms catalyzing important activities in various environments of global relevance.

Materials and Methods

Termite Maintenance

Zootermopsis nevadensis specimens were collected from fallen Jeffrey (*Pinus jeffreyi*) and Ponderosa Pine (*Pinus ponderosa*) at Mt. Pinos in the Los Padres National Forest and at the Chilao Campground in the Angeles National Forest. Colonies were maintained in the laboratory on Ponderosa at 23 °C and at a constant humidity of 96%, achieved via incubation over saturated solutions of KH_2PO_4 within 10-gallon aquaria (55).

PCR on Microfluidic Chips

Microfluidic devices were purchased from Fluidigm Corporation (www.fluidigm.com/didIFC.htm). On-chip multiplex PCR reactions contained 0.05 units μL^{-1} iTaq DNA polymerase (BioRad), iTaq PCR buffer, 200 μM each dNTP, 1.5 mM

MgCl₂, and 0.1% Tween-20. In almost all PCR reactions, primers and probes were used at 400 nM; all bacterial 16S primers were used at 600 nM in on-chip reactions. Primers and probes were purchased from Integrated DNA Technologies and had the following sequences: FTHFS forward, 5'-GAATCACGCGAAGACTGGTTC-3'; reverse, 5'-TTGAGTTACAACCGTGTGCGAT-3'; probe, 5'-CAAGGCGCAATGGCAGCCCT-3' (FAM and Black Hole Quencher 1 labelled), all bacterial rRNA 357 forward 5'-CTCCTACGGGAGGCAGCAG-3' (modified from (32)), 1492 reverse 5'-TACGGYTACCTTGTTACGACTT-3' (modified from (20)); 1389 reverse probe 5'-CTTGTACACACCGCCCGTC-3' (described in (46), labelled with CY5 and Iowa Black quencher). Termite gut spirochete-specific SSU rRNA amplification was achieved using the 1389R probe and 357F primer with a spirochete-specific 1409R primer (sequence 5'-GGGTACCTCCAACCTCGGATGGTG-3').

Zootermopsis hindguts were extracted from worker larvae, suspended in sterile TE (10 mM Tris-HCl, 1 mM EDTA, pH 8), and disrupted via repeated aspiration using a 1 mL Eppendorf pipettor. Suspensions were allowed to stand briefly to sediment large particles, then diluted to working concentrations in TE and mixed 1 to 10 with the PCR reaction mixture (above) for immediate loading onto microfluidic chips.

Chips were loaded using air pressure. 200 µL gel-loading tips were filled with sample and connected to air lines at 12-15 PSI (pounds per square inch) pressure. Control channels were loaded with 35% PEG (polyethylene glycol) 3350 (ca. 50 µL, in gross excess). The 12 sample channels were loaded with 15 µL of PCR reaction (again, in

excess). After loading, sample lines were allowed to reequilibrate to atmospheric pressure. Control valves were closed by the application of 25 PSI air pressure to control lines.

Cycling was carried out on flat-block thermocyclers (MJ Research). Microscope immersion oil (Cargille, Type FF) was applied between the chip and thermocycling block, and the cycling program was as follows: 98 °C 30 s, 97 °C 30 s, 95 °C 2 min, [56 °C 30 s, 58 °C 30 s, 60 °C 30 s, 98 °C 15 s] x 40 cycles, 60 °C for 10 min.

Reaction results were evaluated by fluorescent signal strength as measured using an ArrayWoRx scanner (Applied Precision). Spot intensities were located and retrieved using either ArrayWoRx software or the ScanAlyze program (version 2.50, Michael Eisen). Cutoff values for positive amplification were calculated for each sample panel independently. Chambers in the bottom 25% of the intensity range were assumed to contain no amplification, and positive chambers were defined as chambers whose spot intensity was more than 10 standard deviations above the mean of points in this range for the FTHFS probe. The 16S rRNA gene probe gave a more variable signal, so the threshold for this channel was set at 5 standard deviations above the mean.

Sample Retrieval and Analysis

Single-cell PCR products were retrieved from amplification-positive chambers. Chips were peeled from the backing slide, and pressure was removed from control channels (most valves remained fused despite relief of external pressure). Target chambers were

located using a dissecting microscope, and the tip of a 30 gauge syringe needle was inserted into each chamber through the bottom surface of the chip. Needles were then swirled briefly in 10 μ L of TE to desorb the PCR product.

Retrieval efficiency was checked by real time PCR using the same primers as above in BioRad SYBR Green PCR Master Mix. Reactions were carried out using the Chromo4 system (BioRad), and temperature program 95 °C 3 min, (95 °C 15 s, 60 °C 1 min30 s) x 40 cycles. FTHFS concentration standards contained a 1.2 kb section of “ZA-gut Clone U” type FTHFS gene sequence (42). Termite community DNA was used as a standard for all bacterial 16S rRNA gene PCR, and *T. primitia* ZAS-2 genomic DNA for spirochete-specific reactions. Samples that contained 104 or more gene copies were deemed successful retrievals.

Retrieved PCR products were amplified for cloning and/or sequencing using EXPAND high fidelity polymerase (Roche), Fail-Safe PCR PreMix D (Epicentre), and primers and cycling conditions as above. PCR products were purified using the Qiagen PCR purification kit, and sequenced using the FTHFS PCR primers and 16S rRNA gene internal primers 1100R and 533F (5'-AGGGTTGCGCTCGTTG-3' and 5'-GTGCCAGCMGCCGCGGTAA-3', respectively; modified from ref. (20)). Some samples contained a mixture of 16S rRNA sequences. These sequences were cloned using the TOPO TA cloning kit for sequencing (Invitrogen). Eight colonies from each cloning reaction were picked and used as template for high-fidelity PCR as described above. Ten μ L of each reaction was digested at 37 °C for 2 hr with 3 units HinPII from

New England Biolabs and analyzed by agarose gel electrophoresis. A representative of each RFLP (restriction fragment length polymorphism) type was prepared for sequencing as described above, using recommended T3 and T7 primers. All sequencing reactions were carried out by the California Institute of Technology DNA Sequencing Facility.

Sequences were assembled and edited using the Lasergene software package (DNASTAR). Phylogenetic analysis and alignment of 16S rRNA gene sequences was carried out using the ARB software package (30). FTHFS sequences were translated into protein, and aligned using GenomatixSuite software (Genomatix). Nucleic acid sequences were aligned according to the protein alignment. All 16S rRNA gene sequences were screened using chimera identification programs Bellerophon (16) and Pintail (6). Three chimeric sequences were identified and eliminated from further analysis.

Real-Time PCR Standards and DNA Template Preparation

Plasmid templates were purified from *E. coli* strains from the library of Salmassi and Leadbetter using the Qiaprep Spin Miniprep Kit (Qiagen). Termite gut community DNA was extracted from the pooled gut contents of five termites. Guts were disrupted using the protocol laid out in Salmassi and Leadbetter (42), with the substitution of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8) for the phosphate buffer described in that paper. After bead-beating and phenol extraction, DNA was purified from the aqueous phase using the Qiagen DNeasy Tissue kit, with the protocol described for extraction of DNA from crude lysates (DNeasy Tissue Handbook, July 2003 version). Template concentrations were

measured using the Hoefer DyNAQuant 200 fluorometer and DNA quantification system (Amersham Pharmacia Biotech) using reagents and procedures directed in the manual (DQ200-IM, Rev C1, 5-98). Termite gut cell suspensions were prepared as described in the main body of the paper.

Chapter Four Appendix

- 1. Design and Validation of Primers and Probes for Microfluidic Digital PCR**
- 2. Table 4.1. Sequences used in phylogenetic analysis**

Design and Validation of Primers and Probes for Microfluidic Digital PCR

Amplification of Formyl-tetrahydrofolate Synthetase Genes from Termite Gut Acetogens

Primers and probes were designed to specifically amplify FTHFS genes from “*Clone H Group*” acetogens, which comprised 43% of the *Zootermopsis* FTHFS clones inventoried by Salmassi and Leadbetter (42). These primers are distinct from those previously employed to amplify FTHFS genes from pure cultures and environmental samples (23, 24, 28, 39). The newly designed primers and probes were tested for on-chip amplification and specificity using purified plasmid DNA (Figure 4.6). The copy number as deduced from the number of positive chambers detected (adjusted based on a Poisson distribution of template) fell within 11%–110% of the copy number calculated based on the concentration of double-stranded DNA in the template plasmid preparation. Freeze-thaw and template age may be one variable influencing observed amplification efficiencies; it has been recently reported that amplification efficiency can approach 99% (53). A small amount of amplification was detected from closely related clones (Figure 4.6i), with a signal to background ratio less than half of that detected in positive clones. This low level of amplification from closely related species was also apparent in later experiments, as several FTHFS clones mapping to the “*Clone P Group*” were retrieved from on-chip reactions (see main text). No fluorescent signal was detected from amplification of distant relatives (clostridial and nonacetogenic FTHFS types, Figure 4.6k). FTHFS copies were also detectable within DNA extracted from whole termite guts and from termite gut cell suspensions.

FTHFS simplex experiments used DyNAzyme II polymerase (Finnzymes) at 0.2 units per μl and 1x TaqMan Universal PCR Master Mix (Applied Biosystems) for real-time PCR. Due to the high concentration of detergent in the enzyme storage buffer, only 0.05% Tween-20 (Sigma) was added. All other experiments described used the iTaq system described in the main body of the paper, as this enzyme was found to perform well on the chip at lower concentrations, and had hot-start capabilities to ensure that the enzyme was inactive during the chip loading process.

Design of “All-bacterial” 16S rRNA Primers and Probes

Primers and probes for amplification of bacterial 16S rRNA were also employed. Bacterial 16S rRNA genes detected in on-chip amplification from termite gut community DNA preparation amounted to 1.4×10^5 copies per ng (1 copy every 6.7 MB DNA), which was 5.9-fold higher than the copy number deduced by real-time PCR using *Treponema primitia* ZAS-2 genomic DNA as a standard. Background amplification has been reported in a number of general bacterial 16S real-time assays, and is commonly attributed to DNA fragments present in commercial enzyme preparations (11). In on-chip experiments with the final primer set, negative controls never exceeded 1.2% positive chambers (1.9 copies per μl).

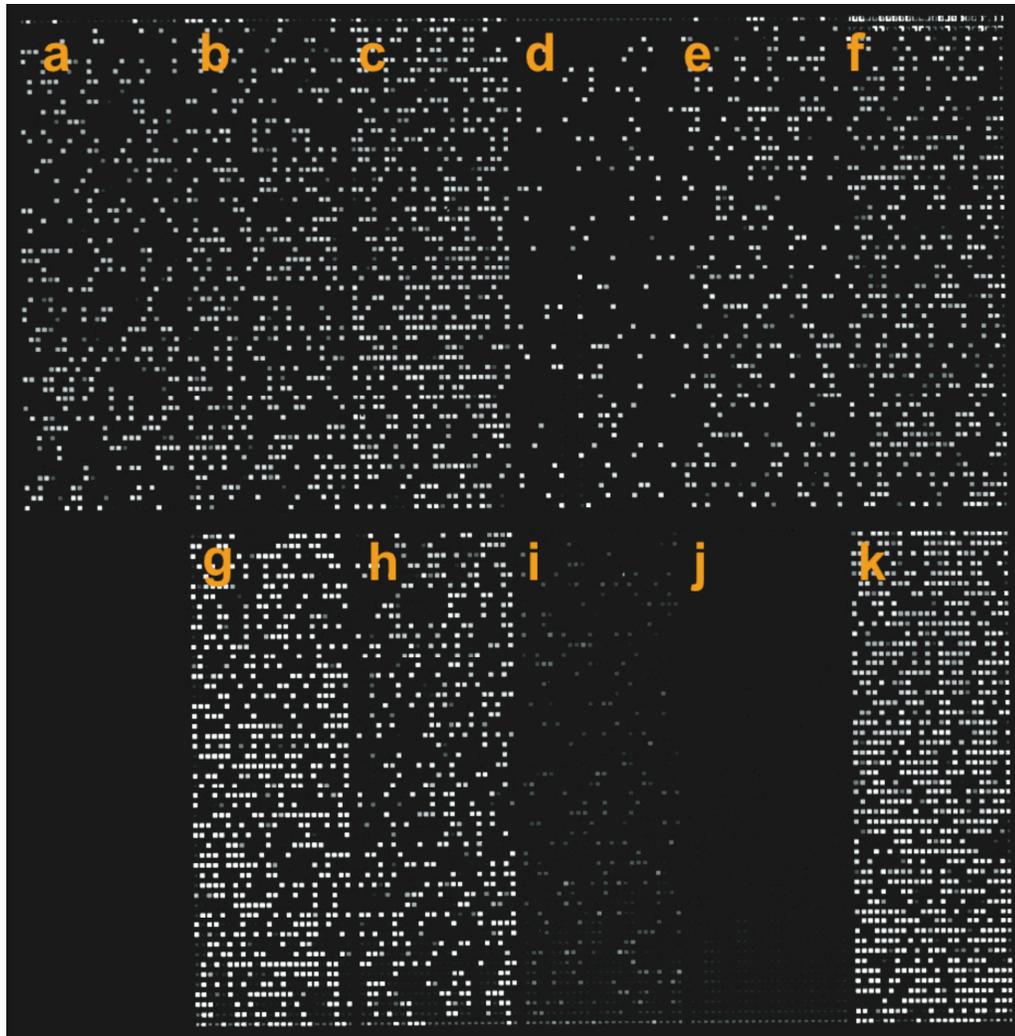


Figure 4.6. FTHFS primer specificity and demonstration of single copy sensitivity. A single microfluidic chip on which the FTHFS primers and probe were tested against purified plasmid templates. Panels a through h and k each show amplification from one of nine different Clone H Group FTHFS genotypes. Panel i contains six pooled non-H type FTHFS genotypes that cluster within the termite *Treponeme* FTHFS cluster. Panel j contains four pooled FTHFS genotypes that do not cluster phylogenetically with termite *treponemes*. All clones (and each clone within pooled templates) were added at DNA concentrations equivalent to ~200 copies per μl . Specific clone types and observed copy number are as follows: a Clone E2, 57 cp/ μl ; b Clone F2, 70 cp/ μl ; c Clone G2, 97 cp/ μl ; d Clone H, 22 cp/ μl ; e Clone I, 51 cp/ μl ; f Clone L, 78 cp/ μl ; g Clone U, 102 cp/ μl ; h.) Clone R, 72 cp/ μl ; I.) Clones G, P, Z, C, N, and A, 11 cp/ μl ; j Clones F, T, Y, E, 0 copies detected; and k Clone M, 145 cp/ μl . To allow cross-comparison of sample panels, a single threshold for positive amplification was calculated for the entire chip; this value was set to 5 standard deviations above the mean of chambers in the lowest 25% of the intensity range.

Specific Detection of Termite Cluster Treponemes Through Use of a Spirochete-specific Reverse Primer.

A 16S rRNA gene reverse primer was designed that matched 41 out of 60 termite gut spirochetes with sequence data covering the primer site. Of the known 16S rRNA sequences that did not match the primer, three were associated with the “*termite gut treponeme*” ribotype cluster (26). The remaining mismatches were with sequences affiliated with “*treponeme subgroup I*” (38), which represents less than 1% of spirochetal 16S clones amplified from *Z. nevadensis* using conventional methods and other spirochete-specific primers (*unpublished data*, primers from Lilburn, Schmidt, and Breznak (26)). Our new primers were tested for specificity and efficiency in simplex and multiplex reactions with FTHFS primers/probes using conventional and real-time PCR methods. In on-chip PCR reactions using purified PCR products as template they detected 11% of the expected copy number.

Table 4.1. Sequences Used for Phylogenetic Analysis

Source/Sequence Type	Designation	Gene	Accession	Reference
<i>T. primitia</i> ZAS-1	ZAS-1	16S	AF093251	(22)
<i>T. primitia</i> ZAS-2	ZAS-2	16S	AF093252	(22)
<i>T. azotonutricium</i> ZAS-9	ZAS-9	16S	AF320287	(25)
<i>T. primitia</i> ZAS-1	ZAS-1a	FTHFS	AY162313	(42)
<i>T. primitia</i> ZAS-2	ZAS-2	FTHFS	AY162315	(42)
<i>T. azotonutricium</i> ZAS-9	ZAS-9	FTHFS	AY162316	(42)
<i>Z. angusticollis</i> Gut Clone	A	FTHFS	AY162294	(42)
<i>Z. angusticollis</i> Gut Clone	C	FTHFS	AY162295	(42)
<i>Z. angusticollis</i> Gut Clone	E	FTHFS	AY162296	(42)
<i>Z. angusticollis</i> Gut Clone	E2	FTHFS	AY162297	(42)
<i>Z. angusticollis</i> Gut Clone	F	FTHFS	AY162298	(42)
<i>Z. angusticollis</i> Gut Clone	F2	FTHFS	AY162299	(42)
<i>Z. angusticollis</i> Gut Clone	G	FTHFS	AY162300	(42)
<i>Z. angusticollis</i> Gut Clone	G2	FTHFS	AY162301	(42)
<i>Z. angusticollis</i> Gut Clone	H	FTHFS	AY162302	(42)
<i>Z. angusticollis</i> Gut Clone	I	FTHFS	AY162303	(42)
<i>Z. angusticollis</i> Gut Clone	L	FTHFS	AY162304	(42)
<i>Z. angusticollis</i> Gut Clone	M	FTHFS	AY162305	(42)
<i>Z. angusticollis</i> Gut Clone	N	FTHFS	AY162306	(42)
<i>Z. angusticollis</i> Gut Clone	P	FTHFS	AY162307	(42)
<i>Z. angusticollis</i> Gut Clone	R	FTHFS	AY162308	(42)
<i>Z. angusticollis</i> Gut Clone	T	FTHFS	AY162309	(42)
<i>Z. angusticollis</i> Gut Clone	U	FTHFS	AY162310	(42)
<i>Z. angusticollis</i> Gut Clone	Y	FTHFS	AY162311	(42)
<i>Z. angusticollis</i> Gut Clone	Z	FTHFS	AY162312	(42)
<i>Z. nevadensis</i> Genomovar	ZEG 10.1	FTHFS	DQ420342	This study
<i>Z. nevadensis</i> Genomovar	ZEG 10.2	FTHFS	DQ420343	This study
<i>Z. nevadensis</i> Genomovar	ZEG 10.3	FTHFS	DQ420344	This study
<i>Z. nevadensis</i> Genomovar	ZEG 10.4	FTHFS	DQ420345	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.1	FTHFS	DQ420346	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.2	FTHFS	DQ420347	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.3	FTHFS	DQ420348	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.4	FTHFS	DQ420349	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.5	FTHFS	DQ420350	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.6	FTHFS	DQ420351	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.7	FTHFS	DQ420352	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.1	FTHFS	DQ420353	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.2	FTHFS	DQ420354	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.3	FTHFS	DQ420355	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.4	FTHFS	DQ420356	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.5	FTHFS	DQ420357	This study
<i>Z. nevadensis</i> Genomovar	ZEG 13.1	FTHFS	DQ420358	This study
<i>Z. nevadensis</i> Genomovar	ZEG 10.1	16S	DQ420325	This study
<i>Z. nevadensis</i> Genomovar	ZEG 10.2	16S	DQ420326	This study
<i>Z. nevadensis</i> Genomovar	ZEG 10.3	16S	DQ420327	This study
<i>Z. nevadensis</i> Genomovar	ZEG 10.4	16S	DQ420328	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.1	16S	DQ420329	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.2	16S	DQ420330	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.3	16S	DQ420331	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.4	16S	DQ420332	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.5	16S	DQ420333	This study
<i>Z. nevadensis</i> Genomovar	ZEG 11.6	16S	DQ420334	This study

Source/Sequence Type	Designation	Gene	Accession	Reference
<i>Z. nevadensis</i> Genomovar	ZEG 11.7	16S	DQ420335	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.1	16S	DQ420336	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.2	16S	DQ420337	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.3	16S	DQ420338	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.4	16S	DQ420339	This study
<i>Z. nevadensis</i> Genomovar	ZEG 12.5	16S	DQ420340	This study
<i>Z. nevadensis</i> Genomovar	ZEG 13.1	16S	DQ420341	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG1	16S	DQ420259	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG2A	16S	DQ420263	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG2B	16S	DQ420264	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG3	16S	DQ420275	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG4	16S	DQ420273	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG5A	16S	DQ420269	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG5C	16S	DQ420270	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG6	16S	DQ420271	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG7A	16S	DQ420266	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG7B	16S	DQ420262	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG8A	16S	DQ420284	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG9	16S	DQ420317	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG10	16S	DQ420319	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG11A	16S	DQ420272	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG11B	16S	DQ420258	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG12	16S	DQ420261	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG13A	16S	DQ420286	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG13B	16S	DQ420287	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG14	16S	DQ420257	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG15A	16S	DQ420277	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG15B	16S	DQ420278	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG15C	16S	DQ420279	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG16A	16S	DQ420280	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG16B	16S	DQ420281	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG17A	16S	DQ420282	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG17B	16S	DQ420283	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG18A	16S	DQ420255	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FG18B	16S	DQ420276	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G1	16S	DQ420256	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G2	16S	DQ420254	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G3	16S	DQ420265	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G4A	16S	DQ420310	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G4B	16S	DQ420311	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G4C	16S	DQ420312	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G5A	16S	DQ420313	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G5B	16S	DQ420314	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G6	16S	DQ420260	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G7	16S	DQ420268	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G8	16S	DQ420267	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G9	16S	DQ420315	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G10	16S	DQ420285	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G11	16S	DQ420274	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G12A	16S	DQ420316	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G12B	16S	DQ420324	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G13	16S	DQ420298	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G14	16S	DQ420299	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G15A	16S	DQ420320	This study

Source/Sequence Type	Designation	Gene	Accession	Reference
<i>Z. nevadensis</i> Gut Clone	Zn-G15B	16S	DQ420321	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G15C	16S	DQ420322	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G16	16S	DQ420300	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G17	16S	DQ420301	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G18	16S	DQ420302	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G19	16S	DQ420303	This study
<i>Z. nevadensis</i> Gut Clone	Zn-G20	16S	DQ420323	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FS1	16S	DQ420288	This study
<i>Z. nevadensis</i> Gut Clone	Zn-FS2	16S	DQ420289	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S1A	16S	DQ420307	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S2	16S	DQ420295	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S3	16S	DQ420308	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S4A	16S	DQ420309	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S5	16S	DQ420296	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S6	16S	DQ420297	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S7A	16S	DQ420304	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S7B	16S	DQ420305	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S8	16S	DQ420290	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S9	16S	DQ420291	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S10	16S	DQ420292	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S11A	16S	DQ420306	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S12	16S	DQ420293	This study
<i>Z. nevadensis</i> Gut Clone	Zn-S13	16S	DQ420294	This study
<i>Acetoneuma longum</i>	APO-1	16S	M61919	(19)
<i>Acholeplasma laidlawii</i>	JA1	16S	M23932	(54)
<i>Clostridium mayombeii</i>	SFC-5	16S	M62421	(18)
Comamonadaceae Clone	C-6	16S	AF523013	(29)
<i>N. koshunensis</i> symbiont	Nk-S93	16S	AB084970	(34)
<i>R. flavipes</i> Gut Clone	RFS88	16S	AF068344	(26)
<i>R. santonensis</i> Gut Clone	RsaHf236	16S	AY571482	(58)
<i>R. santonensis</i> Gut Clone	RsaHf303	16S	AY571478	(58)
<i>R. speratus</i> Gut Clone	Rs-B05	16S	AB088896	(15)
<i>R. speratus</i> Gut Clone	Rs-B10	16S	AB088880	(15)
<i>R. speratus</i> Gut Clone	Rs-B29	16S	AB088891	(15)
<i>R. speratus</i> Gut Clone	Rs-D17	16S	AB089048	(15)
<i>R. speratus</i> Gut Clone	Rs-D39	16S	AB089089	(15)
<i>R. speratus</i> Gut Clone	Rs-D40	16S	AB088874	(15)
<i>R. speratus</i> Gut Clone	Rs-D46	16S	AB088865	(15)
<i>R. speratus</i> Gut Clone	Rs-E47	16S	AB088921	(15)
<i>R. speratus</i> Gut Clone	Rs-F14	16S	AB088939	(15)
<i>R. speratus</i> Gut Clone	Rs-F63	16S	AB088934	(15)
<i>R. speratus</i> Gut Clone	Rs-E64	16S	AB088888	(15)
<i>R. speratus</i> Gut Clone	Rs-K70	16S	AB089106	(15)
<i>R. speratus</i> Gut Clone	Rs-M74	16S	AB089115	(15)
<i>R. speratus</i> Gut Clone	Rs-P59	16S	AB088914	(15)
<i>R. speratus</i> Gut Clone	Rs-Q39	16S	AB089075	(15)
<i>Sporomusa termitida</i>	JSN-2	16S	M61920	(19)
<i>Termitobacter acetivorus</i>	SYR	16S	Z49863	(13)
TM7 phylum Env. Clone	BU080	16S	AF385568	
<i>Treponema amylovorum</i>	HA2P	16S	Y09959	(56)
<i>Treponema denticola</i>	II:11:33520	16S	M71236	(38)
<i>Treponema maltophilum</i>	patient BR	16S	X87140	(57)
<i>Treponema pallidum</i>	Nichols	16S	M88726	(38)
<i>Treponema phagedenis</i>	K5	16S	M57739	(38)

References

1. Assuming a Poisson distribution, if 67% percent of chambers are empty then the expected number of cells per chamber is $-\ln 0.67$ or 0.40. The probability that a chamber contains more than one cell is $1 - 0.67 - ((e^{-0.40}) * (0.40^1)) / (1!)$ or 6.2%
2. Value \pm one standard deviation, 13 termites served as source of cells for n=32 sample panels. All sample panels that met our conditions for single cell separation and contained at least one FTHFS-positive chamber were used in calculation of gut bacterial loads.
3. The binomial distribution function was used to calculate the probability that, in 13 out of 28 trials, a sequence that is present in 5 percent (0 out of 20 16S-only chambers) of chambers would randomly co-localize with Clone H Group FTHFS sequences.
4. **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 constructed from the same sample. *Appl Environ Microbiol* **71**:8966-8969.
5. **Amann, R. I., W. Ludwig, and K. H. Schleifer.** 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* **59**:143-169.
6. **Ashelford, K. E., N. A. Chuzhanova, J. C. Fry, A. J. Jones, and A. J. Weightman.** 2005. At least 1 in 20 16S rRNA sequence records currently held in

- public repositories is estimated to contain substantial anomalies. *Appl Environ Microbiol* **71**:7724-7736.
7. **Béjà, O., L. Aravind, E. V. Koonin, M. T. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich, and E. F. DeLong.** 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**:1902-1906.
 8. **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.
 9. **Breznak, J. A., and H. S. Pankratz.** 1977. In situ morphology of the gut microbiota of wood-eating termites [*Reticulitermes flavipes* (Kollar) and *Coptotermes formosanus* Shiraki]. *Appl Environ Microbiol* **33**:406-426.
 10. **Breznak, J. A., and J. M. Switzer.** 1986. Acetate synthesis from H₂ plus CO₂ by termite gut microbes. *Appl Environ Microbiol* **52**:623-630.
 11. **Corless, C. E., M. Guiver, R. Borrow, V. Edwards-Jones, E. B. Kaczmariski, and A. J. Fox.** 2000. Contamination and sensitivity issues with a real-time universal 16S rRNA PCR. *J Clin Microbiol* **38**:1747-1752.
 12. **Gray, N. D., and I. M. Head.** 2001. Linking genetic identity and function in communities of uncultured bacteria. *Environ Microbiol* **3**:481-492.
 13. **GrechMora, I., M. L. Fardeau, B. K. C. Patel, B. Ollivier, A. Rimbault, G. Prensier, J. L. Garcia, and E. GarnierSillam.** 1996. Isolation and characterization of *Sporobacter termitidis* gen nov sp nov, from the digestive tract

- of the wood-feeding termite *Nasutitermes lujae*. *Int J Syst Evol Microbiol* **46**:512-518.
14. **Hong, J. W., and S. R. Quake.** 2003. Integrated nanoliter systems. *Nat Biotechnol* **21**:1179-1183.
 15. **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.
 16. **Huber, T., G. Faulkner, and P. Hugenholtz.** 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**:2317-2319.
 17. **Hugenholtz, P., B. M. Goebel, and N. R. Pace.** 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* **180**:4765-4774.
 18. **Kane, M. D., A. Brauman, and J. A. Breznak.** 1991. *Clostridium mayombe* sp. nov., an H₂/CO₂ acetogenic bacterium from the gut of the African soil-feeding termite, *Cubitermes speciosus*. *Arch Microbiol* **156**:99-104.
 19. **Kane, M. D., and J. A. Breznak.** 1991. *Acetonema longum* gen. nov. sp. nov., an H₂/CO₂ acetogenic bacterium from the termite, *Pterotermes occidentis*. *Arch Microbiol* **156**:91-98.
 20. **Lane, D. J.** 1991. 16S/23S rRNA sequencing, p. 115-175. *In* M. G. E. Stackebrandt (ed.), *Nucleic Acid Techniques in Bacterial Systematics*. Wiley, New York.

21. **Leadbetter, J. R., and J. A. Breznak.** 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**:3620-3631.
22. **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.
23. **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.
24. **Leaphart, A. B., and C. R. Lovell.** 2001. Recovery and analysis of formyltetrahydrofolate synthetase gene sequences from natural populations of acetogenic bacteria. *Appl Environ Microbiol* **67**:1392-1395.
25. **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.
26. **Lilburn, T. G., T. M. Schmidt, and J. A. Breznak.** 1999. Phylogenetic diversity of termite gut spirochaetes. *Environ Microbiol* **1**:331-345.
27. **Ljungdahl, L. G.** 1986. The autotrophic pathway of acetate synthesis in acetogenic bacteria. *Ann Rev Microbiol* **40**:415-450.
28. **Lovell, C. R., and A. B. Leaphart.** 2005. Community-level analysis: key genes of CO₂-reductive acetogenesis. *Methods Enzymol* **397**:454-469.

29. **Loy, A., W. Beisker, and H. Meier.** 2005. Diversity of bacteria growing in natural mineral water after bottling. *Appl Environ Microbiol* **71**:3624-3632.
30. **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. W. 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. H. Schleifer.** 2004. ARB: a software environment for sequence data. *Nucl Acids Res* **32**:1363-1371.
31. **Manefield, M., A. S. Whiteley, R. I. Griffiths, and M. J. Bailey.** 2002. RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. *Appl Environ Microbiol* **68**:5367-5373.
32. **Nadkarni, M. A., F. E. Martin, N. A. Jacques, and N. Hunter.** 2002. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiol* **148**:257-266.
33. **Nielsen, J. L., D. Christensen, M. Kloppenborg, and P. H. Nielsen.** 2003. Quantification of cell-specific substrate uptake by probe-defined bacteria under *in situ* conditions by microautoradiography and fluorescence *in situ* hybridization. *Environ Microbiol* **5**:202-211.
34. **Noda, S., M. Ohkuma, A. Yamada, Y. Hongoh, and T. Kudo.** 2003. Phylogenetic position and *in situ* identification of ectosymbiotic spirochetes on protists in the termite gut. *Appl Environ Microbiol* **69**:625-633.

35. **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.
36. **Ohkuma, M., and T. Kudo.** 1996. Phylogenetic diversity of the intestinal bacterial community in the termite *Reticulitermes speratus*. *Appl Environ Microbiol* **62**:461-468.
37. **Pang, H., and H. H. Winkler.** 1993. Copy number of the 16S rRNA gene in *Rickettsia prowazekii*. *J Bacteriol* **175**:3893-3896.
38. **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 the spirochetes. *J Bacteriol* **173**:6101-6109.
39. **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.
40. **Rainey, F. A., N. L. Ward-Rainey, P. H. Janssen, H. Hippe, and E. Stackebrandt.** 1996. *Clostridium paradoxum* DSM 7308T contains multiple 16S rRNA genes with heterogeneous intervening sequences. *Microbiol* **142**:2087-2095.
41. **Riesenfeld, C. S., P. D. Schloss, and J. Handelsman.** 2004. Metagenomics: Genomic analysis of microbial communities. *Ann Rev Genet* **38**:525-552.
42. **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*. *Microbiol* **149**:2529-2537.

43. **Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler.** 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**:502-504.
44. **Sogin, S. J., M. L. Sogin, and C. R. Woese.** 1971. Phylogenetic measurement in procaryotes by primary structural characterization. *J Mol Evol* **1**:173-184.
45. **Strous, M., E. Pelletier, S. Mangenot, T. Rattei, A. Lehner, M. W. Taylor, M. Horn, H. Daims, D. Bartol-Mavel, P. Wincker, V. Barbe, N. Fonknechten, D. Vallenet, B. Segurens, C. Schenowitz-Truong, C. Médigue, A. Collingro, B. Snel, B. E. Dutilh, H. J. Op den Camp, C. van der Drift, I. Cirpus, K. T. van de Pas-Schoonen, H. R. Harhangi, L. van Niftrik, M. Schmid, J. Keltjens, J. van de Vossenberg, B. Kartal, H. Meier, D. Frishman, M. A. Huynen, H. W. Mewes, J. Weissenbach, M. S. Jetten, M. Wagner, and D. Le Paslier.** 2006. Deciphering the evolution and metabolism of an anammox bacterium from a community genome. *Nature* **440**:790-794.
46. **Suzuki, M. T., L. T. Taylor, and E. F. DeLong.** 2000. Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl Environ Microbiol* **66**:4605-4014.
47. **Tholen, A., and A. Brune.** 2000. Impact of oxygen on metabolic fluxes and *in situ* rates of reductive acetogenesis in the hindgut of the wood-feeding termite *Reticulitermes flavipes*. *Environ Microbiol* **2**:436-449.
48. **Thorsen, T., S. J. Maerkl, and S. R. Quake.** 2002. Microfluidic large-scale integration. *Science* **298**:580-584.

49. **Tringe, S. G., C. von Mering, A. Kobayashi, A. A. Salamov, K. Chen, H. W. Chang, M. Podar, J. M. Short, E. J. Mathur, J. C. Detter, P. Bork, P. Hugenholtz, and E. M. Rubin.** 2005. Comparative metagenomics of microbial communities. *Science* **308**:554-557.
50. **Tyson, G. W., J. Chapman, P. Hugenholtz, E. E. Allen, R. J. Ram, P. M. Richardson, V. V. Solovyev, E. M. Rubin, D. S. Rokhsar, and J. F. Banfield.** 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**:37-43.
51. **Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Neelson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y. H. Rogers, and H. O. Smith.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**:66-74.
52. **Vogelstein, B., and K. W. Kinzler.** 1999. Digital PCR. *Proc Natl Acad Sci USA* **96**:9236-9241.
53. **Warren, L., D. Bryder, I. L. Weissman, and S. R. Quake.** 2006. Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR. *Proc Natl Acad Sci USA* **103**:17807-17812.
54. **Weisburg, W. G., J. G. Tully, D. L. Rose, J. P. Petzel, H. Oyaizu, D. Yang, L. Mandelco, J. Sechrest, T. G. Lawrence, J. Van Etten, and et al.** 1989. A phylogenetic analysis of the mycoplasmas: basis for their classification. *J Bacteriol* **171**:6455-6467.

55. **Winston, P. W., Bates, Donald H.** 1960. Saturated solutions for the control of humidity in biological research. *Ecology* **41**:232-237.
56. **Wyss, C., B. K. Choi, P. Schüpbach, B. Guggenheim, and U. B. Göbel.** 1997. *Treponema amylovorum* sp. nov., a saccharolytic spirochete of medium size isolated from an advanced human periodontal lesion. *Int J Syst Bacteriol* **47**:842-845.
57. **Wyss, C., B. K. Choi, P. Schüpbach, B. Guggenheim, and U. B. Göbel.** 1996. *Treponema maltophilum* sp. nov., a small oral spirochete isolated from human periodontal lesions. *Int J Syst Bacteriol* **46**:745-752.
58. **Yang, H., D. Schmitt-Wagner, U. Stingl, and A. Brune.** 2005. Niche heterogeneity determines bacterial community structure in the termite gut (*Reticulitermes santonensis*). *Environ Microbiol* **7**:916-932.
59. **Zhang, K., A. C. Martiny, N. B. Reppas, K. W. Barry, J. Malek, S. W. Chisholm, and G. M. Church.** 2006. Sequencing genomes from single cells by polymerase cloning. *Nat Biotechnol* **24**:680-686.
60. **Zuckerkindl, E., and L. Pauling.** 1965. Molecules as documents of evolutionary history. *J Theor Biol* **8**:357-366.
61. **Zwirgmaier, K., W. Ludwig, and K. H. Schleifer.** 2004. Recognition of individual genes in a single bacterial cell by fluorescence *in situ* hybridization - RING-FISH. *Mol Microbiol* **51**:89-96.

Microfluidic Digital PCR with Degenerate Primers: Multiplex Molecular Community Analysis of Acetogenic Bacteria in the Termite Hindgut

Abstract

PCR-based molecular profiling techniques allow in-depth analysis of uncultured environmental microorganisms, but are limited by their single-gene nature. Here, we present a method for multiplex PCR interrogation of uncultured environmental bacteria using degenerate primers that target protein coding genes and 16S rRNA. The use of microfluidic digital PCR to generate environmental gene inventories from parallel analysis of separated bacterial cells should minimize the effect of PCR bias on library composition and eliminate chimeric sequence artifacts. The ability to perform multiplex gene inventories allows the discovery of 16S rRNA gene sequences of organisms carrying a genetic marker of interest, or the association of two or more metabolic markers to single strains of uncultured bacteria. We used this technique to discover the 16S rRNA-based species identities of termite gut bacteria carrying the gene for formyl-tetrahydrofolate synthetase, a key enzyme for CO₂-reductive acetogenesis.

Introduction

The use of molecular community profiling techniques has transformed the field of microbial ecology (16, 27). Assays that target ribosomal RNA genes are routinely used to characterize the species composition of complex environments (9, 20, 36), while assays targeting metabolic genes are used to evaluate the diversity of organisms carrying

a genetic capacity of interest (3, 18, 47). Molecular profiling experiments use PCR to amplify a subset of related sequences from an environmental DNA pool. This generates a PCR product pool sequences that reflects the diversity of organisms within a sample that carry the genetic marker of interest; the diversity of sequences in this pool can be measured indirectly by techniques such as terminal restriction length polymorphism or denaturing gradient gel electrophoresis, or directly through generation and sequencing of clone libraries (25).

Large-scale environmental sequencing projects have proven valuable sources of novel genes and insights into environmental processes (32, 38). Similarly, techniques for single cell genome sequencing show promise for the metabolic characterization of uncultured organisms (24, 37, 48). However, these analyses cannot match the target specificity (and therefore the potential survey depth) of PCR-based techniques. As an example, a recent survey used 16S rRNA primers in conjunction with high-density 454 pyrosequencing to generate 8 sequence libraries from deep sea DNA samples that contained 6,505–22,994 total sequences, 2,656–8,699 of them unique (36). In contrast, the initial Sargasso Sea metagenomic analysis comprised over 1.88 million sequence reads, yet yielded only 1,412 rRNA genes (40).

However, the targeted nature of PCR-based techniques can also represent a drawback. As genes are studied in isolation, it is difficult to establish relationships between phylotypes identified in inventories of diverse genes from a single sample. When metabolic genes are examined, the species identity of the source organisms can only be

hypothesized according to their similarity to gene sequences from previously cultivated and characterized organisms. This type of analysis can result in large clusters of unassigned sequences when cultured representatives are rare (7, 13). We have developed a microfluidics-based technique that allows the association of such sequences with the 16S rRNA gene sequences of the uncultured environmental bacteria that encode them.

In 2006, we demonstrated a technique for multiplex, microfluidic digital PCR-based interrogation of hundreds of bacterial cells in parallel (26). We used a microfluidic device to partition an environmental sample into hundreds of independent reaction chambers. At low sample dilutions, reaction chambers contained no more than one bacterial cell. Each of those cells was then used as template for a multiplex PCR reaction targeting bacterial 16S rRNA genes and a specific functional gene sequence. PCR products from chambers showing amplification of both genes were retrieved and sequenced.

This technique allowed 16S rRNA gene-based identification of uncultured bacteria carrying a genetic marker of interest. However, the utility of the technique as described was limited by the nature of the PCR chemistry utilized. PCR amplification within the microfluidic device was detected using amplicon-specific Taqman probes. While the diversity of sequences amplified in PCR assays can be expanded through the use of degenerate primers, Taqman probes are highly sequence specific. The strong sequence conservation of the 16S rRNA gene allowed the design of an “all bacterial” probe with low degeneracy, yet broad specificity (39). However, protein-encoding genes are

generally less well conserved, and a nondegenerate probe will only detect closely related sequence subgroups.

Here, we present an approach that allows multiplex digital PCR with degenerate primers that target genes encoding clusters of orthologous proteins. We used a universal-template probe strategy developed by Zhang et al. (49), in which a probe-binding sequence is attached to the 5' end of a real-time PCR primer. This sequence is incorporated into the amplicon during the first round of amplification, allowing the Taqman probe to detect amplification of that product. Zhang et al. proposed this approach as a method to reduce the costs associated with real-time PCR. However, we have adapted this strategy to pair nondegenerate Taqman probes with degenerate primers for multiplex PCR.

We used this technique in the context of our on-going efforts to characterize the acetogenic community of the termite hindgut. In wood-feeding termites, CO₂-reducing acetogens are the primary consumers of H₂ generated during the fermentation of wood polysaccharides; the acetate produced by these bacteria powers 22%–26% of the insect's energy metabolism (4, 5, 30). The gene for formyl-tetrahydrofolate synthetase (FTHFS), a key enzyme in the acetyl-CoA pathway, can be used as a genetic marker of acetogenic capability (19, 22). FTHFS-based molecular community analyses have been carried out on a number of termite species, which are dominated by a sequence cluster that includes FTHFS genes from two acetogenic spirochetes, *Treponema primitia* ZAS-1 and ZAS-2 (17, 29, 34). However, many of the recovered sequences are only distantly related to these two isolates, and the termite *Treponeme* cluster as a whole affiliates with FTHFS

sequences from *Firmicute* acetogens. As a result, even phylum-level classification of the bacteria that encode key FTHFS types is ambiguous.

We have designed a multiplex, microfluidic digital PCR assay that allows the use of all-bacterial 16S rRNA gene primers and FTHFS primers that target the Lovell cluster of acetogenic FTHFS types. Using this assay, we have identified FTHFS-bearing organisms from the termite hindgut, and used 16S rRNA gene phylogeny to discover their species identity. To demonstrate the general applicability of our strategy, we also designed and implemented an assay to discover the gene sequences for the ATPase subunit of the Clp protease (ClpX) of uncultured termite *Treponemes*. ClpX was chosen because it a potential target for design of species-specific internal controls for environmental expression analyses.

Materials and Methods

Laboratory Maintenance of Termites and Bacterial Strains

Zootermopsis nevadensis specimens were collected from fallen Ponderosa Pine (*Pinus ponderosa*) at the Chilao Campground in the Angeles National Forest. Colonies were maintained in the laboratory on Ponderosa at 23 °C and at a constant humidity of 96%, achieved via incubation over saturated solutions of KH_2PO_4 within 10-gallon aquaria (43). *Treponema primitia* ZAS-1 was maintained in the laboratory as described in (17).

PCR Primer Design

Degenerate primers were designed using the CODEHOP program (33). FTHFS sequences from acetogenic bacteria and partial termite *Treponeme* FTHFS genes from the *Nasutitermes* metagenome (42) were used to design primers. A 57-60 °C consensus region was found to be optimal for the 60 °C extension/annealing temperature used in real-time PCR experiments. Consensus regions suggested by the CODEHOP program were adjusted to match codon preferences observed in termite *Treponeme* FTHFS sequences.

Table 5.1. Primers Used in Microfluidic Digital PCR

Name	Sequence	Target
357F	CTCCTACGGGAGGCAGCAG	All bacterial 16S rRNA
1492RL2D	TACGGYTACCTTGTTACGACTT	All bacterial 16S rRNA
1409Ra	GGGTACCTCCAACCTCGGATGGTG	Termite <i>Treponeme</i> 16S rRNA
1409Rb	CGGGTACCCTCTACTCGGATGGTG	Termite <i>Treponeme</i> 16S rRNA
533F	AGGGTTGCGCTCGTTG	16S rRNA sequencing
1100R	GTGCCAGCMGCCGCGTAA	16S rRNA sequencing
FTHFS-Fa	GGICIGTITTYGGIGTIAARGG	FTHFS, unprobed
FTHFS-Ra	CCIGGCATIGTCATIATITCICCI	FTHFS, unprobed
FTHFS-Fb	ACCTGCACTTCACCGGAGAYTTYCAYGCIAT	FTHFS, probed
UP149-FTHFS-Fb	GGCGGCGAACCTGCACCTTCACCGGAGAYTTYCAYGCIAT	FTHFS, probed
FTHFS-Rb	ACGCCTTCGCCACCCTTIKCCCAIAC	FTHFS, probed
ClpX_F	CGAAGCGGGCTATGTGCGIGARGAYGT	ClpX, probed
ClpX_R	GATGGGAAGCCTGCCGATGAAYTCIGGDAT	ClpX, probed
UP149-ClpX-R	GGCGGCGAGATGGGAAGCCTGCCGATGAAYTCIGGDAT	ClpX, probed

ClpX protease primers were designed using a similar strategy. ClpX protease sequences were downloaded from the *Nasutitermes* metagenome data set (42) and aligned with sequences from published microbial genomes. A putative termite *Treponeme* cluster of ClpX sequences was identified based on phylogenetic similarity to ClpX sequences from published *Treponeme* genomes (*T. denticola* (35) and *T. pallidum* (11)) and unpublished

ZAS-2 ClpX sequence (Eric Matson, personal communication). Primer sequences are presented in Table 5.1. Inosine base analogues (denoted as “I” in the primer sequence) were used in the place of N to reduce the degeneracy of the primers. Other than the Roche universal probe, all primers and probes were purchased from Integrated DNA Technologies.

Template Preparation

DNA was purified from *T. primitia* ZAS-1 using the Qiagen DNeasy Tissue kit, with the protocol described for extraction of DNA from gram negative bacteria cells (DNeasy Tissue Handbook, July 2003 version). Template concentrations were measured using the Hoefer DyNAQuant 200 fluorometer and DNA quantification system (Amersham Pharmacia Biotech) using reagents and procedures directed in the manual (DQ200-IM, Rev C1, 5-98).

A “synthetic gut fluid” (SGF) salt solution was used for suspension and dilution of bacterial cells prior to addition to the PCR reaction. This solution contained 29.4 mM, K_2HPO_4 , 11.6 mM KH_2PO_4 , 5.6 mM KCl, and 30 mM NaCl. DNase-free RNase (Roche) was added just prior to cell dilution at 0.5 μ g/mL to prevent PCR inhibition by ribosomal RNA. *T. primitia* ZAS-1 cells were collected from late exponential phase cultures and diluted in sterile SGF. Single *Z. nevadensis* hindguts were extracted from worker larvae, suspended in sterile SGF, and physically disrupted by crushing the gut with a sterile pipette tip followed by brief (2–3 s) pulses of vortexing. Suspensions were allowed to stand briefly to sediment large particles, then diluted to working concentrations in SGF

and mixed 1 to 15 v/v with the PCR reaction mixture for immediate loading onto microfluidic chips.

PCR on Microfluidic Chips

Microfluidic devices were purchased from Fluidigm Corporation. On-chip multiplex PCR reactions contained iQ Multiplex Powermix (BioRad, 170-8848), 0.1% Tween-20, and 150 nM ROX standard. 16S rRNA amplifications used primers and probes described in (26): 357F and 1492RL2D at 400nM for all bacterial 16S rRNA, and 357F, 1409Ra, and 1409Rb at 400nM each for “spirochete-specific” 16S rRNA amplification; all 16S rRNA reactions used the 1389 probe (HEX-CTTGTACACACCGCCCGTC-BHQ1) at 267 nM. Unprobed FTHFS reactions used FTHFS-Fa and FTHFS-Ra at 400 nM each. For universal template probe reactions, Roche Universal Probe #149 was included at 267 nM, the unlabeled primer (FTHFS-Rb or ClpX-F) added at 400 nM. The best signal intensity for universal template probe reactions was obtained when the primer with attached binding site was mixed 50:50 with the same primer without the binding site. Reactions contained 200 nM probe-binding primer (149-FTHFS-Fb or 149-ClpX-R) and 200 nM conventional primer (FTHFS-Fb or ClpX-R).

Chips were loaded and PCR performed using the BioMark system as recommended by Fluidigm. The cycling protocol was 95 °C 3 min, (95 °C 15 s, 60 °C 90 s) x 45. When chips were intended for product retrieval, a final extension step of 10min at 60 °C was added. Amplification curves and reaction results were evaluated using BioMark Digital PCR analysis software (Fluidigm, v.2.0.6).

Sample Retrieval and Analysis

Single cell PCR products were retrieved from amplification-positive chambers. Pressure was released from the accumulators, and chips were then peeled from the carrier and silicon heat sink. Target chambers were located using a dissecting microscope, and the tip of a 26 gauge syringe needle was inserted into each chamber through the bottom surface of the chip. Needle tips were then swirled briefly in 10 μ L of TE to release the PCR product.

Retrieved samples were evaluated for the presence of target genes via simplex PCR at a benchtop scale. For functional gene analysis, primers without the probe binding site were used at 400nM, and the probe-binding primer omitted. In some samples, re-amplification with 357 and 1492RL2D resulted in secondary bands. Utilization of 533F in place of 357 eliminated these artifacts. The cycling protocol for conventional PCR was 95 °C 3 min, (95 °C 15 s, 60 °C 60 s, 72 °C 60 s) x 35. The presence or absence of product was evaluated using agarose gel electrophoresis.

PCR products from successful retrievals were purified using the Qiagen PCR purification kit, and sequenced using the FTHFS PCR primers and 16S rRNA gene internal primers 1100R and 533F. Cycle sequencing reactions were carried out by Laragen (Los Angeles, CA) Sequences were assembled and edited using the Lasergene software package (DNASTAR, version 7.2.1). Sequence alignment and phylogenetic analysis were carried out using the ARB software package (23).

Results

We successfully designed primer/probe sets that allowed Taqman-based detection of FTHFS sequences in multiplex reactions with “all-bacterial” 16S rRNA primers on microfluidic chips (Figure 5.1). Low levels of amplification (5-15 chambers per panel) were detected for the FTHFS channel in “no template added” reactions, most likely as a result of primer dimer formation with 16S rRNA primers. In simplex FTHFS reactions, no amplification was observed in “no template added” reactions. Minor amplification (5-10 positives per panel) was also detected for 16S rRNA in template-free controls; this is likely due to low levels of bacterial DNA contamination (common in commercial PCR reagents (6)).

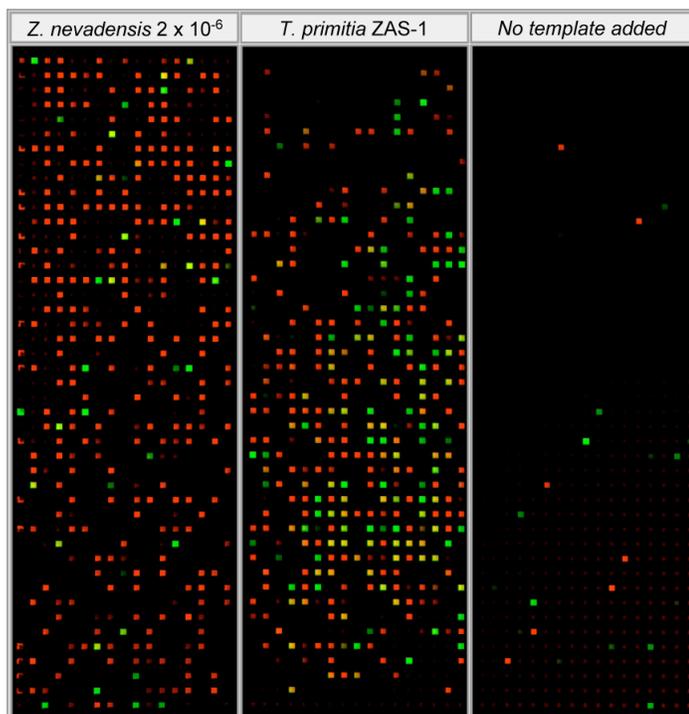


Figure 5.1. Microfluidic Digital PCR for “Lovell cluster” FTHFS and all-bacterial 16S rRNA genes. Three sample panels from a representative chip are shown at amplification cycle 45. FTHFS signal shown in green, 16S rRNA gene signal shown in red. The template source for panel 1 was *Z. nevadensis* hindgut contents, for panel 2 was a late exponential phase *T. primitia* ZAS-1 culture. No template was added to the PCR mixture for panel 3.

Treponema primitia ZAS-1 was chosen for use as a positive control template because it carries two genomic copies each of FTHFS and 16S rRNA (12, 34), allowing direct comparison of amplification efficiency between these two genes. The observed amplification success rate for on-chip, multiplex PCR from purified genomic DNA was 42% for FTHFS and 73% for 16S rRNA (calculated as measured number of copies per μL divided by expected copies/ μL according to DNA concentration). In multiplex PCR reactions from cultured *T. primitia* ZAS-1 cells, approximately 25% of chambers with either FTHFS or 16S rRNA amplification were positive for both genes. This is more than twice the number of colocalizations expected if FTHFS and 16S rRNA were assorting independently. The presence of chambers in which either FTHFS or 16S rRNA amplified alone may be due either to multiplexing failure (where amplification of one gene outcompetes amplification of the other) or to lysis of ZAS-1 cells followed by genome fragmentation.

FTHFS and 16S rRNA genes were successfully amplified in multiplex PCR from hindgut luminal contents from the lower termite *Zootermopsis nevadensis*. *Z. nevadensis* hindgut contents were diluted in SGF salt solution and added to PCR reactions immediately prior to chip loading. The standard for single cell separation was 33% occupancy or less. PCR products were retrieved, reamplified, and sequenced from chambers in which both FTHFS and 16S rRNA genes had amplified. The resultant FTHFS and 16S rRNA sequences were binned with a similarity cutoff of 99.5% and characterized by phylogenetic analysis (Figure 5.2, Table 5.2).

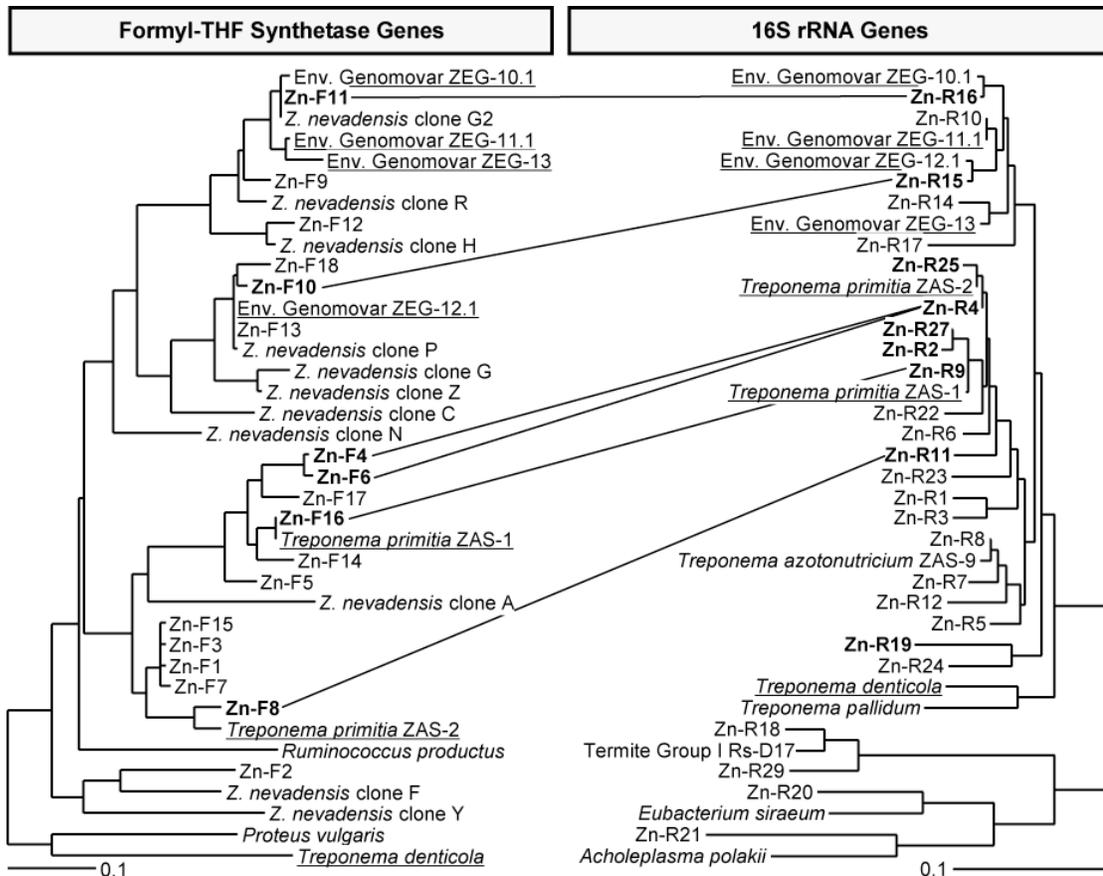


Figure 5.2. Phylogenetic analysis of FTHFS and 16S rRNA gene sequences amplified using microfluidic digital PCR. Trees calculated using Phylip distance Fitch algorithm. **Left,** An FTHFS gene tree constructed using 726 unambiguous, aligned base pairs; short sequences (ZEG sequences and Zn-F18) were added to the finished tree using 192 alignment positions. **Right,** A 16S rRNA gene tree calculated using 722 unambiguous, aligned base pairs. Scale bars represent 0.1 changes per alignment position. Lines identify FTHFS-16S rRNA gene pairs supported by repeated colocalizations or similarity to established associations (Table 5.2). Sequences assigned to an environmental genomovar (Table 5.5) marked in bold.

Table 5.2 lists FTHFS and 16S rRNA gene pairs colocalized using microfluidic digital PCR. In our initial microfluidic digital PCR experiments (26), we found that apparent single cell dilutions sometimes contained multiple 16S rRNA types or a single 16S rRNA type that did not match those found in other experiments. This was attributed to the nature of the dilution method used for cell separation; physically associated cell aggregates sort as one particle. Sorting of cell aggregates followed by a skewed PCR

balance result in false associations. As a result, a single colocalization event was considered insufficient evidence of association; FTHFS-16S rRNA pairs identified in at least two independent colocalizations or supported by similarity to cultured strains or prior-colocalizations are marked in Figure 5.2.

Table 5.2. FTHFS/16S rRNA gene pairs^a

Experiment	FTHFS	16S rRNA	Comment
1	Zn-F1	Zn-R1	
2	Zn-F2	Zn-R2	
3	Zn-F3	Zn-R3	
4	Zn-F4	Zn-R4	Similar to 15; ZEG 14.1
5	Zn-F5	Zn-R5	
6	Zn-F5	Zn-R6	
7	Zn-F5	Zn-R7	
8	Zn-F4	Zn-R8	
9	Zn-F6	Zn-R9	
10	Zn-F6	Zn-R10	
11	Zn-F7	Zn-R7	
12	Zn-F7	Zn-R11	
13	Zn-F8	Zn-R12	
14	Zn-F4	Zn-R4	
15	Zn-F6	Zn-R4	Similar to 4; ZEG 14.2
16	Zn-F8	Zn-R14	
17	Zn-F8	Zn-R10	
18	Zn-F8	Zn-R11	Repeated in 20, 24; ZEG 16.1
19	Zn-F9	Zn-R14	
20	Zn-F8	Zn-R11	Repeated in 20, 24; ZEG 16.1
21	Zn-F10	Zn-R15	Similar to ZEG 12; ZEG 12.5
22	Zn-F11	Zn-R16	Similar to ZEG 10; ZEG 10.5
23	Zn-F12	Zn-R17	
24	Zn-F8	Zn-R11	Repeated in 20, 24; ZEG 16.1
25	Zn-F13	Zn-R18	
26	Zn-F14	Zn-R9	
27	Zn-F15	Zn-R2	
28	Zn-F16	Zn-R19	
29	Zn-F16	Zn-R20	
30	Zn-F16	Zn-R9	Similar to ZAS-1; ZEG 15.1
31	Zn-F17	Zn-R21	

^a Colocalizations 1-17 were collected in experiments using an unprobed FTHFS primer set and termite *Treponeme* specific 16S rRNA primers. Termite gut dilutions were analyzed in side-by-side reactions with all-bacterial 16S rRNA primers to confirm single cell separation. PCR products were retrieved from chambers positive for *Treponeme* 16S rRNA and screened off-chip for FTHFS amplification.

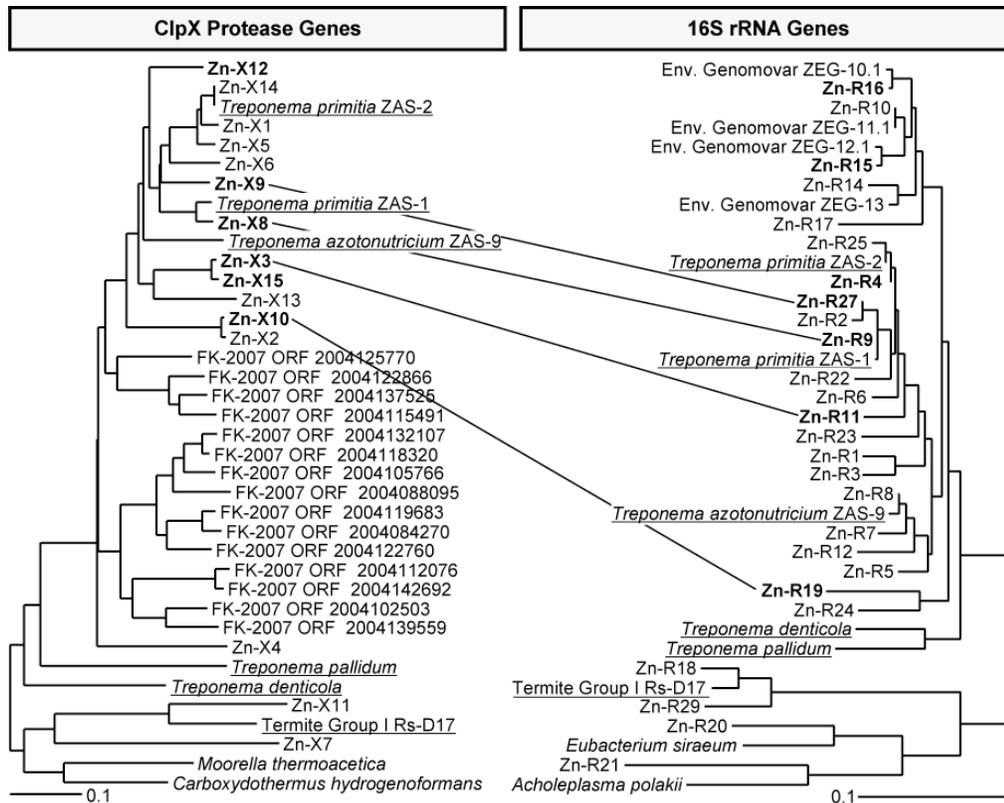


Figure 5.3. Phylogenetic analysis of ClpX and 16S rRNA gene sequences amplified using microfluidic digital PCR. Trees calculated using Phylip distance Fitch algorithm. **Left,** A ClpX gene tree constructed using 397 unambiguous, aligned base pairs. Sequences from the *Nasutitermes* metagenome named as FK-2007 ORF [JGI database GOI] **Right,** 16S rRNA gene tree (calculation described in Figure 5.2). Scale bars represent 0.1 changes per alignment position. Lines identify ClpX-16S rRNA gene pairs supported by repeated colocalizations or similarity to established associations (Table 5.3). Sequences assigned to an environmental genomovar (Table 5.5) marked in bold.

Table 5.3. ClpX/16S rRNA gene pairs

Experiment	ClpX	16S rRNA	Comment
1	Zn-X1	Zn-R22	
2	Zn-X2	Zn-R23	
3	Zn-X3	Zn-R11	Triplex Zn-F8 to Zn-R11 to Zn-X3, X15; ZEG 16.2
4	Zn-X4	Zn-R24	
5	Zn-X5	Zn-R25	
6	Zn-X6	Zn-R18	
7	Zn-X7	Zn-R11	
8	Zn-X8	Zn-R9	Similar to ZAS-1; ZEG 15.2
9	Zn-X9	Zn-R27	Repeated in 10; ZEG 17.1
10	Zn-X9	Zn-R27	Repeated in 9; ZEG 17.1
11	Zn-X10	Zn-R19	Repeated in 13, 14; ZEG 18.1
12	Zn-X11	Zn-R29	Termite Group I
13	Zn-X10	Zn-R19	Repeated in 11, 14; ZEG 18.1
14	Zn-X10	Zn-R19	Repeated in 11, 13; ZEG 18.1

To demonstrate the general applicability of the universal-template probe strategy to degenerate primer based PCR, we designed a primer set to amplify the gene for the ATP-binding subunit of the Clp protease complex (ClpX) from termite gut *Treponemes*. ClpX was chosen as a target because it is highly conserved, a *Treponeme* sequence cluster for this gene was clearly identifiable in the *Nasutitermes* sp. metagenomic dataset, and prior experiments in this laboratory (Matson and Leadbetter, manuscript in preparation) have demonstrated this gene's utility as an internal standard for quantitative PCR-based transcriptional analyses. Multiplex PCR reactions containing *Z. nevadensis* hindgut contents were carried out as described for FTHFS and 16S rRNA. ClpX-16S rRNA associations are presented in Figure 5.3 and Table 5.3. A limited number of FTHFS-ClpX association experiments were also carried out (Figure 5.4, Table 5.4).

Table 5.4. FTHFS/ClpX gene pairs^a

Experiment	FTHFS	ClpX	Comment
1	Zn-F10	Zn-X12	Repeated in 2; ZEG 12.6
2	Zn-F10	Zn-X12	Repeated in 1; ZEG 12.6
3	Zn-F18	Zn-X2	
4	Zn-F13	Zn-X13	
5	Zn-F10	Zn-X6	
6	Zn-F10	Zn-X14	
7	Zn-F8	Zn-X14	
8	Zn-F8	Zn-X15	Triplex Zn-F8 to Zn-R11 to Zn-X3, X15; ZEG 16.3

^a Colocalizations 1-4 were retrieved from reactions using ClpX and the H-group specific FTHFS primer/probe set described in (26). All other colocalizations used the universal template probe FTHFS primer set.

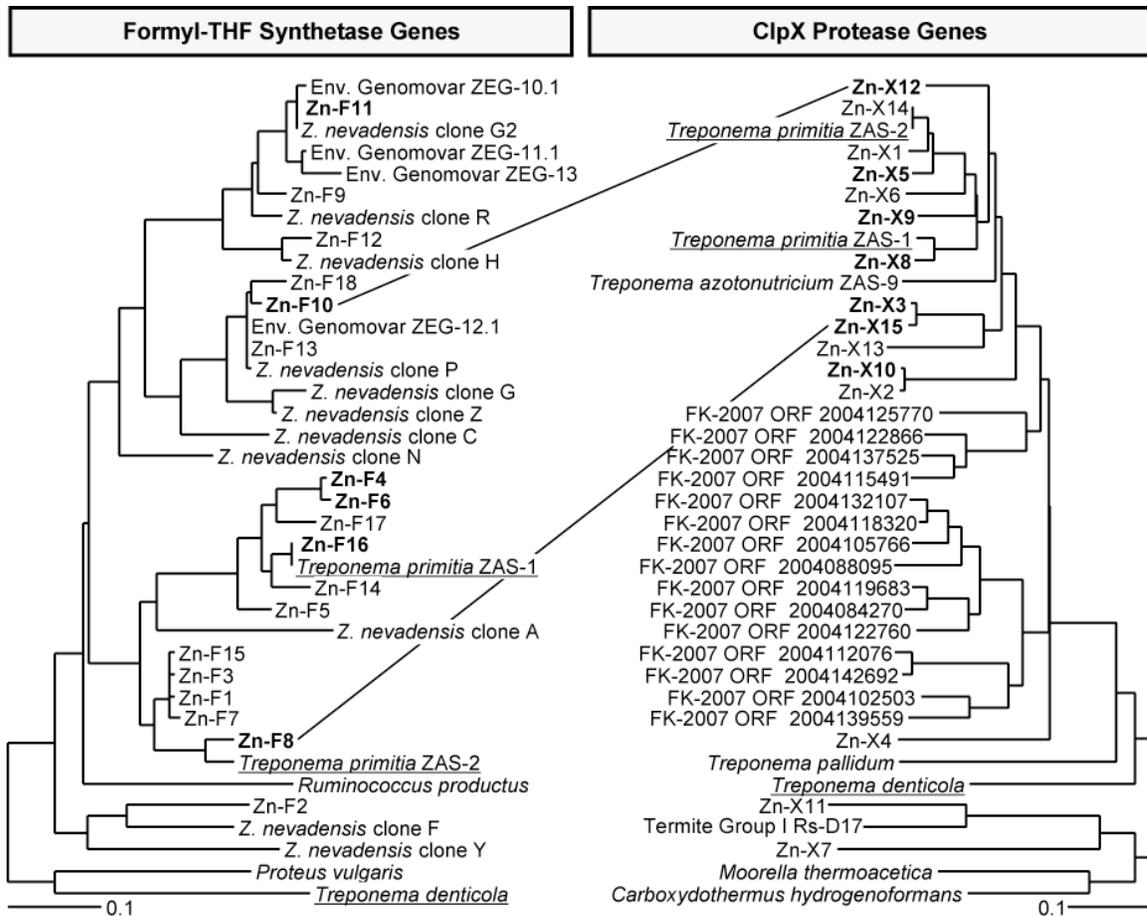


Figure 5.4. Phylogenetic analysis of FTHFS and ClpX gene sequences amplified using microfluidic digital PCR. Trees calculated as described in Figures 5.2 and 5.3. Scale bars represent 0.1 changes per alignment position. Lines identify FTHFS-ClpX gene pairs supported by repeated colocalizations or similarity to established associations (Table 5.4). Sequences assigned to an environmental genomovar (Table 5.5) marked in bold.

In our previous microfluidic experiments, we proposed the term “*environmental genomovar*” to describe uncultured organisms that have been shown to encode particular gene combinations. New environmental genomovars proposed based on the results of these experiments have been assigned monikers ZEG-14 through 18, as listed in Table 5.4. These include the first triplex association, as the FTHFS and 16S rRNA genes from ZEG 16 have been independently associated with highly similar (98%) ClpX sequences.

Table 5.5. Proposed Environmental Genomovars

Name	16S rRNA	FTHFS	ClpX
ZEG 10.5	Zn-R16	Zn-F11	
ZEG 12.6	Zn-R15	Zn-F10	
ZEG 12.7		Zn-F10	Zn-X12
ZEG 14.1	Zn-R4	Zn-F4	
ZEG 14.2	Zn-R4	Zn-F6	
ZEG 15.1	Zn-R9	Zn-F16	
ZEG 15.2	Zn-R9		Zn-X8
ZEG 16.1	Zn-R11	Zn-F8	
ZEG 16.2	Zn-R11		Zn-X3
ZEG 16.3		Zn-F8	Zn-X15
ZEG 17.1	Zn-R27		Zn-X9
ZEG 18.1	Zn-R19		Zn-X10

Discussion

We have developed a strategy for multiplex, microfluidic digital PCR that allows simultaneous amplification and detection of “Lovell cluster” FTHFS genes and bacterial 16S rRNA genes. These primers were used in combination with “all-bacterial” 16S rRNA gene primers to discover the species identity of uncultured, FTHFS-bearing bacteria in the termite gut. An assay targeting the ATP-binding subunit of the Clp protease complex (ClpX) from termite gut *Treponemes* was also developed. This primer/probe set was used to associate ClpX genes with both 16S rRNA gene sequences and FTHFS genes.

The last few decades have seen widespread use of degenerate primers to build environmental inventories of genes encoding enzymes involved in key environmental processes such as nitrogen fixation (*nifH* gene) (45, 46), methanotrophy (*pmoA*, *mmoX*, *mxoF*) (8), and sulfate reduction (*dsrAB*) (41). The ability to carry out such analyses at the level of single environmental cells, however, should greatly expand the information

derived from such techniques. With the use of universal template probes and degenerate primers, microfluidic digital PCR now allows the construction of such inventories in parallel with 16S rRNA analysis, placing the targeted genetic capability within the context of the phylogenetic species identity of the host and the environmental species assemblage. Additionally, the “one template, one reaction” nature of this technique circumvents some of the technical caveats of environmental inventories (1), eliminating chimeric product formation and minimizing the role of PCR bias in determining library composition.

The utility of this approach is not limited to assignment of 16S rRNA species identities to organisms bearing a genetic capacity of interest. In these experiments, we also built associations between FTHFS and ClpX genes. The gene for ClpX protease exhibits steady-state expression during growth of *Treponema primitia* strain ZAS-2, and might therefore be useful as an internal standard for quantitative expression analyses (Matson and Leadbetter, manuscript in preparation). Cross-sample comparison of expression levels among uncultured environmental microbes is currently based on normalization of total RNA concentrations (45), a metric that is highly sensitive to RNA sample quality (10). The incorporation of invariant control transcripts should greatly enhance the resolution of environmental expression analyses.

In conclusion, we have developed a microfluidic digital PCR technique that allows the highly parallel interrogation of individual environmental cells using multiplex, degenerate primers. We have used this technique to simultaneously inventory and

identify (based on rRNA species phylogeny) acetogenic bacteria in termite hindgut samples. The ability to build metabolic gene inventories from environmental samples while simultaneously identifying the ribosomal phylotype of the organisms that carry these genes will greatly enhance the utility of PCR-based molecular community profiling. The ability to carry out in-depth analyses targeting major metabolic guilds is highly complementary to environmental genomic and metagenomic analyses, and will continue as an important element of the microbial ecologist's arsenal.

Chapter Five Appendix

Table 5.6. Sequences used in phylogenetic analysis

Source/Sequence Type	Designation	Accession	Gene	Reference
<i>Z. nevadensis</i> Gut Clone	A	FTHFS	AY162294	(34)
<i>Z. nevadensis</i> Gut Clone	C	FTHFS	AY162295	(34)
<i>Z. nevadensis</i> Gut Clone	F	FTHFS	AY162298	(34)
<i>Z. nevadensis</i> Gut Clone	G	FTHFS	AY162300	(34)
<i>Z. nevadensis</i> Gut Clone	G2	FTHFS	AY162301	(34)
<i>Z. nevadensis</i> Gut Clone	H	FTHFS	AY162302	(34)
<i>Z. nevadensis</i> Gut Clone	N	FTHFS	AY162306	(34)
<i>Z. nevadensis</i> Gut Clone	P	FTHFS	AY162307	(34)
<i>Z. nevadensis</i> Gut Clone	R	FTHFS	AY162308	(34)
<i>Z. nevadensis</i> Gut Clone	Y	FTHFS	AY162311	(34)
<i>Z. nevadensis</i> Gut Clone	Z	FTHFS	AY162312	(34)
<i>Z. nevadensis</i> Genomovar	ZEG 10.1	FTHFS	DQ420342	(26)
<i>Z. nevadensis</i> Genomovar	ZEG 11.1	FTHFS	DQ420346	(26)
<i>Z. nevadensis</i> Genomovar	ZEG 12.1	FTHFS	DQ420353	(26)
<i>Z. nevadensis</i> Genomovar	ZEG 13.1	FTHFS	DQ420358	(26)
<i>Proteus vulgaris</i>		FTHFS	AF295710	(19)
<i>Ruminococcus productus</i>		FTHFS	AF295707	(19)
<i>Treponema denticola</i>		FTHFS	NC_002967	(35)
<i>Treponema primitia</i> ZAS-1	ZAS-1a	FTHFS	AY162313	(34)
<i>Treponema primitia</i> ZAS-2	ZAS-2	FTHFS	AY162315	(34)
<i>Acholeplasma polakii</i>		16S	AF031479	(2)
<i>Eubacterium siraeum</i>	ATCC 29066	16S	L34625	
Termite Group I bacterium	Rs-D17	16S	AB089048	(14)
<i>Treponema azotonutricium</i> ZAS-9	ZAS-9	16S	AF320287	(21)
<i>Treponema denticola</i>	II:11:33520	16S	M71236	(28)
<i>Treponema pallidum</i>	Nichols	16S	M88726	(28)
<i>Treponema primitia</i> ZAS-1	ZAS-1	16S	AF093251	(17)
<i>Treponema primitia</i> ZAS-2	ZAS-2	16S	AF093252	(17)
<i>Z. nevadensis</i> Genomovar	ZEG 10.1	16S	DQ420325	(26)
<i>Z. nevadensis</i> Genomovar	ZEG 11.1	16S	DQ420329	(26)
<i>Z. nevadensis</i> Genomovar	ZEG 12.1	16S	DQ420336	(26)
<i>Z. nevadensis</i> Genomovar	ZEG 13.1	16S	DQ420341	(26)
<i>Carboxydotherrmus hydrogenoformans</i>	Z-2901	ClpX	NC_007503.1	(44)
<i>Moorella thermoacetica</i>	ATCC 39703	ClpX	NC_007644	(31)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004084270	ClpX	JGI GOI_2004084270	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004088095	ClpX	JGI GOI_2004088095	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004102503	ClpX	JGI GOI_2004102503	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004105766	ClpX	JGI GOI_2004105766	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004112076	ClpX	JGI GOI_2004112076	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004115491	ClpX	JGI GOI_2004115491	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004118320	ClpX	JGI GOI_2004118320	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004119683	ClpX	JGI GOI_2004119683	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004122760	ClpX	JGI GOI_2004122760	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004122866	ClpX	JGI GOI_2004122866	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004125770	ClpX	JGI GOI_2004125770	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004132107	ClpX	JGI GOI_2004132107	(42)

Source/Sequence Type	Designation	Accession	Gene	Reference
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004137525	ClpX	JGI GOI 2004137525	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004139559	ClpX	JGI GOI 2004139559	(42)
<i>Nasutitermes</i> sp. FK-2007 metagenome	2004142692	ClpX	JGI GOI 2004142692	(42)
Termite Group I bacterium	Rs-D17	ClpX	AP009510	(15)
<i>Treponema denticola</i>		ClpX	NC_002967	(35)
<i>Treponema pallidum</i>		ClpX	NC_000919	(11)

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 constructed from the same sample. *Appl Environ Microbiol* **71**:8966-8969.
2. **Angulo, A. F., R. Reijgers, J. Brugman, I. Kroesen, F. E. Hekkens, P. Carle, J. M. Bove, J. G. Tully, A. C. Hill, L. M. Schouls, C. S. Schot, P. J. Roholl, and A. A. Polak-Vogelzang.** 2000. *Acholeplasma vituli* sp. nov., from bovine serum and cell cultures. *Int J Syst Evol Microbiol* **50 Pt 3**:1125-1131.
3. **Béjà, O., E. N. Spudich, J. L. Spudich, M. Leclerc, and E. F. DeLong.** 2001. Proteorhodopsin phototrophy in the ocean. *Nature* **411**:786-789.
4. **Breznak, J. A.** 1994. Acetogenesis from carbon dioxide in termite guts, p. 303-330. *In* H. A. Drake (ed.), *Acetogenesis*. Chapman and Hall, New York.
5. **Brune, A.** 2006. Symbiotic associations between termites and prokaryotes, p. 439-474. *In* M. Dworkin, S. Falkow, E. Rosenber, K. H. Schleifer, and E. Stackebrandt (ed.), *The Prokaryotes*, Third ed, vol. I. Springer.
6. **Corless, C. E., M. Guiver, R. Borrow, V. Edwards-Jones, E. B. Kaczmariski, and A. J. Fox.** 2000. Contamination and sensitivity issues with a real-time universal 16S rRNA PCR. *J Clin Microbiol* **38**:1747-1752.
7. **Dhillon, A., A. Teske, J. Dillon, D. A. Stahl, and M. L. Sogin.** 2003. Molecular characterization of sulfate-reducing bacteria in the Guaymas Basin. *Appl Environ Microbiol* **69**:2765-2772.

8. **Dumont, M. G., and J. C. Murrell.** 2005. Community-level analysis: key genes of aerobic methane oxidation. *Methods Enzymol* **397**:413-427.
9. **Elshahed, M. S., N. H. Youssef, A. M. Spain, C. Sheik, F. Z. Najar, L. O. Sukharnikov, B. A. Roe, J. P. Davis, P. D. Schloss, V. L. Bailey, and L. R. Krumholz.** 2008. Novelty and uniqueness patterns of rare members of the soil biosphere. *Appl Environ Microbiol* **74**:5422-5428.
10. **Fleige, S., and M. W. Pfaffl.** 2006. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med* **27**:126-139.
11. **Fraser, C. M., S. J. Norris, G. M. Weinstock, O. White, G. G. Sutton, R. Dodson, M. Gwinn, E. K. Hickey, R. Clayton, K. A. Ketchum, E. Sodergren, J. M. Hardham, M. P. McLeod, S. Salzberg, J. Peterson, H. Khalak, D. Richardson, J. K. Howell, M. Chidambaram, T. Utterback, L. McDonald, P. Artiach, C. Bowman, M. D. Cotton, C. Fujii, S. Garland, B. Hatch, K. Horst, K. Roberts, M. Sandusky, J. Weidman, H. O. Smith, and J. C. Venter.** 1998. Complete genome sequence of *Treponema pallidum*, the syphilis spirochete. *Science* **281**:375-388.
12. **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.
13. **Henckel, T., U. Jackel, S. Schnell, and R. Conrad.** 2000. Molecular analyses of novel methanotrophic communities in forest soil that oxidize atmospheric methane. *Appl Environ Microbiol* **66**:1801-1808.

14. **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.
15. **Hongoh, Y., V. K. Sharma, T. Prakash, S. Noda, T. D. Taylor, T. Kudo, Y. Sakaki, A. Toyoda, M. Hattori, and M. Ohkuma.** 2008. Complete genome of the uncultured Termite Group 1 bacteria in a single host protist cell. *Proc Natl Acad Sci USA* **105**:5555-5560.
16. **Hugenholtz, P., B. M. Goebel, and N. R. Pace.** 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol* **180**:4765-4774.
17. **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.
18. **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.
19. **Leaphart, A. B., and C. R. Lovell.** 2001. Recovery and analysis of formyltetrahydrofolate synthetase gene sequences from natural populations of acetogenic bacteria. *Appl Environ Microbiol* **67**:1392-1395.
20. **Ley, R. E., M. Hamady, C. Lozupone, P. J. Turnbaugh, R. R. Ramey, J. S. Bircher, M. L. Schlegel, T. A. Tucker, M. D. Schrenzel, R. Knight, and J. I.**

- Gordon.** 2008. Evolution of mammals and their gut microbes. *Science* **320**:1647-1651.
21. **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.
22. **Lovell, C. R., and A. B. Leaphart.** 2005. Community-level analysis: key genes of CO₂-reductive acetogenesis. *Methods Enzymol* **397**:454-469.
23. **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. W. 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. H. Schleifer.** 2004. ARB: a software environment for sequence data. *Nucl Acids Res* **32**:1363-1371.
24. **Marcy, Y., C. Ouverney, E. M. Bik, T. Lösekann, N. Ivanova, H. G. Martin, E. Szeto, D. Platt, P. Hugenholtz, D. A. Relman, and S. R. Quake.** 2007. Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proc Natl Acad Sci USA* **104**:11889-11894.
25. **Nocker, A., M. Burr, and A. K. Camper.** 2007. Genotypic microbial community profiling: a critical technical review. *Microb Ecol* **54**:276-289.

26. **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.
27. **Pace, N. R.** 1997. A molecular view of microbial diversity and the biosphere. *Science* **276**:734-740.
28. **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 the spirochetes. *J Bacteriol* **173**:6101-6109.
29. **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.
30. **Pester, M., and A. Brune.** 2007. Hydrogen is the central free intermediate during lignocellulose degradation by termite gut symbionts. *Isme J* **1**:551-565.
31. **Pierce, E., G. Xie, R. D. Barabote, E. Saunders, C. S. Han, J. C. Detter, P. Richardson, T. S. Brettin, A. Das, L. G. Ljungdahl, and S. W. Ragsdale.** 2008. The complete genome sequence of *Moorella thermoacetica* (f. *Clostridium thermoaceticum*). *Environ Microbiol* **10**:2550-2573.
32. **Riesenfeld, C. S., P. D. Schloss, and J. Handelsman.** 2004. Metagenomics: Genomic analysis of microbial communities. *Ann Rev Genet* **38**:525-552.
33. **Rose, T. M., E. R. Schultz, J. G. Henikoff, S. Pietrokovski, C. M. McCallum, and S. Henikoff.** 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res* **26**:1628-1635.

34. **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*. *Microbiol* **149**:2529-2537.
35. **Seshadri, R., G. S. Myers, H. Tettelin, J. A. Eisen, J. F. Heidelberg, R. J. Dodson, T. M. Davidsen, R. T. DeBoy, D. E. Fouts, D. H. Haft, J. Selengut, Q. Ren, L. M. Brinkac, R. Madupu, J. Kolonay, S. A. Durkin, S. C. Daugherty, J. Shetty, A. Shvartsbeyn, E. Gebregeorgis, K. Geer, G. Tsegaye, J. Malek, B. Ayodeji, S. Shatsman, M. P. McLeod, D. Smajs, J. K. Howell, S. Pal, A. Amin, P. Vashisth, T. Z. McNeill, Q. Xiang, E. Sodergren, E. Baca, G. M. Weinstock, S. J. Norris, C. M. Fraser, and I. T. Paulsen.** 2004. Comparison of the genome of the oral pathogen *Treponema denticola* with other spirochete genomes. *Proc Natl Acad Sci USA* **101**:5646-5651.
36. **Sogin, M. L., H. G. Morrison, J. A. Huber, D. Mark Welch, S. M. Huse, P. R. Neal, J. M. Arrieta, and G. J. Herndl.** 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci USA* **103**:12115-12120.
37. **Stepanauskas, R., and M. E. Sieracki.** 2007. Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proc Natl Acad Sci USA* **104**:9052-9057.
38. **Streit, W. R., and R. A. Schmitz.** 2004. Metagenomics--the key to the uncultured microbes. *Curr Opin Microbiol* **7**:492-498.

39. **Suzuki, M. T., L. T. Taylor, and E. F. DeLong.** 2000. Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl Environ Microbiol* **66**:4605-4014.
40. **Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Neelson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y. H. Rogers, and H. O. Smith.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**:66-74.
41. **Wagner, M., A. Loy, M. Klein, N. Lee, N. B. Ramsing, D. A. Stahl, and M. W. Friedrich.** 2005. Functional marker genes for identification of sulfate-reducing prokaryotes. *Methods Enzymol* **397**:469-489.
42. **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**:560-565.
43. **Winston, P. W., Bates, Donald H.** 1960. Saturated solutions for the control of humidity in biological research. *Ecology* **41**:232-237.

44. **Wu, M., Q. H. Ren, A. S. Durkin, S. C. Daugherty, L. M. Brinkac, R. J. Dodson, R. Madupu, S. A. Sullivan, J. F. Kolonay, W. C. Nelson, L. J. Tallon, K. M. Jones, L. E. Ulrich, J. M. Gonzalez, I. B. Zhulin, F. T. Robb, and J. A. Eisen.** 2005. Life in hot carbon monoxide: The complete genome sequence of *Carboxydotherrmus hydrogenoformans* Z-2901. *PLoS Gen* **1**:563-574.
45. **Zehr, J. P., B. D. Jenkins, S. M. Short, and G. F. Steward.** 2003. Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ Microbiol* **5**:539-554.
46. **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.
47. **Zeidner, G., C. M. Preston, E. F. Delong, R. Massana, A. F. Post, D. J. Scanlan, and O. Béjà.** 2003. Molecular diversity among marine picophytoplankton as revealed by psbA analyses. *Environ Microbiol* **5**:212-216.
48. **Zhang, K., A. C. Martiny, N. B. Reppas, K. W. Barry, J. Malek, S. W. Chisholm, and G. M. Church.** 2006. Sequencing genomes from single cells by polymerase cloning. *Nat Biotechnol* **24**:680-686.
49. **Zhang, Y., D. Zhang, W. Li, J. Chen, Y. Peng, and W. Cao.** 2003. A novel real-time quantitative PCR method using attached universal template probe. *Nucleic Acids Res* **31**:e123.

Conclusions

This work focused on the biology and community structure of CO₂-reducing acetogens in the guts of termites. Using the gene for formyl-tetrahydrofolate synthetase (FTHFS) as a genetic marker of acetogenic capability, I explored the diversity of uncultured acetogens present in wood-feeding roaches and diverse termite species. I also used this symbiosis as a platform for the development of microfluidic techniques that allowed molecular characterization of single bacterial cells.

The best-known and longest-studied acetogens are bacteria associated with the phylum *Firmicutes*. These bacteria are widespread in the environment, and can be found even in ecosystems where acetogenesis is not a major H₂ sink. Acetogenic spirochetes, though dominant in the guts of wood-feeding termites, have been found nowhere else on Earth.

In chapter two of this thesis, I present the discovery that the guts of wood-feeding roaches, like those of wood-feeding lower termites, are dominated by acetogenic *Treponemes*. Phylogenetic analysis of roach-derived FTHFS types reveal a cluster of *Treponeme*-like FTHFS genes that represent a basal radiation to the termite *Treponeme* cluster. This suggests that they represent the modern descendants of an ancient divergence, and can be taken as evidence that acetogenic *Treponemes* were present in the last common ancestor of termites and roaches.

In chapter three, I present FTHFS community profiles of higher termite guts. Previous examinations of termite gut acetogens focused on wood-feeding lower termites. To complement these studies, I examined FTHFS genes present in the gut of a wood-feeding

higher termite species (*Nasutitermes* sp. Cost003). This termite was found to be dominated by termite gut *Treponemes*, as were the guts of a palm-feeding *Microcerotermes* sp. and litter-feeding *Rhynchotermes* sp. The story changed, however, when three subterranean termite species were examined. The guts of these species are dominated by a novel group of *Firmicute*-like FTHFS types. Rates of acetogenesis and methanogenesis were not measured for these species, and the exact composition of their diets remains unknown. However, the subterranean lifestyle of these termites suggests a higher degree of exposure to soil. Soil-feeding termites host robust populations of acetogenesis-capable bacteria but generally have low rates of CO₂-reductive acetogenesis. The dramatic alteration of acetogen population structure in subterranean termites (as opposed to wood-feeding termites) suggest environmental conditions that favor acetogenic *Firmicutes* over *Treponemes*. This shift in population structure seems likely to be related to the low rates of CO₂-reductive acetogenesis observed in soil-feeding termites.

Chapters four and five present the development of techniques for microfluidic PCR-based techniques for multiplex PCR from single cells. We developed these techniques in order to facilitate the species-level identification of uncultured acetogens. Using microfluidic devices, we carried out multiplex PCR on hundreds of individual environmental bacteria in parallel. PCR product retrieval and characterization allowed the establishment of FTHFS and 16S rRNA gene pairs derived from uncultured bacteria.

The ability to establish 16S rRNA sequence identities of uncultured, FTHFS-bearing bacteria opens a whole new window into the biology of termite gut acetogens. Much of the information derived from molecular community assays is based on hypotheses derived from phylogenetic inference. However, phylogenetic inference should be taken as circumstantial evidence at best, particularly as regards metabolic genes such as FTHFS. A phylogenetic inference is only as good as your closest cultured representative, which in the case of termite gut *Treponemes* consists of a grand total of two gene sequences (from *T. primitia* ZAS-1 and ZAS-2). Using the microfluidic digital PCR techniques presented herein, it is now possible to establish the species identities of uncultured FTHFS-bearing bacteria.

The first targets for microfluidic digital PCR characterization of environmental acetogens should be the novel sequence clusters we have identified in wood-feeding roaches and higher termites. The evolutionary hypotheses presented in chapter two would be greatly strengthened by definitive evidence that the FTHFS sequences discovered indeed belong to acetogenic *Treponemes*. Of particular interest is the basal “roach group III” cluster, for which the phylogenetic evidence of *Treponemal* derivation is weakest. If this sequence cluster does indeed represent acetogenic spirochetes, it will be interesting to discover where they fall within the termite *Treponeme* 16S rRNA cluster. If these sequences belong to a non-spirochetal organism, it will most likely represent the bacterial lineage from which acetogenic *Treponemes* acquired their FTHFS gene.

The most interesting target for microfluidic digital PCR in higher termites is the “Amitermes clade” of FTHFS sequences that dominates the guts of subterranean higher termites. This group was hypothesized to represent acetogenic *Firmicutes*, as the most closely related cultured organism was *Ruminococcus productus*. However, the distances involved are at least as great as those between *R. productus* and the termite *Treponeme* cluster. An exciting alternate hypothesis would be that this represents a novel lineage of termite *Treponemes* that arose following a lateral gene acquisition of FTHFS from a different acetogenic *Firmicute*.

In summary, this work presents new insights into the evolutionary history of the symbiosis between termites and CO₂-reducing acetogens and the relationship between host diet and acetogen community structure. Furthermore, it presents new microfluidics-based techniques for molecular characterization of uncultured, environmental bacteria. However, the work is not done. The microfluidic approach we developed has great power to expand our understanding of the novel acetogens discovered in studies of acetogen community structure. Furthermore, the stage is now set for expansion into many avenues of scientific research. Ongoing research in this laboratory involves the use of microfluidics for single cell whole genome amplification; any of our newly discovered acetogenic bacteria would make interesting targets for this approach. The dominance of non-spirochetal acetogens in the guts of subterranean targets suggests that these environments might be good targets for cultivation-based characterizations, as attempts to cultivate acetogenic *Firmicutes* have in the past proven more fruitful than those targeting acetogenic *Spirochetes*. Finally, the microfluidic digital PCR approach we have

developed can and should be utilized for other molecular community assays. Likely targets within termite guts include bacterial cellulases discovered in wood-feeding higher termites and genes involved in bacterial nitrogen fixation. Likely targets in other environments include genes involved in sulfate reduction and methanotrophy, both of which feature large clusters of sequences identified in molecular community analyses that cannot be classified based on comparison to cultured strains.