

- I. Formate dehydrogenase gene diversity in
lignocellulose-feeding insect gut microbial
communities**

- II. Metabolic impacts on the hydrogen isotope content of
bacterial lipids**

Thesis by

Xinning Zhang

In Partial Fulfillment of the Requirements for the
degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2010

(Defended April 15, 2010)

© 2010

Xinning Zhang

All Rights Reserved

To my parents,

Xin and Jiajun

ACKNOWLEDGEMENTS

First, I would like to recognize my parents, Xin and Jiajun. This thesis is dedicated to them. There are really no words I can use to express the appreciation and admiration I have for them as parents and human beings. I credit my father for setting me on the path I am on today. From early on, he instilled in me the value of education and emphasized the importance of finding my talents and doing something meaningful with them. I thank my mother for being a great role model, for always encouraging me to follow my dreams, for never giving up on me (especially during those tumultuous teenage years), for always making time for me, and for so much more. I also want to thank my sister, Doris, who has been a consistent source of level-headed advice and encouragement throughout my life.

My mentors at Cornell deserve to be recognized for introducing me to environmental science research. Monroe Weber-Shirk taught me that a little sludge goes a long way, that research doesn't have to utilize fancy instruments (a pressure sensor, filter, and mud work quite well!), and that research doesn't have to cure cancer to be worthwhile. I thank Beth Ahner for believing in my abilities, for encouraging me to apply to Caltech, and for being a great role model for women in science.

I am also grateful to the people I've had the pleasure to learn from and get to know as colleagues and friends over the past 6 years at Caltech.

My advisor, Jared Leadbetter, has and continues to inspire me with his passion for science and his care for the people around him. I was very lucky to have been introduced to

microbiology by Jared. He truly has a gift for microbial enrichment and isolation; I would not be at all surprised if he were known someday as a “microbe whisperer.”

I give credit to Eric Matson for his superb mentorship at the bench over the past 6 years. I really can think of no other person whom I regard as highly with respect to work-ethic and skill at performing meticulous research. Besides being a great role model in science, Eric has also been a good friend over the years. I thank him for all this and more.

Adam Rosenthal is another lab member who deserves a lot of my thanks as he has, in the past year at Caltech, made important impacts on my life as a burgeoning scientist. Adam’s curiosity and enthusiasm for new molecular biology techniques pushed me (or dragged me?) forward to complete the microfluidic work presented in Chapter 4. This chapter contains our combined efforts on a “wouldn’t it be cool...” project thought up during regular coffee sessions.

And I shouldn’t forget Liz Ottesen, the master of the Biomark. Thanks to Liz, I may now be able to design primers for extraterrestrial DNA (or whatever they use to store information). Other members of the Leadbetter lab (past and present) whom I like thank include: Jean Huang, Yajuan Wang, Jian-Yuan Thum, Kasia Gora, Abbie Green, Nick Ballor, and Kaitlyn Lucey. Each has made a positive impact on my life.

I owe a debt of gratitude to numerous people outside the Leadbetter lab at Caltech and around the world. These people include (in no particular order): the Paulot family, Jessamay Howell, Ashley Jones, Gretchen Aleks, Xavier Levine, and Sky Rashby.

With respect to ESE faculty, I thank Alex Sessions for not giving up on our work together when I flooded his lab during my first year at Caltech and for always being available to chat. Dianne Newman is another faculty member who has made a big impact on my life, both scientifically and personally. I will always remember her thoughtfulness after my dad's passing. The rest of the ESE & GPS community has also been influential in my growth as a scientist and person. I couldn't imagine a better atmosphere for research and random musings on life.

Last, but definitely not least, un grand merci à Fabien Paulot. Had I not come to Caltech and had you not decided to venture abroad, we would have never met and I would not be the person I am today.

ABSTRACT

I.) Symbiotic CO₂-reducing acetogens are important bacterial members of lignocellulose-feeding termite and roach gut communities. Acetogens are the major consumers of H₂ derived from lignocellulose fermentation and can contribute up to 1/3 of the acetate that serves as fuel for the insect host. Many acetogens in wood-feeding termites belong to a diverse group of relatively unstudied, uncultured spirochetes within the genus *Treponema*. Here, we used the gene sequence for hydrogenase-linked formate dehydrogenase (FDH_H), an enzyme utilized in sugar fermentation and the acetogenic metabolism of the isolate *T. primitia*, to investigate the diversity, evolution, and activity of uncultured acetogenic spirochetes in lignocellulose-feeding insect guts. To study diversity and evolution, we developed novel degenerate primers for FDH_H genes and constructed gene inventories from the gut communities of taxonomically and nutritionally diverse termites and a wood-feeding roach. Phylogenetic analyses reveal that most genes group with those from *T. primitia*, forming two clades that encode selenium-dependent (Sec) and selenium-independent (Cys) enzymes, respectively. This result implies many uncultured acetogenic spirochetes encode FDH_H genes. Phylogenetic patterns also imply FDH_H gene pool composition between lower and higher termite taxa and termites with different lifestyles varies greatly. We interpret differences as shifts in acetogenic spirochete community structure that occurred during termite evolution. We then investigated activity by sequencing the gut community transcriptome of a termite using high-throughput sequencing techniques and mapping transcript reads to gene inventory and pure culture

data. We discover that FDH_H gene transcription is dominated by relatively few FDH_H phylotypes. Finally, we performed microfluidic digital PCR on gut bacteria to determine the specific 16S rRNA ribotypes of spirochetes that encode FDH_H genes. We report the ribotypes of transcriptionally active spirochetes herein. Our results have implications for the microbial ecology of uncultured acetogenic spirochetes. We suggest that (i) the trace element selenium has shaped the gene content of acetogenic spirochetes in gut communities over evolutionary time scales, (ii) acetogenic spirochete populations have undergone extinctions and radiations associated with an evolutionary bottleneck, convergent evolutions, and possibly even invasion during termite evolution, and (iii) termite gut acetogenesis is largely mediated by only a few spirochete species, which represent a small portion of total acetogenic spirochete diversity.

II.) The hydrogen-stable isotope compositions (D/H) of lipids in the environment vary greatly. All variations have been assumed to result from changes in the D/H of water, a source of lipid hydrogen. However, several studies suggest that water D/H may not be the only influential factor. In this study we report that lipid D/H values can vary by 500‰ in bacterial cultures despite constant water D/H. This indicates variations in lipid/water fractionation need to be considered when interpreting environmental data. More significantly, we demonstrate that lipid D/H values are systematically related to the utilization of different central metabolic pathways in bacteria. Our results suggest that different cellular mechanisms for NADPH synthesis result in lipids with characteristic D/H. We therefore propose that lipid D/H values may be useful isotopic markers of energy metabolism.

TABLE OF CONTENTS

Acknowledgements	iv
Abstract	vii
Table of Contents	ix
List of Figures	xi
List of Tables	xiii
Chapter 1: Introductory Background	1-1
1.1 Preface to topics.	1-1
1.2 Introduction to formate dehydrogenase in insect guts.....	1-3
1.2.1 Early studies of acetogenesis	1-4
1.2.2 Acetogenesis pathway	1-5
1.2.3 Acetogenesis from H ₂ and CO ₂	1-7
1.2.4 Acetogen phylogeny.....	1-8
1.2.5 Acetogen ecology.....	1-9
1.2.6 Ecological impacts of termite gut acetogenesis.....	1-10
1.2.7 Termite biology.....	1-11
1.2.8 Termite gut microbiology	1-15
1.2.9 Termite nutritional ecology.....	1-19
1.2.10 Termite gut acetogens.....	1-24
1.2.11 Formate dehydrogenase in a termite gut acetogen isolate	1-25
1.2.12 Overview of Chapters 2 - 4	1-27
1.3 Introduction to bacterial lipid D/H and metabolism.....	1-29
Overview of Chapter 5.....	1-29
1.4 References.....	1-30
Chapter 2: Formate dehydrogenase gene diversity in acetogenic gut communities of lower, wood-feeding termites and a wood-feeding roach	2-1
Abstract	2-1
Introduction	2-3
Materials and Methods	2-6
Results.....	2-14
Discussion	2-30
Appendix	2-35
References.....	2-47
Chapter 3: Formate dehydrogenase gene phylogeny in higher termites suggests gut microbial communities have undergone an evolutionary bottleneck, convergent evolution, and invasion	3-1
Abstract	3-1
Introduction	3-3

Materials and Methods	3-6
Results	3-12
Discussion	3-34
Appendix	3-41
References	3-56
Chapter 4: RNA-Seq and microfluidic digital PCR identification of transcriptionally active spirochetes in termite gut microbial communities.....	4-1
Abstract	4-1
Introduction	4-3
Materials and Methods	4-6
Results	4-14
Discussion	4-27
Appendix	4-30
References	4-41
Chapter 5: Large D/H variations in bacterial lipids reflect central metabolic pathways.....	5-1
Abstract	5-1
Introduction	5-3
Results	5-5
Discussion	5-17
Conclusions	5-24
Materials and Methods	5-25
References	5-28
Appendix	5-33
Chapter 6: Conclusions.....	6-1

LIST OF FIGURES

Number	Page
1.1. Wood-Ljungdahl pathway for acetogenesis.....	1-6
1.2. Termite family phylogeny	1-14
1.3. Carbon and energy flow in wood-feeding termites.....	1-21
2.1. Degenerate <i>fdhF</i> primer design.....	2-10
2.2. Phylogenetic analysis of FDH from bacterial isolates	2-17
2.3. Predicted SECIS-like elements for inventory sequences	2-22
2.4. Phylogenetic analysis of lower termite FDH _H sequences.....	2-24
2.5. Phylogenetic analysis of Gut spirochete group FDH _H : termites.....	2-26
2.6. Phylogenetic analysis of Gut spirochete group FDH _H : termites and roach.....	2-29
2.7. Phylogenetic analysis of COII from lower termites and roaches	2-45
2.8. Rarefactions curves for lower termite and roach <i>fdhF</i> inventories	2-46
3.1. Phylogenetic analysis of COII from higher termites, lower termites, and roaches.....	3-17
3.2. Phylogenetic analysis of major FDH _H sequence types.....	3-20
3.3. Phylogenetic analysis of Gut spirochete group FDH _H	3-23
3.4. Detailed phylogenetic analysis of higher termite spirochete group FDH _H	3-26
3.5. Phylogenetic analysis of <i>Amitermes-Gnathamitermes-Rhynchotermes</i> clade FDH _H	3-29
3.6. Phylogenetic analysis of Enteric <i>Proteobacteria</i> clade and Unclassified clade FDH _H	3-31
3.7. Principal component analysis of FDH _H phylogeny.	3-33
3.8. Cys-clade specific PCR on insect gut DNA.	3-42
3.9. <i>Amitermes-Gnathamitermes-Rhynchotermes</i> clade specific PCR.	3-43

4.1. Gene inventory, RNA-Seq, and microfluidic PCR work-flow.	4-15
4.2. Microfluidic digital PCR: All Bacteria and Spirochete 16S rRNA.....	4-22
4.3. Phylogenetic analysis of 16S rRNA, <i>fdhF</i> _{Sec} , and <i>fdhF</i> _{Cys} sequences amplified using microfluidic digital PCR.	4-26
5.1. Summary of D/H fractionations.	5-9
5.2. Regressions of δD values versus water for <i>C. oxalaticus</i>	5-12
5.3. Fractionation factor curves for palmitic acid in <i>C. oxalaticus</i>	5-14
5.4. Representative bacterial growth curves.....	5-43
5.5. Relationship between fatty acid and water δD	5-45
5.6. Fractionation curves for hypothetical sets of cultures	5-46
5.7. Fatty acid biosynthetic pathway	5-47
5.8. Summary of major central metabolic pathways highlighting important sources of NADPH	5-49
5.9. NADPH generation in the Embden–Meyerhoff–Parnas Pathway	5-51
5.10. NADPH generation in the oxidative pentose phosphate pathway and the Entner–Doudoroff pathway.....	5-52
5.11. NADPH generation in the glycerate pathway.....	5-52
5.12. NADPH generation in the tricarboxylic acid cycle and the glyoxylate shunt.....	5-53
5.13. NADPH generation from transhydrogenase and the light reactions of photosynthesis.....	5-54

LIST OF TABLES

Number	Page
2.1.	PCR primers for <i>fdhF</i> -type formate dehydrogenases 2-9
2.2.	Sequences used for <i>fdhF</i> primer design..... 2-36
2.3.	Primer combinations for PCR amplifications 2-38
2.4.	Phylotype distribution in lower termite and roach FDH _H libraries..... 2-39
2.5.	Summary of lower termite and roach FDH _H libraries 2-41
2.6.	Alignment of Gut spirochete group amino acid indel..... 2-42
3.1.	Description of higher termites 3-13
3.2.	Summary of <i>fdhF</i> inventories from higher termites 3-18
3.3.	Clone distribution amongst major clades..... 3-21
3.4.	PCR conditions for <i>fdhF</i> inventory construction..... 3-44
3.5.	Phylotype distribution in higher termite FDH _H libraries..... 3-45
3.6.	Sequences used in COII and FDH _H phylogenetic analyses..... 3-48
4.1.	Primer and probes used for microfluidics and inventory analyses 4-13
4.2.	<i>Z. nevadensis</i> gut community FDH gene transcription: RNA-Seq, cDNA inventory, and qRT-PCR 4-19
4.3.	Microfluidic chip retrieval frequencies for <i>fdhF</i> 4-23
4.4.	RNA-Seq nucleotide scaffold data set 4-31
4.5.	Microfluidic digital PCR experiment details 4-40
5.1.	Summary of bacterial D/H culture experiments..... 5-7
5.2.	Relative abundances of fatty acids in bacterial cultures 5-38
5.3.	δD values of fatty acids and culture water..... 5-40
5.4.	Coefficients for regression of R_l on R_w 5-42

