CATECHOL 2,3-DIOXYGENASE-ASSISTED CLEAVAGE OF AROMATICS BY "ANAEROBIC" TERMITE GUT SPIROCHETES AND GENOMIC EVIDENCE OF A COMPLETE *META*-PATHWAY

Thesis by

Kaitlyn Shae Lucey

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2014

(Defended June 13, 2013)

ACKNOWLEDGMENTS

As a young child I grew up in South Boston, or *Southie*. My family and I lived in a brick house that stood on the top of "Pill Hill" (named for the doctors and dentists who built homes and practiced there around the turn of the last century) at the center of the main street of the neighborhood, Broadway. Depending on a Westerly or Easterly direction, I could follow Broadway to the public housing projects or to Boston proper – an epicenter of cultural, artistic, and scholarly diversity. My house was a pivotal point from which my life began, and my early childhood in *Southie* left an indelible mark motivating me to continually challenge myself and strive for better while simultaneously giving back.

Accordingly, first and foremost I would like to thank my mother, father, and younger brother, Gerald, who have believed in my potential, given me the right balance of guidance and freedom to choose my own path, and been genuinely invested in my success from the start. Members of my extended family have also been unparalleled role models and supporters, and I would like to thank my grandparents, William Spencer and Janet Van Dyke, my late grandmother, Betty Lucey, my Godfather, Michael Owens, and my Aunt Carole, Uncle Jim, Aunt Lynn, Uncle Greg, Uncle Dave, and Dave Roberts.

Beginning at Wellesley College, I embarked upon a path in the sciences in hopes of both contributing solutions to the STEM-related challenges our society and

environment is facing, and inspiring others to engage in STEM also. My professors and mentors at Wellesley College, specifically Professor Dave Ellerby and Professor Mary Allen, were integral in helping me initiate my path in the sciences. I would also like to thank Wellesley College, in general, for focusing on undergraduate learning, including a laboratory component with nearly every science class, and providing undergraduates with original research opportunities and prominent responsibility in the lab. My undergraduate experience has played a significant role in where I am today and where I am going.

At Caltech, former lab members Liz Ottesen and Eric Matson also helped kick-star my research by sharing many molecular and physiological skill sets. I would also like to thank both Xinning Zhang and Adam Rosenthal who, in addition to sharing with me much knowledge, have been stellar mentors. And of course, I would like to thank Jared Leadbetter. I have learned so much about being a good scientist and a good teacher working in your lab and with you. TA-ing Microbial Physiology, the Asilomar conference, and preparing my publication have been particularly significant learning experiences. I am grateful, too, to Victoria Oprhan, Sarkis Mazmanian, and Dianne Newman for serving on my thesis committee. It has been an absolute pleasure working with you and with students and staff in your labs, and I am thankful to have had these opportunities.

I also have developed some amazing friendships at Caltech that I hope will extend far beyond our time here. First, I would like to thank Elitza Tocheva, Nathan

Dalleska, and Felica Hunt for being such great mentors and role models, in addition to friends. I would also like to thank Chinny Idigo, Meg Schwamb, Mandy Grantz, and Kate Schilling for their friendship since practically day one here at Caltech, as well as my newest friends Molly and Luna for their encouraging and supportive purrs. Lastly, I would like to thank Los Angeles. From the majestic San Gabriel Mountains, to the unpredictable Pacific Coast, I am very fortunate to have spent the past five years in such an inspirational setting.

I am grateful for the top-notch education and laboratory research opportunities I have had to date, and will continue to strive to move the field of environmental microbial forward as well as teach and mentor the next generation of microbial ecologists. At the very least, I hope to instill in others a technological literacy and appreciation for science even if they do not also pursue STEM careers. Those who have been with me along this journey have been excellent teachers, mentors, scientists, and role models, and I aspire to serve others like you all have inspired me.

ABSTRACT

The termite hindgut microbial ecosystem functions like a miniature lignocellulose-metabolizing natural bioreactor, has significant implications to nutrient cycling in the terrestrial environment, and represents an array of microbial metabolic diversity. Deciphering the intricacies of this microbial community to obtain as complete a picture as possible of how it functions as a whole, requires a combination of various traditional and cutting-edge bioinformatic, molecular, physiological, and culturing approaches. Isolates from this ecosystem, including *Treponema primitia* str. ZAS-1 and ZAS-2 as well as *T. azotonutricium* str. ZAS-9, have been significant resources for better understanding the termite system. While not all functions predicted by the genomes of these three isolates are demonstrated *in vitro*, these isolates do have the capacity for several metabolisms unique to spirochetes and critical to the termite system's reliance upon lignocellulose. In this thesis, work culturing, enriching for, and isolating diverse microorganisms from the termite hindgut is discussed. Additionally, strategies of members of the termite hindgut microbial community to defend against O_2-stress and to generate acetate, the "biofuel" of the termite system, are proposed. In particular, catechol 2,3-dioxygenase and other *meta*-cleavage catabolic pathway genes are described in the "anaerobic" termite hindgut spirochetes *T. primitia* str. ZAS-1 and ZAS-2, and the first evidence for aromatic ring cleavage in the phylum (division) *Spirochetes* is also presented. These results suggest that the potential for O_2-dependent, yet non-respiratory, metabolisms of plant-derived aromatics should be re-evaluated in termite hindgut communities. Potential future work is also illustrated.

TABLE OF CONTENTS

Acknowledgments iii
Abstract vi
Table of Contents vii
List of Tables viii
List of Figures ix

Chapter 1: Research with termite hindgut microorganisms & background ... 1-1
Abstract 1-1
Introduction 1-2
Materials and Methods 1-16
Results 1-24
Discussion 1-34
References 1-39
Appendix 1-44

Chapter 2: Exploring genomic evidence of acetogenic demethylation by *Treponema azotonutricium* str. ZAS-9 2-1
Abstract 2-1
Introduction 2-2
Materials and Methods 2-9
Results 2-10
Discussion 2-17
References 2-23
Appendix 2-26

Chapter 3: Catechol 2,3-dioxygenase and other *meta*-cleavage catabolic pathway genes in the "anaerobic" termite gut spirochete *Treponema primitia* 3-1
Abstract 3-1
Introduction 3-2
Materials and Methods 3-4
Results 3-10
Discussion 3-38
References 3-49
Appendix 3-57

Chapter 4: Conclusions 4-1
References 4-12

LIST OF TABLES

Chapter 1

Table 1-1: Appearance of baker's yeast and appearance and odor of corresponding yeast autolysate and autolysate supernatant.

Table 1-2: Growth rates of *Treponema primitia* str. ZAS-2 on 4YACo medium prepared from yeast autolysates of different brands and packaging types.

Table 1-3: *Treponema primitia* str. ZAS-2 enzymatic defenses against oxidative stress.

Chapter 2

Table 2-1: Methylated compounds tested for growth on by *Treponema azotonutricium* str. ZAS-9.

Chapter 3

Table 3-1: *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-cleavage pathway genes and pfam features.

Table 3-2: Selection pressure analysis of *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-cleavage pathway genes.

Table 3-3: Monoaromatic compound cleavage by *Treponema primitia* str. ZAS-1.

Table 3-4: Average depth of O_2 penetration in *Treponema primitia* str. ZAS-1 and ZAS-2 O_2/aromatic gradient cultures.

LIST OF FIGURES

Chapter 1

Fig. 1-1: A phylogenetically "lower" termite, a dampwood termite *Zootermopsis nevadensis* worker, and a dissected intestinal tract of another *Z. nevadensis* worker specimen.

Fig. 1-2: Phylogram of "higher" and "lower" termite families as well as household and wood roaches.

Fig. 1-3: A phylogenetically "higher" termite, a *Gnathamitermes* sp. worker, and a dissected intestinal tract of another worker *Gnathamitermes* sp. specimen.

Fig. 1-4: An example of the structure of lignin.

Fig. 1-5: General scheme underlying the symbiosis between wood-feeding "lower" termites and their mutualistic hindgut microbiota.

Fig. 1-6: Radial profile of O_2 (and H_2) in the hindgut of the wood-feeding termite *Reticulitermes flavipes* (*figure from* Brune & Friedrich 2000).

Fig. 1-7: *Treponema azotonutricium* str. ZAS-9 **(a)** and *T. primitia* str. ZAS-2 **(b)**.

Fig. 1-8: Spirochete-like morphologies that dominated liquid enrichment cultures for acetogenic spirochetes. **(a)** "Floppy" morphology. **(b)** What would be isolated as *Treponema primitia* str. "ZNS-1."

Fig. 1-9: Proposed mechanisms for combating oxidative stress in the termite hindgut.

Fig. 1-10: Gene neighborhoods of enzymes involved in O_2 transformation.

Chapter 2

Fig. 2-1: Acetyl-CoA (Wood-Ljungdahl) pathway for CO_2-reductive acetogenesis.

Fig. 2-2: *Treponema azotonutricium* str. ZAS-9 gene neighborhoods with both putative methyltransferases and homologs of the acetyl-CoA (Wood-Ljungdahl) pathway of acetogenesis.

Fig. 2-3: Growth of *Treponema azotonutricium* str. ZAS-9 with various meth(ox)ylated compounds.

Chapter 3

Fig. 3-1: Catechol 2,3-dioxygenase-based *meta*-cleavage pathway of aromatic metabolism and shunts.

Fig. 3-2: Gene neighborhoods representing complete catechol 2,3-dioxygenase-based *meta*-cleavage pathways.

Fig. 3-3: (a) Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 catechol 2,3-dioxygenase (PF00903, step 1). **(b)** Example of MUSCLE alignment used to create trees.

Fig. 3-4: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 2-hydroxymuconic semialdehyde hydrolase (PF00561, step 2).

Fig. 3-5: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 2-oxopent-4-enoate hydratase (PF01557, step 3).

Fig. 3-6: Microoxic growth of *Treponema primitia* str. ZAS-1.

Fig. 3-7: *Treponema primitia* str. ZAS-1 O_2/catechol gradient cultures and controls.

CHAPTER 1

Research with termite hindgut microorganisms & background

ABSTRACT:

Microbial ecology research requires a combination of traditional and cutting-edge bioinformatic, culturing, physiological, and molecular, as well as *in vitro*, *in vivo*, and *in situ* approaches. Several of these techniques were utilized to obtain the results presented here, which provide a foundation for research discussed in subsequent chapters. First, to improve growth rates of the *Treponema* termite hindgut isolates that have been continuously passaged for over a decade in 4YACo medium, growth was evaluated on media prepared with yeast autolysate from different brands and packaging of baker's yeast. While doubling time on medium with Red Star yeast autolysate had increased from between 22 and 35hrs to between 54 and 77hrs since the *Treponema* isolates were initially obtained, growth rates on medium with Fleischmann's yeast autolysate markedly improved to 47hrs for *T. primitia* str. ZAS-2. Next, in an effort to obtain additional termite hindgut microbial community members in pure culture, successful enrichment cultures of methanogens were prepared initially as a primer in the enrichment and isolation procedure. Then a spirochete, *T. primitia* str. "ZNS-1," was isolated from the *Zootermopsis nevadensis* termite hindgut. This strain is morphologically similar to *T. primitia* str. ZAS-1 and ZAS-2, shares 100% 16S rRNA sequence identity with *T. primitia* str. ZAS-2 and 99% 16S rRNA sequence identity with *T. primitia* str. ZAS-1, and exhibits preliminary physiological evidence of H_2 + CO_2 acetogenesis. Efforts to isolate the termite hindgut microbial community member encoding the "ZnD2Sec" phylotyope, the

phylotype that is responsible for the majority of the formate dehydrogenase expression in the termite hindgut, however, were unsuccessful. Nevertheless, enrichment cultures targeting this organism were successfully used as templates for microfluidic digital PCR and helped correlate the "ZnD2Sec" phylotype with a deltaproteobacteria and not a spirochete as previously thought. Finally, *T. primitia* str. ZAS-1 and ZAS-2 lack hallmark mechanisms of dealing with oxidative stress, specifically catalase and superoxide dismutase. It is likely that they have other strategies and relationships with O_2 in their environment that are explored here and also relevant to work discussed in Chapter 3. Much of the research presented in this chapter has the potential for continued work in the future.

INTRODUCTION:

Termites and their mutualistic hindgut microbiota

Betraying their small size, termites are organisms with a big impact. Their influence on ecosystems arises from both their unique behavior patterns and their numerical abundance (Wood & Sands 1978). Photosynthetic fixation of CO_2 yields the earth's most abundant form of biomass, lignocellulosic plant material (Breznak & Brune 1994). Termites, as well as only a few other arthropods, have the unique ability to metabolize lignocellulose from living plants and plant materials at various stages of decomposition (Wood 1976; Brune 1998a). This includes wood, grasses, roots, and soil organic mater among other diverse substrates (Wood 1976; Brune 1998a). Moreover, while termites are entirely terrestrial and are restricted to within 45°N and S latitudes, arid regions, and altitudes under 3000m, they cover more than half of the world's land surface (Wood & Johnson 1986). Therefore, termites have been

recognized as ecosystem engineers, facilitating the decomposition of the planet's most abundant yet enzymatically-recalcitrant biopolymers and other aspects of soil function (Breznak & Brune 1994; Sugimoto *et al*. 2000).

Early studies investigating how termites metabolize lignocellulose into nutrients and energy observed that a considerable portion of the termite's biomass is devoted to its intestinal tract as is typical of animals that consume materials that are difficult to digest (Fig. 1-1) (Wood & Johnson 1986).

Fig. 1-1: A phylogenetically "lower" termite, a dampwood termite *Zootermopsis nevadensis* worker, and a dissected intestinal tract of another *Z. nevadensis* worker specimen. The termite intestinal tract represents a large portion of the insect's biomass (Wood & Johnson 1986). Specimen collected from San Gabriel Mountains, Pasadena, CA. Scale bar represents 25mm. Both images are to the same scale. (*Figure courtesy of the Leadbetter Lab*).

Within this intestinal tract, and especially in the largest and most dilated region, the hindgut, a complex, obligate, and nutritionally-symbiotic hindgut microbial

community was discovered (Fig. 1-1) (Leidy 1877; Cleveland 1926; Hungate 1955; Yamin & Trager 1979; Yamin 1980). Although termites are infamous for their ability to devour wood and do contribute enzymes to lignocellulose digestion, it is their complex microbial community comprising upwards of 10^6 microorganisms and representing over 200 species that is integral to metabolizing lignocellulose into nutrients and energy to support the termite system (Leidy 1877; Cleveland 1926; Hungate 1955; Yamin & Trager 1979; Yamin 1980). While the symbiotic association of termites with microorganisms comprises different levels of interaction, ranging from the extracorporal cultivation of fungus gardens to the most intimate associations where bacteria reside intracellularly, the majority of the prokaryotic symbionts of termites are located in the intestinal tract either free-living, attached to the gut epithelium, or associated with intestinal protozoa (Brune 2006).

As termites are recognized for their unusual ability to metabolize lignocellulose, the metabolic capabilities of these microorganisms are also unique. Together, termites and their hindgut microbiota are prime examples of miniature natural bioreactors in the terrestrial environment (Brune 1998a).

"Higher" and "lower" termite lineages

Termites belong to the order Isoptera consisting of over 2000 living species represented by seven families whose biology, behavior, and nutrition are diverse (Breznak & Brune 1994). Members of the first six families are referred to collectively as the "lower" termites and members of the seventh family are referred to as the "higher" termites (Fig. 1-2).

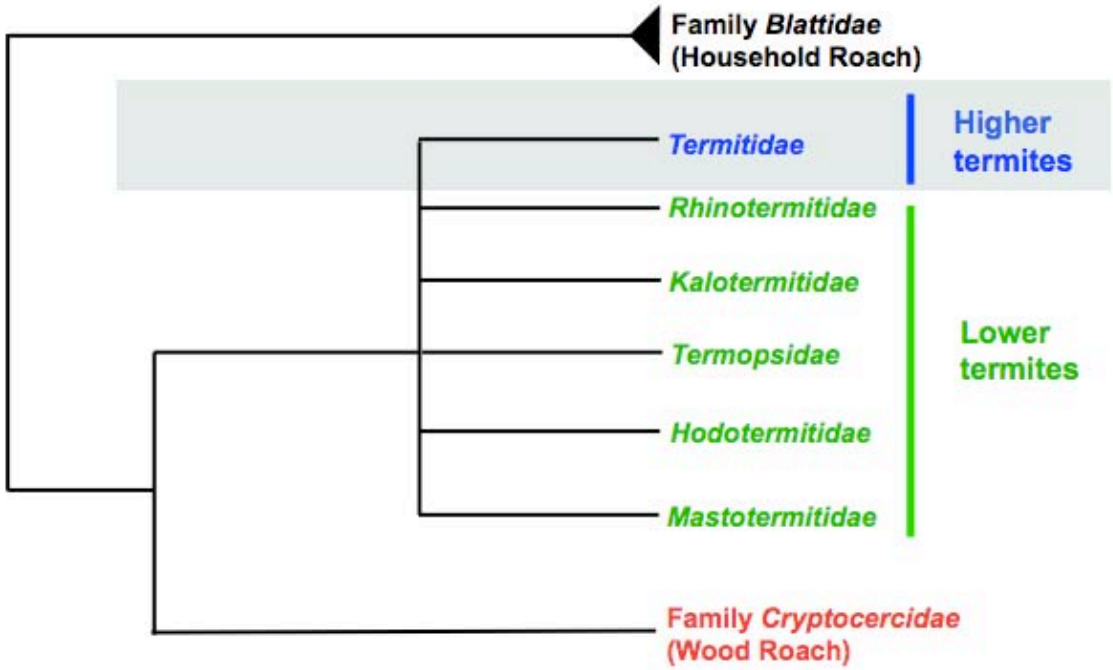


Fig. 1-2: Phylogram of "higher" and "lower" termite families as well as household and wood roaches. Phylogenetically "higher" termites, the most species-rich and abundant lineage of termites, are represented by one family, and "lower" termites are represented by six families. Much more is known about "lower" termites than "higher" termites (Wood & Johnson 1986; Breznak & Brune 1994). Tree adapted from, and based upon, phylogenetic analyses reported by Inward *et al*. 2007.

Although "higher" termites comprise only one family, they are the most species-rich and abundant of the termite lineages. "Higher" termites are more evolved than their "lower" termite counterparts (Fig. 1-2), and distinctions are also seen with respect to intestinal tract architecture (Figs. 1-2 and 1-3), gut microbial community composition, and food preference (Wood & Johnson 1986; Breznak & Brune 1994).

Fig. 1-3: A phylogenetically "higher" termite, a *Gnathamitermes* sp. worker, and a dissected intestinal tract of another worker *Gnathamitermes* sp. specimen. The intestinal tracts of "higher" termites are more compartmentalized than those of "lower" termites (Wood & Johnson 1986; Breznak & Brune 1994). Specimen collected from subterranean nests at Joshua Tree National Park, CA (Ottesen & Leadbetter 2011). Scale bar represents 25mm. Both images are to the same scale. (*Figure courtesy of the Leadbetter Lab*).

First, the gut architecture of "higher" termites is more complex than that of "lower" termites, with the intestinal tracts of "higher" termites being more compartmentalized than those of "lower" termites (Figs. 1-2 and 1-3). Also, the gut microbial community of "lower" termites consists of members representing all three domains of life, the bacteria, archaea, and eukarya such as cellulolytic oxymonad, trichomonad, and hypermastigote protozoa (Wood & Johnson 1986). In contrast, the gut microbial community of most "higher" termites is made up of bacteria and archaea only. These differences are also reflected in general diet trends with "lower" termites generally consuming lignocellulosic plant materials at early stages

of decomposition, and "higher" termites generally consuming lignocellulosic plant materials that have been modified (Wood & Johnson 1986). Some species of phylogenetically "higher" termites also metabolize lignocellulose, but without polysaccharide-fermenting protozoal symbionts (Wood & Johnson 1986; Warnecke *et al*. 2007). Much less is known about "higher" termites than "lower" termites, particularly concerning how they metabolize diverse lignocellulosic substrates.

Lignocellulose

Lignocellulose is the structural polymer in the cell walls and middle lamellae of most higher plants (Breznak & Brune 1994). This compound is also the earth's most abundant form of biomass, and represents a considerable reservoir of carbon. Soil invertebrates, such as termites, help cycle lignocellulose through physical dispersion and initial dissimilation of this polymer. Microorganisms, however, play the largest role in the cycling of lignocellulose via further degradation and, ultimately, respiration (Breznak & Brune 1994).

The metabolism of lignocellulose is nontrivial, as the compound's complex biochemical structure renders it difficult to degrade. The major constituents of lignocellulose are three polymers: cellulose, hemicelluloses, and lignin (Breznak & Brune 1994). In particular, cellulose is a highly-ordered, crystalline, homopolymer of β-linked glucose molecules. As a result of intra- and inter-molecular hydrogen bonding within the molecule, cellulose confers to lignocellulose inelasticity, tensile strength, and resistance to hydrolysis. Hemicellulose is a linear or branched heteropolysaccharide comprised mainly of D-xylose, D-mannose, L-arabinose,

and/or D-galactose. Hemicellulose covalently links with lignin and forms an intricate matrix that surrounds the orderly cellulose microfibrils and impedes the enzymatic degradation of the entire lignocellulose molecule. Lignin, the third and most resistant of the three components to enzymatic degradation, is an aromatic polymer consisting of non-repetitive phenylpropane subunits randomly lined by various C-C and ether bonds. Lignin is intimately interspersed, and covalently linked at various points, with hemicellulose to form a matrix surrounding the orderly cellulose microfibrils and to protect them from enzymatic hydrolysis (Fig. 1-4) (Breznak & Brune 1994).

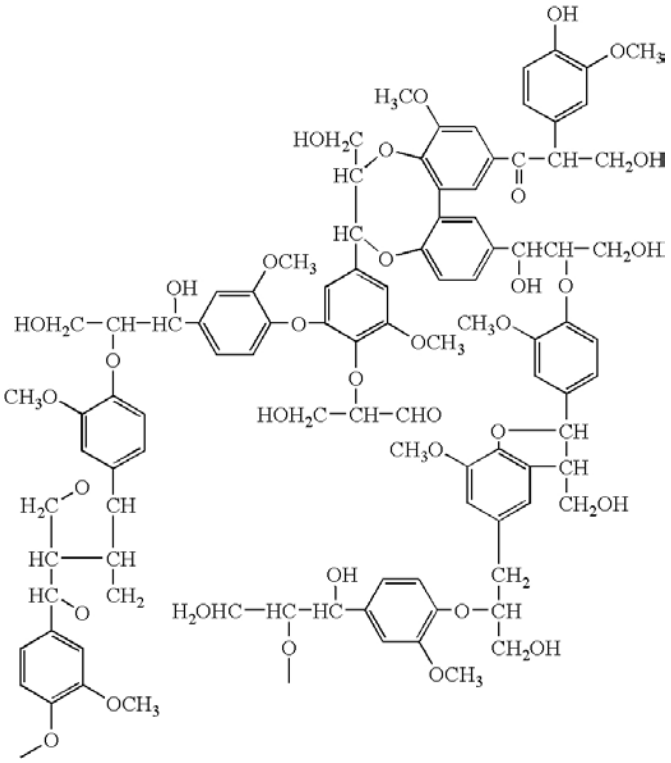


Fig. 1-4: An example of the structure of lignin. Lignin is an aromatic polymer consisting of non-repetitive phenylpropane subunits randomly linked by various C-C bonds. Lignin is also the component of lignocellulose most resistant to enzymatic degradation. This is due to the stability imparted by its aromatic rings (Breznak & Brune 1994) (*figure from* www.freepatentsonline.com).

The proportions of these three constituents vary depending on the type of plant material, and cellulose can represent anywhere from 30-50% of the lignocellulose molecule, with hemicellulose and lignin each constituting 20-30% (Breznak & Brune 1994). Woody plants contain a higher portion of lignin relative to other plant types, and there is an inverse correlation between the lignin content of plant material and its digestibility (Breznak & Brune 1994). The ability to modify and/or metabolize the lignin component of lignocellulose, therefore, is central to carbon turnover in the natural environment.

Lignocellulose metabolism by "lower" termites

Research with wood-feeding, phylogenetically "lower" termites has generated a near complete picture of how "lower" termites and their microbial symbionts collectively metabolize lignocellulose into nutrients and energy (Fig. 1-5).

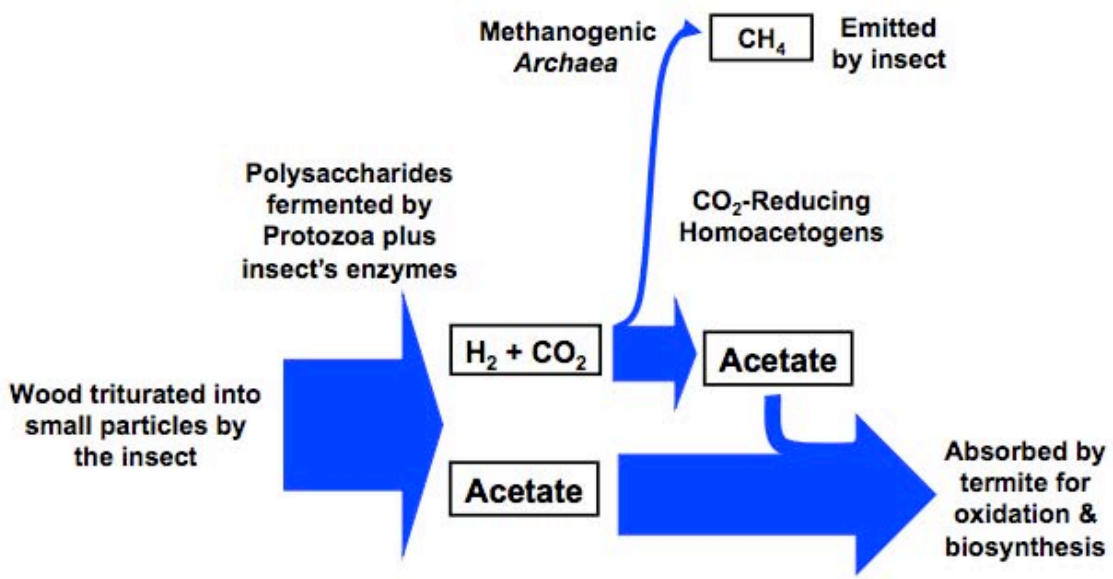


Fig. 1-5: General scheme underlying the symbiosis between wood-feeding "lower" termites and their mutualistic hindgut microbiota. Acetate is the "biofuel" of the termite system (Odelson & Breznak 1983; Breznak & Switzer 1986; Breznak 1994;

Leadbetter *et al*. 1999; Tholen & Brune 2000; Graber & Breznak 2004; Graber *et al*. 2004). First, symbiotic hindgut protozoa hydrolyze wood polysaccharides and ferment the resulting sugar monomers into acetate as well as CO_2 and H_2. Methanogenic archaea convert some this CO_2 and H_2 into methane, but the majority of this CO_2 and H_2 is converted into additional acetate by CO_2-reducing homoacetogens (Cleveland 1926; Hungate 1955; Yamin & Trager 1979; Yamin 1980; Odelson & Breznak 1983; Breznak & Switzer 1986; Breznak 1994; Leadbetter *et al*. 1999). The fate of lignin in the system, however, remains contentious (Butler & Buckerfield 1979; Cookson 1987; Pasti *et al*. 1990; Kuhnigk *et al*. 1994). (*Figure courtesy of the Leadbetter Lab*).

Initially, symbiotic hindgut protozoa hydrolyze wood polysaccharides, such as cellulose and hemicellulose, and ferment the resulting sugar monomers into acetate, CO_2, and H_2 according to the reaction:

$$C_6H_{12}O_6 + 2H_2O \rightarrow 2CH_3COOH + 2CO_2 + 4H_2$$

(Fig. 1-5) (Odelson & Breznak 1983; Breznak & Brune 1994; Brune 1998a). Next, prokaryotic microorganisms in the hindgut either use some of the resulting H_2 and CO_2 to produce methane (as is the case with methanogenic archaea), or convert most of that H_2 and CO_2 into an additional acetate molecule (as is the case with homoacetogenic bacteria such as spirochetes) according to the following:

$$4H_2 + 2CO_2 \rightarrow CH_3OOH + 2H_2O$$

(Fig. 1-5) (Brauman *et al*. 1992; Leadbetter *et al*. 1999). These products are then absorbed and oxidized by the termite host for both nutrition and energy (Odelson & Breznak 1983; Brauman *et al*. 1992; Brune 1998a). Acetate is an oxidizable energy source as well as a precursor of amino acids, hydrocarbons, and terpenes (Brauman *et al*. 1992).

Although the termite system has been found to dissimilate 75-99% of cellulose and 65%-90% of the hemicellulose, much less is known about the fate of lignin in the termite system, and research to that end is contentious (Butler & Buckerfield 1979; Cookson 1987; Pasti *et al*. 1990; Kuhnigk *et al*. 1994). Reliably measuring C^{14}-labelled lignins before, during, and after passage through the termite gut tract is challenging (Butler & Buckerfield 1979; Cookson 1987; Pasti *et al*. 1990; Kuhnigk *et al*. 1994). Many studies have hypothesized that O_2 is a necessary co-substrate for the complete degradation of lignin aromatic monomers (Brune 1998b).

Physical parameters of termite hindgut environment

Environmental conditions inside the termite intestinal tract affect microbial metabolism and lignocellulose digestion (Zimmer & Brune 2005). Likewise, microbial activity also partly determines physiological gut conditions such as pH level (Appel 1993; Zimmer & Topp 1997), redox potential (Bignell 1984; Kappler & Brune 2002), and oxygen concentration (Bignell 1984; Brune *et al.* 1995a and b). Combined, physical parameters innate to the termite intestinal tract and imparted by the inhabiting complex microbial community may have adapted over evolutionary time for efficient metabolism of lignocellulose (Zimmer & Brune 2005).

Initial investigations into the physiochemical conditions of the gut deemed the environment an anoxic habitat in which anaerobic microorganisms ferment wood polysaccharides into nutrients and energy (Cleveland 1926; Hungate 1955; Brune 1998a). Nevertheless, a termite hindgut is surrounded by tissues aerated by the insect's tracheal system, and subsequent studies determined that as a result of its

small size and relatively large surface-area-to-volume ratio, not only is the hindgut environment partially oxic, but the majority of its volume, approximately 60%, contains some level of oxygen between 50 to 100μM at the hindgut epithelium before reaching anoxia within about 150 to 200μm of the wall (Fig. 1-6) (Brune *et al*. 1995a and b; Tholen *et al*. 1997; Brune 1998a and b; Graber & Breznak 2004; Zimmer & Brune 2005).

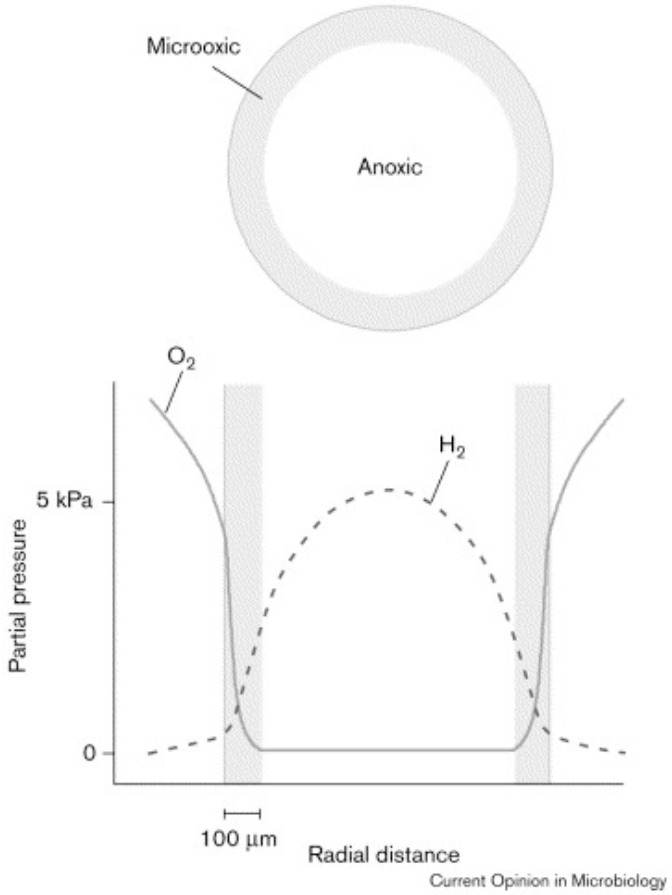


Fig. 1-6: Radial profile of O_2 (and H_2) in the hindgut of the wood-feeding termite *Reticulitermes flavipes* (*figure from* Brune & Friedrich 2000). Approximately 60% of the hindgut's volume contains some level of oxygen between 50 to 100μM at the hindgut epithelium before reaching anoxia within about 150 to 200μm of the wall. This characteristic has implications to microbial metabolisms in the hindgut (Brune *et al*. 1995a and b; Tholen *et al*. 1997; Brune 1998a and b; Graber & Breznak 2004; Zimmer & Brune 2005).

In particular, a radial oxygen gradient spans from the hindgut's periphery to its center, with oxygen concentrations near the hindgut epithelium 50 to 100μM and decreasing steeply to anoxia within about 150 to 200μm of the wall (Fig. 1-6) (Brune *et al*. 1995a). The discovery that the hindgut contains various levels of oxygen and only central portions of the hindgut are anoxic opened the possibility that the termite gut community consists of microorganisms with varying degrees of oxygen tolerance and metabolisms (Brune *et al*. 1995b; Tholen *et al*. 1997; Wenzel *et al*. 2002; Zimmer & Brune 2005; Wertz & Breznak 2007a; Wertz & Breznak 2007b).

More-recent studies of the gut microflora have confirmed the presence of large numbers of microorganisms with varying degrees of O_2-tolerance (Tholen *et al*. 1997; Wenzel *et al*. 2002; Wertz & Breznak 2007a; Wertz & Breznak 2007b). Given that products of incomplete reduction of O_2, such as superoxide radical ($O_2^{-\cdot}$), hydrogen peroxide (H_2O_2), hydroxyl radical ($HO\cdot$), and singlet oxygen ($^*O_2$), can seriously damage intracellular macromolecules as well as whole cells, some of the more aerotolerant members may possess enzymatic adaptations to combat oxidative stress and maintain anoxic conditions necessary for the survival and functioning of the microbial hindgut community (Brune 1998a and b; Brioukhanov & Netrusov 2007).

"Lower" termite hindgut isolates *Treponema primitia* str. ZAS-1 and ZAS-2 and *T. azotonutricium* str. ZAS-9

Spirochetes are highly motile, spiral or undulate bacteria and are one of the most abundant, consistently present, and morphologically distinct groups of prokaryotes in termite hindguts (Breznak & Brune 1994; Graber & Breznak 2004). The hindguts of termites, in fact, host a diversity of spirochetes unparalleled by any other habitat on earth (Leadbetter *et al.* 1999). Three spirochetes from phylogenetically "lower" termites have been isolated in pure culture: *Treponema primitia* str. ZAS-1 and ZAS-2 and *T. azotonutricium* str. ZAS-9, from which metabolic activities hitherto unknown to be performed by spirochetes have been found (Fig. 1-7a and b) (Leadbetter *et al*. 1999; Graber *et al*. 2004).

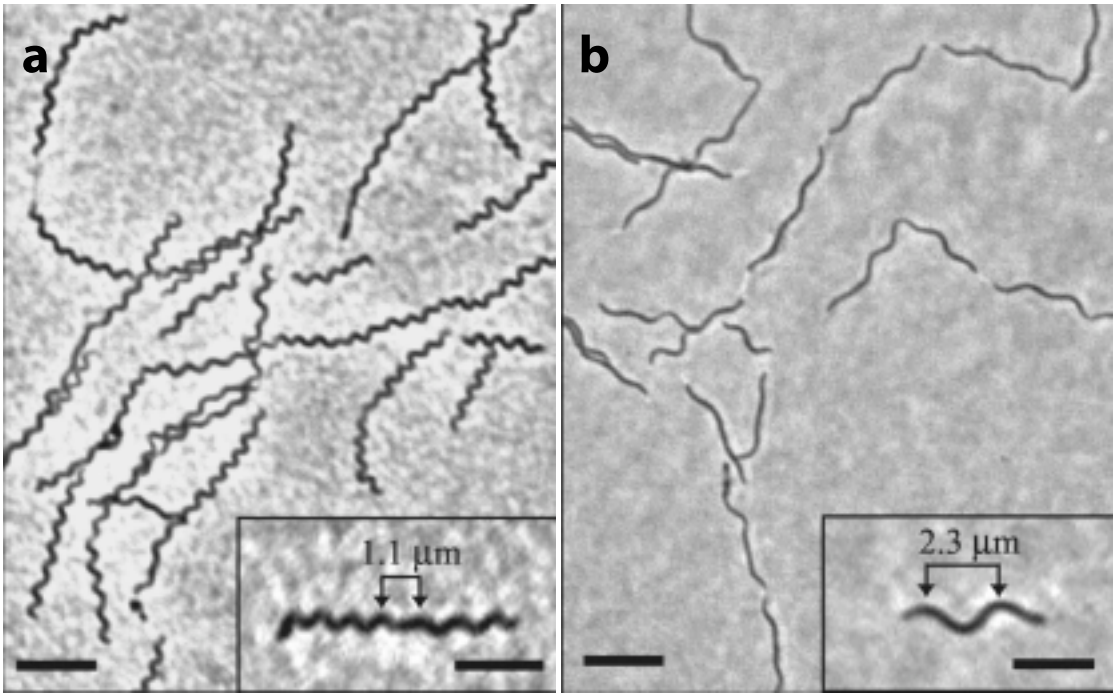


Fig. 1-7: *Treponema azotonutricium* str. ZAS-9 **(a)** and *T. primitia* str. ZAS-2 **(b)**. *T. primitia* str. ZAS-1 is morphologically indistinguishable from *T. primitia* str. ZAS-2. Scale bars represent 5µm (*figure from* Graber *et al*. 2004).

T. primitia str. ZAS-1 and ZAS-2 are able to derive acetate from H_2 plus CO_2 according to the Wood-Ljungdahl pathway of carbon fixation, and *T. azotonutricium* str. ZAS-9 is capable of N_2 fixation (Graber & Breznak 2004). Both of these processes are important in the delivery of carbon, nitrogen, and energy to termites (Odelson & Breznak 1983). Equally significant are the implications for lignocellulose degradation as well as carbon turnover in the environment as a result of these organisms' metabolisms (Warnecke *et al*. 2007).

Leveraging diverse laboratory tools to study complex microbial communities

It is widely known that the vast majority of microorganisms in the natural environment have yet to be obtained in pure culture. Consequently, in the last decades the microbiology field has witnessed an upsurge in the development of bioinformatic and molecular techniques and tools to better understand who these microorganisms are and how they function in their natural environments (Ottesen *et al*. 2006).

Multiplex microfluidic digital-PCR is an excellent tool developed in the Leadbetter Lab for linking functional genes to 16S rRNA genes for bacterial identification to better understand *who* in a mixed sample is capable of doing *what* (Ottesen *et al*. 2006). In addition, the 62Mb P3 hindgut region metagenome from the "higher" termite, *Nasutitermes* sp. has been a significant resource in investigating the functions potentially imparted by the mutualistic hindgut microbial community in the "higher" termite system (Warnecke *et al*. 2007). Moreover, this metagenomic dataset can be used to provide hints at obtaining microorganisms from the "higher"

termite in pure culture, as well as compared and contrasted against what is known about the "lower" termite system (Warnecke *et al*. 2007).

As technologies continue to develop, many researchers are combining various bioinformatic, molecular, physiological, and culturing tools to ask and answer *who, what, where, when, how, how much, with whom, and why* (Rosenthal *et al*. *in review*).

MATERIALS & METHODS:

Termite collection and storage

Worker specimens of the phylogenetically "lower" dampwood termite, *Zootermopsis nevadensis*, were collected in the San Gabriel Mountains, Pasadena, CA. In lab these specimens were maintained in plastic boxes with wood from which they were collected, and boxes were stored at room temperature, at 95% humidity, and in foil-covered glass aquaria.

Phylogenetically "higher" termite specimens were collected from Joshua Tree Natural Park, CA. In lab, these specimens were also maintained in plastic boxes with wood and/or soil from which they were collected, and boxes were stored at room temperature and humidity, and in the dark. If an isolate from the gut microbial community of one of these "higher" termites had been obtained, the termites would have been classified morphologically and their identity would have been determined via cytochrome oxidase gene sequencing (Warnecke *et al*. 2007).

Media and cultivation

Routine *in vitro* growth and maintenance of *Treponema primitia* str. ZAS-1 and ZAS-2 as well as *T. azotonutricium* str. ZAS-9, was in 5mL 4YACo liquid medium in 25mL butyl rubber-stoppered Balch tubes as previously described (Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004). After isolation, *T. primitia* str. "ZNS-1," (*see below*) was grown in the same media and under the same conditions as *T. primitia* str. ZAS-1 and ZAS-2. All cultures were incubated in the dark, at room temperature, in a horizontal position, and without agitation.

Yeast autolysate

While maximum growth yields of *Treponema primitia* str. ZAS-1 and ZAS-2 as well as *T. azotonutricium* str. ZAS-9, reaching between OD_{600nm} = 0.8 and 1.0 in 4YACo medium under normal culture conditions, had not changed markedly since their isolation growth rates had increased, and thus improving the growth rates of these isolates was the focus of the work presented here (Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004). In an effort to improve the growth rates of these isolates in 4YACo medium, growth rates were compared on media comprised of yeast autolysate prepared from different brands and different packaging types within the same brand of baker's yeast. Several yeast autolysate preparations were considered because yeast from different brands and different packaging types within the same brand looked and smelled different (Table 1-1). Growth rates of *T. primitia* str. ZAS-2, specifically, were evaluated on the assumption that yeast autolysate preparations that led to growth rate improvements of this strain would also lead to growth rate improvements of the other *Treponema* isolates.

Autolysate was prepared as previously described and the different batches were prepared simultaneously and identically (Leadbetter *et al.* 1999). Yeast autolysate was prepared from Fleischmann's brand active dry yeast (from large medium, and small packages), SAF brand active dry yeast (from large and small packages), and Red Star brand active dry yeast (from large and small packages). Large packages ranged from 454g to 908g, small packages ranged from 7g to 21g, and medium packaging describes a jar of 113g. All brands and package types contained sorbitan monostearate.

Enrichments and isolation

Methanogens were enriched for from phylogenetically "lower" (*Zootermopsis nevadensis* workers) termites, and acetogenic spirochetes were enriched for from both "lower" and "higher" termites. 4YACo liquid medium was used for both spirochete and methanogen enrichments, and was prepared with yeast autolysate made from jars of Fleischmann's brand active dry yeast as this brand and packaging type of baker's yeast was found to improve growth rates of *Treponema primitia* str. ZAS-2 (Table 1-2) (Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004). This liquid medium was supplemented with all *Treponema* media vitamins and cofactors, a mixture of B12 vitamin formulations (100mg/300mL each of vitamin B12 (ICN Biomedicals Inc.), hydroxocobalamin acetate (Sigma), and hydroxocobalamin-HCl (Sigma)), 1.875g/50mL xylan, and a sugar cocktail of 3.25g/L cellobiose and 1g each of D-maltose, D-sucrose, D-trehalose, D-glucose, D-fructose, D-xylose, D-mannose, L-rhamnose, D-galactose, L-arabinose, D-arabinose, L-sorbose, D-mannitol, D-ribose, L-fucose, sodium D-glucuronate, and sodium D-

galactoronate per liter (Leadbetter & Breznak 1996; Leadbetter *et al*. 1998; Graber & Breznak 2004; Graber *et al*. 2004).

Several dilution-to-extinction enrichments from whole gut contents of "lower" (*Zootermopsis nevadensis* workers) and "higher" termites under 80% H_2/20% CO_2 (vol/vol) were performed. Enrichment cultures from whole gut contents of "higher" termites could only be successfully transferred two or three times, however, and "higher" termite specimens ran out or expired (typically after three weeks in lab) before successful enrichments, and subsequently isolates, could be obtained. Even the addition of 200uL of the supernatant of spun down *Dysgonomonas* str. JT5-1 cultures (a recent isolate from the "higher" termite hindgut) to enrichment cultures did not lead to successful enrichments nor isolations.

Successful dilution-to-extinction liquid enrichments for methanogens and spirochetes from the gut contents of *Z. nevadensis* were followed by three successive single-colony picks from agar dilution series for isolation (Leadbetter & Breznak 1996, Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004). Agar dilution series for isolation were performed in 10mL of the 4YACo medium described above (also in 25mL butyl rubber-stoppered Balch tubes) solidified by incorporating 3% Ultrapure agarose. Sterile, de-oxygenated needles were used to transfer subsurface, well-isolated colonies from one successive isolation to another.

All enrichment and isolation cultures were incubated in the dark, at room temperature, in a horizontal position, and without agitation.

Methanogens

To enrich for methanogens specifically from *Zootermopsis nevadensis* gut contents, the 4YACo medium used for enrichments described above was amended with the antibacterial drugs rifamycin SV and cephalothin (Leadbetter & Breznak 1996). Many methanogens are naturally resistant to these antibiotics (Leadbetter & Breznak 1996).

In liquid cultures methanogens were observed as a result of their unique F420 fluorescence. Negative headspace pressure generated in liquid cultures under an 80% H_2/20% CO_2 headspace, as well as increased growth close to the 80% H_2/20% CO_2 headspace in agar dilution series, also suggested growth of methanogens.

Acetogenic spirochetes

To enrich for acetogenic spirochetes specifically from *Zootermopsis nevadensis* gut contents, the 4YACo medium used for enrichments described above was amended with the antibacterial drugs rifamycin, phosphomycin, and bromoethanesulfonate, as was used to isolate *Treponema primitia* str. ZAS-1 and ZAS-2 as well as *T. azotonutricium* str. ZAS-9 (Leadbetter *et al*. 1999; Graber *et al*. 2004). Many spirochetes are naturally resistant to these antibiotics, and the latter is known to inhibit growth of H_2-consuming methanogens (Leadbetter *et al*. 1999).

In liquid cultures spirochetes were observed as a result of their distinct morphology. Negative headspace pressure generated in liquid cultures under an 80% H_2/20% CO_2 headspace, as well as increased growth close to the 80% H_2/20% CO_2 headspace in agar dilution series, also suggested growth of acetogens.

A spirochete *T. primitia* str. "ZNS-1" was isolated from a dilution-to-extinction enrichment tube that received the equivalent of 0.1 gut contents. After isolation *T. primitia* str. "ZNS-1" was cultivated under the same conditions as described for routine maintenance of *T. primitia* str. ZAS-1 and ZAS-2 (*see above*) (Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004).

Identifying "ZnD2Sec"

To identify the termite hindgut microbial community member encoding the "ZnD2Sec" phylotyope, the phylotype that is responsible for the majority of the formate dehydrogenase expression in the termite hindgut, enrichments for this organism were prepared in the same manner as enrichments for acetogenic spirochetes as described above, on the assumption that because this organism is responsible for the majority of formate dehydrogenase expression that it is an acetogen (Rosenthal *et al. in review*). It was hoped that this organism could be isolated and that its genome could be obtained for further research. Although an isolate was not obtained, these enrichment cultures were successfully used as templates for microfluidic multiplex digital PCR experiments. Before using these enrichments as templates for microfluidic multiplex digital PCR, however, qRT-PCR was employed to examine relative abundance of organisms with the "ZnD2Sec" phylotype relative to the other organisms in the enrichment culture.

qRT-PCR

qRT-PCR reactions (25μL total) contained iTaq SYBR Green Supermix with ROX (Bio-Rad Laboratories, Irvine, CA) (12.5μL/rxn), ZNO 1636F (1μL/rxn), ZNO 1729R (1μuL/rxn), PCR water (9.5μL/rxn), and enrichment culture template (1μL/rxn). The SYBR Green protocol was used on the Bio-Rad DNAEngine thermocycler (Chromo4 real time detector) and was hot start 95°C for 3min followed by 44 cycles of 95°C for 15sec and 60°C for 30sec.

Microfluidic multiplex digital PCR

Microfluidic chip experiments were performed as described previously, and with degenerate PCR primer and probe sets targeting FDH (*fdhF*) and bacterial SSU rRNA (Appendix Table 1A-1) (Ottesen *et al*. 2006; Tadmor *et al*. 2011; Rosenthal *et al. in review*). Samples were manually retrieved from chip chambers for sequencing, also as described previously (Ottesen *et al*. 2006; Tadmor *et al*. 2011; Rosenthal *et al. in review*).

16S rRNA and *fdhF* PCR

For 16S rRNA identification in enrichment and isolate cultures, once cultures reached OD_{600nm} = 0.5, 200μL of culture fluid was removed, centrifuged at 13,000rpm for 1min, and resuspended in 50μL 1x TAE. Alternatively, cultures were harvested at OD_{600nm} = 0.5 for DNA extraction using a DNeasy extraction kit (QIAGEN, Valencia, CA). This sample was then used as a template for simplex PCR reactions on a Mastercycler Model 5331 thermocycler and with agarose gel electrophoresis. PCR reactions (20μL total) contained FailSafe PCR Premix D

(Epicentre Bioechnologies, Madison, WI) (10μL/rxn), Prok 27F primer (1μL/rxn), Prok 1492R primer (1μL/rxn), Hi Fidelity Taq Polymerase (0.2μL/rxn), PCR water (5.8μL/rxn), and template (2μL/rxn). Benchtop thermocycling conditions were lid 105°C, hot start 95°C for 2min, (denaturation 95°C for 30sec, annealing 50°C for 30sec, and extension 72°C for 1.5min repeated 30X), and final extension 72°C for 5min.

Samples retrieved from microfluidic chip experiments were screened for both 16S rRNA and *fdhF* gene products also via simplex PCR on a Mastercycler Model 5331 thermocycler and with agarose gel electrophoresis. PCR reactions (50μL rxn total, brought to final volume with PCR water) contained iQ Multiple Powermix (Bio-Rad Laboratories), 200-300nM of each primer (Prok 533F and Gen 1100R for 16S rRNA, ZNO 1204F and ZNO 1792R for *fdhF*; Appendix Table 1A-1), and 2.5μL of template. Benchtop thermocycling conditions were lid 105°C, hot start 95°C for 2min, (denaturation 95°C for 15sec, annealing 60°C for 1min, and extension 72°C for 1min repeated 30 or 35X), and final extension 72°C for 10min (*Rosenthal et al. in review*).

Products from samples that yielded 16S rRNA amplicons (and *fdhF* amplicons as was the case for microfluidic chip-derived samples) were PCR purified (QIAquick PCR purification, QIAGEN). 16S rRNA PCR products were cloned in TOPO-TA vectors (TOPO-TA cloning kit, Invitrogen). Plasmids from randomly chosen clones were purified (QIAprep Spin Miniprep, QIAGEN). 16S rRNA PCR products and plasmids were sequenced with generic T3 and T7 primers, and *fdhF* products were sequenced with ZNO 1204F and ZNO 1729R (Appendix Table 1A-1). All sequencing

reactions were performed at Laragen, Inc. (Los Angeles, CA) and BLAST hits to sequences were obtained to identify microfluidic chip-derived samples (Geer *et al*. 2010).

Treponema primitia enzymatic defenses against O_2-stress

For identification of enzymes representing defenses against O_2-stress in *Treponema primitia* str. ZAS-1 and ZAS-2, genomic data for the strains were obtained from RAST (Aziz *et al*. 2008), JGI IMG/M (Markowitz *et al*. 2006), and NCBI BLAST (Geer *et al*. 2010). Genomes were then searched for enzymes recorded in the literature as being active against O_2-stress (Dolla *et al.* 2006; Brioukhanov & Netrusov 2007; Rocha *et al*. 2007; Imlay 2008b; LeFourn *et al*. 2008; Sund *et al*. 2008; Riebe *et al*. 2009; Lakhal *et al*. 2011; Xiao *et al*. 2011; Figueiredo *et al.* 2012).

RESULTS:

Yeast autolysate for media and cultivation

The isolation of *Treponema primitia* str. ZAS-1 and ZAS-2 as well as *T. azotonutricium* str. ZAS-9, was non-trivial, and one media component that aided in their acquisition in pure culture was a yeast autolysate prepared fresh from baker's yeast (Leadbetter *et al.* 1999). Interestingly, commercial yeast extracts could not be used *in lieu* of this freshly prepared yeast autolysate (Leadbetter *et al.* 1999). From the *Treponema* organisms' initial isolation in "YACo" medium, however, after being passaged for over a decade, its growth rates had increased from a doubling time initially of 22, 29, and 35hrs for *T. primitia* str. ZAS-1, *T. primitia* str. ZAS-2, and *T. azotonutricium* str. ZAS-9, respectively, to a doubling time between 54 to 77hrs

currently for each *Treponema* isolate (Leadbetter *et al.* 1999; Graber & Breznak 2004).

In an effort to improve the growth rates of the *Treponema* isolates in 4YACo medium, one factor that warranted re-evaluation was the vital yeast autolysate media component (Leadbetter *et al*. 1999). For example, store-bought baker's yeast contains sorbitan monostearate – a surfactant with emulsifying properties – in addition to the yeast itself. This additive can inhibit microbial growth, and consequently, 4YACo medium comprised of baker's yeast autolysate prepared from different brands, and different packaging within the same brand, were prepared and growth rates of *T. primitia* str. ZAS-2 were compared on each as a proxy for growth by the other isolates. As a first observation, yeast from different brands and different packaging types within the same brand looked and smelled different so autolysate prepared from each was tested (Table 1-1).

Table 1-1: Appearance of baker's yeast and appearance and odor of corresponding yeast autolysate and autolysate supernatant.

Brand	Packaging	Dry yeast appearance	Autolysate appearance	Autolysate odor	Autolysate supernatant
Red Star[a]	large	long cylindrical pellet; tan	light tan	strong sour	orange/brown
SAF	small	small cylindrical pellet; dark tan	tan	sweet	dark yellow
SAF	large	long cylindrical pellet; cream	cream	very sweet	orange
Fleischmann's	small	small cylindrical pellet; light tan	cream	bread (sweet)	yellow
Fleischmann's	large	small spheres; light brown	dark tan	sweet	dark yellow

[a]Red Star brand active dry yeast was used for original yeast autolysate preparations (Leadbetter *et al*. 1999)

After examining growth rates of *T. primitia* str. ZAS-2 obtained on 4YACo medium prepared from Fleischmann's brand active dry yeast (from large medium, and small packages), SAF brand active dry yeast (from large and small packages), and Red Star brand active dry yeast (from large and small packages), growth rates improved the most on media with yeast autolysate prepared from jars of Fleischmann's brand active dry yeast (Table 1-2). Specifically, this was 47hrs compared to between 54 and 77hrs previously, and 29hrs upon isolation (Table 1-2) (Leadbetter *et al.* 1999; Graber & Breznak 2004). From this point on, therefore, routine growth as well as enrichment and isolation attempts occurred in 4YACo liquid medium prepared with yeast autolysate made from jars of Fleischmann's brand baker's yeast.

Table 1-2: Growth rates of *Treponema primitia* str. ZAS-2 on 4YACo medium prepared from yeast autolysates of different brands and packaging types.

Brand	Packaging	Doubling time[b] +/- std. error (hrs)
Red Star[a]	large	74 +/- 1
SAF	small	85 +/- 1
SAF	large	111 +/- 12
Fleischmann's	small	47 +/- 0
Fleischmann's	large	50 +/- 4

[a]Red Star brand active dry yeast was used for original yeast autolysate preparations (Leadbetter *et al*. 1999)
[b]Doubling time is an average of triplicate cultures

Enrichment

Methanogens

As an exercise in the enrichment and isolation procedure, methanogens were targeted from phylogenetically "lower" termites (workers of *Zootermopsis nevadensis*). This is because, similar to acetogens, they are also H_2 + CO_2 consumers and would demonstrate similar relationships with their 80% H_2/20% CO_2 (vol/vol) headspace atmosphere as acetogens. In general, however, from what is known about spirochete and methanogen isolates obtained from the termite system thus far, methanogens appear to grow faster than spirochetes (35 to 40hrs) in the medium used here and success or problems in the procedure are able to be detected faster (Leadbetter & Breznak 1996; Leadbetter *et al*. 1998).

Methanogens were enriched for in liquid culture and noted by their F420 fluorescence. In liquid culture, methanogens appeared both filamentous as well as cocci and rod-shaped. Negative headspace pressure was also generated by these liquid cultures under an 80% H_2/20% CO_2 headspace. A dilution-to-extinction

enrichment tube that received the equivalent of 0.1 gut contents was used to inoculate an agar dilution series for subsequent isolation. Within the first of three agar dilutions preferential growth of colonies was observed close to the meniscus of the cultures at the agar/headspace interface. Presumably this was because methanogens were in the cultures and were growing close to their substrate, the 80% H_2/20% CO_2 (vol/vol) headspace. Distinct colonies could be readily transferred from these agar cultures that displayed F420 fluorescence and were consistent in morphology with what had already been observed in enrichments as methanogens. Because methanogen enrichments and isolations were simply trials, however, efforts were directed towards acetogenic spirochete enrichments and isolations.

Acetogenic spirochetes

Acetogenic spirochetes were enriched for in liquid culture from both "lower" and "higher" termites, and noted by their unique morphologies. In liquid culture, spirochetes appeared diverse in morphology with two morphologies dominating the cultures (Fig. 1-8a and b). Negative headspace pressure was also generated by these liquid cultures under an 80% H_2/20% CO_2 headspace. A dilution-to-extinction enrichment tube that received the equivalent of 0.1 gut contents was used to inoculate an agar dilution series for subsequent isolation. Within the first of three agar dilutions, preferential growth of colonies was observed close to the meniscus of the cultures at the agar/headspace interface. Presumably this was because acetogenic spirochetes were in the cultures and were growing close to their substrate, the 80% H_2/20% CO_2 (vol/vol) headspace. Distinct colonies could be

readily transferred from these agar dilution series cultures and were consistent in morphology with what had already been thought to be spirochetes in liquid cultures. Colonies were disc-like in appearance. No acetogenic spirochetes, nor any other microorganisms for that matter, were obtained from dilution-to-extinction enrichments from Joshua Tree National Park "higher" termite samples.

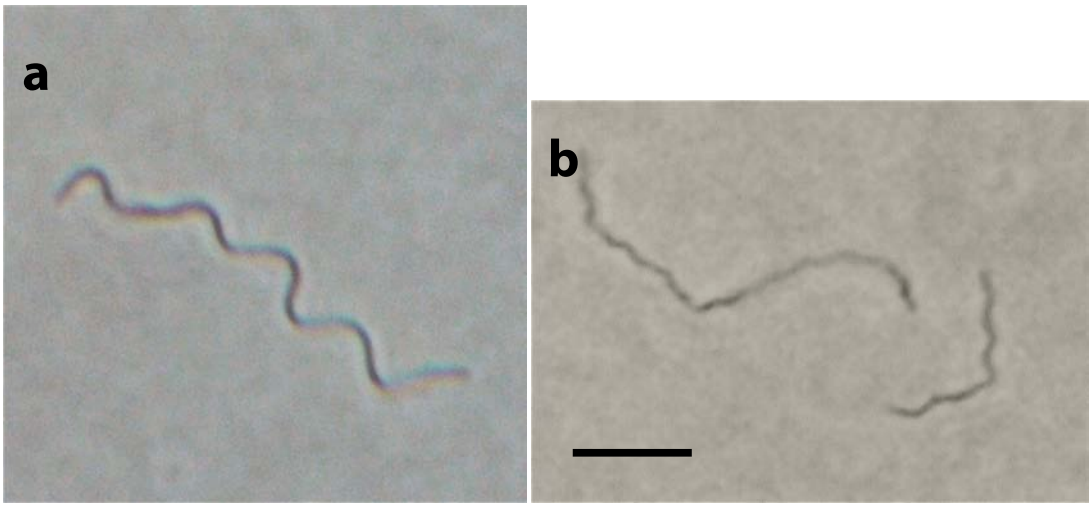


Fig. 1-8: Spirochete-like morphologies that dominated liquid enrichment cultures for acetogenic spirochetes. **(a)** "Floppy" morphology. **(b)** What would be isolated as *Treponema primitia* str. "ZNS-1." 4YACo medium for liquid enrichment cultures was prepared with yeast autolysate made from jars of Fleischmann's brand active dry yeast supplemented with all *Treponema* media vitamins and cofactors, a mixture of B12 vitamin formulations (100mg/300mL each of vitamin B12 (ICN Biomedicals Inc.), hydroxocobalamin acetate (Sigma), and hydroxocobalamin-HCl (Sigma)), 1.875g/50mL xylan, and a sugar cocktail of 3.25g/L cellobiose and 1g each of D-maltose, D-sucrose, D-trehalose, D-glucose, D-fructose, D-xylose, D-mannose, L-rhamnose, D-galactose, L-arabinose, D-arabinose, L-sorbose, D-mannitol, D-ribose, L-fucose, sodium D-glucuronate, and sodium D-galactoronate per liter (Graber & Breznak 2004; Graber *et al*. 2004). To enrich for acetogenic spirochetes the media was amended with the antibacterial drugs rifamycin, phosphomycin, and bromoethanesulfonate. Dilution-to-extinction enrichments from whole gut contents of a *Zootermopsis nevadensis* worker under 80% H_2/20% CO_2 (vol/vol) was performed. Then, three successive single-colony picks from agar dilution series for isolation were undertaken (Leadbetter & Breznak 1996, Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004). Agar dilution series for isolation were performed in 10mL of 4YACo medium (also in 25mL butyl rubber-stoppered Balch

tubes) solidified by incorporating 3% Ultrapure agarose. Sterile, de-oxygenated needles were used to transfer subsurface, well-isolated colonies from one successive isolation to another. All enrichment and isolation cultures were incubated in the dark, at room temperature, in a horizontal position, and without agitation. Scale bar represents 5μm.

The spirochete successfully isolated shares 100% 16S rRNA sequence identity with *T. primitia* str. ZAS-2 and 99% 16S rRNA sequence identity with *T. primitia* str. ZAS-1, and phylogenetically its closest cultured relatives are these *T. primitia* strains (Appendix Item 1A-1). Given, however, that this isolate was obtained from the a *Zootermopsis nevadensis* worker termite, as opposed to str. ZAS-1 and ZAS-2 that were obtained from *Z. angusticollis*, this isolate was named *T. primitia* str. "ZNS-1." Most cells are approximately 0.2μm in diameter by 3 to 15μm long, with a wave-length or body pitch of approximately 2μm similar to str. ZAS-1 and ZAS-2 (Fig. 1-7b). From their successful growth in 4YACo medium with a 80% H_2/20% CO_2 (vol/vol) headspace, colony-establishment preference close to the agar/headspace interface in agar tubes, and consumption of headspace as indicated by negative pressure, str. "ZNS-1" is very probably an acetogen (at least deriving acetate from H_2 + CO_2) like str. ZAS-1 and ZAS-2. *T. primitia* str. "ZNS-1" also grows well at 25°C and a pH of 7.2 (Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004).

Identifying "ZnD2Sec"

To identify the termite hindgut microbial community member encoding the "ZnD2Sec" phylotyope, the phylotype that is responsible for the majority (40%) of the formate dehydrogenase expression in the termite hindgut, enrichments for the organism that encode this phylotype were prepared the same as enrichments for acetogenic spirochetes, on the assumption that that because this organism is

responsible for the majority of formate dehydrogenase expression that it is an acetogen (Rosenthal *et al*. *in review*). While attempts to isolate the termite hindgut microbial community member encoding the "ZnD2Sec" phylotyope were unsuccessful, in enrichments, uniform small cocci were observed, 16S rRNA analysis confirmed the presence of "ZnD2Sec" in the sample, and qRT-PCR confirmed its relative abundance. These enrichments were used as templates for microfluidic digital PCR to have a better chance at targeting both FDH (*fdhF*) and bacterial SSU rRNA given their relative abundance in the enrichment vs. gut contents. Over 50% of the samples that co-localized 16S rRNA and *fdhF* were deltaproteobacteria. Non-matches to deltaproteobacteria were often uncultured organisms.

Treponema primitia enzymatic defenses against O_2-stress

Preliminary work suggests that cultures of the termite hindgut spirochetes *Treponema primitia* maintained growth after the addition of as much as 0.5% (vol/vol) O_2 to their headspace atmosphere, and exhibited both NAD(P)H peroxidase and NAD(P)H oxidase activities (Table 1-3) (Graber & Breznak 2004). Although neither catalase nor superoxide dismutase, which together constitute a system to combat reactive oxygen species common among aerobes, were detected in *T. primitia* (Table 1-3), other enzymes that scavenge superoxide, hydrogen peroxide, and oxygen abound in the commonly-studied anaerobe, *Clostridium acetobutylicum*, and may provide insight into relevant enzymes in the termite system (Graber & Breznak 2004; Brioukhanov & Netrusov 2007; Imlay 2008a).

Table 1-3: *Treponema primitia* str. ZAS-2 enzymatic defenses against oxidative stress.

Enzyme	**Activity**[a]		**Genomic Evidence**	
	str. ZAS-1	**str. ZAS-2**	**str. ZAS-1**	**str. ZAS-2**
NADH oxidase	+	++	-	Y
NADPH oxidase	+	+	-	-
NADH peroxidase	+	+	-	-
NADPH peroxidase	+	+	-	-
Catalase	ND	ND	-	-
Superoxide dismutase	ND	ND	-	-

[a]From Graber & Breznak 2004

+, low activity; ++, high activity; ND, not detected; -, no evidence; Y, evidence

For example, O_2 can steal two electrons from a variety of reduced flavoenzymes, ultimately generating H_2O_2 (Brioukhanov & Netrusov 2007; Imlay 2008a). In turn, H_2O_2 can react with unincorporated intracellular iron to produce the powerful oxidant, HO·, according to the Fenton reaction:

$$H_2O_2 + Fe_2^+ \rightarrow HO\cdot + OH^- + Fe_3^+$$

In *C. acetobutylicum*, when H_2O_2 directly oxidizes a mononuclear iron atom within a peroxide stress regulator (PerR), a repressor of peroxide stress responses, the oxidized PerR protein loses the capacity to bind DNA. Whereas PerR would normally bind to DNA to repress the synthesis of proteins that suppress Fenton chemistry, these proteins are formed and then hinder the conversion of H_2O_2 to HO· (Fig. 1-9). Further, rubrerythrins have been associated with both peroxidase and oxygen-scavenging activities with rubredoxins serving as electron donors in *C.*

acetobutylicum (Fig. 1-9). In addition, in low-oxygen environments low-potential electron carriers such as flavodoxins, rubredoxins, and ferredoxins will oftentimes reduce resting enzymes (Brioukhanov & Netrusov 2007; Imlay 2008a).

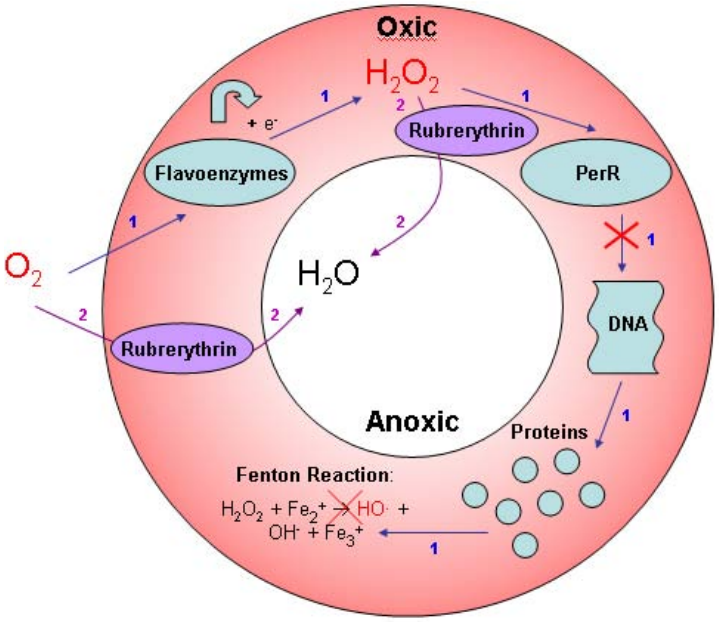


Fig. 1-9: Proposed mechanisms for combating oxidative stress in the termite hindgut. (1) Molecular oxygen can steal two electrons from a variety of reduced flavoenzymes, ultimately generating H_2O_2. When H_2O_2 directly oxidizes a mononuclear iron atom within the peroxide stress regulator (PerR), the oxidized PerR protein loses the capacity to bind DNA. Whereas PerR would normally bind to DNA to represses the synthesis of proteins that suppress Fenton chemistry, these proteins are formed and hinder the conversion of H_2O_2 to HO·. (2) In addition, rubrerythrins have been associated with both peroxidase and oxygen-scavenging activities, converting H_2O_2 and O_2 to H_2O, respectively (Dolla *et al.* 2006; Brioukhanov & Netrusov 2007; Rocha *et al*. 2007; Imlay 2008b; LeFourn *et al*. 2008; Sund *et al*. 2008; Riebe *et al*. 2009; Lakhal *et al*. 2011; Xiao *et al*. 2011; Figueiredo *et al.* 2012).

As with *C. acetobutylicum*, rubrerythrin and rubredoxin in both *T. primitia* strains are oftentimes found within the same gene neighborhoods (Fig. 1-10a). So, too, are the genes regulating several flavodoxins and ferredoxins in these organisms and

also found in close association with the aforementioned enzymes (Fig. 1-10b). Preliminary observations, therefore, suggest that in addition to NAD(P)H peroxidase and NAD(P)H oxidase, other enzymes important in combating oxidative stress in the termite hindgut may include PerRs, rubrerythrins, rubredoxins, flavodoxins, and ferredoxins (Brioukhanov & Netrusov 2007; Imlay 2008b).

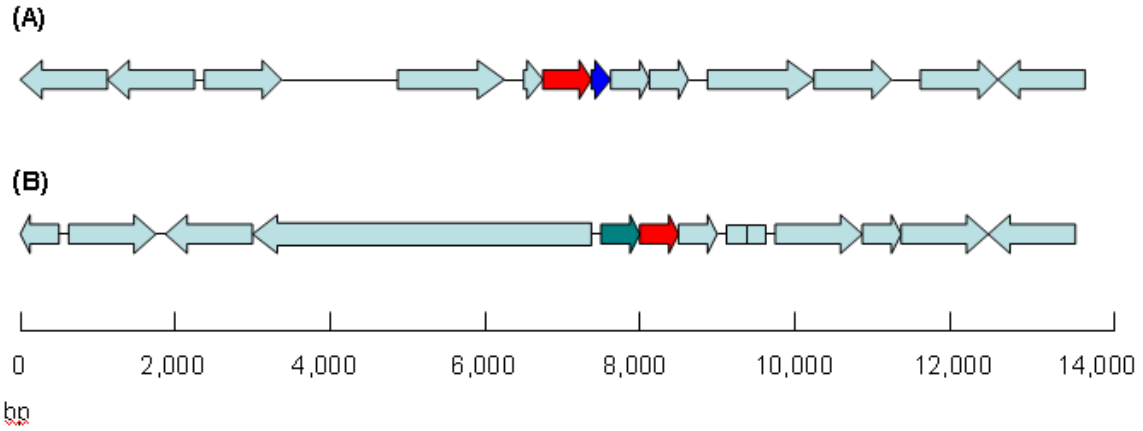


Fig. 1-10: Gene neighborhoods of enzymes involved in O_2 transformation. Predicted ORFS include **(a)** rubrerythrin (red) and rubredoxin (dark blue) as well as **(b)** peroxide stress regulator (green) and rubrerythrin (red) in *Treponema primitia* str. ZAS-2.

DISCUSSION:

Yeast autolysate for media and cultivation

While growth rates of the *Treponema* isolates on medium with Red Star yeast autolysate were 54 to 77hrs, growth rates of *T. primitia* str. ZAS-2 on medium with Fleischmann's yeast autolysate markedly improved to 47hrs (Table 1-2). Growth rates of *T. primitia* str. ZAS-2 on yeast autolysate prepared from different brands and packaging of baker's yeast were also quite disparate (Table 1-2). Although yeast autolysate prepared from fresh baker's yeast is an integral component of the 4YACo medium on which *T. primitia* str. ZAS-1 and ZAS-2 as well as *T.*

azotonutricium str. ZAS-9 grow, that there is such a range of growth rates on yeast autolysate prepared from different brands and packages of yeast is noteworthy (Table 1-2) (Leadbetter *et al*. 1999). These results suggest that one must pay careful attention when cultivating organisms in pure culture to each of the media components as they can have considerable effects on culture growth. This includes paying attention to their source, how they may or may not change, their age, and their storage, among other factors.

Moreover, while an improvement in growth rates was seen with new yeast autolysate prepared from jars of Fleischmann's brand active dry yeast, the termite hindgut turns over every 24 hours (Bignell 1984). Consequently, if *in situ* the *Treponema* isolates were to double every 47hrs, presumably they would be lost from the termite system. These observations, therefore, hint that *Treponema* grows with a doubling time under 24hrs *in situ* and that current culture conditions are suboptimal. This disparity could be a result of the isolates evolving as they have been passaged over time to double slower and/or due to a limiting factor in the media (*see Chapter 4*).

Enrichments and isolation

Methanogens

Methanogens appeared to have been enriched for successfully as a primer in enrichment and isolation techniques. This was because colonies in agar dilution series cultures were growing close to methanogen substrates, the 80% H_2/20% CO_2 (vol/vol) headspace, and upon picking colonies and resuspending them in liquid for

microscopic observation, they displayed F420 fluorescence and morphologies consistent with those in liquid enrichments. Additionally, because the agar media was homogenous, that these colonies developed closer to the headspace was likely due to their preference for H_2 and CO_2 and not a result of any gradient of an inhibitory factor in the agar media. With additional time and attention it would be worthwhile to pursue methanogen enrichments. This is because termite methanogens are cited as a small but significant source of the atmospherically relevant trace gas and there have been relatively few studies on these termite symbionts (Brauman *et al*. 1992). Moreover, previous molecular analyses, as well as studies involving the isolation and cultivation of these organisms, have shown that the methanogens inhabiting the termite gut are distinct from other known methanogen species (Ohkuma *et al*. 1999).

Acetogenic spirochetes

A spirochete, *Treponema primitia* str. "ZNS-1," was isolated from the *Zootermopsis nevadensis* termite hindgut (Fig. 1-8b). This strain exhibits preliminary evidence of H_2 + CO_2 acetogenesis, as colonies of the strain preferentially grow at the agar/80% H_2/20% CO_2 headspace and negative pressure generated from cultures growing under 80% H_2/20% CO_2 is observed (Leadbetter *et al*. 1999; Graber & Breznak 2004, Graber *et al*. 2004). Similar to the methanogen agar cultures, because the agar media was homogenous, that *T. primitia* str. "ZNS-1" colonies developed closer to the headspace was likely due to their preference for H_2 and CO_2 and not a result of any gradient of an inhibitory factor in the agar media.

Identifying "ZnD2Sec"

Efforts to isolate the termite hindgut microbial community member encoding the "ZnD2Sec" phylotyope, were unsuccessful. Nevertheless, enrichment cultures were successfully used as templates for microfluidic digital PCR and helped correlate the "ZnD2Sec" phylotype with a deltaproteobacteria and not a spirochete as previously thought (Rosenthal *et al*. *in review*). Originally, it was thought that "ZnD2Sec" was a spirochete because spirochetes are one of the most abundant, consistently present prokaryotes in termite hindguts (Breznak & Brune 1994; Graber & Breznak 2004; Rosenthal *et al*. *in review*). Moreover, it was assumed that because this organism is responsible for the majority of formate dehydrogenase expression that it is an acetogen, and since two of the three spirochetes already isolated from the termite hindgut demonstrate bona fide H_2 + CO_2 acetogenic activities, it was also assumed this organism is an acetogenic spirochete. Nevertheless, multiplex microfluidic digital PCR experiments confirmed that the termite hindgut microbial community member encoding the "ZnD2Sec" phylotype is a deltaproteobacteria (Rosenthal *et al*. *in review*). This finding supports that one must remain open to the unexpected (*see Chapter 4*). It is also noteworthy that the phylotype that is responsible for the second-most formate dehydrogenase expression in the termite hindgut (14% in contrast to 40% for "ZnD2Sec") is encoded by a free-living spirochete (Rosenthal *et al*. *in review*).

***Treponema primitia* enzymatic defenses against O_2-stress**

Treponema primitia str. ZAS-1 and ZAS-2 lack hallmark mechanisms for dealing with oxidative stress, specifically catalase and superoxide dismutase (Table 1-3) (Graber

Breznak 2004). Although they do not have catalase and superoxide dismutase, this does not necessarily mean that str. ZAS-1 and ZAS-2 are especially sensitive to O_2. Likely rather, they have other strategies and relationships with O_2 in their environment, which is relevant to work discussed in Chapter 3. Possible mechanisms are also listed below and should be investigated further (Appendix Table 1A-2).

REFERENCES:

Appel HM (1993) Phenolics in ecological interactions: the importance of oxidation. *Journal of Chemical Ecology*, **19**, 1521-1552.

Aziz RK, Bartels D, Best AA *et al*. (2008) The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, **9**, 75.

Bignell DE (1984) The Arthropod Gut as an Environment for Microorganisms. In: *Invertebrate-Microbial Interactions* (eds Anderson JM, Rayner ADM, Walton DWH), pp. 205-227. Cambridge University Press, Cambridge.

Brauman A, Kane MD, Labat M, Breznak JA (1992) Genesis of acetate and methane by gut bacteria of nutritionally diverse termites. *Science*, **257**, 1384-1387.

Breznak JA, Brune A (1994) Role of microorganisms in the digestion of lignocellulose by termites. *Annual Review of Entomology*, **39**, 453-487.

Brioukhanov AL, Netrusov AI (2007) Aerotolerance of strictly anaerobic microorganisms and factors of defense against oxidative stress: a review. *Applied Biochemistry and Microbiology*, **43**, 567-582.

Brune A (1998a) Termite guts: the world's smallest bioreactors. *Trends in Biotechnology*, **16**, 16-21.

Brune A (1998b) Microbial degradation of aromatic compounds: aerobic versus anaerobic processes. *Mitteilungen der Deutschen Bodenkundlichen Gesellschaft*, **87**, 65-78.

Brune A (2006) Symbiotic Associations Between Termites and Prokaryotes. In: *Prokaryotes* (eds Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E), pp. 439-474. Springer, New York.

Brune A, Emerson D, Breznak JA (1995a) The termite gut microflora as an oxygen sink: microelectrode determination of oxygen and pH gradients in guts of lower and higher termites. *Applied and Environmental Microbiology*, **61**, 2681-2687.

Brune A, Friedrich M (2000) Microecology of the termite gut: structure and function on a microscale. *Current Opinion in Microbiology*, **3**, 263-269.

Brune A, Miambi E, Breznak JA (1995b) Roles of oxygen and the intestinal microflora in the metabolism of lignin-derived phenylpropanoids and other monoaromatic compounds by termites. *Applied and Environmental Microbiology*, **61**, 2688-2695.

Butler JHA, Buckerfield JC (1979) Digestion of lignin by termites. *Soil Biology & Biochemistry*, **11**, 507-513.

Cleveland LR (1926) Symbiosis among animals with special reference to termites and their intestinal flagellates. *The Quarterly Review of Biology*, **1**, 51-60.

Cookson LJ (1987) ^{14}C-lignin degradation by three Australian termite species. *Wood Science and Technology*, **21**, 11-25.

Dolla A, Fournier M, Dermoun Z (2006) Oxygen defense in sulfate-reducing bacteria. *Journal of Biotechnology*, **126**, 87-100.

Figueiredo MC, Lobo SA, Carita JN *et al*. (2012) Bacterioferritin protects the anaerobe *Desulfovibrio vulgaris* Hildenborough against oxygen. *Anaerobe*, **18**, 454-458.

Geer LY, Marchler-Bauer A, Geer RC *et al*. (2010) The NCBI BioSystems database. *Nucleic Acids Research*, **38**, D492-D496.

Graber JR, Breznak JA (2004) Physiology and nutrition of *Treponema primitia*, an H_2/CO_2-acetogenic spirochete from termite hindguts. *Applied and Environmental Microbiology*, **70**, 1307-1314.

Graber JR, Leadbetter JR, Breznak JA (2004) Description of *Treponema azotonutricium* sp. nov. and *Treponema primitia* sp. nov., the first spirochetes isolated from termite guts. *Applied and Environmental Microbiology*, **70**, 1315-1320.

Hungate R (1955) Mutualistic Intestinal Protozoa. In: *Biochemistry and Physiology of Protozoa* (eds Hunter S, Lwoff A), pp. 159-199. Academic Press, New York.

Imlay J (2008a) How obligatory is anaerobiosis? *Molecular Microbiology*, **68,** 801 804.

Imlay J (2008b) Cellular defenses against superoxide and hydrogen peroxide. *Annual Reviews of Biochemistry*, **77**, 755-776.

Inward D, Vogler A, Eggleton P (2007) A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspect of their evolutionary biology. *Molecular Phylogenetics and Evolution*, **44**, 953-967.

Kappler A, Brune A (2002) Dynamics of redox potential and changes in redox state of iron and humic acids during gut passage in soil-feeding termites (*Cubitermes* spp.). *Soil Biology & Biochemistry*, **34**, 221-227.

Kuhnigk T, Borst EM, Ritter A *et al*. (1994) Degradation of lignin monomers by the hindgut flora of xylophagous termites. *Systematic and Applied Microbiology*, **17**, 76-85.

Lakhal R, Auria R, Davidson S *et al*. (2011) Oxygen uptake rates in the hyperthermophilic anaerobe *Thermotoga maritima* grown in a bioreactor under controlled oxygen exposure: clues to its defense strategy against oxidative stress. *Archives of Microbiology*, **193**, 429-438.

Leadbetter JR, Breznak JA (1996) Physiological ecology of *Methanobrevibacter cuticularis* sp. nov. and *Methanobrevibacter curvatus* sp. nov., isolated from the hindgut of the termite *Reticulitermes flavipes*. *Applied and Environmental Microbiology*, **62**, 3620-3621.

Leadbetter JR, Crosby LD, Breznak JA (1998) *Methanobrevibacter filiformis* sp. nov., a filamentous methanogen from termite hindguts. *Archives of Microbiology*, **169**, 287-292.

Leadbetter JR, Schmidt TM, Graber JR, Breznak JA (1999) Acetogenesis from H_2 plus CO_2 by spirochetes from termite guts. *Science*, **283**, 686-689.

LeFourn C, Fardeau ML, Ollivier B *et al*. (2008) The hyperthermophilic anaerobe *Thermotoga maritima* is able to cope with limited amount of oxygen: insights into its defense strategies. *Environmental Microbiology*, **10**, 1877-1887.

Leidy J (1877) On intestinal parasites of *Termes flavipes*. *Proceedings of the Academy of Natural Sciences of Philadelphia*, **29**, 146-149.

Markowitz VM, Korzeniewski F, Palaniappan K *et al*. (2006) The integrated microbial genomes (IMG) system. *Nucleic Acids Research*, **34**, D344-D348.

Odelson DA, Breznak JA (1983) Volatile fatty acid production by the hindgut microbiota of xylophagous termites. *Applied and Environmental Microbiology*, **45**, 1602-1613.

Ohkuma M, Noda S, Kudo T (1999) Phylogenetic relationships of symbiotic methanogens in diverse termites. *FEMS Microbiology Letters*, **171**, 147-153.

Ottesen EA, Hong JW, Quake SR, Leadbetter JR (2006) Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science*, **314**, 1464-1467.

Ottesen EA, Leadbetter JR (2011) Formyltetrahydrofolate synthetase gene diversity in the guts of higher termites with different diets and lifestyles. *Applied and Environmental Microbiology*, **77**, 3461-3467.

Pasti MB, Pometto III AL, Nuti MP, Crawford DL (1990) Lignin-solubilizing ability of actinomycetes isolated from termite (Termitidae) gut. *Applied and Environmental Microbiology*, **56**, 2213-2218.

Riebe O, Fischer RJ, Wampler DA *et al*. (2009) Pathway for H_2O_2 and O_2 detoxification in *Clostridium acetobutylicum*. *Microbiology*, **155**, 16-24.

Rocha ER, Tzianabos AO, Smith CJ (2007) Thioredoxin reductase is essential for thiol/disulfide redox control and oxidative stress survival of the anaerobe *Bacteroides fragilis*. *Journal of Bacteriology*, **189**, 8015-8023.

Rosenthal AZ, Zhang X, Lucey KS *et al*. Localizing transcripts to single cells suggests an ecological role for the deltaproteobacteria living on termite gut protozoa. *In review (PNAS)*

Sugimoto A, Bignell DE, MacDonald JA (2000) Global Impact of Termites on the Carbon Cycle and Atmospheric Trace Gases. In: *Termites: Evolution, Sociality, Symbioses, Ecology* (eds Abe T, Bignell DE, Higashi M), pp. 409-435. Kluwer Academic Publishers, Netherlands.

Sund CJ, Rocha ER, Tzinabos AO *et al*. (2008) The *Bacteroides fragilis* transcriptome response to oxygen and H_2O_2: the role of OxyR and its effect on survival and virulence. *Molecular Microbiology*, **67**, 129-142.

Tadmor AD, Ottesen EA, Leadbetter JR, Philips R (2011) Probing individual environmental bacteria for viruses by using microfluidic digital PCR. *Science*, **333**, 58-62.

Tholen A, Schink B, Brune A (1997) The gut microflora of *Reticulitermes flavipes*, its relation to oxygen, and evidence for oxygen-dependent acetogenesis by the most abundant *Enterococcus* sp. *FEMS Microbiology Ecology*, **24**, 137-149.

Warnecke F, Luginbühl P, Ivanova N *et al*. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, **450**, 560-565.

Wenzel M, Schonig I, Berchtold M *et al*. (2002) Aerobic and facultatively anaerobic cellulolytic bacteria from the gut of the termite *Zootermopsis angusticollis*. *Journal of Applied Microbiology*, **92**, 32-40.

Wertz JT, Breznak JA (2007a) *Stenoxybacter acetivorans* gen. nov., sp. nov., an acetate-oxidizing obligate microaerophile among diverse O_2-consuming bacteria from termite guts. *Applied and Environmental Microbiology*, **73**, 6819-6828.

Wertz JT, Breznak JA (2007b) Physiological ecology of *Stenoxybacter acetivorans,* an obligate microaerophile in termite guts. *Applied and Environmental Microbiology*, **73**, 6829-6841.

Wood TG (1976) The Role of Termites (Isoptera) in Decomposition Processes. In: *The Role of Terrestrial and Aquatic Organisms in Decomposition Processes* (eds Anderson JA, McFadden A), pp. 145-168. Blackwell Scientific Publications, Oxford.

Wood TG, Johnson RA (1986) The Biology, Physiology, and Ecology of Termites. In: *Economic Impact and Control of Social Insects* (ed Vinson SB), pp. 1-68. Praeger Publications, New York.

Wood TG, Sands WA (1978) The Role of Termites in Ecosystems. In: *Production Ecology of Ants and Termites* (ed Brian MV), pp. 245-292. Cambridge University Press, Cambridge.

Xiao M, Xu P, Zhao J *et al*. (2011) Oxidative stress-related responses of *Bifidobacterium longum* subsp. *longum* BBMN68 at the proteomic level after exposure to oxygen. *Microbiology*, **157**, 1573-1588.

Yamin MA (1980) Cellulose metabolism by the termite flagellate *Trichomitopsis termopsidis*. *Applied and Environmental Microbiology*, **39**, 859-863.

Yamin MA, Trager W (1979) Cellulolytic activity of an axenically-cultivated termite flagellate, *Trichomitopsis termopsidis*. *Journal of General Microbiology*, **113**, 417-420.

Zimmer M, Brune A (2005) Physiological properties of the gut lumen of terrestrial isopods (Isopoda: Oniscidea): adaptive to digesting lignocellulose? *Journal of Comparative Physiology* B, **175**, 275-283.

Zimmer M, Topp W (1997) Homeostatic responses in the gut of *Porcellio scaber* (Isopoda: Oniscidea) optimize litter degradation. *Journal of Comparative Physiology B*, **167**, 582-285.

CHAPTER 1: APPENDIX

Table 1A-1: 16S rRNA and *fdhF* multiplex and simplex primer and probe designs.

Table 1A-2: Enzymatic defenses against O_2-stress.

Item 1A-1: *Treponema primitia* str. "ZNS-1" 16S rRNA sequence.

Table 1A-1: 16S rRNA and *fdhF* multiplex and simplex primer and probe designs.

Primer or Probe	Sequence	Target	Experiments
Prok 27F	5' - AGA GTT TGA TCC TGG CTC AG - 3'	Gen bac 16S rRNA	Isolate identification
Prok 1492R	5' – TAC GGY TAC CTT GTT ACG ACT T – 3'	Gen bac 16S rRNA	Isolate identification; Microfluidic digital PCR on chip
Prok 357F	5' – CTC CTA CGG GAG GCA GCA G – 3'	Gen bac 16S rRNA	Microfluidic digital PCR on chip
GenBac HEX-1389Prb	5' - HEX – CTT GTA CAC ACC GCC CGT C-3BHQ1 – 3'	Gen bac 16S rRNA (probe)	Microfluidic digital PCR on chip
FAM-1636Prb	5'– ACT ATG ACC GGC AAT TGT CGC CTG TT –3'	ZnD2Sec *fdhF* (probe)	Microfluidic digital PCR on chip
ZNO 1204F	5' - AAC GAA CAT GAC GGC GTC TAC TCT - 3'	ZnD2Sec *fdhF*	To look for in enrichments; Microfluidic digital PCR on chip; chip well pick check
ZNO 1792R	5' - TCA GAC CCA TAT CAC GGC AAA GTT - 3'	ZnD2Sec *fdhF*	To look for in enrichments; Microfluidic digital PCR on chip; chip well pick check
Prok 533F	5' - GTG CCA GCM GCC GCG GTA A - 3'	Gen bac 16S rRNA	chip well pick check
Gen 1100R:	5' - AGG	Gen bac 16S	chip well pick

	GTT GCG CTC GTT G -3’	rRNA	check

Table 1A-2: Possible enzymatic defenses against O_2-stress.

Citation	Enzymes to investigate
Rocha *et al*. 2007	glutathione/glutaredoxin; thioredoxin peroxidase
Sund *et al*. 2008	alkyl hyperoxide reductase; cytochrome C peroxidase/oxidase; ferritin; thiol peroxidase scavangase; "oxygen-induced starch utilization" genes; "global oxidative response" genes
Le Fourn *et al*. 2008	any oxidoreductase; any oxidase; any peroxidase
Riebe *et al*. 2009	rubrerythrin, rubredoxin, NADH: rubredoxin oxidoreductase
Xiao *et al*. 2011	alkyl hyperoxide reductase; pyridine nucleotide-disulfide reductase; DNA oxidative damage protective proteins (DNA-binding); ribonucleotide reductase; nucleotide triphosphate; pyrophosphohydrolase; polynucleotide phosphorylase, enolase
Lakhal *et al*. 2011	neelaredoxin
Figueiredo *et al*. 2012	bacterioferritin
Dolla *et al*. 2006	superoxide reductase; nigerythrin

Item 1A-1: *Treponema primitia* str. "ZNS-1" 16S rRNA sequence.

TTACCTTGTTACGACTTCACCCTCCTCACTAAACGTACCTTCGACAGCGCGCTCCTTGCGG
TTACGCTACCGGCTTCGGGTACCTCCAACTCGGATGGTGTGACGGGCGGTGTGTACAAGG
CCCGGGAACGTATTCACCGCGCCATGCTGATGCGCGATTACTAGCGATTCCAACTTCATG
AAGTCGGGTTTCAGACTTCAATCCGAACTACGGGCGGCTTTTTGCGCTTCGCTTTGACCT
CGCGGTTTCGCGTCGCTTTGTACCGCCCATTGTAGCACGTGTGTAGCCCTGGACATAAGG
GCCATGATGACTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTTCCGCCAG
AGTCCTCAGCGTTACCTGTTAGTAACTGGCAGTAGGGGTTGCGCTCGTTGCGGGACTTAA
CCCAACACCTCACGGCACGAGCTGACGACAGCCATGCAGCACCTGTGATAGCGCGTATTG
CTACGCTGATACATCTCTGCATCATTCACTACCATGTCAAACCCAGGTAAGGTTCCTCGC
GTATCATCGAATTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGA
GTTTCACCCTTGCGGGCATACTCCCCAGGCGGTGCACTTATCACGTTCGCTTCGGCACTG
AGCCGCTTGGCCCAACACCTAGTGCACATCGTTTATAGTGCGGACTACCAGGGTATCTAA
TCCTGTTTGCTCCCCGCACTTTCGCGCCTCAGCGTCAGTCATTGGCTAGTAGTTCGCCTTC
GCCACCGGTGTTCTTCCGCATATCTACAGATTTCACCCCTACACGCGGAATTCCAACTAC
CCCTCCATGA

CHAPTER 2

Exploring genomic evidence of acetogenic demethylation by *Treponema azotonutricium* str. ZAS-9

ABSTRACT:

Unlike the acetogenic termite hindgut spirochete isolates *Treponema primitia* str. ZAS-1 and ZAS-2, the isolate *T. azotonutricium* str. ZAS-9 does not display acetogenic activity *in vitro*. Moreover, *T. azotonutricium* str. ZAS-9 lacks formate dehydrogenase and methylene-tetrahydrofolate reductase, enzymes integral to the methyl branch of the acetyl-CoA (Wood-Ljungdahl) pathway of acetogenesis found in *T. primitia*. Together, these observations suggest that the strain is incapable of contributing to the termite system's acetate pool via acetogenesis. Nevertheless, *T. azotonutricium* str. ZAS-9's genome does have two carbon monoxide dehydrogenase homologs representing the other branch of the Wood-Ljungdahl pathway, the carboxyl branch, as well as 51 genes annotated as either methylases or methyltransferases which – with available methyl- or methoxylated substrates – can generate the same end product as the methyl branch of the pathway. Potentially, therefore, *T. azotonutricium* str. ZAS-9 can contribute to acetogenesis in the termite hindgut by coupling its methyl(transfer)ase and carboxyl branch capabilities. That the termite hindgut is replete with wood-derived meth(ox)ylated compounds also supports this notion. To test this hypothesis, *T. azotonutricium* str. ZAS-9 was grown on the meth(ox)ylated compounds DL-methionine, DMSO, methanol, methylamine, dimethylamine, trimethylamine, betaine, and trimethoxybenzoate. Specifically, since acetate is the fuel for both the termite host and its microbial

symbionts alike, possible increases in growth rate and yield of *T. azotonutricium* str. ZAS-9 were monitored as a proxy for metabolism of these meth(ox)ylated compounds into acetate. Despite convincing genomic evidence, however, no significantly different growth rates and yields were observed *in vitro*. While there is currently no *in vitro* evidence in support of acetogenic demethylation by *T. azotonutricium* str. ZAS-9, this does not mean that this function is not occurring *in situ* and in conjunction with other members of the termite hindgut microbial community. Therefore, this hypothesis could be re-evaluated using a combination of diverse microbial ecology research techniques discussed here.

INTRODUCTION:

Acetate is an important metabolite in the termite-hindgut microbial community mutualism

Hungate was the first to detect acetate in termite hindguts, specifically in the hindguts of the phylogenetically "lower" termites *Zootermopsis* sp. and *Reticulitermes claripennis* (1939, 1943). Decades later, it has become well-established that acetate is the "biofuel" of the termite host itself as well as its mutualistic hindgut microbial community (Odelson & Breznak 1983; Breznak & Switzer 1986; Breznak 1994; Leadbetter *et al*. 1999; Tholen & Brune 2000; Graber & Breznak 2004; Graber *et al*. 2004). In particular, acetate serves as the primary oxidizable energy source and a precursor of amino acids, hydrocarbons, and terpenes in the termite system (Blomquist *et al*. 1979; Mauldin 1982; Odelson & Breznak 1983; Breznak & Switzer 1986; Brauman *et al*. 1992; Breznak & Brune 1994). In *Reticuliterms flavipes*, for example, acetate generated by the hindgut

microbiota supports 77 to 100% of the respiratory requirement of the insect system (Mauldin 1982; Breznak & Switzer 1986).

In the wood-feeding, phylogenetically "lower" termites, symbiotic cellulolytic hindgut protozoa are known to hydrolyze wood polysaccharides and ferment the resulting sugar monomers into acetate, carbon dioxide, and hydrogen (*please see Chapter 1 Fig 1-2*) (Hungate 1955; Yamin 1980; Odelson & Breznak 1983; Odelson & Breznak 1985a; Odelson & Breznak 1985b; Breznak & Brune 1994). Odelson and Breznak (1983) hypothesized that additional acetate is generated from the H_2 + CO_2 fermentation byproducts, and Breznak and Switzer (1986) validated H_2 + CO_2 acetogenic activity in the termite hindgut. Specifically, acetogenic bacteria are responsible for transforming this H_2 + CO_2 into additional acetate (*please see Chapter 1 Fig 1-2*) (Breznak & Switzer 1986; Brauman *et al*. 1992; Leadbetter *et al*. 1999; Pester & Brune 2007). This bacteria mediated synthesis of acetate from fermentation-generated H_2 + CO_2 is significant and could account for up to approximately 1/3 of all the acetate produced during the hindgut fermentation (Breznak & Switzer 1986).

Acetogens isolated from the termite hindgut

Several H_2 + CO_2 acetogenic bacteria have been isolated from the hindguts of nutritionally diverse termites. Reflective of the importance of H_2 + CO_2 acetogenesis to termite nutrition, these organisms are diverse and include *Sporomusa termitida* obtained from the wood-feeding "higher" termite *Nasutitermes nigriceps* (Breznak *et al*. 1988), *Acetonema longum* from the dry wood-feeding "lower" termite

Pterotermes occidentis (Kane & Breznak 1991), *Clostridium mayombei* from the African soil-feeding "higher" termite *Cubitermes speciosus* (Kane *et al*. 1991), and *Treponema primitia* str. ZAS-1 and ZAS-2 from *Zootermopsis angusticollis* (Leadbetter *et al*. 1999).

While it is known that these organisms perform H_2 + CO_2 acetogenesis *in vitro*, there is not a good understanding of their contribution to acetogenesis and, in turn, to nutrition and energy in the termite hindgut (Brune 2006). Leadbetter's enlightening isolation of, and work with, *T. primitia* str. ZAS-1 and ZAS-2 illustrated that spirochetes - prominent members of termite hindgut microbial communities - could be significant contributors to the termite system's acetate pool given their numbers and acetogenic capabilities (1999). In addition to H_2 + CO_2 acetogenesis, both *T. primitia* strains were able to ferment homoacetogenically various hexoses, pentoses, and disaccharides (Graber & Breznak 2004). *T. primitia* str. ZAS-2 was found to also utilize methoxylated aromatic compounds (Graber & Breznak 2004).

Homoacetogenic growth on meth(ox)ylated compounds

In observing enrichment cultures from sewage sludge, Fischer was the first to report the formation of acetate from H_2 + CO_2 (1932). Then Wieringa reported the first bacterium to derive energy for growth by acetate synthesis from H_2 + CO_2, *Clostridium aceticum* (1936). Most acetogenic bacteria derive acetate from H_2 + CO_2 with H_2 serving as the electron donor for CO_2 reduction to acetate (Diekert 1992). Specifically, the methyl group of acetate is formed from CO_2 via formate and the carboxyl group is derived from CO, synthesized from O_2 by carbon monoxide

dehydrogenase (Diekert 1992). Nevertheless, acetogens are a very metabolically versatile group of anaerobic bacteria (Diekert 1992; Schink 1995). Most are able to grow on a variety of different substrates including sugars, C-1 compounds, methylated compounds, non-aromatic or aromatic methoxylated compounds, and alcohols (Schink 1995).

Acetogens that grow with meth(ox)ylated compounds are known as "demethylating acetogens" and include *Butyribacterium methylotrophicum* (Lynd & Zeikus 1983), *Eubacterium limosum* (Genthner *et al*. 1981), *Acetobacterium woodii* (Bache & Pfennig 1981), *Sporomusa* sp. (Breznak 1988), *Clostridium formicoaceticum,* and *Acetobacterium carbinolicum* (Diekert 1992). Commonly used methyl substrates are methanol (Zeikus *et al*. 1980), methoxylated aromatic compounds such as vanillate or syringate (Bache & Pfennig 1981), and methyl chloride (Traunecker *et al*. 1991). These and similar compounds may be formed as intermediates of wood decomposition in the termite hindgut and offer an alternative method of generating acetate (Bache & Pfennig 1981).

Experimental motivation

Two termite hindgut spirochete isolates, *Treponema primitia* str. ZAS-1 and ZAS-2, are bona fide H_2 + CO_2 acetogens (Leadbetter *et al.* 1999). In particular, acetogenesis from H_2 + ^{14}C-labelled CO_2 was found to support most of the growth of *T. primitia* str. ZAS-1 and ZAS-2 in accordance with acetyl-CoA (Wood-Ljungdahl) stoichiometries (Leadbetter *et al.* 1999). Both strains also demonstrated the use of other organic substrates homoacetogenically (Graber & Breznak 2004).

A third spirochete isolate from the same termite species hindgut, *T. azotonutricium* str. ZAS-9, however, does not display acetogenic physiology in pure culture (Graber *et al.* 2004). After first being obtained in pure culture, the genomes of the three *Treponema* species have become available for analysis and comparison (Rosenthal *et al.* 2011; Ballor *et al*. 2012). Although the physiology is not present *in vitro*, *T. azotonutricium*'s genome is not completely devoid of Wood-Ljungdahl pathway homologs (Rosenthal *et al.* 2011; Ballor *et al*. 2012).

As background, the acetyl-CoA (Wood-Ljungdahl) pathway of acetogenesis consists of a carboxyl branch and a methyl branch (Fig. 2-1) (Diekert 1992; Drake 1994; Ragsdale 1997). In the carboxyl branch, carbon monoxide dehydrogenase reduces CO_2 to CO forming the carboxyl group of acetate (Fig. 2-1). In the methyl branch, CO_2 is reduced to formate, which is subsequently bound to tetrahydrofolate generating formyl-tetrahydrofolate (Fig. 2-1). This formyl-tetrahydrofolate is then reduced to methyl-H4 folate via methenyl- and methylene-H4 folate (Fig. 2-1). The resulting methyl group is next transferred to a corrinoid protein that becomes the methyl group of acetate (Fig. 2-1) (Ljungdahl *et al*. 1966; Diekert 1992; Drake 1994; Ragsdale 1997).

T. azotonutricium str. ZAS-9 has genes with the requisite Pfam domains suggesting functionality for the carboxyl branch of acetogenesis (represented by carbon monoxide dehydrogenase, although initial studies did not detect activity for this enzyme; Graber & Breznak 2004) and homologs also with requisite Pfams for four of the six enzymes responsible for the methyl branch of acetogenesis (Fig. 2-1)

(Rosenthal *et al.* 2011; Ballor *et al*. 2012). Specifically, *T. azotonutricium* str. ZAS-9 has all of the genes for the Wood-Ljungdahl pathway of acetogenesis that *T. primitia* str. ZAS-1 and ZAS-2 have, with the exception of formate dehydrogenase and methylene-tetrahydrofolate reductase (Fig. 2-1) (Rosenthal *et al.* 2011; Ballor *et al*. 2012).

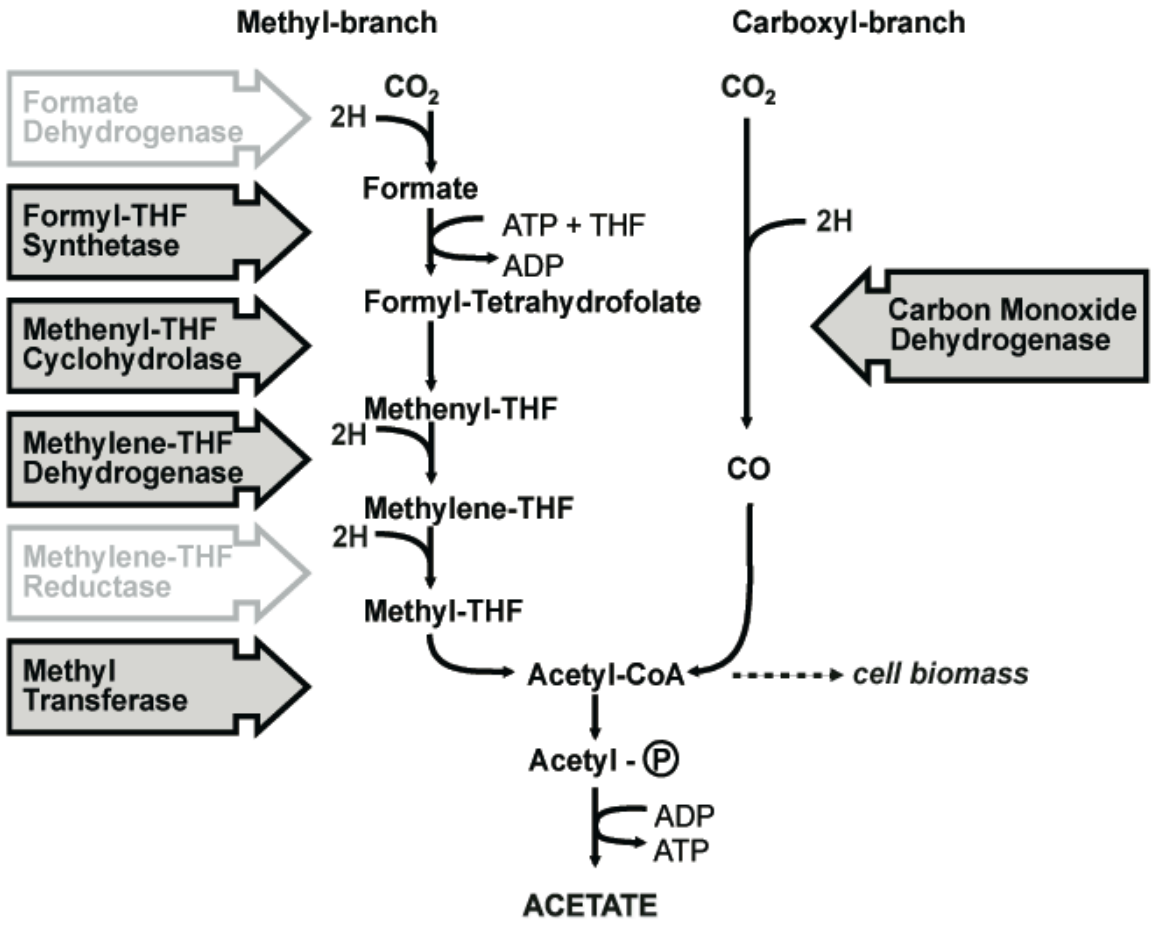


Fig. 2-1: Acetyl-CoA (Wood-Ljungdahl) pathway for CO_2-reductive acetogenesis. THF, tetrahydrofolate; Acetyl-P, acetyl-phosphate. Pathway enzymes with homologs in *Treponema azotonutricium* str. ZAS-9's genome are dark whereas pathway enzymes without homologs in *T. azotonutricium* str. ZAS-9's genome are light. Dashed line indicates intermediate can be instead utilized for biosynthetic processes. (*Figure courtesy of the Leadbetter Lab*).

Formate dehydrogenase is responsible for the first step of the methyl branch of the Wood-Ljungdahl pathway for the formation of the methyl group of acetate (Fig. 2-1) (Diekert 1992). It catalyzes the reduction of CO_2 to formate and is an oxygen sensitive enzyme (Fig. 2-1) (Diekert 1992). Given that the termite hindgut is actually an environment with varying levels of oxygen, it is plausible that *T. azotonutricium* str. ZAS-9 has lost this enzyme in favor of a route to generate acetate that can withstand an environment of varying oxygen levels (Brune *et al*. 1995).

T. azotonutricium str. ZAS-9 lacks another key methyl branch enzyme, methylene-tetrahydrofolate reductase (Fig. 2-1). While missing both formate dehydrogenase and methylene-tetrahydrofolate reductase means that *T. azotonutricium* str. ZAS-9 cannot generate the methyl group of acetate via this methyl branch, *T. azotonutricium* str. ZAS-9's genome contains 51 annotated methyl(transfer)ases (Appendix Table 2A-1). It is possible, therefore, that *T. azotonutricium* str. ZAS-9 can directly shunt methyl groups from meth(ox)ylated compounds in the termite hindgut environment into the Wood-Ljungdahl pathway to contribute to the methyl group of acetate, bypassing the majority of the methyl branch (Fig. 2-1) (Ljungdahl *et al*. 1966; Diekert 1992; Drake 1994; Ragsdale 1997). Although *T. azotonutricium* str. ZAS-9 does derive acetate as well as ethanol, H_2, and CO_2 from non-homoacetogenic fermentation of carbohydrates, this ability would allow the strain to contribute even more to termite nutrition via acetogenesis from demethylation of organic compounds. To test this hypothesis liquid cultures of *T. azotonutricium* str. ZAS-9 were grown with a variety of meth(ox)ylated compounds. Growth rates and yields were measured as a proxy for acetate production.

MATERIALS AND METHODS:

Media and cultivation

Routine *in vitro* growth and maintenance of *Treponema azotonutricium* str. ZAS-9 in 4YACo liquid medium under a headspace of 80% N_2/20% CO_2 was carried out as previously described (*please see Chapter 1 Materials and Methods*) (Graber *et al*. 2004). For examination of growth of *T. azotonutricium* str. ZAS-9 on several meth(ox)ylated compounds, the growth medium was maintained at 4YACo or reformulated to 1YACo (yeast autolysate for YACo media prepared from jars of Fleischmann's brand yeast, *please see Chapter 1*), and maltose was either increased to 40mM, retained at 20mM, decreased to 5mM, 2.5mM, or 1mM final concentration, or omitted. In some instances 100% H_2 was added sterily to the culture headspace of 80% N_2/20% CO_2 (to achieve a final concentration of 0.5% H_2, vol/vol) prior to or at the time of inoculation. Cultures were incubated in the dark, at room temperature, in a horizontal position, and without agitation.

Growth in the presence of meth(ox)ylated compounds

Individual 50mM stock solutions of DL-methionine, DMSO, methanol, methylamine, dimethylamine, trimethylamine, betaine, and trimethoxybenzoate were prepared sterily, under N_2, and stored in the dark (Table 2-1). As needed, stock solutions were neutralized with NaOH during the preparations.

To investigate the potential toxicity of these meth(ox)ylated compounds on *Treponema azotonutricium* str. ZAS-9, initial studies compared growth of the strain on each substrate at 1mM, 5mM, or 10mM final concentration. Controls were liquid

cultures of the same experimental conditions except with water instead of meth(ox)ylated substrate. As a result of preliminary studies, cultures of *T. azotonutricium* str. ZAS-9 were screened for metabolism of meth(ox)ylated compounds at a final concentration of 5mM. Changes in optical density of growing cultures was a proxy for meth(ox)ylated compound utilization and were measured at 600_{nm} with a Spectronic 20 colorimeter.

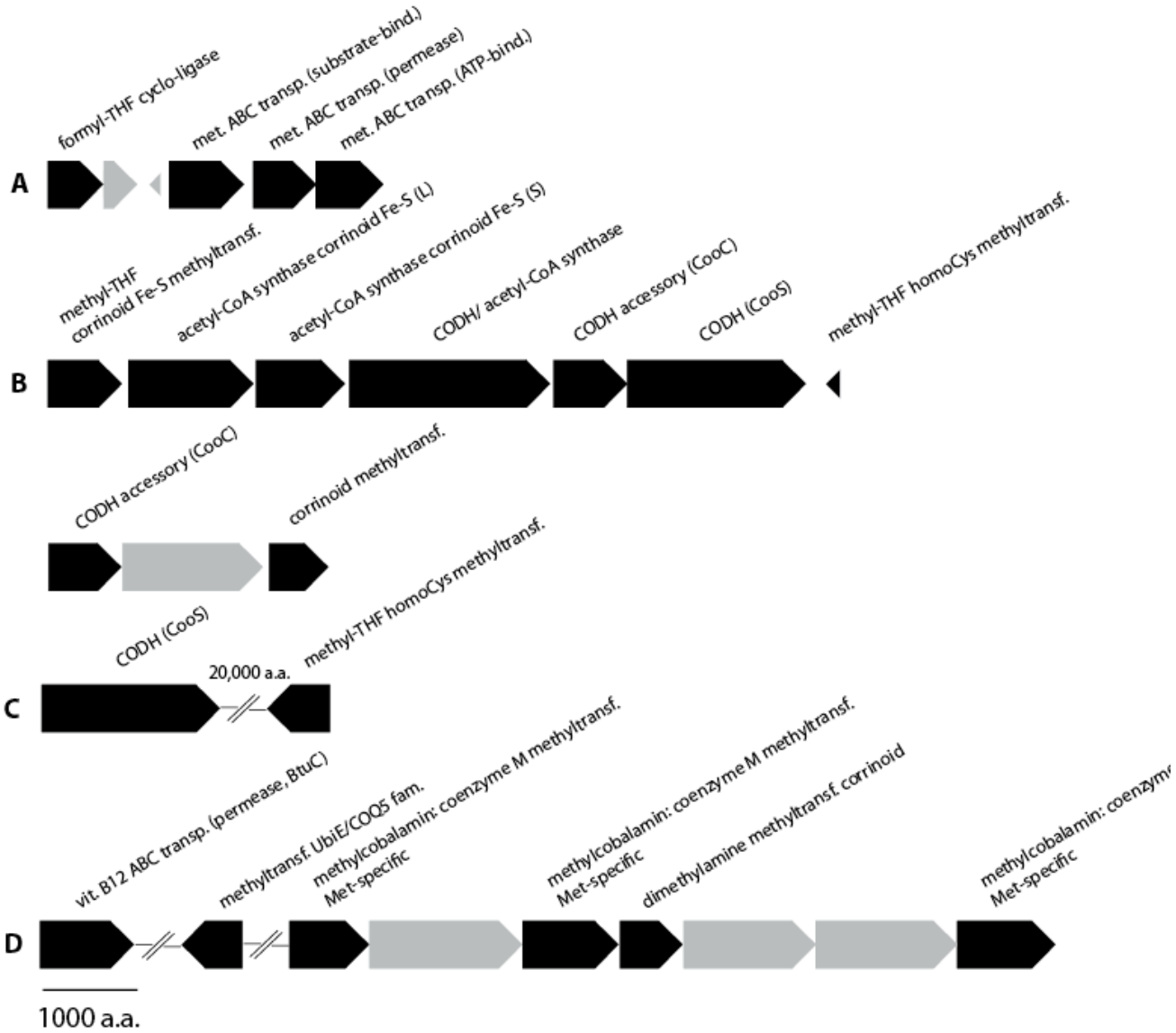
Bioinformatics

For identification of methylases and methyltransferases in *Treponema azotonutricium* str. ZAS-9's genome, data were obtained from RAST (Aziz *et al*. 2008), JGI IMG/M (Markowitz *et al*. 2006), and NCBI BLAST (Geer *et al*. 2010).

RESULTS:

***Treponema azotonutricium* str. ZAS-9 has putative methyl(transfer)ases in neighborhoods of acetyl-CoA (Wood-Ljungdahl) pathway homologs**

Several methyl(transfer)ases are in the gene neighborhoods of Wood-Ljungdahl pathway homologs in the genome of *T. azotonutricium* str. ZAS-9, and this arrangement might have implications to functionality (Fig. 2-2). This localization supports the hypothesis that *T. azotonutricium* str. ZAS-9 can bypass the methyl branch of the Wood-Ljungdahl pathway of acetogenesis and, instead, directly transfer methyl groups from wood-derived meth(ox)ylated compounds in the termite hindgut directly to carboxylic acid generated from the carboxyl branch of



the pathway to form acetate.

Fig. 2-2: *Treponema azotonutricium* str. ZAS-9 gene neighborhoods with both putative methyl(transfer)ases and homologs of the acetyl-CoA (Wood-Ljungdahl) pathway of acetogenesis. Relevant genes are in black, and unrelated or hypothetical proteins are in grey. THF, tetrahydrofolate; Met, methionine; transp., transporter; bind., binding; (L), large subunit; (S), small subunit; CODH, carbon monoxide dehydrogenase; Cys, Cysteine; methyltransf., methyltransferase; vit., vitamin; fam., family.

Treponema azotonutricium str. ZAS-9 growth on meth(ox)ylated compounds is inconclusive

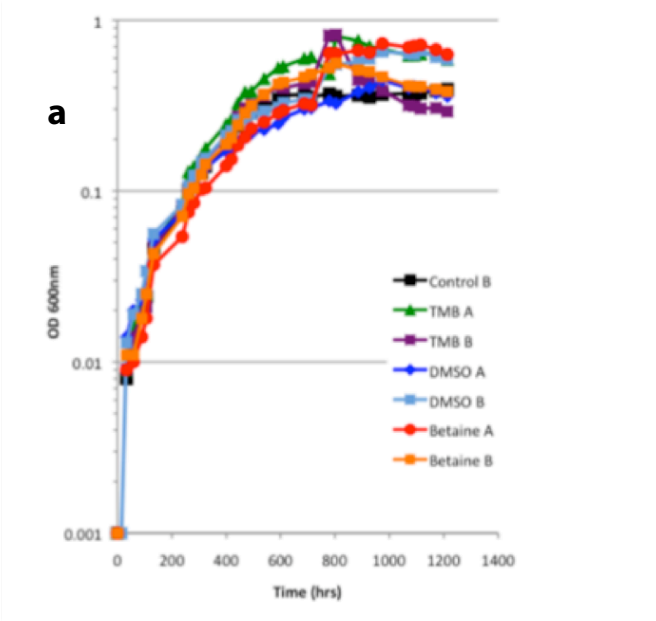
As a preliminary screen of potential toxicity of meth(ox)ylated compounds on *Treponema azotonutricium* str. ZAS-9, growth of the strain was evaluated on each substrate at 1mM, 5mM, or 10mM final concentration in 1YACo medium with 20mM maltose (data not shown). Growth was successful and growth yields were greatest on compounds provided at 5mM concentration compared to 1 and 10mM (data not shown). From this point on, substrates were added to cultures to 5mM final concentration for growth experiments.

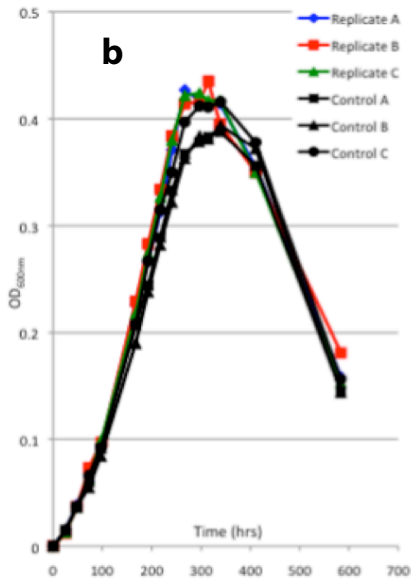
Next, as a first pass at comparison of growth of *T. azotonutricium* str. ZAS-9 on 5mM of each meth(ox)ylated substrate, liquid cultures of the strain were prepared in duplicate in 1YACo medium with 20mM maltose (Fig. 2-3a; Appendix Fig. 2A-1). A small but significant increase in growth yield was observed in *T. azotonutricium* str. ZAS-9 cultures grown with DMSO, betaine, and trimethoxybenzoate relative to growth on other substrates or just water as a control (Fig. 2-3a; Appendix Fig. 2A-1, Table 2A-2). No discernable differences in growth rate were observed (Fig. 2-3a; Appendix Fig. 2A-1).

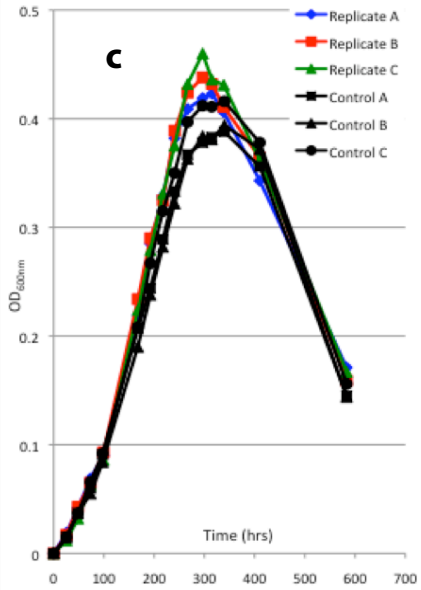
As a result of this initial comparison, betaine, DMSO, and trimethoxybenzoate appeared to be promising compounds for further testing, but significant increases in growth yield on these substrates relative to controls were small (Fig. 2-3a; Appendix Fig. 2A-1, Table 2A-2). Therefore, growth on all of the aforementioned meth(ox)ylated compounds was re-evaluated at 5mM concentration and with or without 5mM maltose. The amount of maltose added to experimental cultures was

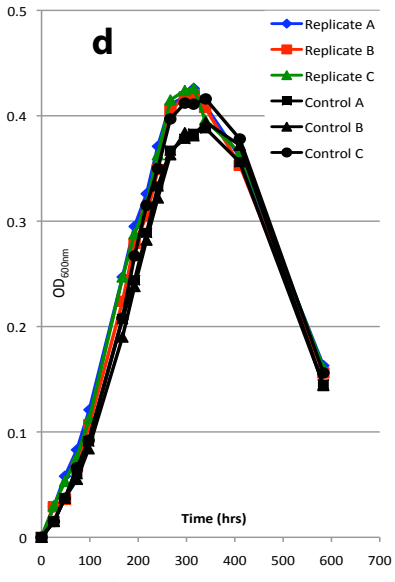
reduced on the assumption that 20mM maltose, as examined earlier, provided too much substrate for growth. Perhaps, *T. azotonutricium* str. ZAS-9 was preferentially utilizing maltose instead of the meth(ox)ylated compounds. Therefore, by reducing maltose to 5mM or omitting it, it was supposed that str. ZAS-9 would metabolize the meth(ox)ylated compounds and more striking growth yield differences would be seen. This time, "sparking" amounts of 100% H_2 were also added to the 80% N_2/20% CO_2 headspace in case H_2 had been a limiting factor in the previous experiment. A small but significant increase in growth yield was observed in *T. azotonutricium* str. ZAS-9 cultures grown with methylamine, dimethylamine, and trimethoxybenzoate relative to growth on other substrates or just water as a control (Fig. 2-3b, c, and d; Appendix Fig. 2A-2a, b, c, and d, Table 2A-3, Table 2A-4). No discernable differences in growth rate were seen (Fig. 2-3b, c, and d; Appendix Fig. 2A-2a, b, c, and d).

Because the most significantly different increase in growth yield relative to controls was observed with growth on dimethylamine, dimethylamine was examined with various YACo, maltose, and H_2 regimes in an effort to find the combination of YACo, maltose, and H_2 that would "kickstart" growth and prompt optimal utilization of the meth(ox)ylated compound. The increases in growth yield or rate with dimethylamine, however, were not significant despite evaluating all of these parameters (Fig. 2-3e; Appendix Fig. 2A-3a and b). The results obtained did not warrant further investigation using these approaches.


a
OD 600nm
Time (hrs)
Control B
TMB A
TMB B
DMSO A
DMSO B
Betaine A
Betaine B


b
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
Time (hrs)


c
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
Time (hrs)


d
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
Time (hrs)

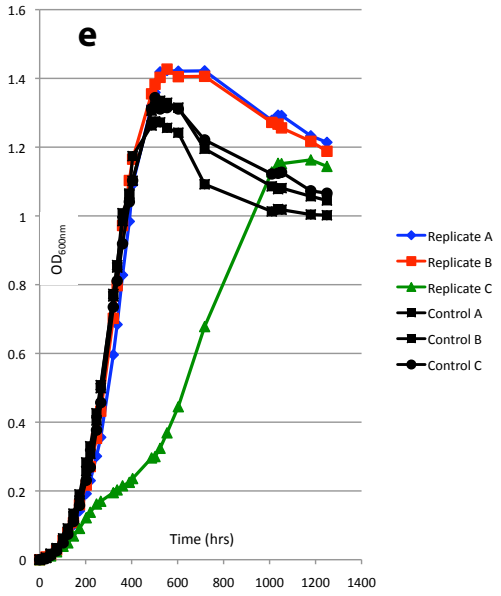
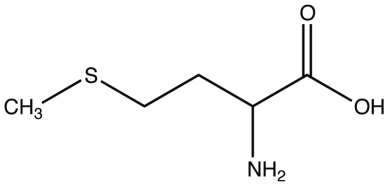
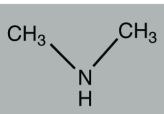
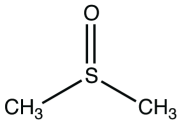
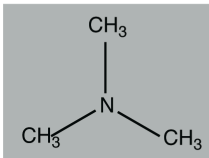
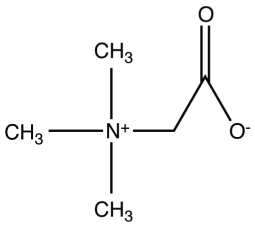
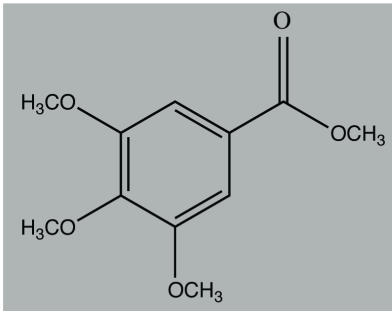

e
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

Fig. 2-3: Growth of *Treponema azotonutricium* str. ZAS-9 with various meth(ox)ylated compounds. Small but significant increases in growth yield were observed from 1YACo liquid cultures of *T. azotonutricium* str. ZAS-9 under a headspace of 80% N_2/20% CO_2 growing with **(a**) 5mM DMSO, betaine, and trimethoxybenzoate (TMB) and supplemented with 20mM maltose (initial experiment) and **(b)** 5mM methylamine, **(c**) 5mM dimethylamine, and **(d)** 5mM trimethoxybenzoate all supplemented with 5mM maltose and 100% H_2 to achieve a final concentration of 0.5% vol/vol (all second experiment results) compared to controls with water instead of the meth(ox)ylated compounds. Inconclusive increases in growth yield were observed from **(e)** liquid cultures of *T. azotonutricium* str. ZAS-9 growing in 4YACo medium with 5mM dimethylamine, 40mM maltose, and 100% H_2 to achieve a final concentration of 0.5% vol/vol compared to controls with water instead of the meth(ox)ylated compound. All cultures were incubated in the dark, at room temperature, in a horizontal position, and without agitation. Changes in optical density of growing cultures was a proxy for meth(ox)ylated compound utilization and were measured at 600_{nm} with a Spectronic 20 colorimeter. No discernable differences in growth rate were seen.

Table 2-1: Methylated compounds tested for growth on by *Treponema azotonutricium* str. ZAS-9

Monoaromatic Compound	Evidence of significantly different growth yield on by *T. azotonutricium* str. ZAS-9
$H_3C—OH$ methanol	neutral
$H_3C—NH_2$ methylamine	+
 DL-methionine	-
 dimethylamine	+
 dimethylsulfoxide	+
 trimethylamine	neutral
 betaine	+
 3, 4, 5-trimethoxybenzoate	+

DISCUSSION:

***Treponema azotonutricium* str. ZAS-9 growth on meth(ox)ylated compounds is inconclusive**

No significant differences in growth rates of liquid cultures of *Treponema azotonutricium* str. ZAS-9 with meth(ox)ylated compounds relative to controls were seen (Fig. 2-3a, b, c, d, and e; Appendix Fig. 2A-1, Fig 2A-2a, b, c, and d, Fig. 2A-3a and b). Any significant growth yield increases in experimental cultures compared to controls were small and not reproducible among the different meth(ox)ylated substrates tested (Fig. 2-3a, b, c, d, and e; Appendix Fig. 2A-1, Table 2A-2, Fig 2A-2a, b, c, and d, Table 2A-3, Table 2A-4, Fig. 2A-3a and b). Specifically, in 1YACo liquid medium under an 80% N_2/20% CO_2 headspace and supplemented with 20mM maltose, cultures of *T. azotonutricium* str. ZAS-9 displayed significant but relatively small increases in growth yield with 5mM DMSO, betaine, or trimethoxybenzoate (TMB) (Fig. 2-3a, b, c, d, and e; Appendix Fig. 2A-1, Table 2A-2). On the other hand, in 1YACo liquid medium under an 80% N_2/20% CO_2 headspace with 100% H_2 added to achieve 0.5% vol/vol final concentration, and supplemented with 5mM maltose or no maltose, cultures of *T. azotonutricium* str. ZAS-9 displayed significant but relatively small increases in growth yield with 5mM methylamine, dimethylamine, or TMB (Fig. 2-3b, c, and d; Appendix Fig. 2A-2a, b, c, and d, Table 2A-3, Table 2A-4). Although from these experiments dimethylamine appeared to be the most promising substrate, experiments focusing on dimethylamine with various regimes of YACo, maltose, and H_2 in an effort to observe greater differences in growth yield between experimental and control cultures were inconclusive (Fig. 2-3e; Appendix Fig. 2A-3a and b).

Future work:

Analyze Treponema azotonutricium str. ZAS-9 methyl(transfer)ases genes and gene neighborhoods

First, while methyl(transfer)ases were observed in the same gene neighborhoods as homologs of the acetyl-CoA (Wood-Ljungdahl) pathway of acetogenesis, it would be worthwhile to examine those neighborhoods for evidence of promoters, operons, and regulations to gauge function. Further, a simple exercise to also evaluate function would be to look for Pfam domains and key residues known to be important for functionality in these gene and protein sequences. The presence of key Pfam domains and amino acid residues could support the hypothesis of acetogenic demethylation by *T. azotonutricium* str. ZAS-9. The absence of key Pfam domains and amino acid residues would not necessarily rule out this hypothesis, but could re-direct studies keeping these differences in mind.

Decrease experimental concentration of meth(ox)ylated substrate

That dimethylamine was the substrate that led to the most promising possible increases in *T. azotonutricium* str. ZAS-9 growth yield corroborates earlier studies demonstrating that *T. primitia* str. ZAS-2 is able to use methoxylated aromatic compounds, such as TMB, as energy sources (Graber & Breznak 2004). Similar to the work with *T. azotonutricium* str. ZAS-9 presented herein, *T. primitia* str. ZAS-2 did not grow on methanol (Graber & Breznak 2004). In the work with *T. primitia* str. ZAS-2, however, meth(ox)ylated compounds were supplied at concentrations less than or equal to 2.5mM because higher concentrations inhibited growth (Graber & Breznak 2004). Perhaps testing growth with 5mM of meth(ox)ylated

aromatic compounds here was inhibitory. The maximum concentration of meth(ox)ylated substrates that did not lead to inhibitory effects was chosen for experimental evaluation under the assumption that it would lead to most product formation and observable differences between experimental and control cultures. However, perhaps growth was not visibly inhibited by 5mM concentrations but compound utilization was because enzyme substrate binding sites were over-loaded. In the future one could test a range of concentrations between the 1mM and 5mM concentrations examined here and simultaneously measure how much of the meth(ox)ylated is used. The experiment could then be repeated using that particular concentration in order to not overwhelm the system. Another potential experiment would be to re-evaluate this hypothesis after "priming" *T. azotonutricium* str. ZAS-9 that has been sub-cultured away from its original environment for over a decade with meth(ox)ylated compounds to "kickstart" this putative metabolism.

Directly measure meth(ox)ylated compound usage and acetate production

That significant but small boosts in growth yield were observed suggests that, under the experimental conditions examined herein, any possible acetate generation by *T. azotonutricium* str. ZAS-9 via the proposed acetogenic demethylation mechanism does not result in an increase in growth rate and yield such as can be detected by optical density measurements. While it is still undetermined whether or not *T. azotonutricium* str. ZAS-9 can generate acetogenesis via acetogenic demethylation, perhaps different experimental approaches such as radio-labeling and tracing of the methyl groups of the examined meth(ox)ylated compounds, or GC-MS, ICS, or LC-MS

measurements of culture media for acetate production over time of culture growth, will provide more conclusive answers.

Feeding experiments and co-culturing

Earlier and subsequent bioinformatic and genomic work with the *T. primitia meta*-pathway homologs also indicates that growth yield increases are not necessarily the best proxy for an acetate-yielding function (*please see Chapter 3*). This may be because *in situ*, perhaps, this acetate is not used by the generating organism, but is instead given to another member of the hindgut microbial community. For example, it has been suggested that related to their symbiotic relationships with the termite gut microbial community and the termite host itself, *T. primitia* str. ZAS-1 and ZAS-2 and *T. azotonutricium* str. ZAS-9 do not convert substrates to products for rapid growth and efficient generation of cell material (Graber & Breznak 1994; Graber *et al*. 1994). Rather, they give the acetate they produce to the rest of the hindgut microbial community and their host, and display limited production of biomass (Graber & Breznak 1994; Graber *et al*. 1994). This transfer of acetate may be able to be measured by conducting feeding experiments whereby cell-free *T. azotonutricium* str. ZAS-9 culture fluid is fed to other cultures that would demonstrate a marked increase in growth rate and yield from acetate. Success of co-culture experiments with *T. azotonutricium* and this organism could also be informative (*Please see Chapters 2 and 3*) (Rosenthal *et al*. 2011).

What is known ecologically about the termite hindgut ecosystem suggests that the ability to perform acetogenic demethylation would not only contribute to the

acetate pool in the termite system, but may also be a competitive strategy in a complex ecosystem. Because homoacetogens are specialists in C1 metabolism and compete with methanogens and sulfate reducers, methyl groups of meth(ox)ylated carbon compounds are a food source that is relatively less competed for (Schink 1995). Moreover, lignin is abundant in the termite hindgut and lignin derivatives have been postulated to act as methyl donors for methyl-group formation by homoacetogens (Schink 1995). In light of subsequent work with the *T. primitia meta*-cleavage pathway, if *T. azotonutricium* utilizes the methyl groups of lignin-derived aromatics, this could help prepare these substrates for the *T. primitia* catechol 2,3-dioxygenases (*please see Chapter 3*). This hypothesis could be further tested with *T. azotonutricium* and *T. primitia* co-culture experiments (per Rosenthal *et al*. 2011). If further *in vitro* analyses still do not demonstrate acetogenic demethylation by *T. azotonutricium* str. ZAS-9, the culture media conditions should be re-examined.

Enzyme activity assays and expression work

Lastly, it might be worthwhile to **clone** and express these genes in an expression vector such as ***E. coli*** to evaluate enzyme functionality and activity without variables in the *T. azotonutricium* cultures that may be negatively impacting these genes'/enzymes' functions. These genes can be expressed and evaluated individually or together and enzyme activity assays could b performed. qRT-PCR with primers designed to target relevant genes could also hint at expression or lack thereof.

If with additional work the hypothesis that *T. azotonutricium* str. ZAS-9 is able to bypass the methyl branch of the Wood-Ljungdahl pathway and generate acetate with methyl groups it takes from wood is not supported, at least these experiments demonstrate that an organism, even with a complete set of genes representing a particular function, does not necessarily perform that function.

REFERENCES:

Aziz RK, Bartels D, Best AA *et al*. (2008) The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, **9**, 75.

Bache R, Pfennig N (1981) Selective isolation of *Acetobacterium woodii* on methoxylated aromatic acids and determination of growth yields. *Archives of Microbiology*, **130**, 255-261.

Ballor NR, Paulsen I, Leadbetter JR (2012) Genomic analysis reveals multiple [FeFe] hydrogenases and hydrogen sensors encoded by treponemes from the H_2-rich termite gut. *Microbial Ecology*, **63**, 282-294.

Blomquist GJ, Howard RW, McDaniel CA (1979) Biosynthesis of the cuticular hydrocarbons of the termite *Zootermopsis angusticollis* (Hagen). Incorporation of propionate into dimethylalkanes. *Insect Biochemistry*, **9**, 371-374.

Brauman A, Kane MD, Labat M, Breznak JA (1992) Genesis of acetate and methane by gut bacteria of nutritionally diverse termites. *Science*, **257**, 1384-1387.

Breznak JA (1994) Acetogenesis from Carbon Dioxide in Termite Guts. In: *Acetogenesis* (ed Drake HL), pp. 303-330. Chapman & Hall, New York.

Breznak JA, Brune A (1994) Role of microorganisms in the digestion of lignocellulose by termites. *Annual Review of Entomology*, **39**, 453-487.

Breznak JA, Switzer JM (1986) Acetate synthesis from H_2 plus CO_2 by termite gut microbes. *Applied and Environmental Microbiology*, **52**, 623-630.

Breznak JA, Switzer JM, Seitz HJ (1988) *Sporomusa termitida* sp. nov., an H_2/CO_2-utilizing acetogen isolated from termites. *Archives of Microbiology*, **150**, 282-288.

Brune A (2006) Symbiotic Associations Between Termites and Prokaryotes. In: *Prokaryotes* (eds Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E), pp. 439-474. Springer, New York.

Brune A, Emerson D, Breznak JA (1995) The termite gut microflora as an oxygen sink: microelectrode determination of oxygen and pH gradients in guts of lower and higher termites. *Applied and Environmental Microbiology*, **61**, 2681-2687.

Diekert G (1992) The Acetogenic Bacteria. In: *The Prokaryotes* (eds Balows A, Trüper HG, Dworkin H, Harder W, Schleifer KH), pp. 517-533. Springer, New York.

Drake HL (1994) Acetogenesis, Acetogenic Bacteria, and the Acetyl-CoA "Wood-Ljungdahl" Pathway: Past and Current Perspectives. In: *Acetogenesis* (ed Drake HL), pp. 3-60. Chapman & Hall, New York.

Fischer F, Lieske R, Winzer K (1932) Biologische Gasreaktionen. II: Über die Bildung von Essigsäure bei der biologischen Umsetzung von Kohlenoxyd und Kohlensäure mit Wasserstoff zu Methan. *Biochemische Zeitschrift*, **245**, 2-12.

Geer LY, Marchler-Bauer A, Geer RC *et al*. (2010) The NCBI BioSystems database. *Nucleic Acids Research*, **38**, D492-D496.

Genthner BRS, Davis CL, Bryant MP (1981) Features of rumen and sewage sludge strains of *Eubacterium limosum*, a methanol and H_2-CO_2-utilizing species. *Applied and Environmental Microbiology*, **42**, 12-19.

Graber JR, Breznak JA (2004) Physiology and nutrition of *Treponema primitia*, an H_2/CO_2-acetogenic spirochete from termite hindguts. *Applied and Environmental Microbiology*, **70**, 1307-1314.

Graber JR, Leadbetter JR, Breznak JA (2004) Description of *Treponema azotonutricium* sp. nov. and *Treponema primitia* sp. nov., the first spirochetes isolated from termite guts. *Applied and Environmental Microbiology*, **70**, 1315-1320.

Hungate RE (1939) Experiments on the nutrition of *Zootermopsis*. III. The anaerobic carbohydrate dissimilation by the intestinal protozoa. *Ecology*, **20**, 230-245.

Hungate RE (1943) Quantitative analyses on the cellulose fermentation by termite protozoa. *Annals of the Entomological Society of America*, **36**, 730-739.

Hungate R (1955) Mutualistic Intestinal Protozoa. In: *Biochemistry and Physiology of Protozoa* (eds Hunter S, Lwoff A), pp. 159-199. Academic Press, New York.

Kane MD, Brauman A, Breznak JA (1991) *Clostridium mayombei* sp. nov., an H_2/CO_2 acetogenic bacterium from the gut of the African soil-feeding termite, *Cubitermes speciosus*. *Archives of Microbiology*, **156**, 99-104.

Kane MD, Breznak JA (1991) *Acetonema longum* gen. nov. sp. nov., an H_2/CO_2 acetogenic bacterium from the termite, *Pterotermes occidentis*. *Archives of Microbiology*, **156**, 91-98.

Leadbetter JR, Schmidt TM, Graber JR, Breznak JA (1999) Acetogenesis from H_2 plus CO_2 by spirochetes from termite guts. *Science*, **283**, 686-689.

Ljungdahl LG, Irion E, Wood HG (1966) Role of corrinoids in the total synthesis of acetate from CO_2 by *Clostridium thermoaceticum. Federation of American Societies for Experimental Biology Journal*, **25**, 1642-1648.

Lynd LH, Zeikus JG (1983) Metabolism of H_2-CO_2, methanol, and glucose by *Butyribacterium-methylotrophicum*. *Journal of Bacteriology*, **153**, 1415-1423.

Markowitz VM, Korzeniewski F, Palaniappan K *et al*. (2006) The integrated microbial genomes (IMG) system. *Nucleic Acids Research*, **34**, D344-D348.

Mauldin JK (1982) Lipid synthesis from [^{14}C]-acetate by two subterranean termites, *Reticulitermes flavipes* and *Coptotermes formosanus*. *Insect Biochemistry*, **12**, 193-199.

Odelson DA, Breznak JA (1983) Volatile fatty acid production by the hindgut microbiota of xylophagous termites. *Applied and Environmental Microbiology*, **45**, 1602-1613.

Odelson DA, Breznak JA (1985a) Nutrition and growth characteristics of *Trichomitopsis termopsidis*, a cellulolytic protozoan from termites. *Applied and Environmental Microbiology*, **49**, 614-621.

Odelson DA, Breznak JA (1985b) Cellulase and other polymer-hydrolyzing activities of *Trichomitopsis termopsidis,* a symbiotic protozoan from termites. *Applied and Environmental Microbiology*, **49**, 622-626.

Pester M, Brune A (2007) Hydrogen is the central free intermediate during lignocellulose degradation by termite gut symbionts. *ISME Journal*, **1**, 551-565.

Ragsdale SW (1997) The eastern and western branches of the Wood/Ljungdahl pathway: how the east and west were won. *BioFactors*, **6**, 3-11.

Rosenthal AZ, Matson EG, Eldar A, Leadbetter JR (2011) RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. *ISME Journal*, **5**, 1133-1142.

Schink B (1995) Diversity, Ecology, and Isolation of Acetogenic Bacteria. In: *Acetogenesis* (ed Drake HL), pp. 197-235. Chapman & Hall, New York.

Tholen A, Brune A (2000) Impact of oxygen on metabolic fluxes and *in situ* rates of reductive acetogenesis in the hindgut of the wood-feeding termite *Reticulitermes flavipes*. *Environmental Microbiology*, **2**, 436-449.

Traunecker J, Preuss A, Diekert G, Preuss A (1991) Isolation and characterization of a methyl-chloride utilizing, strictly anaerobic bacterium. *Archives of Microbiology*, **156**, 416-421.

Wieringa KT (1940) The formation of acetic acid from carbon dioxide and hydrogen by anaerobic spore-forming bacteria. *Antonie van Leeuwenhoek*, **6**, 251-262.

Yamin MA (1980) Cellulose metabolism by the termite flagellate *Trichomitopsis termopsidis*. *Applied and Environmental Microbiology*, **39**, 859-863.

CHAPTER 2: APPENDIX

Table 2A-1: Genes annotated as methyl(transfer)ases in the genome of *Treponema azotonutricium* str. ZAS-9.

Fig. 2A-1: Growth of *Treponema azotonutricium* str. ZAS-9 in 1YACo liquid medium under a headspace of 80% N_2/20% CO_2 with 5mM various meth(ox)ylated compounds (specified on graphs) supplemented with 20mM maltose.

Table 2A-2: Student's T-test P-value results (corresponding to Fig. 2A-1).

Fig 2A-2 a, b, c, and d: Growth of *Treponema azotonutricium* str. ZAS-9 in 1YACo liquid medium under a headspace of 80% N_2/20% CO_2 supplemented with 100% H_2 to achieve a final concentration of 0.5% vol/vol and 5mM various meth(ox)ylated compounds. Cultures represented in **c** and **d** also have 5mM maltose.

Table 2A-3: Student's T-test P-value results (corresponding to Fig. 2A-2 C and D).

Table 2A-4: Average highest OD600nm obtained of triplicate cultures.

Fig 2A-3 a and b: Growth of *Treponema azotonutricium* str. ZAS-9 in 1YACo (**a**) or 4YACo (**b**) liquid medium under a headspace of 80% N_2/20% CO_2 and 5mM dimethylamine with 100% H_2 to achieve a final concentration of 0.5% vol/vol and different maltose additions (specified on graphs).

Table 2A-2: Genes annotated as methyl(transfer)ases in the genome of *Treponema azotonutricium* str. ZAS-9.

Table 2A-1: Genes annotated as methyl(transfer)ases in the genome of *Treponema azotonutricium* str. ZAS-9.

Accession #	Annotation
YP_004527701	23S rRNA (guanine-N-2-) - methyltransferase rlmL EC 2.1.1.-)
YP_004527670	Ribosomal RNA small subunit methyltransferase F (EC 2.1.1.-)
YP_004527552	23S rRNA (Uracil-5-) - methyltransferase rumA (EC 2.1.1.-)
YP_004527273	tRNA (guanine46-N7-)-methyltransferase (EC 2.1.1.33)
YP_004527257	prophage LambdaSo, DNA modification methyltransferase, putative
YP_004527217	Ribosomal RNA small subunit methyltransferase D (EC 2.1.1.-)
YP_004527214	23S rRNA methyltransferase and Florfenicol/chloramphenicol resistance protein /Radical SAM family enzyme, UPF0063 family
YP_004527011	Methyltransferase/methylase
YP_004526961	S-adenosyl-methyltransferase mraW (EC 2.1.1.-)
YP_004526539	COG1092 family predicted tRNA methylase
YP_004526508	tetrapyrrole methylase family protein/MazG family protein
YP_004526426	Methylcobalamin:coenzyme M methyltransferase, methylamine-specific
YP_004526424	Methylcobalamin:coenzyme M methyltransferase, methylamine-specific
YP_004526423	Dimethylamine methyltransferase corrinoid protein
YP_004526420	Methylcobalamin:coenzyme M methyltransferase, methylamine-specific
YP_004526382	Predicted N6-adenine-specific RNA methylase containing THUMP domain
YP_004526350	5-methyltetrahydrofolate--homocysteine methyltransferase (EC 2.1.1.13)
YP_004526265	5-methyltetrahydrofolate:corrinoid

	iron-sulfur protein methyltransferase
YP_004526258	5-methyltetrahydrofolate--homocysteine methyltransferase (EC 2.1.1.13)
YP_004526233	Corrinoid methyltransferase protein
YP_004526169	Uroporphyrinogen-III methyltransferase (EC 2.1.1.107) /Uroporphyrinogen-III synthase (EC 4.2.1.75)
YP_004526097	Putative RNA 2'-O-ribose methyltransferase mtfA (EC 2.1.1.-)
YP_004526087	Tetrapyrrole (Corrin-Porphyrin) methylase family protein UPF0011
YP_004526074	TRNA/rRNA methyltransferase
YP_004526031	hypothetical tRNA/rRNA methyltransferase yfiF [EC:2.1.1.-]
YP_004526023	Methyltransferase (EC 2.1.1.-); N6-adenine-specific DNA methylase
YP_004525911	rRNA methylases
YP_004525892	Serine hydroxymethyltransferase (EC 2.1.2.1)
YP_004525883	Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)
YP_004525782	methyltransferase
YP_004525756	methylated-DNA--protein-cysteine methyltransferase-related protein
YP_004525659	Methylcobalamin:coenzyme M methyltransferase, methanol-specific
YP_004525658	Methylcobalamin:coenzyme M methyltransferase, methanol-specific
YP_004525617	Methylcobalamin:coenzyme M methyltransferase, methanol-specific
YP_004525614	RNA methyltransferase, TrmH family
YP_004525608	DNA modification methylase (Adenine-specific methyltransferase) (EC 2.1.1.72)
YP_004529019	Methyl-directed repair DNA adenine methylase (EC 2.1.1.72)
YP_004529009	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)

YP_004528983	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)
YP_004528953	tRNA (Guanine37-N1) -methyltransferase (EC 2.1.1.31)
YP_004528906	5-methyltetrahydrofolate--homocysteine methyltransferase (EC 2.1.1.13)
YP_004528849	Methyltransferase gidB (EC 2.1.-.-)
YP_004528840	HEN1 C-terminal domain; double-stranded RNA 3'-methylase
YP_004528749	Methylase of polypeptide chain release factors
YP_004528723	tRNA:Cm32/Um32 methyltransferase
YP_004528722	Small ribosomal subunit 16S rRNA methyltransferase ## U1498-specific in E.coli
YP_004528495	Aminomethyltransferase (glycine cleavage system T protein) (EC 2.1.2.10)
YP_004528467	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (EC 2.1.1.61)
YP_004528453	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)
YP_004528275	Glycine N-methyltransferase (EC 2.1.1.20)
YP_004528036	Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)

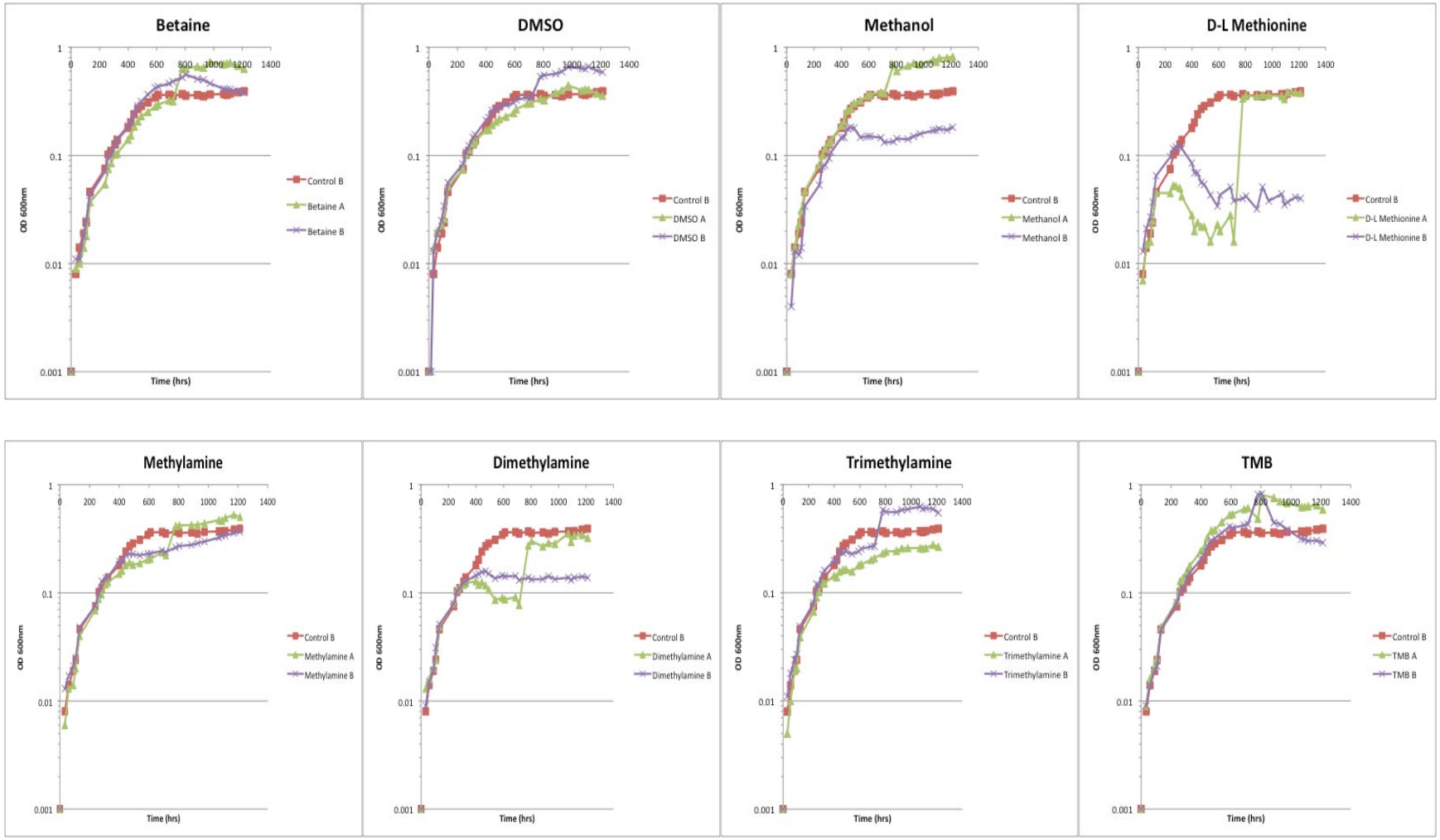

Betaine
Control B
Betaine A
Betaine B
DMSO
Control B
DMSO A
DMSO B
Methanol
Control B
Methanol A
Methanol B
D-L Methionine
Control B
D-L Methionine A
D-L Methionine B
Methylamine
Control B
Methylamine A
Methylamine B
Dimethylamine
Control B
Dimethylamine A
Dimethylamine B
Trimethylamine
Control B
Trimethylamine A
Trimethylamine B
TMB
Control B
TMB A
TMB B
OD 600nm
Time (hrs)
1
0.1
0.01
0.001
0
200
400
600
800
1000
1200
1400

Fig. 2A-1: Growth of *Treponema azotonutricium* str. ZAS-9 in 1YACo liquid medium under a headspace of 80% N_2/20% CO_2 with 5mM various meth(ox)ylated compounds (specified on graphs) supplemented with 20mM maltose. Controls have water instead of the meth(ox)ylated compounds. Small but significant increases in growth yield were observed from cultures of *T. azotonutricium* str. ZAS-9 growing with 5mM DMSO, betaine, or trimethoxybenzoate (TMB). These cultures were incubated in the dark, at room temperature, in a horizontal position, and without agitation. Changes in optical density of growing cultures was a proxy for meth(ox)ylated compound utilization and were measured at 600_{nm} with a Spectronic 20 colorimeter.

Table 2A-2: Student's T-test P-value results

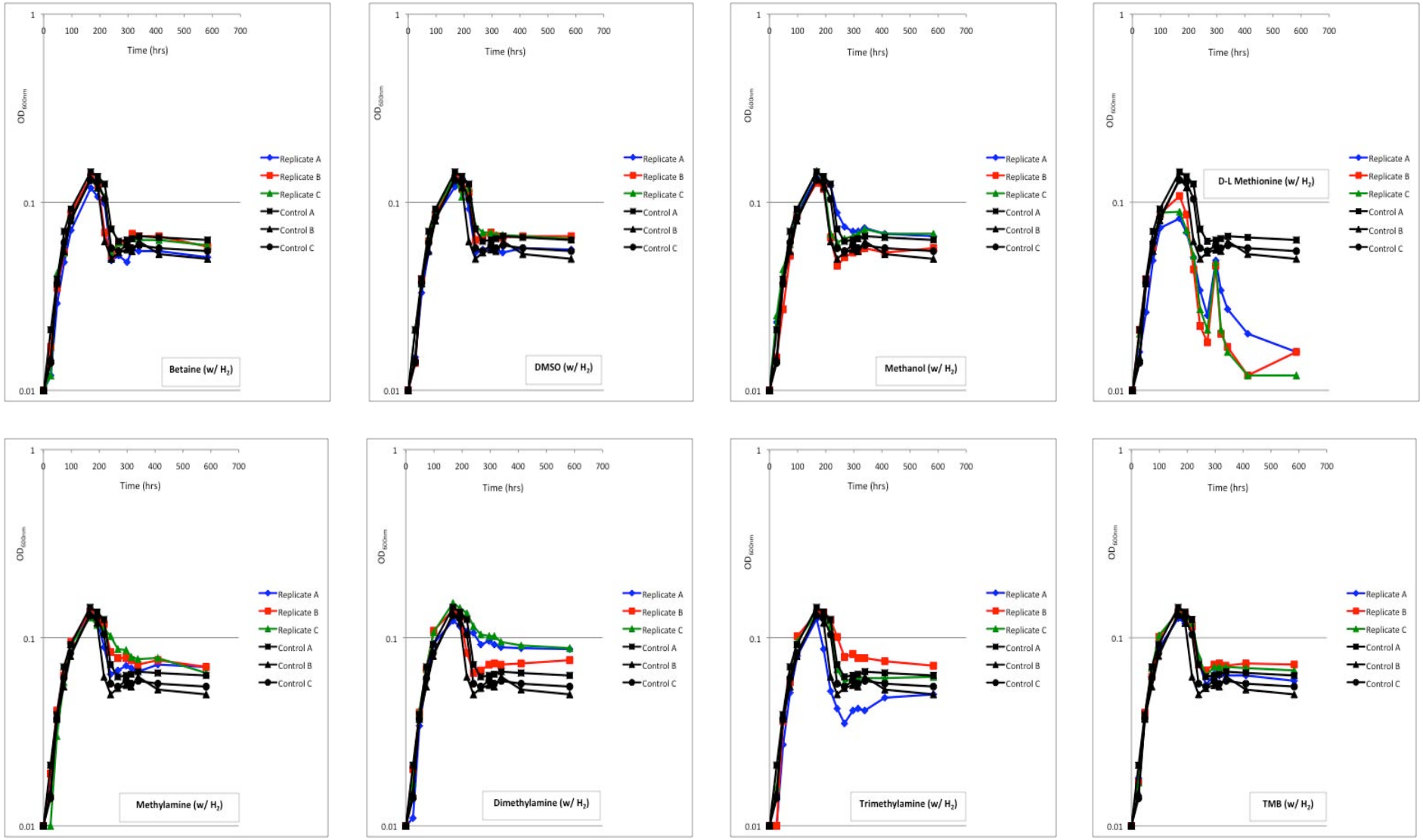
Meth(ox)ylated Compound	P-Value
Betaine	**0.0081**
DMSO	**0.037**
Methanol	0.37
D-L Methionine	0.57
Dimethylamine	0.52
Methylamine	0.11
Trimethylamine	0.29
TMB	**0.005**

Results correspond to Fig. 2A-1.

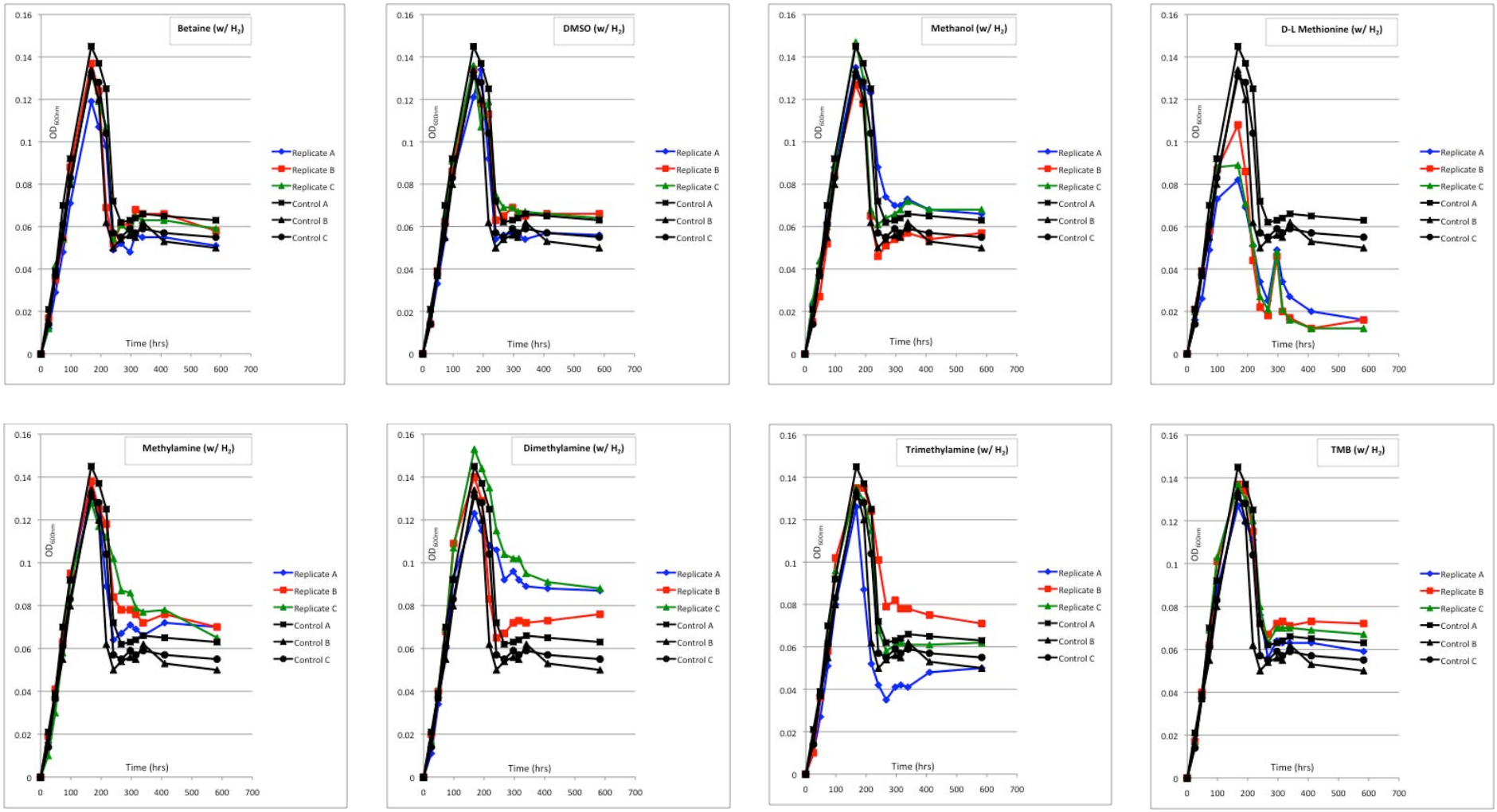
P-value compares highest growth yield (measured OD_{600nm}) achieved by *Treponema azotonutricium* str. ZAS-9 grown with 5mM of various meth(ox)ylated compounds, compared to highest growth yield achieved by controls with water *in lieu* of meth(ox)ylated substrate. Specifically, the means of replicate data representing each group are compared. All cultures were in 1YaCo liquid medium supplemented with 20mM maltose under an 80% N_2/20% CO_2 headspace.

95% confidence level was used, therefore a P-value < 0.05 signifies that the means of the data from each meth(ox)ylated treatment replicate is significantly different from that of the controls (bold).

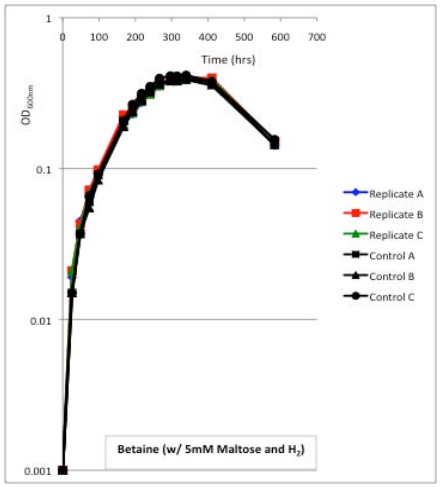
a

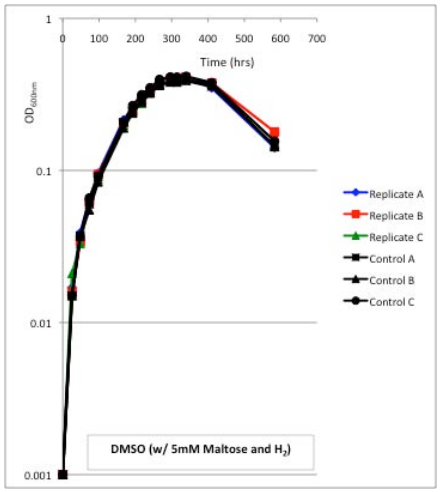

Time (hrs)
OD600nm
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
Betaine (w/ H2)
DMSO (w/ H2)
Methanol (w/ H2)
D-L Methionine (w/ H2)
Methylamine (w/ H2)
Dimethylamine (w/ H2)
Trimethylamine (w/ H2)
TMB (w/ H2)

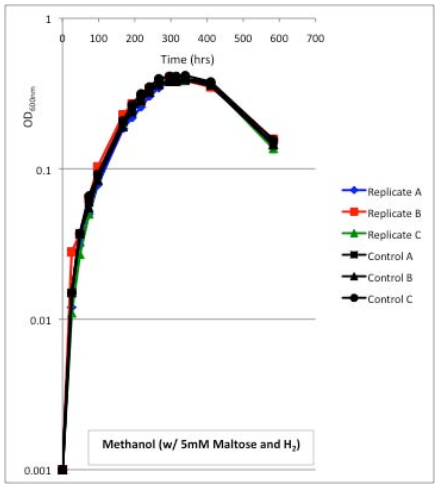
b

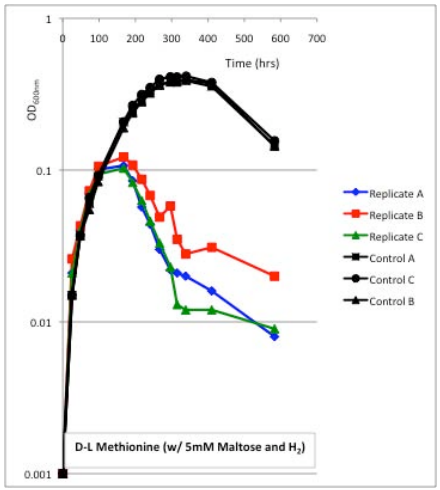

Betaine (w/ H_2)
DMSO (w/ H_2)
Methanol (w/ H_2)
D-L Methionine (w/ H_2)
Methylamine (w/ H_2)
Dimethylamine (w/ H_2)
Trimethylamine (w/ H_2)
TMB (w/ H_2)
OD_{600nm}
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

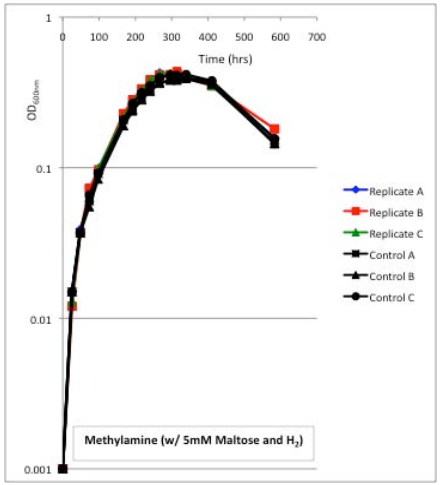
C

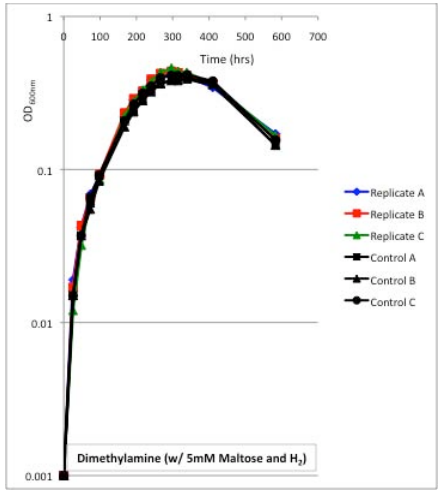

Time (hrs)
OD600nm
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
Betaine (w/ 5mM Maltose and H2)

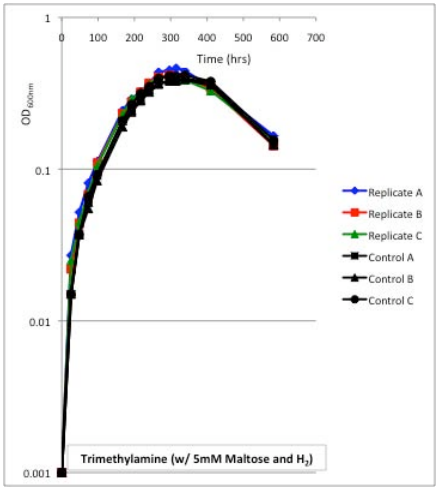

Time (hrs)
OD600nm
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
DMSO (w/ 5mM Maltose and H2)

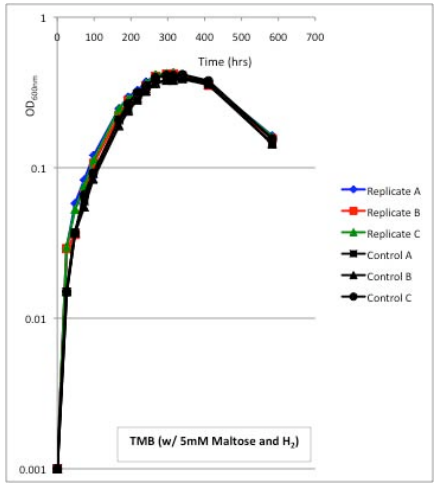

Time (hrs)
OD600nm
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
Methanol (w/ 5mM Maltose and H2)


Time (hrs)
OD600nm
Replicate A
Replicate B
Replicate C
Control A
Control C
Control B
D-L Methionine (w/ 5mM Maltose and H2)

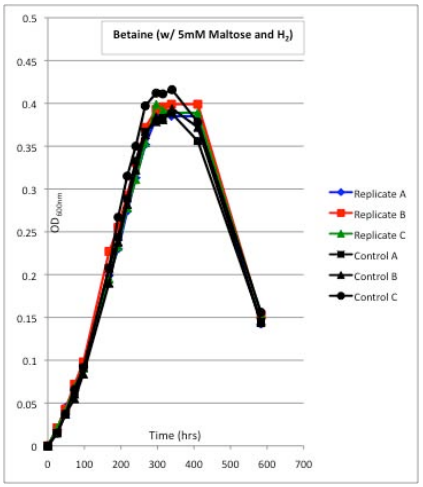

Time (hrs)
OD600nm
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
Methylamine (w/ 5mM Maltose and H2)

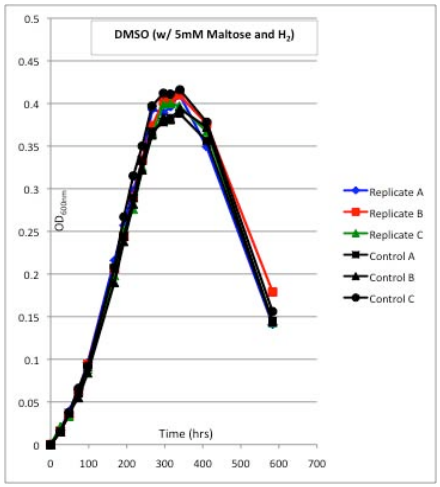

Time (hrs)
OD600nm
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
Dimethylamine (w/ 5mM Maltose and H2)

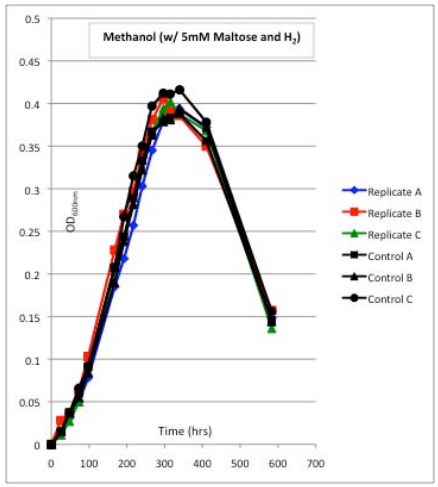

Time (hrs)
OD600nm
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
Trimethylamine (w/ 5mM Maltose and H2)

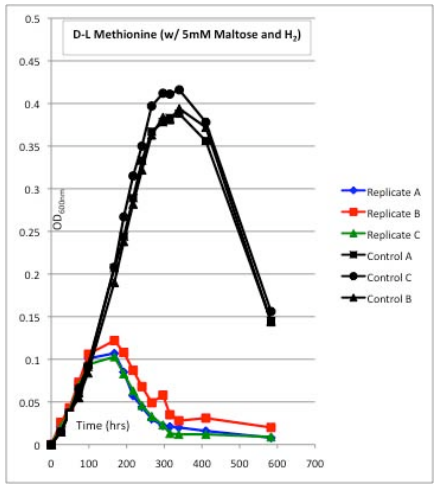

Time (hrs)
OD600nm
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C
TMB (w/ 5mM Maltose and H2)

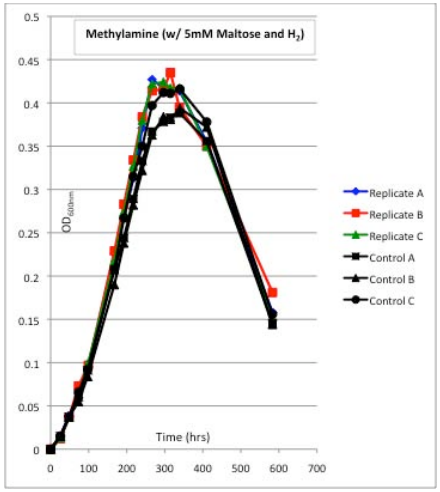
d

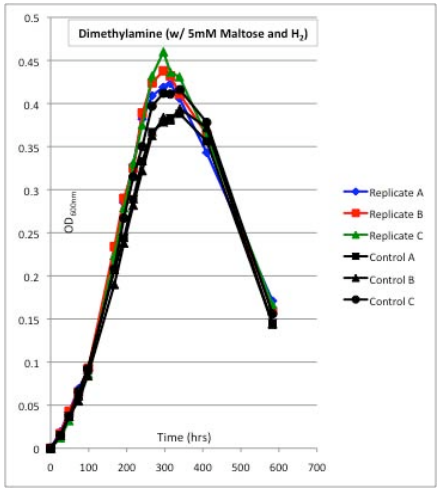

Betaine (w/ 5mM Maltose and H2)
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

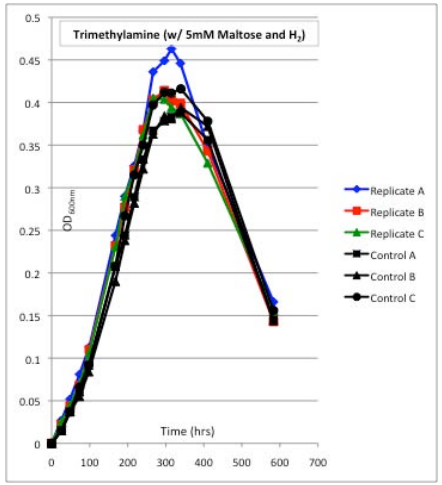

DMSO (w/ 5mM Maltose and H2)
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C


Methanol (w/ 5mM Maltose and H2)
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C


D-L Methionine (w/ 5mM Maltose and H2)
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control C
Control B


Methylamine (w/ 5mM Maltose and H2)
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C


Dimethylamine (w/ 5mM Maltose and H2)
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C


Trimethylamine (w/ 5mM Maltose and H2)
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

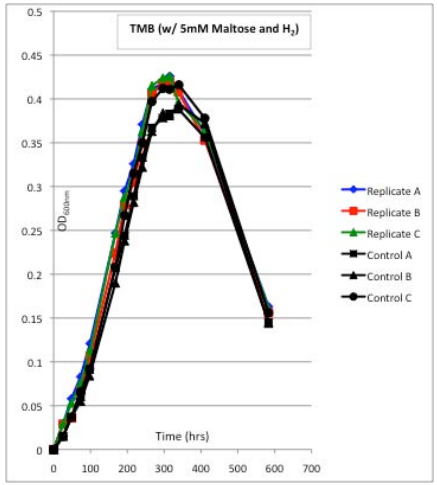

TMB (w/ 5mM Maltose and H2)
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

Fig. 2A-2: Growth of *Treponema azotonutricium* str. ZAS-9 in 1YACo liquid medium under a headspace of 80% N_2/20% CO_2 supplemented with 100% H_2 to achieve a final concentration of 0.5% vol/vol and 5mM various meth(ox)ylated compounds (specified on graphs). Controls have water instead of the meth(ox)ylated compounds. Some cultures also have 5mM maltose (**c** and **d**). **a** and **c** graphs are the logarithmic plots of **b** and **d** graphs, respectively. Small but significant increases in growth yield were observed from cultures of *T. azotonutricium* str. ZAS-9 growing with 5mM methylamine, dimethylamine, or trimethoxybenzoate (TMB) and with 5mM maltose. These cultures were incubated in the dark, at room temperature, in a horizontal position, and without agitation. Changes in optical density of growing cultures was a proxy for meth(ox)ylated compound utilization and were measured at 600_{nm} with a Spectronic 20 colorimeter.

Table 2A-3: Student's T-test P-value results

Meth(ox)ylated Compound	P-Value
Betaine	0.65
DMSO	0.45
Methanol	0.89
D-L Methionine	**< 0.0001**
Methylamine	**0.032**
Dimethylamine	**0.040**
Trimethylamine	0.23
TMB	**0.048**

Results correspond to Fig. 2A-2 **c** and **d**.

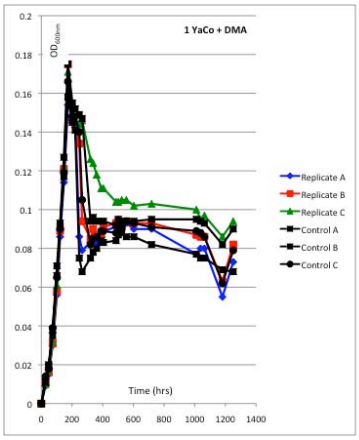
P-value compares highest growth yield (measured OD_{600nm}) achieved by *Treponema azotonutricium* str. ZAS-9 grown with 5mM of various meth(ox)ylated compounds, compared to highest growth yield achieved by controls with water *in lieu* of meth(ox)ylated substrate. Specifically, the means of replicate data representing each group are compared. All cultures were in 1YaCo liquid medium supplemented with 5mM maltose under an 80% N_2/20% CO_2 headspace to which 100% H_2 was added to 0.5% vol/vol final concentration.

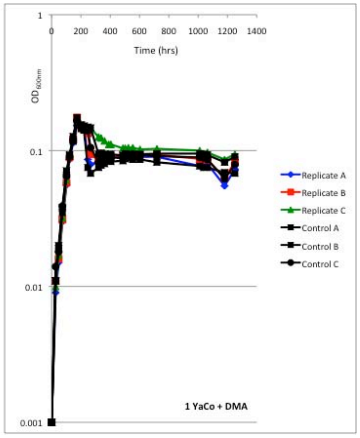
95% confidence level was used, therefore a P-value < 0.05 signifies that the means of the data from each meth(ox)ylated treatment replicate is significantly different from that of the controls (bold).

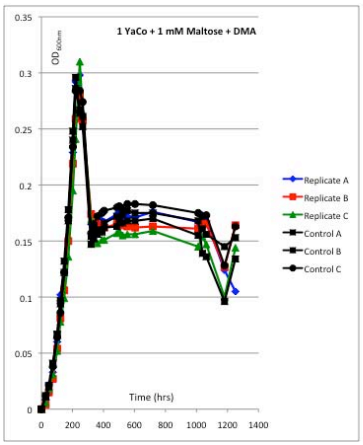
Table 2A-4: Average highest OD_{600nm} obtained of triplicate cultures

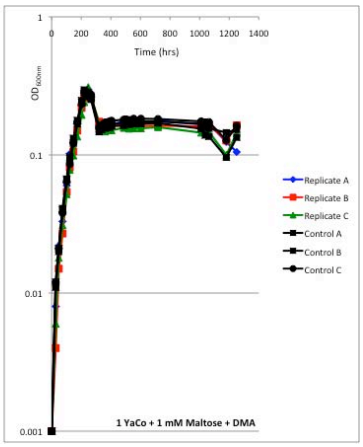
Sample	**Highest OD_{600nm} Obtained**
4 YaCo, 40mM Maltose	1.31
4 YaCo, 40mM Maltose (w/H_2)	1.23
1 YaCo, 20mM Maltose	0.484
1 YaCo, 20mM Maltose (w/H_2)	0.479
1 YaCo, 5mM Maltose	0.399
1 YaCo, 5mM Maltose (w/H_2)	0.399
1 YaCo	0.126
1 YaCo (w/H_2)	0.137
All of the following have H_2:	
1 Yaco, 5mM Betaine	0.130
1 YaCo, 5mM Betaine, 5mM Maltose	0.395
1 YaCo, 5mM DMSO	0.135
1 YaCo, 5mM DMSO, 5mM Maltose	0.407
1 YaCo, 5mM Methanol	0.136
1 YaCo, 5mM Methanol, 5mM Maltose	0.401
1 YaCo, 5mM D-L Methionine	0.093
1 YaCo, 5mM D-L Methionine, 5mM Maltose	0.111
1 YaCo, 5mM Methylamine	0.134
1 Yaco, 5mM Methylamine, 5mM Maltose	0.429
1 YaCo, 5mM Dimethylamine	0.139
1 YaCo, 5mM Dimethylamine, 5mM Maltose	0.440
1 YaCo, 5mM Trimethylamine	0.132
1 YaCo, 5mM Trimethylamine, 5mM Maltose	0.427
1 YaCo, 5mM TMB	0.134
1 YaCo, 5mM TMB, 5mM Maltose	0.424

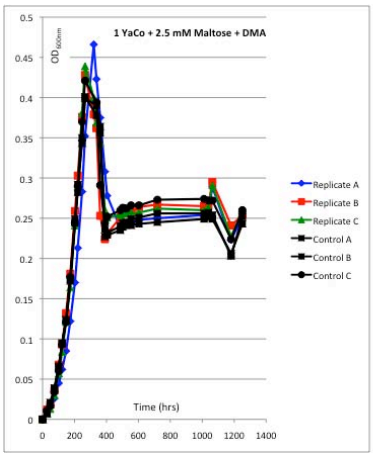
a

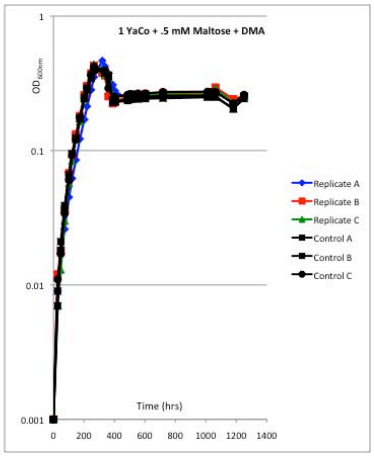

1 YaCo + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

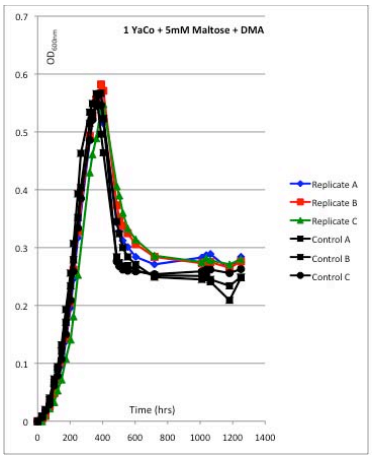

1 YaCo + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

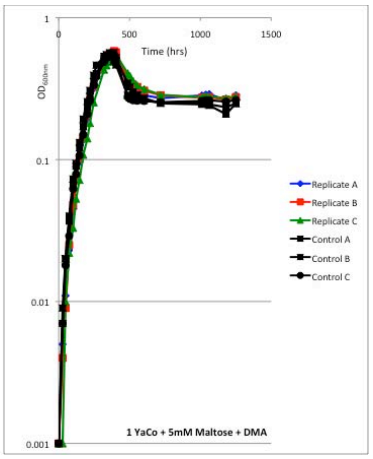

1 YaCo + 1 mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C


1 YaCo + 1 mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

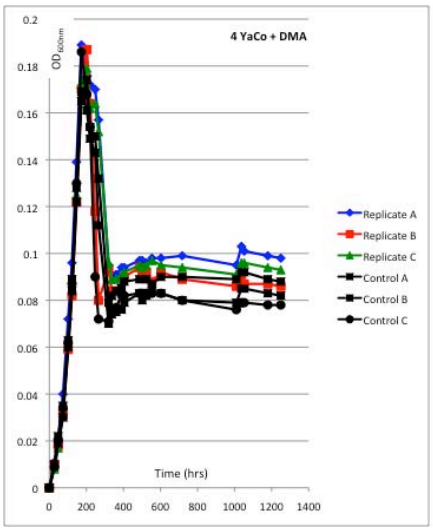

1 YaCo + 2.5 mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

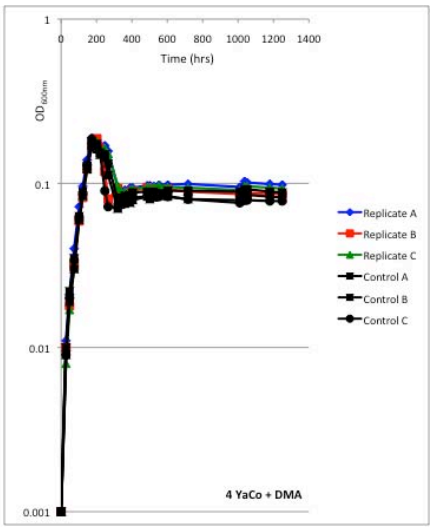

1 YaCo + .5 mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

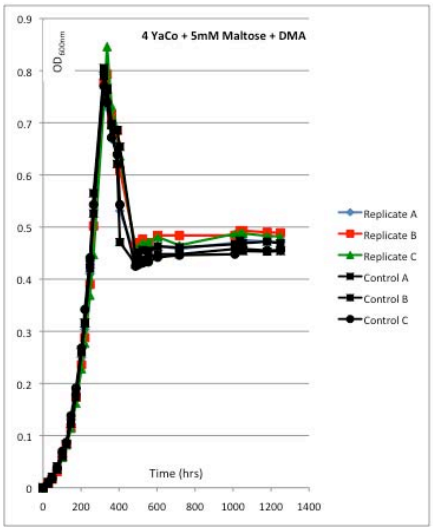

1 YaCo + 5mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

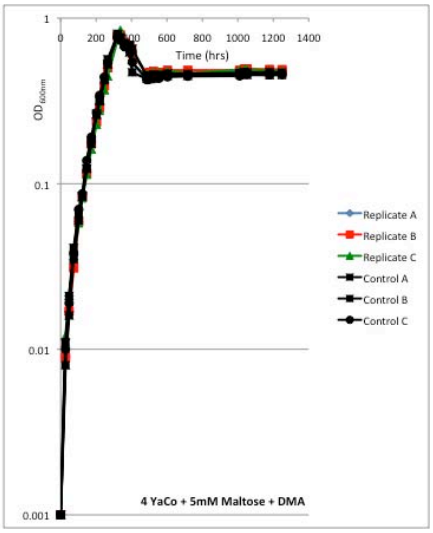

1 YaCo + 5mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

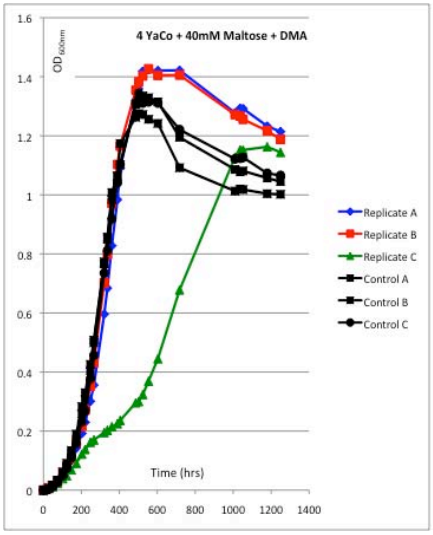
b


4 YaCo + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C


4 YaCo + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C


4 YaCo + 5mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C


4 YaCo + 5mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C


4 YaCo + 40mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

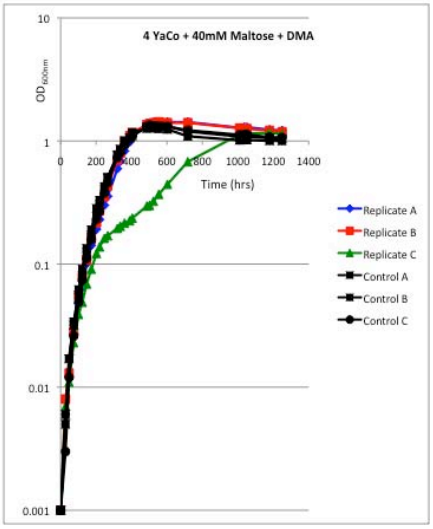

4 YaCo + 40mM Maltose + DMA
OD600nm
Time (hrs)
Replicate A
Replicate B
Replicate C
Control A
Control B
Control C

Fig. 2A-3: Growth of *Treponema azotonutricium* str. ZAS-9 in (**a**) 1YACo or (**b**) 4YACo liquid medium under a headspace of 80% N_2/20% CO_2 supplemented with 100% H_2 to achieve a final concentration of 0.5% vol/vol and dimthylamine with different maltose additions (specified on graphs). Controls have water instead of the meth(ox)ylated compounds. These cultures were incubated in the dark, at room temperature, in a horizontal position, and without agitation. Changes in optical density of growing cultures was a proxy for meth(ox)ylated compound utilization and were measured at 600_{nm} with a Spectronic 20 colorimeter.

CHAPTER 3

Catechol 2,3-dioxygenase and other *meta*-cleavage catabolic pathway genes in the "anaerobic" termite gut spirochete *Treponema primitia*

ABSTRACT:

Catechol 2,3-dioxygenases were observed in the genomes of strains of the symbiotic termite hindgut "anaerobe," *Treponema primitia*. Bioinformatics confirmed the presence of a cassette encoding genes, PFAM domains, and expected key residues for enzymes associated with a complete aromatic *meta*-cleavage pathway typical to aerobic pseudomonads. Curiously, transcripts for each gene were even observed expressed in strictly anaerobic cultures of *T. primitia*. Phylogenetic analyses suggest that the dioxygenase and several other essential genes of the *meta*-pathway were acquired by *T. primitia* from an alphaproteobacterium in the distant past, to augment several genes acquired from anaerobic firmicutes that do not directly catabolize aromatics but can contribute to final *meta*-pathway steps. To examine if the genes were functional, cultures of *T. primitia* were adapted to grow in unreduced media containing a microoxic headspace. The addition of catechol resulted in the transient accumulation of trace amounts of a yellow intermediate having the spectrophotometric characteristics of the expected ring cleavage product, hydroxymuconic semialdehyde. This is the first evidence for aromatic ring cleavage in the phylum (division) *Spirochetes*. Catechol 2,3-dioxygenase *meta*-cleavage pathways are not found widespread across the bacterial line of descent, previously having been observed, via physiology or genomics, restricted to select species representing only three of the major bacterial phyla (divisions):

Proteobacteria, *Firmicutes*, and *Actinobacteria*. However, *T. primitia* did not consume catechol to exhaustion nor exhibit any marked stimulations in growth or acetate yield. Nevertheless, the results suggest that the potential for O_2-dependent yet non-respiratory metabolisms of plant-derived and other aromatics should be re-evaluated in termite hindgut communities.

INTRODUCTION:

Many termites thrive on one of the planet's most abundant forms of biomass, wood, comprised primarily of cellulose, hemicellulose, and lignin (Breznak & Brune 1994). The presence and many roles of the complex, obligate, and nutritionally-mutualistic symbiotic microbial communities in termite guts during the metabolic processing of lignocellulose has long been established (Leidy 1877; Cleveland 1926; Hungate 1955; Yamin & Trager 1979; Yamin 1980). In the wood-feeding, phylogenetically "lower" termites, symbiotic hindgut protozoa are known to hydrolyze wood polysaccharides and ferment the resulting sugar monomers into acetate, carbon dioxide, and hydrogen (Hungate 1955; Yamin 1980; Odelson & Breznak 1983; Odelson & Breznak 1985a; Odelson & Breznak 1985b; Breznak & Brune 1994). Homoacetogenic bacteria then transform the H_2 + CO_2 into additional acetate (Breznak & Switzer 1986; Brauman *et al*. 1992; Leadbetter *et al*. 1999; Pester & Brune 2007). Polysaccharide-derived acetate serves as the oxidizable energy source as well as a precursor of amino acids, hydrocarbons, and terpenes in the wood-feeding host insect (Blomquist *et al*. 1979; Mauldin 1982; Odelson & Breznak 1983; Breznak & Switzer 1986; Brauman *et al*. 1992). Together, termites and their gut

microbiota are prime examples of tiny and natural bioreactors operative on a massive scale in many terrestrial environments (Brune 1998a).

As reviewed previously (Breznak & Brune 1994), the fate of the most inaccessible component of lignocellulose in the diet of wood feeding termites, lignin, has been and remains comparatively unclear and often controversial (Butler & Buckerfield 1979; Cookson 1987; Pasti *et al*. 1990; Breznak & Brune 1994; Kuhnigk *et al*. 1994; Scharf & Tartar 2008; Tartar *et al*. 2009; Scharf & Boucias 2010; Sethi *et al*. 2013). Reliably measuring the oxidation of C^{14}-labelled lignins before, during, and after passage through the termite gut tract has been challenging (Butler & Buckerfield 1979; Cookson 1987; Pasti *et al*. 1990; Breznak & Brune 1994; Kuhnigk *et al*. 1994; Brune *et al*. 1995a). In addition, studies have hypothesized that O_2 is a necessary co-substrate for the complete degradation of lignin aromatic monomers (Brune *et al*. 1995a). Studies over the past two decades have revealed that the hindguts of termites contain various levels of oxygen in the periphery, and that only the central portions of the hindgut are totally anoxic (Brune *et al*. 1995b; Brune 1998a; Kappler & Brune 1999; Kappler & Brune 2002; Zimmer & Brune 2005; Pester & Brune 2007). Cultivation and gene-based results have revealed that microorganisms with varying degrees of oxygen tolerance and oxygen metabolisms are operative as bona fide members of termite gut communities (Leadbetter & Breznak 1996; Shima *et al*. 1999; Shima *et al*. 2001; Boga & Brune 2003; Boga *et al*. 2003; Tholen *et al*. 2007; Wertz & Breznak 2007a; Wertz & Breznak 2007b; Wertz *et al.* 2012).

In an effort to glean more about roles of the gut microbiota in the symbiosis, the genomes of three strains of sugar and H_2 metabolizing termite gut symbiotic spirochetes and the partial termite gut metagenome, all previously described, were compared (Warnecke *et al*. 2007; Rosenthal *et al*. 2011; Ballor *et al*. 2012). Many genes of interest in the metagenomic dataset were represented in the genomes of the isolates, confirming another level of their validity as models for *in vitro* studies (Warnecke *et al*. 2007; Rosenthal *et al*. 2011; Ballor *et al*. 2012). Perhaps more surprising, though, were numerous examples of genes present in the isolates that were not represented in the larger (albeit not exhaustive) metagenomic dataset. Here, a subset of such genes that have relevance to oxygen and aromatic metabolism in the termite hindgut from two strains of *Treponema primitia* are described.

MATERIALS AND METHODS:

Media and cultivation

Routine *in vitro* growth and maintenance of *Treponema primitia* str. ZAS-1 and ZAS-2 in 2YACo medium was as previously described (Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004). For examination of the microoxic growth of *T. primitia* str. ZAS-1 and ZAS-2, the growth medium was reformulated to contain no dithiothreitol (DTT) or other reducing agent, no resazurin (a redox indicator), and no maltose (unless used as a substrate for positive controls). The headspace was 80% N_2/20% CO_2. To serum-stoppered culture tubes of this unreduced medium, room air was injected through a 0.2μm filter into the culture headspace (to achieve a final concentration of 0.5% O_2, vol/vol) prior to or at the time of inoculation.

These cultures were incubated at 25°C in a horizontal position without agitation. Because *T. primitia* str. ZAS-1 grew more successfully than *T. primitia* str. ZAS-2 under these conditions, *T. primitia* str. ZAS-1 was used for further liquid culture physiological experiments with this microorxic regime.

It is noteworthy that the "YACo" media formulation contains cofactors required of all the *meta*-cleavage pathway enzymes (catechol 2,3-dioxygenase requires Fe^{2+}; 2-oxopent-4-enoate hydratase requires Mn^{2+} or Mg^{2+} (Izumi *et al*. 2007); 4-hydroxy-2-oxopentanoate aldolase rquires Mn^{2+} (Lei *et al*. 2008); acetaldehyde dehydrogenaserequires CoASH and NAD^+ (Lei *et al*. 2008)).

Growth in the presence of aromatic compounds

Individual 50mM stock solutions of catechol, guaiacol, veratrole, protocatechualdehyde, protocatechuic acid, vanillic acid, gallic acid, homoprotocatechuic acid, homoveratric acid, hydrocaffeic acid, caffeic acid, and ferulic acid were prepared under N_2 and stored in the dark. As needed, stock solutions were neutralized with NaOH during the preparations.

Liquid cultures of *T. primitia* str. ZAS-1 (prepared as described above for examination of microoxic growth and with 0.5mM aromatic compound) were screened for increases in growth rate and yield as a proxy for acetate generation from *meta*-cleavage pathway metabolism of aromatic compounds. Changes in optical density of growing cultures were measured at 600_{nm} with a Spectronic 20 colorimeter. UV/Vis absorbance spectra were obtained using a Cary WinUV Spectrophotometer.

For the examination of growth in oxygen gradients, 1mL of molten 3% agarose containing 10mM catechol, protocatechuic acid, hydrocaffeic acid, or caffeic acid (or water as a negative control) was dispensed into 25mL Balch tubes. After hardening, this plug was overlaid with 10mL of a molten agarose medium. The medium used was the unreduced 2YACo medium supplemented with low melting point agarose to a final concentration of 0.8%, and resazurin as an oxygen and redox indicator. Upon agarose additions to the Balch tubes, care was taken to ensure residual agarose did not stick to the walls of the tube so that the level of agarose in the tubes was consistent among the tubes. Once the agarose hardened, sterile tuberculin syringes with 6in., 18ga. needles were used to inoculate a consistent volume (400μL) of *T. primitia* cells (or water as a control) along the vertical length of the agarose. Room air was then injected through a 0.2μm filter into the culture headspace to achieve a headspace concentration of 4% O_2, 70% N_2, 10% CO_2, vol/vol. Cultures were cultivated at 25°C in a vertical position. *T. primitia* str. ZAS-2 was used for these initial gradient tube experiments in addition to *T. primitia* str. ZAS-1 to see if it could establish itself in an optimal O_2/aromatic gradient and grow more successfully than in liquid media. Still, however, *T. primitia* str. ZAS-1 performed better under these conditions and was used for subsequent gradient tube experiments.

Enrichments

Enrichment cultures to obtain a fresh, O_2/aromatic-metabolizing isolate from the termite hindgut environment were prepared from the whole gut contents of a *Zootermopsis nevadensis* worker termite. Enrichment culture media was prepared

as described above for examination of microoxic growth, also with 0.5mM aromatic compound (*see Chapter 1 for enrichment and isolation protocol*). Enrichment cultures with ferulic acid or vanillic acid did transiently turn various shades of yellow for the first two sets of transfers, but not after, suggestive that at some early point in enrichment work O_2/aromatic-metabolizing organisms might have been present in these enrichments. This bodes promising for future attempts at enrichment and isolation.

Bioinformatics

Sequences were obtained from a variety of sources: JGI IMG/M (Markowitz *et al*. 2006); NCBI BLAST (Geer *et al*. 2010); Sanger Pfam (Punta *et al*. 2012); CAMERA (Sun *et al*. 2011); and JCVI CMR (Davidsen *et al*. 2010) to collect as complete of a dataset as possible. For the construction of phylogenetic trees, initially individual datasets ranging from approximately 500 to 1500 top BLAST sequence hits corresponding to each of the *Treponema primitia meta*-cleavage pathway proteins, or Pfam domains, were assembled. Datasets were collected by looking for natural cut-offs in sequence identity and similarity in BLAST results, instead of choosing an arbitrary number of the top BLAST hits, as to not miss potentially relevant sequences.

For each protein or Pfam CLUSTALW, MAFFT, DIALIGN, T-Coffee, and MUSCLE alignments were generated on the Mobyle Server and compared (Néron *et al*. 2009). Consistently, MUSCLE generated the best alignments and therefore sequences were unambiguously aligned using this program (Fig. 3-3b). Conservative filters were

used to disregard columns of data containing sequence gaps or ambiguous alignments. After initial analyses, phylograms using fewer and more relevant sequences from the original large datasets were prepared using packages within ARB v. 5.2 (Ludwig *et al*. 2004) and Mr. Bayes v. 3.2 using mixed-model settings (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Additional details of tree construction are found in the tree legends. Accession numbers are reported in Appendix (Appendix Table 3A-1).

Putative promoter regions within or adjacent to the *meta*-cleavage pathway gene neighborhoods were determined using Virtual Footprint v. 3.0 (Münch *et al*. 2005). For examination of the genes using selection pressure models, MEGA v. 5.05 (Tamura *et al*. 2007) was used. The several models tested included a positive selection hypothesis (dS > dN) via a codon-based Z-test: NG86 (Nei-Gojobori method with Jukes-Cantor nucleotide substitution model); modified NG86 (modified Nei-Gojobori methods with the Jukes-Cantor nucleotide substitution model); LWL85 (Li-Wu-Lou method); PBL85 (Pamilo-Bianchi-Li method); and Kumar (Kumar method). Results are reported as the overall average of pair-wise analyses of the following number of unambiguous codon positions without gaps for each gene: ferredoxin-like peptide, 73; catechol 2,3-dioxygenase, 291; 2-hydroxymuconic semialdehyde hydrolase, 264; 2-oxopent-4-enoate hydratase, 260; 4-hydroxy-2-oxopentanoate aldolase, 313; and acetaldehyde dehydrogenase, 261. All analyses were run with 1000 bootstrap replicates. Z must be less than zero (dS > dN) in a statistically significant manner to indicate positive selection. “ND” denotes cases in which it was not possible to estimate evolutionary distances.

For examination of key residues and other features of the putative *meta*-cleavage pathway proteins in *T. primitia* str. ZAS-1 and ZAS-2, as well as in other organisms, data were obtained from RAST (Aziz *et al*. 2008), Meta-Cyc (Caspi *et al*. 2012), and KEGG (Kanehisa & Goto 2000), in addition to the databases noted above. These analyses, coupled with extensive literature review, were also used to determine where across the bacterial line of descent catechol 2,3-dioxygenase *meta*-cleavage pathways are found.

For examination of relevant gene expression under routine cultivation conditions, data obtained in a previous study was re-analyzed (Rosenthal *et al*. 2011). First, expression data from all *T. primitia* str. ZAS-2 genes were compared to gauge the range of reads/kb measurements (Appendix Table 3A-2) (Rosenthal *et al*. 2011). Next, expression data for each *meta*-cleavage pathway gene was evaluated relative to other genes including those representing house keeping and other and metabolic functions (Appendix Table 3A-3).

Enzyme activity assays

Activity assays for catechol 2,3-dioxygenase were conducted with 0.5mM catechol, 50mM potassium phosphate buffer, and 10μL crude cell lysate in a final volume of 1mL at room temperature in the dark. Catechol was added to begin reaction and reaction was exposed to room air and inverted for O_2 exposure. Although many iterations of assays were performed, none yielded results. Iterations were a combination of:

- Adding catechol and O_2 to cultures at various time points prior to the assay to allow cultures to begin to express catechol 2,3-dioxygenase in anticipation of assay
- Not adding catechol or O_2 until the time of the enzyme assay so that maximum activity has not already occurred prior to assay
- Preparing crude lysate from 100x concentrated exponential phase *Treponema primitia* str. ZAS-1 cultures
- Disrupting *T. primitia* str. ZAS-1 cells via CellLytic (microscope check confirmed approximately 50% of cells disrupted)
- Disrupting *T. primitia* str. ZAS-1 cells via sonnication (microscope check confirmed approximately 80% of cells disrupted)
- Disrupting *T. primitia* str. ZAS-1 cells via microsonnication (on ice) (microscope check confirmed approximately 80% of cells disrupted)
- Preparing crude cell lysate under N_2
- Adding 3μL of 2.1g/100mL $FeS0_4$ to reaction mix to ensure enough Fe^{2+} cofactor was available

An appropriate positive control would be *Pseudomonas putida* and an appropriate negative control would be *T. azotonutricium* str. ZAS-9, but these were not examined here (Kita *et al*. 1999).

RESULTS:

Reciprocal BLAST revealed catechol 2,3-dioxygenase in *Treponema primitia*

To compare the metabolic potentials of phylogenetically "higher" and "lower" termite hindgut microbes using available datasets, reciprocal BLAST analyses were conducted between genomes from three "lower" termite hindgut isolates, *Treponema primitia* str. ZAS-1 (3.8Mb) and ZAS-2 (4.1Mb), and *T. azotonutricium* str. ZAS-9 (3.9Mb) (Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004; Rosenthal *et al*. 2011; Ballor *et al*. 2012), and the P3 hindgut region

metagenome from the "higher" termite, *Nasutitermes* sp. (62Mb) (Warnecke *et al*. 2007). Specifically, genes present in any of the three "lower" termite hindgut isolates' genomes, but not in the "higher" termite hindgut metagenome, and vice versa, were identified. 21%, 20%, and 17% of the *T. primitia* str. ZAS-1, *T. primitia* str. ZAS-2, and *T. azotonutricium* str. ZAS-9 genomes, respectively, had no orthologs in the metagenomic dataset (*see Discussion for why metagenome dataset does not represent full-coverage of "higher" termite sample*) (Appendix Table 3A-4). Among those non-orthologs, a putative "catechol 2,3-dioxygenase" gene was found in each of the *T. primitia* str. ZAS-1 and ZAS-2 genomes, but not in the *T. azotonutricium* str. ZAS-9 genome (Appendix Table 3A-4). No catechol 2,3-dioxygenase homologs were found in the "higher" termite metagenome either, thus corroborating earlier analyses (Appendix Table 3A-4) (Warnecke *et al*. 2007). Because catechol 2,3-dioxygenase is a key enzyme of the *meta*-cleavage pathway of many aerobic bacteria and both strains are considered anaerobes - and preliminary gene function assignments are often prone to error - a more in depth evaluation of the genome and metabolic potentials for O_2-dependent aromatic metabolism by *T. primitia* was initiated.

Treponema primitia **strains ZAS-1 and ZAS-2 have genes representing complete catechol 2,3-dioxygenase *meta*-cleavage pathways**

As background, during the metabolism of many aromatic compounds numerous preparative or "upper" pathways generate (di)hydroxylated intermediates from a wider range of structurally diverse aromatic compounds. Then, "lower" pathways transform the resulting aromatics into non-aromatic, central cell metabolites

(Viggiani *et al*. 2004; Siani *et al*. 2006). Catechol 2,3-dioxygenases are often employed to catalyze the key ring cleavage step in such "lower" pathways of monoaromatic metabolism (Fig. 3-1).

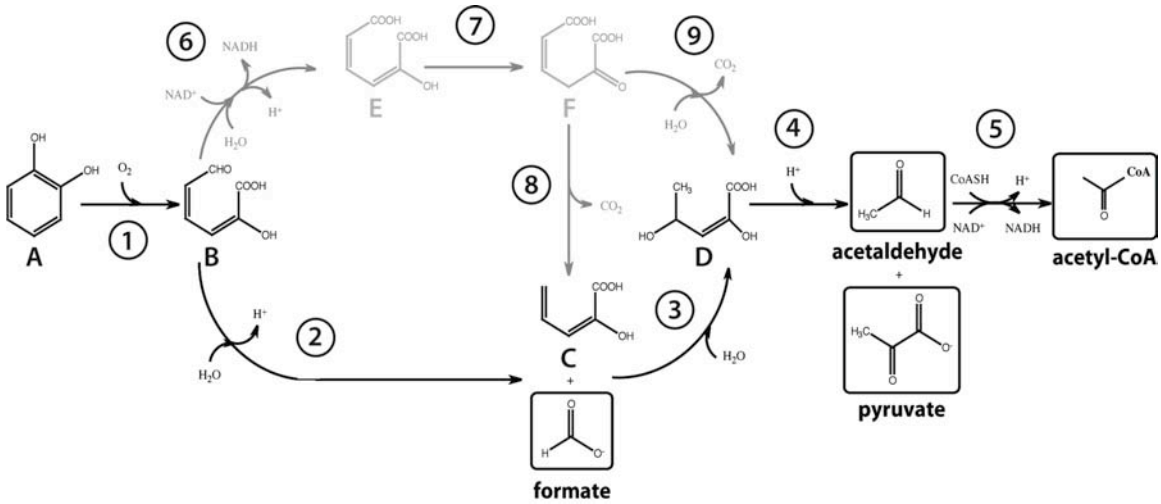


Fig. 3-1: Catechol 2,3-dioxygenase-based *meta*-cleavage pathway of aromatic metabolism and shunts. *Treponema primitia* str. ZAS-1 and ZAS-2 have genes for a complete *meta*-cleavage pathway represented by steps 1, 2, 3, 4, and 5 (black). Two alternative routes (grey) are, Shunt A: steps 1, 6, 7, 8, 3, 4, and 5 (skips step 2) Shunt B: steps 1, 6, 7, 9, 4, and 5 (skips steps 2 and 3). Substrates and products are A: catechol; B: 2-hydroxymuconic semialdehyde; C: 2-oxopent-4-enoate; D: 4-hydroxy-2-oxopentanoate, E: 2-hydroxymuconic acid, and F: γ-oxalocrotonate. Central cell metabolite products are outlined in boxes. Enzymes represented are: 1, catechol 2,3-dioxygenase; 2, 2-hydroxymuconic semialdehyde hydrolase; 3, 2-oxopent-4-enoate hydratase; 4, 4-hydroxy-2-oxopentanoate aldolase; 5, acetaldehyde dehydrogenase; 6, 2-hydroxymuconic semialdehyde dehydrogenase; 7, 4-oxalocrotonate tautomerase; 8, 4-oxalocrotonate decarboxylase; and 9, 4-oxalocrotonate carboxy-lyase.

Catechol 2,3-dioxygenases are classified as extradiol dioxygenases, thus the "lower" pathways they initiate are often referred to as "*meta*-cleavage" pathways (Bugg & Winfield 1998; Vaillancourt *et al*. 2006). Catechol 2,3-dioxygenases, therefore,

typically function in conjunction with a suite of up to eight other *meta*-pathway enzymes including: 2-hydroxymuconic semialdehyde hydrolase (PF00561, step 2); 2-oxopent-4-enoate hydratase (PF01557, step 3); 4-hydroxy-2-oxopentanoate aldolase (PF00682 and PF07836, step 4); acetaldehyde dehydrogenase (PF01118 and PF09290, step 5); 2-hydroxymuconic semialdehyde dehydrogenase (PF00171, step 6); 4-oxalocrotonate tautomerase (PF01361, step 7); 4-oxalocrotonate decarboxylase (PF01557, step 8); and 4-oxalocrotonate carboxy-lyase (PF01557, step 9) (Table 3-1, Fig. 3-1) (Harayama *et al*. 1987; Harayama & Rekik 1990; Furukawa *et al*. 1993; Suenaga *et al*. 2007). An associated ferredoxin-like peptide (PF00111, step 1') plays an indirect pathway role, reactivating O_2-inactivated catechol 2,3-dioxygenase by reducing its oxidized iron cofactor (Table 3-1, Fig. 3-1) (Polissi & Harayama 1993; Hugo *et al*. 2000). Catechol 2,3-dioxygenase initiated *meta*-cleavage pathways generally proceed via one route (Fig. 3-1, black, steps: 1, 2, 3, 4, and 5), but two alternative shunts are known (Fig. 3-1, grey, Shunt A: steps 1, 6, 7, 8, 3, 4, and 5, Shunt B: steps 1, 6, 7, 9, 4, and 5) (Khajamohiddin *et al*. 2008). In an effort to determine whether the annotation of genes in *Treponema primitia* str. ZAS-1 and ZAS-2 as being "catechol 2,3-dioxygenases" is valid, their gene neighborhoods (and the rest of their genomes) were analyzed for the presence of other key genes for *meta*-cleavage pathways.

Table 3-1: *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-cleavage pathway genes and pfam features

		Present in		Protein Size (a.a.) in			Pfam Residues in	
Step	Gene Annotation	str. ZAS-1	str. ZAS-2	str. ZAS-1	str. ZAS-2	Pfam(s) Represented (PF#, family, domain)	str. ZAS-1	str. ZAS-2
1'	ferredoxin-like peptide	+	+	98	94	PF00111, Fer2, [2Fe-2S] cluster binding	13-85	14-85
1	catechol 2,3-dioxygenase	+	+	308	308	PF00903, Glyoxalase, glyoxalase/bleomycin resistance protein/dioxygenase	12-76	12-68
						PF00903, Glyoxalase, glyoxalase/bleomycin resistance protein/dioxygenase	151-266	151-266
2	2-hydroxymuconic semialdehyde hydrolase	+	+	274	274	PF00561, Abhydrolase 1, α/β hydrolase fold	32-266	32-266
3	2-oxopent-4-enoate hydratase	+	+	260	260	PF01557, FAA hydrolase, fumarylacetoacetate hydrolase	61-259	60-259
4	4-hydroxy-2-oxopentanoate aldolase	+	+	340	335	PF00682, HMGL-like, HMGL-like	19-250	14-245
						PF07836, DmpG comm., DmpG-like communication	274-339	269-334
5	acetaldehyde dehydrogenase	+	+	291	288	PF01118, Semialdehyde dh, semialdehyde dehydrogenase NAD binding	7-116	5-114
						PF09290, Acetdehyd dimer, prokaryotic acetaldehyde dehydrogenase dimerisation	126-262	124-260
6	2-hydroxymuconic semialdehyde dehydrogenase	-	-			PF00171 Aldedh, aldehyde dehydrogenase family		
7	4-oxalocrotonate tautomerase	-	-			PF01361 Tautomerase, tautomerase enzyme		
8	4-oxalocrotonate decarboxylase	-	-			PF01557 FAA hydrolase, fumarylacetoacetate hydrolase		
9	4-oxalocrotonate carboxy-lyase	-	-			PF01557 FAA hydrolase, fumarylacetoacetate hydrolase		

In both *T. primitia* str. ZAS-1 and ZAS-2, directly upstream of the putative "catechol 2,3-dioxygenase" are four other putative *meta*-cleavage pathway genes (Fig. 3-2). Collectively, in order they are: 2-hydroxymuconic semialdehyde hydrolase; 2-oxopent-4-enote hydratase; acetaldehyde dehydrogenase; 4-hydroxy-2-oxopentanoate aldolase; ferredoxin-like peptide; and catechol 2,3-dioxygenase (Fig. 3-2). Thus, *T. primitia* str. ZAS-1 and ZAS-2 have genes representing all key enzymatic steps known for the better studied *meta*-pathway route, but not for either of the alternative shunts (Fig. 3-1 in black).

Pfams and conserved residues in *Treponema primitia meta*-pathway genes suggest functionality

The 2-hydroxymuconic semialdehyde hydrolase; 2-oxopent-4-enote hydratase; acetaldehyde dehydrogenase; 4-hydroxy-2-oxopentanoate aldolase; ferredoxin-like peptide; and catechol 2,3-dioxygenase in *Treponema primitia* str. ZAS-1 and ZAS-2,

are consistent in size with their equivalent, reported, functional *meta*-cleavage pathway genes, and each genes' corresponding protein has the expected Pfam (protein family) domains and conserved amino acid residues suggestive of *meta*-pathway functionality (Table 3-1; Appendix Fig. 3A-1) (Díaz & Timmis 1995; Kita *et al*. 1999; Nardini & Dijkstra 1999; Holmquist 2000; Nandhagopal *et al*. 2001; Rea *et al*. 2005; Izumi *et al*. 2007; Lei *et al*. 2008).

The *T. primitia* str. ZAS-1 ferredoxin-like peptide has one Family Fer2, [2Fe-2S] cluster binding domain (PF00111) spanning 73 residues (13 to 85) of the 98 amino acid-long protein consistent with other studied *meta*-pathway ferredoxins (Table 3-1; Appendix Fig. 3A-1) (Aziz *et al*. 2008; Punta *et al*. 2012). Likewise the *T. primitia* str. ZAS-2 ferredoxin-like peptide has one Family Fer2, [2Fe-2S] cluster binding domain (PF00111) spanning 72 residues (14 to 85) of the 94 amino acid-long protein (Table 3-1; Appendix Fig. 3A-1) (Aziz *et al*. 2008; Punta *et al*. 2012). The *T. primitia* str. ZAS-1 ferredoxin-like peptide has an active site comprised of Val-36, Ser-37, Gly-43, Gly-46, Ala-47, and Leu-80, and an iron-binding site that also contributes to activity comprised of Cys-40, Cys-45, Cys-48, and Cys-81 (Appendix Fig. 3A-1) (Aziz *et al*. 2008). The *T. primitia* str. ZAS-2 ferredoxin-like peptide has the same features except for a Thr-37 instead of a Ser-37 in the active site (Appendix Fig. 3A-1) (Aziz *et al*. 2008). The four cysteine residues are highly conserved among catechol 2,3-dioxygenase-associated ferredoxins, and have a conserved role also as ligands for the [2Fe-2S] cluster (Hugo *et al*. 2000).

Next, both the *T. primitia* str. ZAS-1 and ZAS-2 catechol 2,3-dioxygenases are 308 amino acids in length, consistent with the residue span of previously reported catechol 2,3-dioxygenases (Table 3-1; Appendix Fig. 3A-1) (Kita *et al*. 1999; Viggiani *et al* 2004). Also like known catechol 2,3-dioxygenases, the *T. primitia* str. ZAS-1 and ZAS-2 genes each contain an N- and C-terminal domain that both represent the glyoxalase/bleomycin resistance protein/dioxygenase superfamily of proteins (PF00903) (Table 3-1; Appendix Fig. 3A-1). The sizes of each of those domains in both the *T. primitia* str. ZAS-1 and ZAS-2 gene also correlate with published lengths (Table 3-1; Appendix Fig. 3A-1).

Previous primary structure analysis of nearly 40 diverse catechol 2,3-dioxygenase sequences suggests that conserved residues include three metal ligands, His-146, His-210, and Glu-260, and three additional active site residues, His-195, His-241, and Tyr-250. The other strictly-conserved residues, Gly-28, Leu-165, and Pro-254, are remote from the active site, and thus are likely to play structural or folding roles (Suenaga *et al*. 2009). These conserved residues are all found in the *T. primitia* str. ZAS-1 and ZAS-2 catechol 2,3-dioxygenase, and like other well-studied catechol 2,3-dioxygenases the C-terminal domain of the *T. primitia* str. ZAS-1 and ZAS-2 gene appears to contain the active site (Appendix Fig. 3A-1).

Regarding the 2-hydroxymuconic semialdehyde hydrolase, both the *T. primitia* 2-hydroxymuconic semialdehyde hydrolases have one Family Abhydrolase 1, α/β hydrolase fold domain (PF00561) spanning 235 residues (32 to 266) of the 274

amino-acid long proteins like other well-studied 2-hydroxymuconic semialdehyde hydrolases (Table 3-1; Appendix Fig. 3A-1) (Aziz *et al*. 2008; Punta *et al*. 2012). In addition, both the *T. primitia* 2-hydroxymuconic semialdehyde hydrolases each have an active site comprised of Ser-105, Asp-226, and His-254 (Appendix Fig. 3A-1) (Aziz *et al*. 2008). These three residues reflect the catalytic triad configuration, typically nucleophile-acid-histidine or Ser-107, Asp-228, and His-256, conserved among other functional α/β hydrolase fold 2-hydroxymuconic semialdehyde hydrolases (Díaz & Timmis 1995; Nandhagopal *et al*. 2001). Further, a "nucleophilic elbow" consensus sequence, Sm-X-Nu-X-Sm-Sm, typically surrounds the Ser-107 nucleophile in α/β hydrolase fold enzymes with Sm representing small amino acids (generally glycine) and X representing any amino acid (Díaz & Timmis 1995). This "nucleophilic elbow" in *Pseudomonas putida*, is Gly-105, Asn-106, Ser-107, Phe-108, Gly-109, and Gly-110 (Díaz & Timmis 1995). Similarly, both the *T. primitia* 2-hydroxymuconic semialdehyde hydrolases have a "nucleophilic elbow" of Gly-103, Asn-104, Ser-105, Phe-106, Gly-107, and Gly-108 (Appendix Fig. 3A-1).

Moreover, a conserved motif among *meta*-cleavage product hydrolases is His-35, Gly-36, X-37, Gly-38, Pro-39, and Gly-40 with X representing a small amino acid and both His-35 and Gly-36 involved in the formation of an oxyanion hole critical to catalytic function (Nandhagopal *et al*. 2001). Indeed both the *T. primitia* 2-hydroxymuconic semialdehyde hydrolases have the motif His-34, Gly-35, Ser-36, Gly-37, Pro-38, and Gly-39 with the catalytic His-34 and Gly-35 residues (Appendix Fig. 3A-1). Other conserved features among *meta*-cleavage product hydrolases are

located in the bottom of the substrate-binding pocket. These are Asn-46, Asn-109, and Gln-266 (Nandhagopal *et al*. 2001). Correspondingly, the *T. primitia* str. ZAS-1 2-hydroxymuconic semialdehyde hydrolase has residues Asn-47, Asn-104, and Gln-262, and the *T. primitia* str. ZAS-2 2-hydroxymuconic semialdehyde hydrolase has residues Asn-45, Asn-104, and Gln-257 (Appendix Fig. 3A-1).

Next, the *T. primitia* str. ZAS-1 2-oxopent-4-enoate hydratase has one Family FAA hydrolase, fumarylacetoacetate hydrolase domain (PF01557) spanning 199 residues (61 to 259) of the 260 amino acid-long protein as reported for other 2-oxopent-4-enoate hydratases (Table 3-1; Appendix Fig. 3A-1) (Aziz *et al*. 2008; Punta *et al*. 2012). The *T. primitia* str. ZAS-2 2-oxopent-4-enoate hydratase has one Family FAA hydrolase, fumarylacetoacetate hydrolase domain (PF01557) spanning 200 residues (60 to 259) of the 260 amino acid-long protein (Table 3-1; Appendix Fig. 3A-1) (Aziz *et al*. 2008; Punta *et al*. 2012). Three conserved residues of these hydratases, Glu-106, Glu-108, and Glu-139, coordinate the metal ion (Izumi *et al*. 2007). Both the *T. primitia* 2-oxopent-4-enoate hydratases have these three conserved residues, represented by Glu-104, Glu-106, and Glu-137 (Appendix Fig. 3A-1) (Aziz *et al*. 2008). Other conserved residues of *meta*-cleavage pathway 2-oxopent-4-enoate hydratases are Lys-61, Leu-64, Asp-79, and Asn-168 that comprise the active site (Izumi *et al*. 2007). In *T. primitia* these conserved active site residues are Lys-60, Leu-63, Asp-78, and Asn-168 in str. ZAS-1 and Asn-157 in str. ZAS-2 (Appendix Fig. 3A-1) (Aziz *et al*. 2008).

The *T. primitia* 4-hydroxy-2-oxopentanoate aldolases each have one HMGL-like family domain (PF00682) spanning 232 residues (19 to 250 in str. ZAS-1 and 14-245 in str. ZAS-2) and one Family DmpG comm., DmpG-like communication domain (PF07836) spanning 66 residues (274-339 in str. ZAS-1 and 269-334 in str. ZAS-2) of the 340 and 335 amino acids-long proteins in *T. primitia* str. ZAS-1 and ZAS-2, respectively like other researched 4-hydroxy-2-oxopentanoate aldolases (Table 3-1; Appendix Fig. 3A-1) (Aziz *et al*. 2008; Punta *et al*. 2012).

Conserved metal ion ligands of 4-hydroxy-2-oxopentanoate aldolases are Asp-18, His-200, and His-202. Accordingly, *T. primitia* str. ZAS-1's 4-hydroxy-2-oxopentanoate aldolase has conserved metal ion ligands at Asp-20, His-200, and His-202 whereas *T. primitia* str. ZAS-2's 4-hydroxy-2-oxopentanoate aldolase has conserved metal ion ligands at Asp-15, His-195, and His-197 (Appendix Fig. 3A-1) (Aziz *et al*. 2008). These conserved metal ion ligands interact with the conserved active site residues Arg-17, His-21, Gly-52, and Tyr-291 which are represented by Arg-19, His-23, Gly-54, and Tyr-291 in *T. primitia* str. ZAS-1, and Arg-14, His-18, Gly-54, and Tyr-286 in str. ZAS-2 (Appendix Fig. 3A-1) (Aziz *et al*. 2008).

The *T. primitia* acetaldehyde dehydrogenases each have one Family Semialdehyde dh/semialdehyde dehydrogenase, NAD binding domain (PF01118) spanning 110 residues (7-116 in str. ZAS-1 and 5-114 in str. ZAS-2) and one Family Acetdehyd dimmer, prokaryotic acetaldehyde dehydrogenase dimerisation domain (PF09290) spanning 137 residues (126-262 in str. ZAS-1 and 124-260 in str. ZAS-2) of the 291 and 288 amino acid-long proteins in *T. primitia* str. ZAS-1 and ZAS-2, respectively

consistent with previously reported acetaldehyde dehydrogenases (Table 3-1; Appendix Fig. 3A-1) (Aziz *et al*. 2008; Punta *et al*. 2012). The active site of acetaldehyde dehydrogenases is located between its NAD+-binding and dimerization domains and has a conserved residue, Cys-132. Likewise in the *T. primitia* acetaldehyde dehydrogenase this residue is Cys-126 (str. ZAS-1) and Cys-124 (str. ZAS-2) (Appendix Fig. 3A-1).

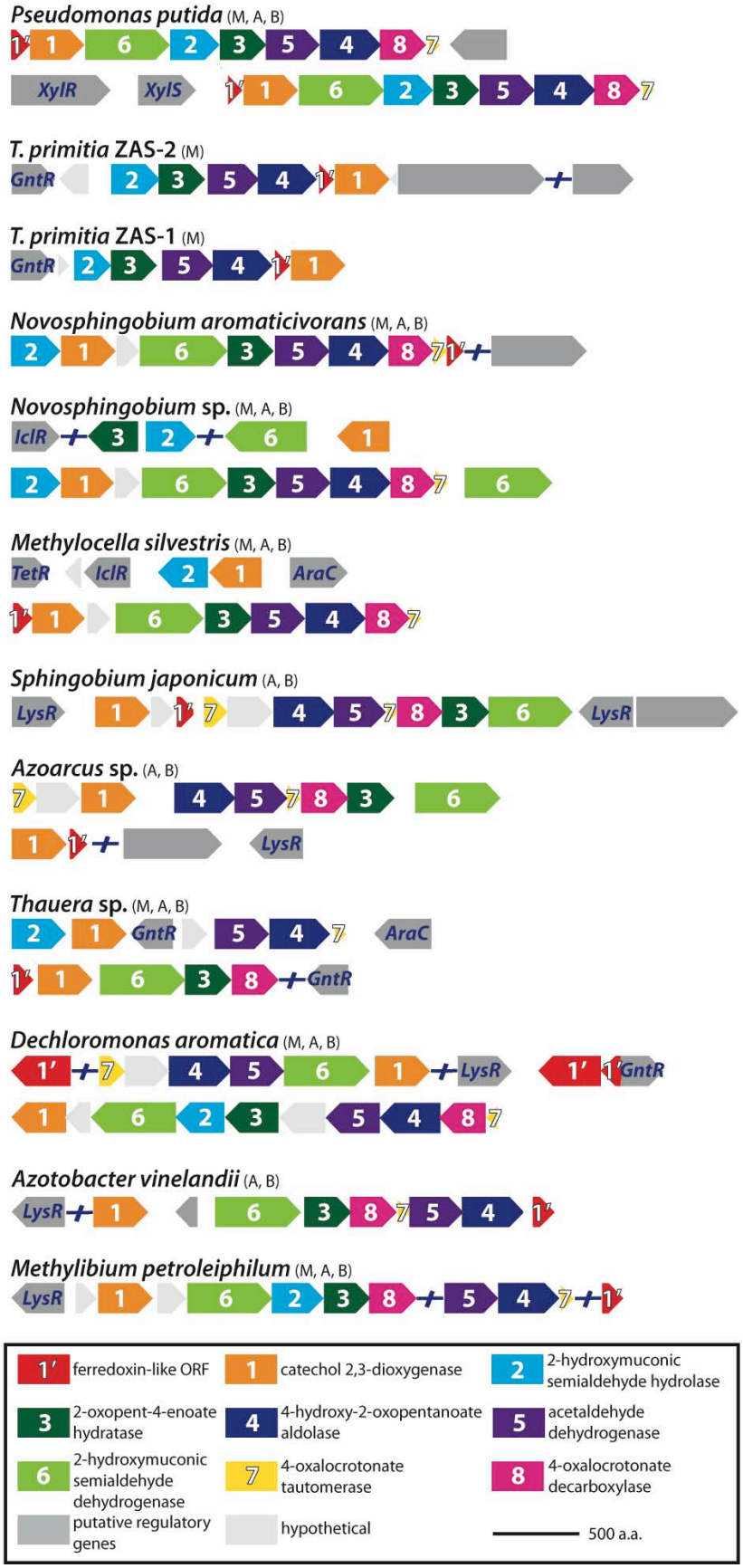

Pseudomonas putida (M, A, B)
XylR
XylS
T. primitia ZAS-2 (M)
GntR
T. primitia ZAS-1 (M)
GntR
Novosphingobium aromaticivorans (M, A, B)
Novosphingobium sp. (M, A, B)
IclR
Methylocella silvestris (M, A, B)
TetR
IclR
AraC
Sphingobium japonicum (A, B)
LysR
LysR
Azoarcus sp. (A, B)
LysR
Thauera sp. (M, A, B)
GntR
AraC
GntR
Dechloromonas aromatica (M, A, B)
LysR
GntR
Azotobacter vinelandii (A, B)
LysR
Methylibium petroleiphilum (M, A, B)
LysR
1′ ferredoxin-like ORF
1 catechol 2,3-dioxygenase
2 2-hydroxymuconic semialdehyde hydrolase
3 2-oxopent-4-enoate hydratase
4 4-hydroxy-2-oxopentanoate aldolase
5 acetaldehyde dehydrogenase
6 2-hydroxymuconic semialdehyde dehydrogenase
7 4-oxalocrotonate tautomerase
8 4-oxalocrotonate decarboxylase
putative regulatory genes
hypothetical
500 a.a.

Fig. 3-2: Gene neighborhoods representing complete catechol 2,3-dioxygenase-based *meta*-cleavage pathways. In parentheses near the names of organisms with genes representing complete *meta*-cleavage pathways, "M," "A," and/or "B," represents having genes for the main pathway, Shunt A, or Shunt B, respectively. Ferredoxin-like peptide (step 1', red), catechol 2,3-dioxygenase (step 1, orange), 2-hydroxymuconic semialdehyde hydrolase (step 2, light blue), 2-oxopent-4-enoate hydratase (step 3, dark green), 4-hydroxy-2-oxopentanoate aldolase (step 4, dark blue), acetaldehyde dehydrogenase (step 5, purple), 2-hydroxymuconic semialdehyde dehydrogenase (step 6, light green), 4-oxalocrotonate tautomerase (step 7, yellow), 4-oxalocrotonate decarboxylase (step 8, pink), putative sregulatory genes (dark grey), and hypothetical proteins (light grey) are depicted. Regulatory families are noted in blue print when known.

The *T. primitia* str. ZAS-1 and ZAS-2 *meta*-cleavage pathway gene neighborhoods are arranged similarly between their respective genomes (Fig. 3-2). In contrast, their arrangement is dissimilar from the order of the *meta*-cleavage pathway enzymatic steps they represent (Fig. 3-1 and 3-2). Relative to catechol 2,3-dioxygenase *meta*-cleavage pathway gene neighborhoods in other organisms possessing this pathway, the *T. primitia* strains' gene arrangement is distinct (Fig. 3-2).

No evidence for "upper" preparatory pathway genes in *Treponema primitia*

Many *meta*-cleavage pathway-containing organisms perform preparatory reactions before proceeding to cleave aromatic substrates, and often have genes for monooxygenases and dioxygenases. These are often found directly upstream, downstream, or among the *meta*-cleavage pathway-encoding gene cluster. Nevertheless, no evidence for preparatory reactions was observed in the genomes of the *Treponema primitia* isolates.

Evolutionary origins of the *Treponema primitia meta*-cleavage pathway

In order to learn more about the evolutionary origins and history of the pathway genes in *Treponema primitia*, a phylogenetic analysis was performed on each by comparing each gene in the pathway to databases of their respective homologs. The results of that analysis suggest that the genes for the pathway likely originate from at least two, if not three sources: an alphaproteobacterium, a member of the clostridiales, and possibly a member of the bacillales (Fig. 3-3, 3-4, and 3-5; Appendix Fig 3A-2, 3A-3, and 3A-4). However, gene transfer events seem to have occurred in the distant past, inasmuch as relatively long branch lengths are observed for each *T. primitia* homolog, with amelioration of the GC and codon usage.

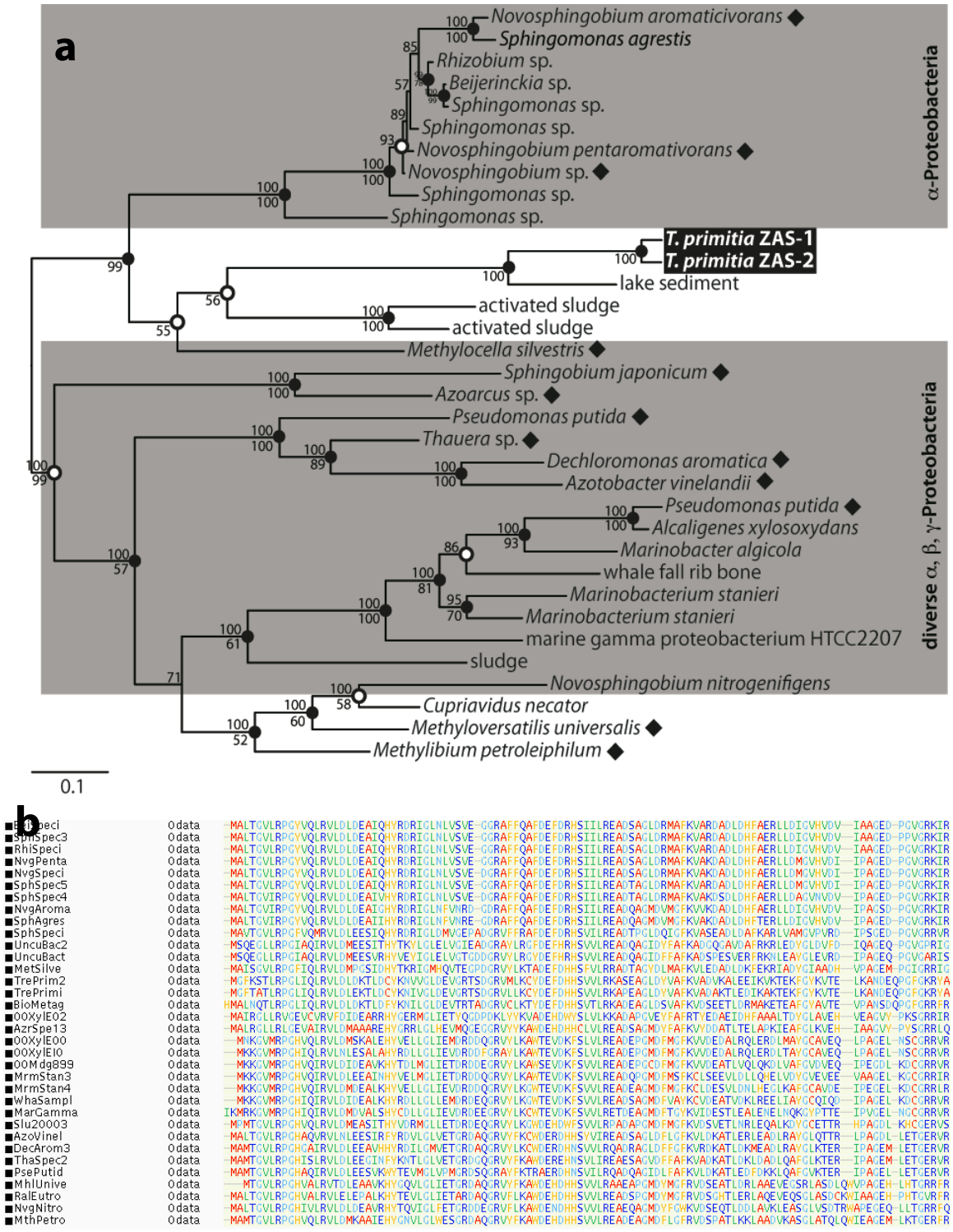


Fig. 3-3: (a) Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 catechol 2,3-dioxygenase (PF00903, step 1). Bayesian protein phylogenetic analysis (90 trees from 36,000 generations; PSRF = 0.999; average standard deviation of split frequencies = 0.008789) is based on 297 unambiguously aligned amino acid

positions of a 308 amino acid-long protein. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Diamonds (◆) indicate organisms with genes for a complete catechol 2,3-dioxygenase-based *meta*-cleavage pathway. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position. Accession numbers are reported in Appendix (Appendix Table 3A-1). **(b)** Example of MUSCLE alignments used to create trees. For each protein or Pfam CLUSTALW, MAFFT, DIALIGN, T-Coffee, and MUSCLE alignments were generated on the Mobyle Server and compared (Néron *et al*. 2009). Consistently, MUSCLE generated the best alignments and therefore sequences were unambiguously aligned using this program. Conservative filters were used to disregard columns of data containing sequence gaps or ambiguous alignments.

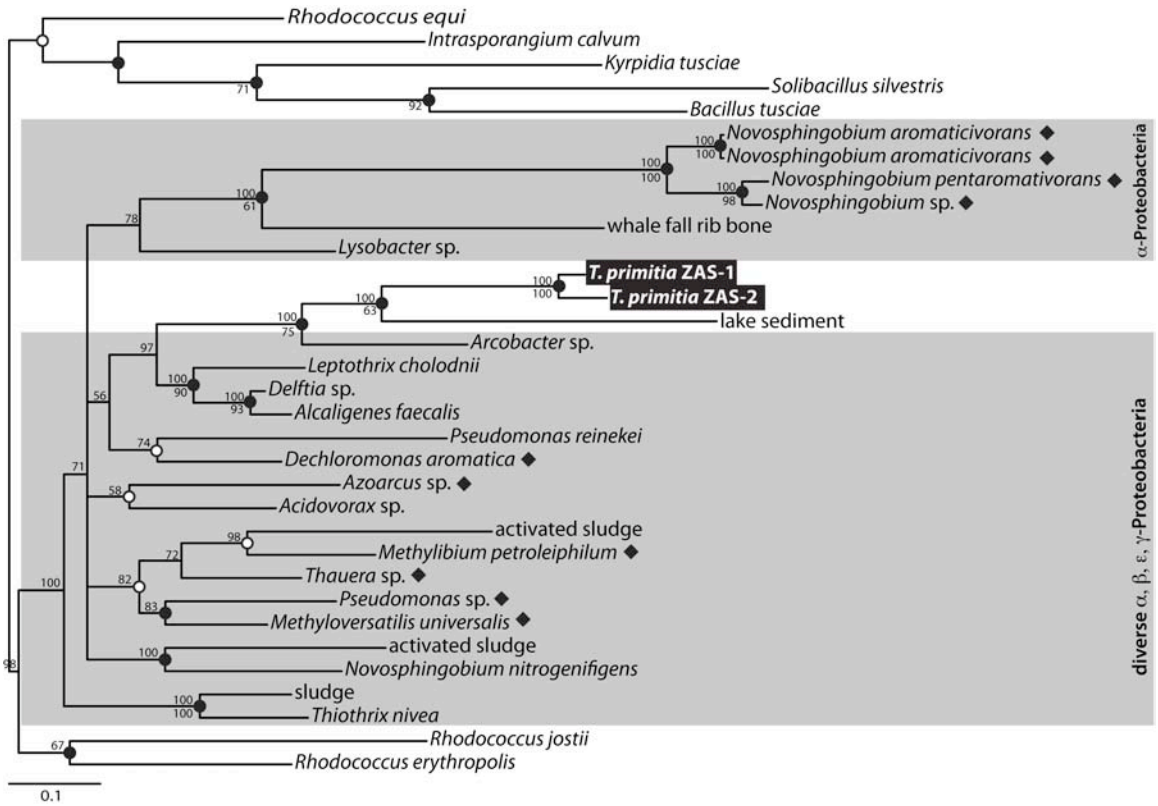


Fig. 3-4: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 2-hydroxymuconic semialdehyde hydrolase (PF00561, step 2). Bayesian protein phylogenetic analysis (2540 trees from 1,016,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.011955) is based on 222 unambiguously aligned amino acid positions of a 274 amino acid-long protein. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Diamonds (◆) indicate organisms with genes for a complete catechol 2,3-dioxygenase-based *meta*-cleavage pathway. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position. Accession numbers are reported in Appendix (Appendix Table 3A-1).

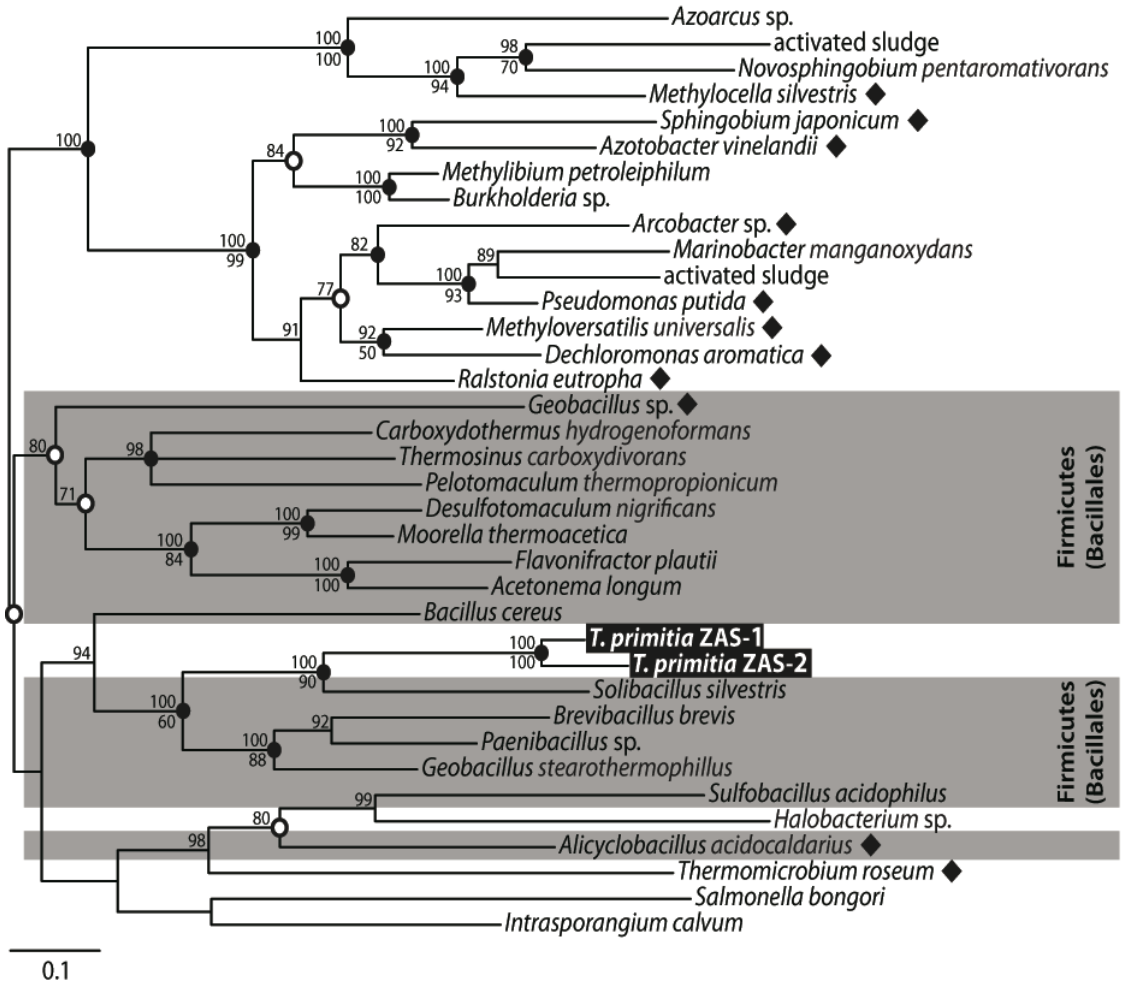


Fig. 3-5: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 2-oxopent-4-enoate hydratase (PF01557, step 3). Bayesian protein phylogenetic analysis (313 trees from 125,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.009239) is based on 243 unambiguously aligned amino acid positions of a 260 amino acid-long protein. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Diamonds (◆) indicate organisms with genes for a complete catechol 2,3-dioxygenase-based *meta*-cleavage pathway. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position. Accession numbers are reported in Appendix (Appendix Table 3A-1).

The catechol 2,3-dioxygenases from the two *T. primitia* strains are each other's closest relative, and demonstrate a strong affiliation with homologs of catechol 2,3-dioxygenases from α-, β-, and γ-proteobacteria, as well as several environmental sequences (Fig. 3-3). Specifically, the majority of the closest-cultured relatives are alphaproteobacteria, specifically "sphingomonads" (Fig. 3-3). Similarly, the short ferredoxin-like peptides from the *T. primitia* strains affiliate phylogenetically with those associated with *meta*-cleavage pathways in diverse proteobacteria (Appendix Fig. 3A-2). The protein that catalyzes the second step in the pathway, 2-hydroxymuconic semialdehyde hydrolase, also supports a proteobacterial origin (Fig. 3-4).

In contrast, the proteins for the remainder of the steps in the pathway, 2-oxopent-4-enoate hydratase, 4-hydroxy-2-oxopentanoate aldolase, and acetaldehyde dehydrogenase, demonstrate strong affiliations with those from firmicutes (Fig. 3-5; Appendix Fig 3A-3 and 3A-4, respectively). The hydratase is related to other hydratases from aerobic species of the bacillales, whereas the aldolase and acetaldehyde dehydrogenase affiliate phylogentically with homologs that are often found widespread across many anaerobic clostridiales, including many homoacetogens (Fig. 3-5; Appendix Fig. 3A-3 and 3A-4, respectively), none of which encode the enzymes catalyzing the key early steps in the pathway. Extensive phylogenetic analyses on each of the specific Pfam (protein family) domains or modules present within catechol 2,3-dioxygenase, 4-hydroxy-2-oxopentanoate aldolase, and acetaldehyde dehydrogenase yielded results consistent with all of the

above conclusions (Appendix Fig. 3A-5, 3A-6, 3A-7, 3A-8, 3A-9, 3A-10, 3A-11, 3A-12, 3A-13, and 3A-14).

Selection pressure and expression analyses

None of the pathway related ORFs in the *Treponema primitia* strains appear to be pseudogenes, and all appear to encode functional enzymes. In order to examine how important these genes may be over more recent evolutionary time, selection pressure analyses were performed to test the hypothesis that these genes might demonstrate signs of positive selection. All of the pathway genes with the exception of 2-hydroxymuconic semialdehyde hydrolase appear to be under positive selection, as their determined Z values were significantly less than zero (i.e. the numbers of synonymous changes (dS) in each were significantly greater than the number of nonsynonymous changes (dN)) (Table 3-2). Curiously, the 2-hydroxymuconic semialdehyde hydrolase genes from the two spirochetes each appear to be under negative selection pressure (Table 3-2).

Table 3-2: Selection pressure analysis of *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-cleavage pathway genes

Gene	Method	Z	P-value	Selection Pressure
ferredoxin-like peptide	NG86	-2.932	1	+
	Modified NG86	-0.34	1	+
	LWL85	-4.003	1	+
	PBL85	-3.882	1	+
	Kumar	-3.273	1	+
catechol 2,3-dioxygenase	NG86	-14.09	1	+
	Modified NG86	-11.326	1	+
	LWL85	-18.94	1	+
	PBL85	-20.956	1	+
	Kumar	-17.675	1	+
2-hydroxymuconic semialdehyde hydrolase	NG86	4.594	0	-
	Modified NG86	7.219	0	-
	LWL85	4.304	0	-
	PBL85	5.136	0	-
	Kumar	6.827	0	-
2-oxopent-4-enoate hydratase	NG86	-8.802	1	+
	Modified NG86	-5.276	1	+
	LWL85	-10.717	1	+
	PBL85	-8.89	1	+
	Kumar	-7.787	1	+
4-hydroxy-2-oxopentanoate aldolase	NG86	-8.732	1	+
	Modified NG86	-5.237	1	+
	LWL85	-10.717	1	+
	PBL85	-9.003	1	+
	Kumar	-8.073	1	+
acetaldehyde dehydrogenase	NG86	-1.795	1	+
	Modified NG86	NA	NA	+
	LWL85	-2.282	1	+
	PBL85	-1.466	1	+
	Kumar	-0.282	1	+

In a re-analysis of data from a published transcriptomics study (Rosenthal *et al*. 2011), the expression of genes in the pathway were evaluated for *T. primitia* str. ZAS-2 grown under anaerobic, homacetogenic growth conditions as a pure-culture or as co-culture with *T. azotonutricium* str. ZAS-9 (Appendix Table 3A-3). Although both pure- and co-cultures of *T. primitia* str. ZAS-2 had been grown in an anoxic, DTT-reduced medium without added aromatic substrates or O_2, all of the *meta*-

cleavage pathway genes were observed as being expressed (Appendix Table 3A-2 and Table 3A-3). Catechol 2,3-dioxygenase expression was at a higher level than the other *meta*-cleavage pathway genes, and at a level comparable to that of a key gene in this species' anaerobic metabolism - formate dehydrogenase (Appendix Table 3A-2 and Table 3A-3). Comparisons of expression levels of the genes in the pathway with other reference genes are presented in the supplementary materials (Appendix Table 3A-2 and Table 3A-3).

Regulation

Expression of the *meta*-cleavage pathway genes appears to be regulated by a GntR family transcripton regulator upstream of the 2-hydroxymuconic semialdehyde hydrolase – and thus all genes representing the *meta*-cleavage pathway – in both *Treponema primitia* str. ZAS-1 and ZAS-2 (Appendix Fig. 3A-15). The GntR family of transcriptional regulators includes more than 1,000 members and is distributed among diverse bacterial groups and biological processes (Rigali *et al*. 2002; Hillerich & Westpheling 2006). This family was names after the repressor of the gluconate operon in *Bacillus subtilis* (Haydon & Guest 1991). GntR family members are known to also regulate the degradation of aromatic compounds, functioning as transcriptional repressors in the absence of pathway substrates such as aromatics (Tropel & van der Meer 2004). Other regulatory families involved in aromatic degradation pathways include LysR, IclR, AraC, Xyl, TetR, and XylRS, and are found near catechol 2,3-dioxygenase-based *meta*-cleavage pathway neighborhoods in other organisms examined here (Tropel & van der Meer 2004). Preliminary phylogenetic analyses suggest the *T. primitia* GntRs' closest cultured relative is a

member of the Phylum *Spirochaetes* (Appendix Fig. 3A-16 and Fig. 3A-17). Other close cultured relatives are representatives of the Phylum *Firmicutes* (Appendix Fig. 3A-16 and Fig. 3A-17).

In addition to a GntR family transcriptional regulator, two promoter regions might be located within the *T. primitia* str. ZAS-1 and ZAS-2 gene neighborhoods upstream of the GntR and before the acetaldehyde dehydrogenase in each (Appendix Fig. 3A-15). Also, although selenocysteine insertion sequence (SECIS) elements - that recode in-frame UGA codons which normally function as stop signals to serve as selenocysteine codons - are found in *T. primitia* formate dehydrogenases (FDH), none were found in the *T. primitia meta*-cleavage pathway genes or gene neighborhoods (Zhang *et al*. 2011).

Metabolism of catechol under microoxic conditions

Neither of the genomes of the *Treponema primitia* strains have any obvious genes for oxygen respiration, for cytochrome-associated proteins or metabolism, or for electron transport phosphorylation. Therefore, any function of the *meta*-cleavage pathway genes in *T. primitia* must be very different from all of the canonical aerobes that typically encode these pathways and which mineralize the ring carbon to CO_2 during oxidative phosphorylation. However, an earlier study by Graber and Breznak documented that *T. primitia* cultures have a number of enzyme activities relevant to O_2 detoxification, and that this species can consume low amounts of O_2 (2004). In cultures spiked with air, growth ceased until the oxygen had been fully consumed, whereupon growth re-initiated (Graber & Breznak 2004). Those results

raise the possibility that aromatic metabolism might be used by the spirochete as an accessory mechanism to consume low amounts of oxygen, and possibly to generate intermediates that can feed into its (or the surrounding microbial community's) oxygen independent energy metabolism. Here, culture conditions were modified allowing the observation of good growth and successive serial transfer of *T. primitia* str. ZAS-1 cultures in an unreduced medium under a 0.5% O_2 vol/vol headspace. Such media was competent for culturing this strain if the cultures were incubated statically. Other than occasionally exhibiting longer initial lag phases, growth progressed at the same rates and yields as those observed in reduced, anoxic media (Fig. 3-6).

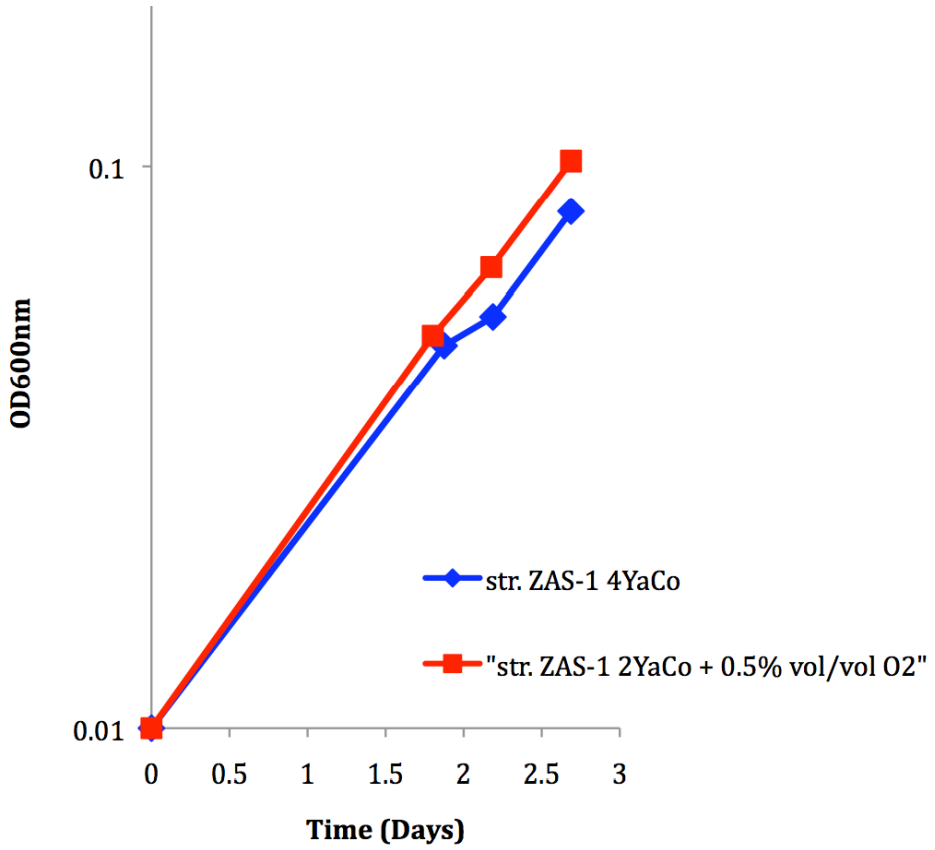


Fig. 3-6: Microoxic growth of *Treponema primitia* str. ZAS-1. *T. primitia* str. ZAS-1 grew in 5mL of 2YACo, DTT-free media, with an 80% N_2/20% CO_2 headspace to

which air was added sterilely to achieve 0.5% vol/vol O_2 (red) and 5mL of 4YACo, DTT-reduced media with an 80% H_2/20% CO_2 headspace (blue). Both treatment types were supplemented with a vitamin mixture (12-vitamin, vitamin B12, and folinic acid solutions, Graber & Breznak, 2004). Cultures were contained in 25mL butyl rubber-stoppered Balch tunes and incubated at 25°C horizontally and still. First three days of growth are shown.

Stoichiometric predictions of catechol conversion to acetate via this *meta*-cleavage pathway suggest an increase in acetate yield than what is obtained from H_2 + CO_2 acetogenesis:

H_2 + CO_2 Acetogenesis:

$$4H_2 + 2CO_2 \rightarrow CH_3OOH + 2H_2O$$

Catechol → Formic Acid + Pyruvic Acid + Acetic Acid:

$$C_6H_6O_2 + O_2 + 3H_2O \rightarrow H_2CO_2 + C_3H_3O_3H + CH_3COOH + H_2 \text{ (+ ATP)}$$

Formic Acid + Pyruvic Acid → Acetic Acid:

$$H_2CO_2 + C_3H_3O_3H + 2[2e^- + 2H^+] \rightarrow 2CH_3COOH + H_2O \text{ (+ 2ATP)}$$

Catechol → Acetic Acid:

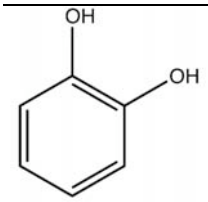
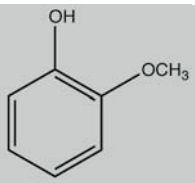
$$C_6H_6O_2 + O_2 + 2H_2O + 2[2e^- + 2H^+] \rightarrow 3CH_3COOH + H_2 \text{ (+ 3ATP)}$$

Therefore, potential increases in growth yields of cultures to which catechol-like aromatics and O_2 were added were evaluated as a proxy for acetate production via this *meta*-cleavage pathway. When *T. primitia* str. ZAS-1 cultures were transferred into this microoxic medium amended with catechol to a final concentration of 0.5mM, no significant stimulations of growth or organic acid yields were observed, and the majority of the substrate went unreacted over time. However, the culture fluids transiently turned a distinct lemon-yellow color after several days of growth, having reached an OD_{600nm} between 0.2 and 0.3 (Table 3-3; Appendix Fig. 3A-18 and Fig. 3A-19). Analysis of culture fluids to which 0.5mM aromatic substrate and 0.5%

vol/vol O_2 were added revealed that the catechol-dependent yellow material had an absorbance maximum at 375nm (Table 3-3; Appendix Fig. 3A-19), consistent with the presence of the expected dioxygenase-mediated ring cleavage product, 2-hydroxymuconic semialdehyde. This peak was not seen in cultures to which only 0.5mM aromatic substrate or only 0.5% vol/vol O_2 was added (Appendix Fig. 3A-19). Cultures transferred into media containing 0.5mM guaiacol also generated this same yellow intermediate (Table 3-3). Thus, the dioxygenase in this spirochete appears to be functional on catechol and guaiacol. No generation of obvious intermediates was observed when any of the following compounds were tested: veratrole, protocatechualdehyde, protocatechuic acid, vanillic acid, gallic acid, homoprotocatechuic acid, homoveratric acid, hydrocaffeic acid, caffeic acid, or ferulic acid.

As a back-of-the-envelope calculation, that 0.5mM catechol added to the cultures examined by UV/Vis means that there were 275μmoles of catechol in each 5mL culture. Therefore, if catechol generates the ring cleavage product, 2-hydroxymuconic semialdehyde, in a ratio of 1:1 according to the soichiometry above, then one woud expect all of this catechol to be converted to 275μmoles of ring cleavage product. Nevertheless, approximately 22μmoles of the ring cleavage product were formed (ε = 36,000 L mol^{-1} cm^{-1} and culture sample diluted 1:100 for UV/Vis spec readings) indicating roughly 10% of the original catechol input to the cutures was transformed into the ring cleavage product (Appendix Fig. 3A-19).

Table 3-3: Monoaromatic compound cleavage by *Treponema primitia* str. ZAS-1

Monoaromatic Compound	Abs. Max (nm)	Ring Cleavage Product Abs. Max (nm)	Ring Cleavage
 catechol (1,2-dihydroxybenzene)	275	375	+
 guaiacol (2-methoxyphenol)	275	375	+

Growth on veratrole, protocatechualdehyde, protocatechuic acid, vanillic acid, gallic acid, homoprotocatechuic acid, homoveratric acid, hydrocaffeic acid, caffeic acid, and ferulic acid was tested, but no evidence of ring cleavage was detected.

As an alternative to growing *T. primitia* in a liquid media under a defined microoxic headspace, cultures were also initiated in tubes containing an agarose-solidified media under a microoxic headspace in order to establish oxygen gradient cultures. Under these conditions it became clear that both strains of *T. primitia* influence the extent of the inward diffusion of oxygen into the media, with or without an aromatic substrate present. Moreover, the distribution of the growth in the anaerobic portions of the agar was not uniform. Rather, cells accumulated upwards towards the oxic anoxic interface as visualized by the transition of the redox indicator dye resazurin from pink in color to clear. Because the balance of the microoxic headspace was N_2/CO_2, the results suggest that the presence of oxygen stimulates the yield of electron donors present in the media (yeast autolysate and, when present, utilizable aromatic substrates). In several cases, the addition of certain aromatic substrates (catechol, protocatechuic acid, hydrocaffeic acid, and caffeic acid) further stimulated the ability of *T. primitia* to impact the degree of oxygen

penetration into the agar, suggesting that these compounds are being metabolized in oxygen sink reactions (Table 3-4). The impact of the presence or absence of aromatics and/or viable cells on oxygen penetration is summarized in Table 3-4 and Fig. 3-7 as well as Appendix Fig. 3A-20a and b and 3A-21a and b. These analyses provide additional support that *T. primitia* str. ZAS-1 is perhaps the better candidate of the two strains for future studies on oxygen and aromatic metabolisms in this "anaerobic" species (Appendix Fig. 3A-20a and b).

Table 3-4: Average depth of O_2 penetration in *Treponema primitia* str. ZAS-1 and ZAS-2 O_2/aromatic gradient cultures

Aromatic Substrate	Depth (mm)	
	str. ZAS-1	str. ZAS-2
	Ave. +/- Std. Error	Ave. +/- Std. Error
catechol	5 +/-1[ab]	20 +/- 1
protocatechuic acid	6 +/-3[ab]	24 +/- 1
hydrocaffeic acid	6 +/- 1[ab]	13 +/- 1[a]
caffeic Acid	7 +/- 2[b]	18 +/-1
H_2O	9 +/- 2[b]	19 +/- 1

[a]Significantly different from corresponding H_2O control as determined by Student's t-test p-value calculations at 95% confidence.
[b]Significantly different from corresponding *T. primitia* str. ZAS-2 cultures as determined by Student's t-test p-value calculations at 95% confidence.
After 7 days of incubation, distance of O_2 penetration into semi-solid agar culture media determined by resazurin front.
Data is average of triplicate cultures. For comparison, O_2 reached the bottom of cell-free control tubes (22mm) by day 3 (Fig. 3-6a).

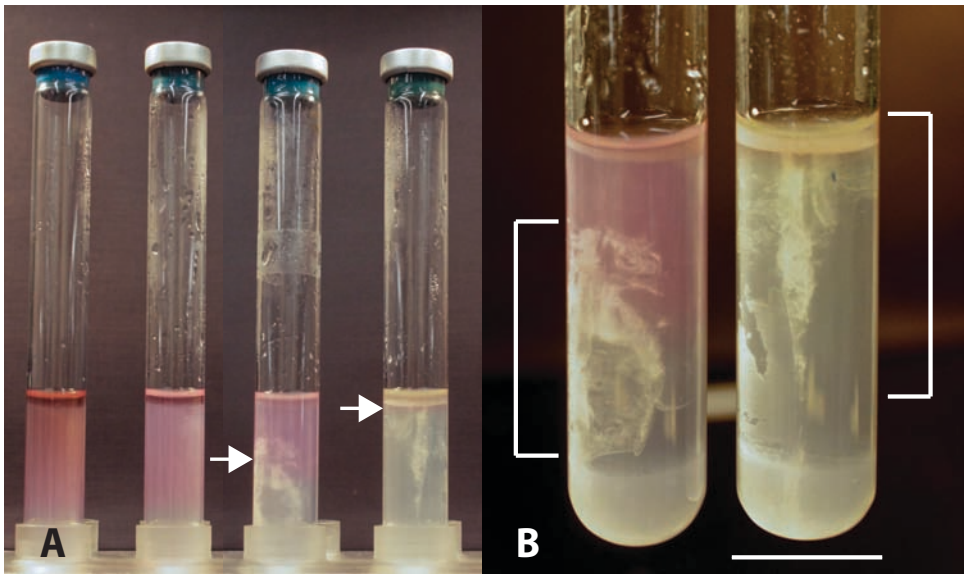


Fig. 3-7: *Treponema primitia* str. ZAS-1 O_2/catechol gradient cultures and controls. Oxygen gradient tubes under a headspace of $N_2/CO_2/O_2$ (70:10:4) after 7 days of incubation. (a) From left to right: uninoculated O_2 and catechol control; uninoculated O_2-only control; inoculated O_2-only control; and inoculated O_2 and catechol. (b) Left: closer view of inoculated O_2-only control; Right: closer view of inoculated O_2 and catechol. Arrows indicate the redox transition zone as represented by resazurin decolorization. Brackets highlight accumulations of cell density. Media and cultivation conditions are detailed in the Materials and Methods. Bar = 15mm.

With a catechol gradient only and no O_2 available, *T. primitia* str. ZAS-1 cultures do not display robust growth. Rather, they appear like their unoculated control (Appendix Fig. 3A-21a).

DISCUSSION:

***Treponema primitia* strains ZAS-1 and ZAS-2 have genes representing complete catechol 2,3-dioxygenase *meta*-cleavage pathways**

That reciprocal BLAST analyses conducted between genomes from three "lower" termite hindgut isolates, *Treponema primitia* str. ZAS-1 and ZAS-2, and *T. azotonutricium* str. ZAS-9 (Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004; Rosenthal *et al*. 2011; Ballor *et al*. 2012), and the P3 hindgut region

metagenome from the "higher" termite, *Nasutitermes* sp. (Warnecke *et al*. 2007) revealed a putative "catechol 2,3-dioxygenase" gene in each of the *T. primitia* str. ZAS-1 and ZAS-2 genomes, but not in the *T. azotonutricium* str. ZAS-9 genome nor higher termite metagenome, was unexpected but helps reconcile O_2 in the termite hindgut, strategies of "anaerobic" organisms to deal with O_2-related stress, and O_2-requiring metabolisms related to lignocellulose degradation.

It should also be noted that the "higher" termite metagenome coverage was less than 1X the entire hindgut population (Warnecke *et al*. 2007). Likely coverage was more than 1X for the most abundant species but less than 1X for lower abundance members. Because the termite gut has a constantly changing composition of members as well as community members in relatively low abundance, it was difficult to make sure a reliable "coverage" is obtained since an "*a-priori*" knowledge of what is in the gut – when this genome was generated - was not available. Nevertheless, current technology could aquire a larger dataset, and with this "higher" termite metagenomic dataset as a reference, another similar project to obtain more coverage could be conducted.

Genomic hints of *meta*-pathway substrate preference

Key residues of the *T. primitia meta*-cleavage pathway genes provide hints about their functions (Díaz & Timmis 1995; Kita *et al*. 1999; Nardini & Dijkstra 1999; Holmquist 2000; Nandhagopal *et al*. 2001; Rea *et al*. 2005; Izumi *et al*. 2007; Lei *et al*. 2008). For example, in *Pseudomonas putida* str. mt-2, catechol and O_2 are thought to enter the active site through a channel comprised mainly of large

residues: His-153, His-199, Tyr-255, and Phe-302. These relatively bulky residues make the channel is narrower than that of other dioxygenases, consistent with the fact that catechol is a relatively small aromatic substrate (Kita *et al*. 1999). Similarly, these bulky residues are represented in the *T. primitia* catechol 2,3-dioxygenases (His-154, His-200, Tyr-256, Phe-303) suggesting these enzymes also prefer smaller substrates such as catechol.

Moreover, the fact that *T. primitia* has genes representing one, but not the other, *meta*-pathway routes has bearing on understanding the aromatic substrates for the pathway. Past studies have indicated that the 2-hydroxymuconic semialdehyde hydrolase (step 2), for instance, prefers ketone group-containing ring cleavage products, in comparison to the 2-hydroxymuconic semialdehyde dehydrogenase (step 6) that is active only on aldehyde group-containing ring cleavage products (Fig. 3-1) (Khajamohiddin *et al.* 2006). This suggests that *T. primitia* catechol 2,3-dioxygenases prefer substrates with *meta*-cleavage products that contain ketone groups.

The *T. primitia meta*-cleavage pathway gene neighborhood arrangement is distinct from other organisms possessing this pathway, which is also unique as there are only a few conserved gene neighborhood arrangements among the *meta*-cleavage pathway containing organisms (Fig. 3-2) (Suenaga *et al.* 2009).

No evidence for "upper" preparatory pathway genes in *Treponema primitia*

In addition to the unique arrangement of *meta*-pathway genes, there is no evidence of genes representing "upper" preparatory pathways in *Treponema primitia*,

commonly found in other organisms possessing *meta*-pathway homologs (Suenaga *et al.* 2009). This suggests that only a limited number of candidate aromatic substrates might serve as possible substrates. Alternatively, given that *T. primitia* str. ZAS-1 and ZAS-2 naturally are part of a complex microbial ecosystem, perhaps another community member has "upper" pathway capabilities and provides "lower" pathway substrates to *T. primitia*. One way to test whether or not other organisms provide "upper" pathway products to *T. primitia* to fuel their "lower" pathway metabolism, would be to feed *T. primitia* str. ZAS-1 and ZAS-2 cell-free culture fluid from an organism known to perform the "upper" pathway in lieu of aromatic-like substrates and filter-sterilized termite hindgut fluid (which would require a lot of termites) and/or cell-free enrichment culture fluid that was prepared to select for organisms with "upper" pathway capabilities. *Azotobacter vinelandii*, *Novosphingobium aromaticivorans*, *Novosphingobium* sp., *Methylocella silvestris*, *Sphingobium japonicum*, *Azoarcus* sp., *Pseudomonas putida*, *Thauera* sp., *Dechloromonas aromatica*, and *Methylibium petroleiphilum* are known to perform "upper" pathways. Distinct *meta*-cleavage products could be detected, for example, colorimetrically and with UV/Vis.

Evolutionary origins of the *Treponema primitia meta*-cleavage pathway

Interestingly, in light of the fact that the majority of the closest-cultured relatives phylogenetically associating with the *Treponema primitia* catechol 2,3-dioxygenases are "Sphingomonad" alphaproteobacteria, examination of gene neighborhoods and genomes in the databases reveals that catechol 2,3-dioxygenases and other key pathway genes in these "sphingomonads" are typically found on plasmids and other

mobile genetic elements, and this is possibly how catechol 2,3-dioxygenase was acquired by *T. primitia* (Harayama & Rekik 1990; Stillwell *et al*. 1995) (Fig. 3-3). Because phylogenetic analyses suggest that the genes for the pathway likely originate from at least two, if not three sources - an alphaproteobacterium, a member of the clostridiales, and possibly a member of the bacillales - it is plausible that a pathway found in many anaerobes, but that does not involve the O_2-dependent utilization of aromatic compounds, was augmented with genes acquired by lateral transfer.

Moreover, analyses of the gene and genome databases indicate that evidence for *meta*-cleavage pathways has previously been restricted to species representing three bacterial phyla: *Proteobacteria*, *Firmicutes*, and *Actinobacteria*. Thus, the current results extend the distribution of the genes for this pathway to a fourth phylum, the *Spirochetes*. The available data suggests *meta*-cleavage pathways are absent in all other spirochetes that have been studied.

Selection pressure and expression analyses

That none of the pathway related ORFs in the *Treponema primitia* strains appear to be pseudogenes, and all appear to encode functional enzymes, and all of the pathway genes with the exception of 2-hydroxymuconic semialdehyde hydrolase appear to be under positive selection, suggests that the majority of the *meta*-cleavage pathway have been used over time and have provided benefit to *T. primitia* (Table 3-2).

Further, that all of the *meta*-pathway genes are expressed even after cultivation without O_2 or aromatic substrate since they were isolated over a decade ago suggests the *meta*-pathway genes are constitutively expressed (Appendix Table 3A-2 and Table 3A-3). Nevertheless, perhaps trace levels of O_2 are in the media, O_2 is a metabolic by-product generated in the media, and aromatic compounds are available in the undefined yeast autolysate leading to the *meta*-pathway genes being expressed. If this were the case, however, the distinct ring cleavage yellow may have been seen which has not been observed before this experiment.

Metabolism of catechol under microoxic conditions

It is concluded that both aromatic substrate and oxygen are needed to be added to cultures to observe ring cleavage products. This is not a phenomenon observed in cultures to which aromatic substrate is added alone (Appendix Fig. 3A-19). Further, the ring cleavage product peak at 375nm is only observed when both 0.5mM aromatic substrate and 0.5% vol/vol O_2 is added to cultures. This peak was not seen in cultures to which either 0.5mM aromatic substrate or 0.5% vol/vol O_2 was added (Appendix Fig. 3A-19).

In liquid culture of *T. primitia* str. ZAS-1 with catechol, the fact that the *meta*-cleavage first product was detected colorimetrically and with UV/Vis but was transient suggests that the intermediate is being metabolized further by the other *meta*-cleavage pathway enzymes. The accumulation of the expected final products was not measured as they would be at the limit of detection, taking into account the small amounts of the substrate metabolized.

With a catechol gradient only and no O_2 available, *T. primitia* str. ZAS-1 cultures do not display robust growth. Rather, they appear like their unoculated control (Appendix Fig. 3A-21a). This could be because these aromatic compounds are inhibitory to *T. primitia* or that in this media that lacks a carbon source other than the aromatic substrate (except what my be minimally provided in the 2YACo) *T. primitia* starves because it cannot metabolize the aromatic substrate. However, when water is provided as a control substrate and no O_2 is available growth (albeit small) is still seen (Appendix Fig. 3A-21b). Therefore it is likely these aromatic compounds are inhibitory and perhaps especially so when cells are starved to some degree.

This trend could mean that in cultures with both aromatic and O_2 gradients in which less O_2 penetration is seen (Table 3-4, Fig. 3-7) that cells are preferentially growing towards the headspace to move away from the aromatic. Decreasing the concentration of aromatic in the gradient tube and generating aromatic gradient only cultures with robust growth throughout could establish an aromatic concentration regime to re-investigate these O_2 and aromatic gradient tubes.

Perhaps, however, with O_2 available, *T. primitia* can cleave the inhibitory aromatic substrates and generate presumably less toxic ring cleavage products. This could be why growth is seen in aromatic gradient tubes to which O_2 is added. Other work in this study verifies catechol ring cleavage by *T. primitia* str. ZAS-1, and this ability could be a strategy to detoxify the hindgut environment of lignocellulose-derived aromatic monomers (Brune *et al*. 1995a).

Conclusions

The genomes of two strains of the "anaerobic" spirochete *T. primitia* both contain all genes for a complete and possibly functional catechol 2,3-dioxygenase *meta*-cleavage pathway. Moreover, *T. primitia* str. ZAS-1 performs at least the first enzymatic step of the *meta*-pathway *in vitro*, utilizing O_2 to cleave the ring of catechol to the expected degradation intermediate. The diet of wood-feeding termites is rich in aromatics, albeit in a form that is recalcitrant to degradation (the aromatic polymer lignin). These results suggest that any monomeric aromatics possibly liberated by insect and microbial metabolism in the gut might become substrates in non-canonical oxygen co-utilizing reactions. Indeed, many past studies have focused on aromatic degradation activities and culture strategies under strictly aerobic or anaerobic conditions (Brune 1998b; Lovely 2000; Gibson & Harwood 2002; Fuchs *et al*. 2011), conditions that might have overlooked the activities of oxygen dependent reactions that do not result in the respiratory mineralization of the ring carbon. In the future, it will likely be important to re-examine any potential for aromatic metabolism in termite hindgut communities with this in mind.

Future work

Although we do not see an improvement of growth by the *Treponema primitia* strains cultured with aromatic and O_2, this does not necessarily indicate that they are not generating acetate with these two substrates via the *meta*-cleavage pathway. Rather, it can be concluded that any potential acetate production is not resulting in an increase in growth rates or yield (Graber & Breznak 2004; Graber *et al*. 2004).

Therefore, in the future other techniques could be pursued to evaluate *meta*-cleavage pathway capabilities by *T. primitia*.

Analytical Methods

If appropriate and sensitive protocols can be established, a combination of gas chromatography, liquid chromatography-mass spectrometry, and ion chromatography could be employed to evaluate different aspects of putative *meta*-cleavage pathway functionality in *Treponema primitia* str. ZAS-1. First, gas chromatography could be used to measure O_2 draw down by *T. primitia* str. ZAS-1 cultures, indicative of the first step of the *meta*-pathway performed by catechol 2,3-dioxygenase, to which both an aromatic substrate and O_2 have been added. Moreover, liquid chromatography-mass spectrometry could potentially detect substrates, intermediates, and products along the *meta*-cleavage pathway route. To evaluate the formation of the *meta*-cleavage pathway end product, acetate, ion chromatography could be utilized. Tracing radio-labeled methyl groups into possible acetate generated by the strain is another more-direct method.

Revise media design

Similarly, while the concentrations of aromatic substrates and O_2 added to cultures thus far do not appear to be inhibitory to growth, they could still be inhibitory to enzyme function seeing as not all aromatic added to the cultures is utilized in this set-up (and the draw down of O_2 in culture headspace is unknown). Therefore, a decrease in aromatic and O_2 substrates would be a good way to start amending media design.

Feeding experiments and co-culturing

One way to test if the *Treponema* strains are generating acetate which is used by other organisms would be to prepare "feeding" experiments and look for marked increases in growth rate and yield of the "fed" organism. For instance, instead of acetate as a substrate, cell-free *Treponema* culture fluid could be provided to other cultures – organisms known to metabolize acetate and/or other organisms isolated from the termite hindgut environment. Co-cultures between the *Treponema* strains and these organisms could also be informative if successful (Rosenthal *et al.* 2011).

Molecular techniques

Under these, and normal culture conditions, qRT-PCR of relevant genes (Appendix Methods 3A-1) and cloning/expression/ enzyme activity assay work could also measure activity, expression, or lack thereof of these relevant *meta*-cleavage pathway genes. Genome-wide transcriptomics could lead to a better understand of the *Treponema primitia* strains' relationship with O_2.

Visualization

A *Zootermopsis nevadensis* worker termite hindgut was successfully sectioned. Protozoa were prominent, including *Streblomastix strix* is an oxymonad symbiont of the termite *Zootermopsis angusticoli* (Appendix Fig. 3A-22b) (Leander & Keeling 2004). While artifacts can be induced during the fixation, dehydration, and embedding protocols, HCR-FISH to map mRNA expression could co-localize the *Treponema* 16S rRNA gene with relevant *meta*-cleavage pathway genes. Due to peristaltic movement of the gut, members of the gut microbial community are likely

constantly being disoriented and re-establishing themselves in preferential locations. While a single time-point section may not capture *T. primitia*'s ideal location, co-localizing 16S rRNA gene with relevant *meta*-cleavage pathway genes *in situ* could provide useful information.

REFERENCES:

Aziz RK, Bartels D, Best AA *et al*. (2008) The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, **9**, 75.

Ballor NR, Paulsen I, Leadbetter JR (2012) Genomic analysis reveals multiple [FeFe] hydrogenases and hydrogen sensors encoded by treponemes from the H_2-rich termite gut. *Microbial Ecology*, **63**, 282-294.

Blomquist GJ, Howard RW, McDaniel CA (1979) Biosynthesis of the cuticular hydrocarbons of the termite *Zootermopsis angusticollis* (Hagen). Incorporation of propionate into dimethylalkanes. *Insect Biochemistry*, **9**, 371-374.

Boga HI, Brune A (2003) Hydrogen-dependent oxygen reduction by homoacetogenic bacteria isolated from termite guts. *Applied and Environmental Microbiology*, **69**, 779-786.

Boga HI, Ludwig W, Brune A (2003) *Sporomusa aerivorans* sp. nov., an oxygen-reducing homoacetogenic bacterium from the gut of a soil-feeding termite. *International Journal of Systematic and Evolutionary Microbiology*, **53**, 1397-1404.

Brauman A, Kane MD, Labat M, Breznak JA (1992) Genesis of acetate and methane by gut bacteria of nutritionally diverse termites. *Science*, **257**, 1384-1387.

Breznak JA, Brune A (1994) Role of microorganisms in the digestion of lignocellulose by termites. *Annual Review of Entomology*, **39**, 453-487.

Breznak JA, Switzer JM (1986) Acetate synthesis from H_2 plus CO_2 by termite gut microbes. *Applied and Environmental Microbiology*, **52**, 623-630.

Brune A (1998a) Termite guts: the world's smallest bioreactors. *Trends in Biotechnology*, **16**, 16-21.

Brune A (1998b) Microbial degradation of aromatic compounds: aerobic versus anaerobic processes. *Mitteilungen der Deutschen Bodenkundlichen Gesellschaft*, 87, 65-78.

Brune A, Miambi E, Breznak JA (1995a) Roles of oxygen and the intestinal microflora in the metabolism of lignin-derived phenylpropanoids and other monoaromatic compounds by termites. *Applied and Environmental Microbiology*, **61**, 2688-2695.

Brune A, Emerson D, Breznak JA (1995b) The termite gut microflora as an oxygen sink: microelectrode determination of oxygen and pH gradients in guts of lower and higher termites. *Applied and Environmental Microbiology*, **61**, 2681-2687.

Bugg TDH, Winfield CJ (1998) Enzymatic cleavage of aromatic rings: mechanistic aspects of the catechol dioxygenases and later enzymes of bacterial oxidative cleavage pathways. *Natural Product Reports*, **15**, 513-530.

Butler JHA, Buckerfield JC (1979) Digestion of lignin by termites. *Soil Biology & Biochemistry*, **11**, 507-513.

Caspi R, Altman T, Dreher K *et al*. (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Research*, **40**, D742-D753.

Choi HMT, Chang JY, Trinh LA *et al*. (2010) Programmable *in situ* amplification for multiplexed imaging of mRNA expression. *Nature Biotechnology*, **28**, 1208-1214.

Cleveland LR (1926) Symbiosis among animals with special reference to termites and their intestinal flagellates. *The Quarterly Review of Biology*, **1**, 51-60.

Cookson LJ (1987) ^{14}C-lignin degradation by three Australian termite species. *Wood Science and Technology*, **21**, 11-25.

Davidsen T, Beck E, Ganapathy A *et al*. (2010) The comprehensive microbial resource. *Nucleic Acids Research*, **38**, D340-D345.

Díaz E, Timmis KN (1995) Identification of functional residues in a 2-hydroxymuconic semialdehyde hydrolase. *Journal of Biological Chemistry*, **270**, 6403-6411.

Eltis LD, Bolin JT (1996) Evolutionary relationships among extradiol dioxygenases. *Journal of Bacteriology*, **178**, 5930-5937.

Fuchs G, Boll M, Heider J (2011) Microbial degradation of aromatic compounds – from one strategy to four. *Nature Reviews Microbiology*, **9**, 803-816.

Furukawa K, Hirose J, Suyama A *et al.* (1993) Gene components responsible for discrete substrate specificity in the metabolism of biphenyl (*bph* operon) and toluene (*tod* operon). *Journal of Bacteriology*, **175**, 5224-5232.

Geer LY, Marchler-Bauer A, Geer RC *et al*. (2010) The NCBI BioSystems database. *Nucleic Acids Research*, **38**, D492-D496.

Gibson J, Harwood CS (2002) Metabolic diversity in aromatic compound utilization by anaerobic microbes. *Annual Review of Microbiology*, **56**, 345-369.

Glauert AM, Lewis PR (1998) Biological Specimen Preparation for Transmission Electron Microscopy. In: *Practical Methods in Electron Microscopy* (ed Glauert AM). Clare Hall, Cambridge.

Graber JR, Breznak JA (2004) Physiology and nutrition of *Treponema primitia*, an H_2/CO_2-acetogenic spirochete from termite hindguts. *Applied and Environmental Microbiology*, **70**, 1307-1314.

Graber JR, Leadbetter JR, Breznak JA (2004) Description of *Treponema azotonutricium* sp. nov. and *Treponema primitia* sp. nov., the first spirochetes isolated from termite guts. *Applied and Environmental Microbiology*, **70**, 1315-1320.

Harayama S, Mermod N, Rekik M *et al.* (1987) Roles of the divergent branches of the *meta*-cleavage pathway in the degradation of benzoate and substituted benzoates. *Journal of Bacteriology*, **169**, 558-564.

Harayama S, Rekik M (1990) The *meta*-cleavage operon of TOL degradative plasmid pWW0 comprises 13 genes. *Molecular and General Genetics*, **221**, 113-120.

Haydon DJ, Guest JR (1991) A new family of bacterial regulatory proteins. *FEMS Microbiology Letters*, **79**, 291-296.

Hillerich B, Westpheling J (2006) A new GntR family transcriptional regulator in *Streptomyces coelicolor* is required for morphogenesis and antibiotic production and controls transcription of an ABC transporter in response to carbon source. *Journal of Bacteriology*, **188**, 7477-7487.

Holmquist M (2000) Alpha/beta-hydrolase fold enzymes: structures, functions and mechanisms. *Current Protein & Peptide Science*, **1**, 209-235.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754-755.

Hugo N, Meyer C, Armengaud J *et al.* (2000) Characterization of three XylT-like [2Fe-2S] ferredoxins associated with catabolism of cresols or naphthalene: evidence for their involvement in catechol dioxygenase reactivation. *Journal of Bacteriology*, **182**, 5580-5585.

Hungate R (1955) Mutualistic Intestinal Protozoa. In: *Biochemistry and Physiology of Protozoa* (eds Hunter S, Lwoff A), pp. 159-199. Academic Press, New York.

Izumi A, Rea D, Adachi T *et al.* (2007) Structure and mechanism of HpcG, a hydratase in the homoprotocatechuate degradation pathway of *Escherichia coli*. *Journal of Molecular Biology*, **370**, 899-911.

Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, **28**, 27-30.

Kappler A, Brune A (1999) Influence of gut alkalinity and oxygen status on mobilization and size-class distribution of humic acids in the hindgut of soil-feeding termites. *Applied Soil Ecology*, **13**, 219-229.

Kappler A, Brune A (2002) Dynamics of redox potential and changes in redox state of iron and humic acids during gut passage in soil-feeding termites (*Cubitermes* spp.). *Soil Biology & Biochemistry*, **34**, 221-227.

Khajamohiddin S, Babu PS, Chakka D *et al.* (2006) A novel *meta*-cleavage product hydrolase from *Flavobacterium* sp. ATCC27551. *Biochemical and Biophysical Research Communications*, **351**, 675-681.

Khajamohiddin S, Repalle ER, Pinjari AB *et al.* (2008) Biodegradation of aromatic compounds: an overview of *meta*-fission product hydrolases. *Critical Reviews in Microbiology*, **34**, 13-31.

Kim D, Chae JC, Jang JY *et al*. (2005) Functional characterization and molecular modeling of methylcatechol 2,3-dioxygenase from *o*-xylene-degrading *Rhodococcus* sp. strain DK17. *Biochemical and Biophysical Research Communications*, **326**, 880-886.

Kita A, Kita S, Fujisawa I *et al*. (1999) An archetypical extradiol-cleaving catecholic dioxygenase: the crystal structure of catechol 2,3-dioxygenase (metapyrocatechase) from *Pseudomonas putida* mt-2. *Structure*, **7**, 25-34.

Kobayashi T, Ishida T, Horiike K *et al*. (1995) Overexpression of *Pseudomonas putida* catechol 2,3-dioxygenase with high specific activity by genetically engineered *Escherichia coli*. *Journal of Biochemistry*, **117**, 614-622.

Kuhnigk T, Borst EM, Ritter A *et al*. (1994) Degradation of lignin monomers by the hindgut flora of xylophagous termites. *Systematic and Applied Microbiology*, **17**, 76-85.

Leadbetter JR, Breznak JA (1996) Physiological ecology of *Methanobrevibacter cuticularis* sp. nov. and *Methanobrevibacter curvatus* sp. nov., isolated from the hindgut of the termite *Reticulitermes flavipes*. *Applied and Environmental Microbiology*, **62**, 3620-3631.

Leadbetter JR, Schmidt TM, Graber JR, Breznak JA (1999) Acetogenesis from H_2 plus CO_2 by spirochetes from termite guts. *Science*, **283**, 686-689.

Leander BS, Keeling PJ (2004) Symbiotic innovation in the oxymonad *Streblomastix strix*. *Journal of Eukaryotic Microbiology*, **3**, 291-300.

Lei Y, Pawelek PD, Powlowski J (2008) A shared binding site for NAD^+ and coenzyme A in an acetaldehyde dehydrogenase involved in bacterial degradation of aromatic compounds. *Biochemistry*, **47**, 6870-6882.

Leidy J (1877) On intestinal parasites of *Termes flavipes*. *Proceedings of the Academy of Natural Sciences of Philadelphia*, **29**, 146-149.

Lovley DR (2000) Anaerobic benzene degradation. *Biodegradation*, **11**, 107-116.

Ludwig W, Strunk O, Westram R *et al*. (2004) ARB: a software environment for sequence data. *Nucleic Acids Research*, **32**, 1363-1371.

Markowitz VM, Korzeniewski F, Palaniappan K *et al*. (2006) The integrated microbial genomes (IMG) system. *Nucleic Acids Research*, **34**, D344-D348.

Mauldin JK (1982) Lipid synthesis from [^{14}C]-acetate by two subterranean termites, *Reticulitermes flavipes* and *Coptotermes formosanus*. *Insect Biochemistry*, **12**, 193-199.

Münch R, Hiller K, Grote A *et al*. (2005) Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics*, **21**, 4187-4189.

Nandhagopal N, Yamada A, Hatta T *et al*. (2001) Crystal structure of 2-hydroxyl-6-oxo-6-phenylhexa-2,4-dienoic acid (HPDA) hydrolase (BphD enzyme) from the *Rhodococcus* sp. strain RHA1 of the PCB degradation pathway. *Journal of Molecular Biology*, **309**, 1139-1151.

Nardini M, Dijkstra BW (1999) a/b hydrolase fold enzymes: the family keeps growing. *Current Opinion in Structural Biology*, **9**, 732-737.

Néron B, Ménager H, Maufrais C *et al*. (2009) Mobyle: a new full web bioinformatics framework. *Bioinformatics*, **25**, 3005-3011.

Odelson DA, Breznak JA (1983) Volatile fatty acid production by the hindgut microbiota of xylophagous termites. *Applied and Environmental Microbiology*, **45**, 1602-1613.

Odelson DA, Breznak JA (1985a) Nutrition and growth characteristics of *Trichomitopsis termopsidis*, a cellulolytic protozoan from termites. *Applied and Environmental Microbiology*, **49**, 614-621.

Odelson DA, Breznak JA (1985b) Cellulase and other polymer-hydrolyzing activities of *Trichomitopsis termopsidis,* a symbiotic protozoan from termites. *Applied and Environmental Microbiology*, **49**, 622-626.

Pasti MB, Pometto III AL, Nuti MP, Crawford DL (1990) Lignin-solubilizing ability of actinomycetes isolated from termite (Termitidae) gut. *Applied and Environmental Microbiology*, **56**, 2213-2218.

Pester M, Brune A (2007) Hydrogen is the central free intermediate during lignocellulose degradation by termite gut symbionts. *ISME Journal*, **1**, 551-565.

Polissi A, Harayama S (1993) *In vivo* reactivation of catechol 2,3-dioxygenase mediated by a chloroplast-type ferredoxin: a bacterial strategy to expand the substrate specificity of aromatic degradative pathways. *EMBO Journal*, **12**, 3339-3347.

Punta M, Coggill PC, Eberhardt RY *et al*. (2012) The Pfam protein families database. *Nucleic Acids Research*, **40**, D290-D301.

Rea D, Fülöp V, Bugg TDH, Roper DI (2005) Expression, purification and preliminary crystallographic analysis of 2,4-dihydroxy-hepta-2-ene-1,7-dioate aldolase (HpcH) from *Escherichia coli C. Acta Crystallographica Section F*, **61**, 821-824.

Rigali S, Derouaux A, Giannotta F, Dusart J (2002) Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies. *The Jounral of Biological Chemistry*, **277**, 12507-12515.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572-1574.

Rosenthal AZ, Matson EG, Eldar A, Leadbetter JR (2011) RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. *ISME Journal*, **5**, 1133-1142.

Scharf ME, Boucias DG (2010) Potential of termite-based biomass pre-treatment strategies for use in bioethanol production. *Insect Science*, **17**, 166-174.

Scharf ME, Tartar A (2008) Termite digestomes as sources for novel lignocellulases. *Biofuels, Bioproducts and Biorefining*, **2**, 540-552.

Sethi A, Slack JM, Kovaleva ES, Buchman GW, Scharf ME (2013) Lignin-associated metagene expression in a lignocellulose-digesting termite. *Insect Biochemistry and Molecular Biology*, **43**, 91-101.

Shima S, Netrusov A, Sordel M *et al*. (1999) Purification, characterization, and primary structure of a monofunctional catalase from *Methanosarcina barkeri*. *Archives of Microbiology*, **171**, 317-323.

Shima S, Sordel-Klippert M, Brioukhanov A *et al*. (2001) Characterization of a heme-dependent catalase from *Methanobrevibacter arboriphilus. Applied and Environmental Microbiology*, **67**, 3041-3045.

Siani L, Viggiani A, Notomista E, *et al*. (2006) The role of residue Thr249 in modulating the catalytic efficiency and substrate specificity of catechol-2,3-dioxygenase from *Pseudomonas stutzeri* OX1. *FEBS Journal*, **273**, 2963-2976.

Stillwell LC, Thurston SJ, Schneider RP *et al*. (1995) Physical mapping and characterization of a catabolic plasmid from the deep-subsurface bacterium *Sphingomonas* sp. strain F199. *Journal of Bacteriology*, **177**, 4537-4539.

Suenaga H, Koyama Y, Miyakoshi M *et al*. (2009) Novel organization of aromatic degradation pathway genes in a microbial community as revealed by metagenomic analysis. *ISME Journal*, **3**, 1335-1348.

Suenaga H, Ohnuki T, Miyazaki K (2007) Functional screening of a metagenomic library for genes involved in microbial degradation of aromatic compounds. *Environmental Microbiology*, **9**, 2289-2297.

Sun S, Chen J, Weizhong L *et al*. (2011) Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis: the CAMERA resource. *Nucleic Acids Research*, **39**, D546-D551.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596-1599.

Tartar A, Wheeler MM, Zhou X *et al*. (2009) Parallel metatranscriptome analyses of host and symbiont gene expression in the gut of the termite *Reticulitermes flavipes. Biotechnology for Biofuels*, **2**, 25.

Tholen A, Pester M, Brune A (2007) Simultaneous methanogenesis and oxygen reduction by *Methanobrevibacter cuticularis* at low oxygen fluxes. *FEMS Microbiology Ecology*, **62**, 303-312.

Tropel D, van der Meer JR (2004) Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiology and Molecular Biology Reviews*, **68**, 474-500.

Vaillancourt FH, Bolin JT, Eltis LD (2006) The ins and outs of ring-cleaving dioxygenases. *Critical Reviews in Biochemistry and Molecular Biology*, **41**, 241-267.

Viggiani A, Siani L, Notomista E *et al*. (2004) The role of the conserved residues His-246, His-199, and Tyr-255 in the catalysis of catechol 2,3-dioxygenase from *Pseudomonas stutzeri* OX1. *Journal of Biological Chemistry*, **279**, 48630-48639.

Warnecke F, Luginbühl P, Ivanova N *et al*. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature*, **450**, 560-565.

Wertz JT, Breznak JA (2007a) *Stenoxybacter acetivorans* gen. nov., sp. nov., an acetate-oxidizing obligate microaerophile among diverse O_2-consuming bacteria from termite guts. *Applied and Environmental Microbiology*, **73**, 6819-6828.
Wertz JT, Breznak JA (2007b) Physiological ecology of *Stenoxybacter acetivorans,* an obligate microaerophile in termite guts. *Applied and Environmental Microbiology*, **73**, 6829-6841.

Wertz JT, Kim E, Breznak JA *et al*. (2012) Genomic and physiological characterization of the *Verrucomicrobia* isolate *Diplosphaera colitermitum* gen. nov., sp. nov., reveals microaerophily and nitrogen fixation genes. *Applied and Environmental Microbiology*, **78**, 1544-1555.

Yamin MA (1980) Cellulose metabolism by the termite flagellate *Trichomitopsis termopsidis*. *Applied and Environmental Microbiology*, **39**, 859-863.

Yamin MA, Trager W (1979) Cellulolytic activity of an axenically-cultivated termite flagellate, *Trichomitopsis termopsidis*. *Journal of General Microbiology*, **113**, 417-420.

Zhang X, Matson EG, Leadbetter JR (2011) Genes for selenium dependent and independent formate dehydrogenase in the gut microbial communities of three lower, wood-feeding termites and a wood-feeding roach. *Environmental Microbiology*, **13**, 307-323.

Zimmer M, Brune A (2005) Physiological properties of the gut lumen of terrestrial isopods (Isopoda: Oniscidea): adaptive to digesting lignocellulose? *Journal of Comparative Physiology* B, **175**, 275-283.

CHAPTER 3: APPENDIX

Table 3A-1: GenBank accession numbers and references of organisms and proteins presented in text and figures.

Table 3A-2: Gene expression patterns of *Treponema primitia* str. ZAS-2 in mono-culture or co-culture with *T. azotonutricium* str. ZAS-9.

Table 3A-3: Expression of *Treponema primitia* str. ZAS-2 genes.

Table 3A-4: Genes unique to *Treponema* genomes relative to metagenome of *Nasutitermes* sp. P3 hindgut region.

Fig. 3A-1: PFAM domains and conserved functional residues and sequence motifs in *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-cleavage pathway proteins.

Fig. 3A-2: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 ferredoxin-like peptide (PF00111, Step 1').

Fig. 3A-3: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 4-hydroxy-2-oxopentanoate aldolase (PF00682 and PF07836, Step 4).

Fig. 3A-4: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 acetaldehyde dehydrogenase (PF01118 and PF09290, Step 5).

Fig. 3A-5: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 N-terminal Domain (PF00903).

Fig. 3A-6: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 N-terminal Domain (PF00903) with extra-domain region.

Fig. 3A-7: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 N-terminal Domain (PF00903) with intra-domain region.

Fig. 3A-8: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 C-terminal Domain (PF00903).

Fig. 3A-9: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 C-terminal Domain (PF00903) with extra-domain region.

Fig. 3A-10: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 C-terminal Domain (PF00903) with intra-domain region.

Fig. 3A-11: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 4-hydroxy-2-oxopentanoate aldolase HMGL-like domain (PF00682).

Fig. 3A-12: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 4-hydroxy-2-oxopentanoate aldolase DmpG-like communication domain (PF07836).

Fig. 3A-13: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 acetaldehyde dehydrogenase, NAD binding domain (PF01118).

Fig. 3A-14: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 acetaldehyde dehydrogenase dimerisation domain (PF09290).

Fig. 3A-15: Transcriptional regulatory elements within the *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-cleavage pathway gene neighborhoods.

Fig 3A-16: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-pathway associated GntR family transcriptional regulator (PF00392).

Fig. 3A-17: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-pathway associated GntR family transcriptional regulator, Bacterial regulatory proteins gntR family domain (PF00392).

Fig. 3A-18: Yellow *Treponema primitia* str. ZAS-1 cultures.

Fig. 3A-19: Absorbance spectra of *Treponema primitia* str. ZAS-1 cultures.

Fig. 3A-20: *Treponema primitia* str. ZAS-1 and ZAS-2 O_2/catechol gradient cultures and controls. **(a)** uninoculated catechol and O_2 control and triplicates of O_2/catechol gradient tubes inoculated with of *T. primitia* str. ZAS-2 **(b)** O_2/catechol gradient tubes inoculated with of *T. primitia* str. ZAS-2 compared to *T. primitia* str. ZAS-1.

Fig. 3A-21: *Treponema primitia* str. ZAS-1 **(a)** catechol gradient cultures and controls and **(b)** water gradient cultures and controls.

Method 3A-1: Visualizing *Treponema primitia* in termite hindgut.

Method 3A-2: qRT-PCR of each *meta*-cleavage pathway gene.

Table 3A-1: GenBank accession numbers and references of organisms and proteins presented in text and figures

Figure	Organisms Name	Organism Accession Number	Reference(s)
			Keil *et al.* 1985; Keil *et al.* 1987a; Keil *et al.* 1987b; Osborne *et al.* 1988; Assinder *et al.* 1992; Assinder *et al.* 1993; Gallegos *et al.* 1997; Sentchilo *et al.* 2000; Tsuda and Genka 2001; Yano *et al.* 2007;
Fig. 3-2	*Pseudomonas putida*	AB238971	Miyakoshi *et al.* 2012
Fig. 3-2	*Treponema primitia* str. ZAS-2	CP001843	Rosenthal *et al.* 2011
Fig. 3-2	*Treponema primitia* str. ZAS-1	CP001843	Ballor *et al.* 2011
Fig. 3-2	*Novosphingobium aromaticivorans*	AF079317	Romine *et al.* 1999
Fig. 3-2	*Novosphingobium* sp.	FR856862	D'Argenio *et al.* 2011
Fig. 3-2	*Methylocella silvestris*	CP001280	Chen *et al.* 2010
Fig. 3-2	*Sphingobium japonicum*	AP010803	Nagata *et al.* 2010
Fig. 3-2	*Azoarcus* sp.	AM406670	Krause *et al.* 2006
Fig. 3-2	*Thauera* sp.	CP001281	NA
Fig. 3-2	*Dechloromonas aromatica*	CP000089	NA
Fig. 3-2	*Azotobacter vinelandii*	CP001157	Setubal *et al.* 2009
Fig. 3-2	*Methylibium petroleiphilum*	CP000555	Kane *et al.* 2007
		Catechol 2,3-dioxygenase Accession Number	
Fig. 3-3	*Novosphingobium aromaticivorans*	NP_049202	Romine *et al.*1999
Fig. 3-3	*Sphingomonas agrestis*	AAB03075	Yrjala *et al.*1994
Fig. 3-3	*Rhizobium* sp.	ABF82226	NA
Fig. 3-3	*Beijerinckia* sp.	B57264	Kim and Zylstra 1995
Fig. 3-3	*Sphingomonas* sp.	AAM14600	NA
Fig. 3-3	*Sphingomonas* sp.	ADK27485	NA
Fig. 3-3	*Novosphingobium pentaromativorans*	ZP_09195387	NA
Fig. 3-3	*Novosphingobium* sp.	YP_004534265	D'Argenio *et al.* 2011
Fig. 3-3	*Sphingomonas* sp.	AAD11448	NA
Fig. 3-3	*Sphingomonas* sp.	AAD11452	NA
Fig. 3-3	*Treponema primitia* str. ZAS-1	ZP_09718346	Ballor *et al.* 2011
Fig. 3-3	*Treponema primitia* str. ZAS-2	YP_004531633	Rosenthal *et al.* 2011
Fig. 3-3	lake sediment		
Fig. 3-3	activated sludge	BAH89314	Suenaga *et al.* 2007; Suenaga *et al.* 2009
Fig. 3-3	activated sludge	BAH90343	Suenaga *et al.* 2007; Suenaga *et al.* 2009
Fig. 3-3	*Methylocella silvestris*	YP_002361794	Chen *et al.* 2010
Fig. 3-3	*Sphingobium japonicum*	YP_003544642	Nagata *et al.* 2010

Fig. 3-3	*Pseudomonas putida*	BAB62050	NA
Fig. 3-3	*Thauera* sp.	YP_002890083	NA
Fig. 3-3	*Dechloromonas aromatica*	YP_287003	NA
Fig. 3-3	*Azotobacter vinelandii*	YP_002800217	Setubal *et al.* 2009
Fig. 3-3	*Pseudomonas putida*	YP_709322	Keil *et al.* 1985; Keil *et al.* 1987a; Keil *et al.* 1987b; Osborne *et al.* 1988; Assinder *et al.* 1992; Assinder *et al.* 1993; Gallegos *et al.* 1997; Sentchilo *et al.* 2000; Tsuda and Genka 2001; Yano *et al.* 2007; Miyakoshi *et al.* 2012
Fig. 3-3	*Alcaligenes xylosoxydans*		
Fig. 3-3	*Marinobacter algicola*	ZP_01893284	NA
Fig. 3-3	whale fall rib bone		
Fig. 3-3	*Marinobacterium stanieri*	ZP_09507225	NA
Fig. 3-3	*Marinobacterium stanieri*	ZP_09506167	NA
Fig. 3-3	marine gamma proteobacterium HTCC2207	ZP_01224201	NA
Fig. 3-3	sludge		
Fig. 3-3	*Novosphingobium nitrogenifigens*	ZP_08210144	NA
Fig. 3-3	*Cupriavidus necator*		
Fig. 3-3	*Methyloversatilis universalis*	ZP_08506618	NA
Fig. 3-3	*Methylibium petroleiphilum*	YP_001021468	Kane *et al.* 2007
		2-hydroxymuconic hydrolase Accession Number	
Fig. 3-4	*Rhodococcus equi*	YP_004004919	Letek *et al.* 2008; Letek *et al.* 2010
Fig. 3-4	*Intrasporangium calvum*	YP_004098338	Del Rio *et al.* 2010
Fig. 3-4	*Kyripidia tusciae*	YP_003588846	NA
Fig. 3-4	*Solibacillus silvestris*	YP_006460640	Morohoshi *et al.* 2012
Fig. 3-4	*Bacillus tusciae*	YP_003590143	NA
Fig. 3-4	*Novosphingobium aromaticivorans*	NP_049203	Romine *et al.* 1999
Fig. 3-4	*Novosphingobium pentaromativorans*	ZP_09195386	NA
Fig. 3-4	*Novosphingobium* sp.	YP_004534264	D'Argenio *et al.* 2011
Fig. 3-4	whale fall rib bone		
Fig. 3-4	*Lysobacter* sp.	BAH80176	NA
Fig. 3-4	*Treponema primitia* str. ZAS-1	ZP_09718351	Ballor *et al.* 2011
Fig. 3-4	*Treponema primitia* str. ZAS-2	YP_004531638	Rosenthal *et al.* 2011
Fig. 3-4	lake sediment		
Fig. 3-4	*Arcobacter* sp.	YP_005554631	Toh *et al.* 2011
Fig. 3-4	*Leptothrix cholodnii*	YP_001792367	NA
Fig. 3-4	*Delftia* sp.	ZP_11254636	NA
Fig. 3-4	*Alcaligenes faecalis*	ABV30923	NA

			Nikodem *et al.* 2003; Camara *et al.* 2007a;
Fig. 3-4	*Pseudomonas reinekei*	ABH07023	Camara *et al.* 2007b
Fig. 3-4	*Dechloromonas aromatica*	YP_286985	NA
Fig. 3-4	*Azoarcus* sp.	YP_933472	Krause *et al.* 2006
Fig. 3-4	*Acidovorax* sp.	YP_984551	NA
Fig. 3-4	activated sludge	BAH89584	Suenaga *et al.*, 2007; Suenaga *et al.*, 2009
Fig. 3-4	*Methylibium petroleiphilum*	YP_001021465	Kane *et al.* 2007
Fig. 3-4	*Thauera* sp.	YP_002890084	NA
			Keil *et al.* 1985; Keil *et al.* 1987a; Keil *et al.* 1987b; Osborne *et al.* 1988; Assinder *et al.* 1992; Assinder *et al.* 1993; Gallegos *et al.* 1997; Sentchilo *et al.* 2000; Tsuda and Genka 2001; Yano *et al.* 2007;
Fig. 3-4	*Pseudomonas putida*	YP_709324	Miyakoshi *et al.* 2012
Fig. 3-4	*Methyloversatilis universalis*	ZP_08506603	NA
Fig. 3-4	activated sludge	BAH90404	Suenaga *et al.*, 2007; Suenaga *et al.*, 2009
Fig. 3-4	*Novosphingobium nitrogenifigens*	ZP_08210141	NA
Fig. 3-4	sludge		
Fig. 3-4	*Thiothrix nivea*	WP_002710562	NA
Fig. 3-4	*Rhodococcus jostii*	YP_707286	McLeod *et al.* 2006
Fig. 3-4	*Rhodococcus erythropolis*	NP_898789	Stecker *et al.* 2003
		2-oxopent-4-enoate Hydratase Accession Number	
Fig. 3-5	*Azoarcus* sp.	YP_933478	Krause *et al.* 2006
Fig. 3-5	activated sludge	BAH89311	Suenaga *et al.*, 2007; Suenaga *et al.*, 2009
Fig. 3-5	*Novosphingobium pentaromativorans*	ZP_09195390	NA
Fig. 3-5	*Methylocella silvestris*	YP_002361797	Chen *et al.* 2010
Fig. 3-5	*Sphingobium japonicum*	YP_003544633	Nagata *et al.* 2010
Fig. 3-5	*Azotobacter vinelandii*	YP_002800201	Setubal *et al.* 2009
Fig. 3-5	*Methylibium petroleiphilum*	YP_001021464	Kane *et al.* 2007
Fig. 3-5	*Burkholderia* sp.		
Fig. 3-5	*Arcobacter* sp.	YP_005554636	Toh *et al.* 2011
Fig. 3-5	*Marinobacter manganoxydans*	ZP_09159763	NA
Fig. 3-5	activated sludge	BAH89583	Suenaga *et al.*, 2007; Suenaga *et al.*, 2009
Fig. 3-5	*Pseudomonas putida*	YP_709325	Keil *et al.* 1985; Keil *et al.* 1987a; Keil *et al.* 1987b; Osborne *et al.* 1988; Assinder *et al.* 1992;

			Assinder *et al.* 1993; Gallegos *et al.* 1997; Sentchilo *et al.* 2000; Tsuda and Genka 2001; Yano *et al.* 2007; Miyakoshi *et al.* 2012
Fig. 3-5	*Methyloversatilis universalis*	ZP_08506604	NA
Fig. 3-5	*Dechloromonas aromatica*	YP_286984	NA
Fig. 3-5	*Ralstonia eutropha*	YP_728710	Pohlmann *et al.* 2006
Fig. 3-5	*Geobacillus* sp.	YP_003988985	NA
Fig. 3-5	*Carboxydothermus hydrogenoformans*	YP_360112	Wu et al., 2005
Fig. 3-5	*Thermosinus carboxydivorans*	ZP_01665862	NA
Fig. 3-5	*Pelotomaculum thermopropionicum*	YP_001211031	Kosaka et al., 2006; Kosaka et al., 2008
Fig. 3-5	*Desulfotomaculum nigrificans*	ZP_08114002	NA
Fig. 3-5	*Moorella thermoacetica*	YP_430621	Pierce et al., 2008
Fig. 3-5	*Flavonifractor plautii*	ZP_09382281	NA
Fig. 3-5	*Acetonema longum*	ZP_08624290	NA
Fig. 3-5	*Bacillus cereus*	**ZP_03107092**	
Fig. 3-5	*Treponema primitia* str. ZAS-1	ZP_09718350	Ballor *et al.* 2011
Fig. 3-5	*Treponema primitia* str. ZAS-2	YP_004531637	Rosenthal *et al.* 2011
Fig. 3-5	*Solibacillus silvestris*	YP_006460632	Morohoshi et al., 2012
Fig. 3-5	*Brevibacillus brevis*	YP_002772457	NA
Fig. 3-5	*Paenibacillus* sp.	BAH79101	Kasai et al., 2009
Fig. 3-5	*Geobacillus stearothermophilus*	AAZ76889	NA
Fig. 3-5	*Sulfobacillus acidophilus*	YP_005258131	NA
Fig. 3-5	*Halobacterium* sp.	ZP_09027988	NA
Fig. 3-5	*Alicyclobacillus acidocaldarius*	YP_003185022	Mavromatis et al., 2010
Fig. 3-5	*Thermomicrobium roseum*	YP_002522592	Wu et al., 2009
Fig. 3-5	*Salmonella bongori*	YP_004729834	NA
Fig. 3-5	*Intrasporangium calvum*	YP_004098344	Del Rio et al., 2010

Table 3A-2: Gene expression patterns of *Treponema primitia* str. ZAS-2 in mono-culture or co-culture with *T. azotonutricium* str. ZAS-9

Cohort	Expression Range	Mono-Culture		Co-Culture	
	(reads/kb)	No. genes	%	No. genes	%
1[a]	0	344	9	395	10
2	1-9	503	13	556	14
3[b]	10-99	1642	43	1548	40
4[c]	100-999	1248	32	1234	32
5	> 1000	110	3	117	3

Analysis complements Table 3A-3.

Analysis based on expression dataset from Rosenthal *et al.* 2011.

[a]In addition to having 0 reads/kb, all genes in Cohort 1 have 0 expression.

[b]Cohort includes *meta*-pathway 2-hydroxymuconic semialdehyde hydrolase, 2-oxopent-4-enoate hydratase, 4-hydroxy-2-oxopentanoate aldolase, acetaldehyde dehydrogeanse, and the associated ferredoxin-like peptide.

[c]Cohort includes *meta*-pathway catechol 2,3-dioxygenase.

Table 3A-3: Expression of *Treponema primitia* str. ZAS-2 genes

Gene	Pure-Culture Expression (reads/kb)	Co-Culture Expression (reads/kb)
Meta-pathway		
Ferredoxin-like peptide[a]	53	13
catechol 2,3-dioxygenase[b]	219	170
2-hydoxymuconic semialdehyde hydrolase[a]	43	60
2-oxopent-4-enoate hydratase[a]	60	43
4-hydroxy-2-oxopentanoate aldolase[a]	61	36
acetaldehyde dehydrogenase[a]	59	62
House-keeping		
ATP-dependent Clp protease ATP-binding subunit ClpX	314	345
Metallo-beta-lactamase family protein, RNA-specific	140	144
RNA polymerase sigma factor RpoD	72	77
RNA polymerase sigma factor RpoD	183	135
RNA polymerase sigma factor RpoD	188	162
RNA polymerase sigma factor RpoD	94	92
RNA polymerase sigma factor RpoD	30	94
Acetogenesis		
Formate dehydrogenase chain D	62	2
Formate dehydrogenase chain D	399	21
Formate dehydrogenase; cysteine-containing variant	601	33
Formate dehydrogenase; selenocysteine-containing variant	168	331
Carbon monoxide dehydrogenase CooS subunit	2698	2524
Formate-tetrahydrofolate ligase (FTHFS)	1463	1600
Glycolysis		
Hexokinase	168	190
Glucose-6-phosphate isomerase	169	170
6-phosphofructokinase	121	139
Fructose-bisphosphate aldolase class II	703	947
Triosephosphate isomerase	344	418
NAD-dependent glyceraldehyde-3-phosphate dehydrogenase	134	173
NAD-dependent glyceraldehyde-3-phosphate dehydrogenase	152	167
2,3-bisphosphoglycerate-independent phosphoglycerate mutase	440	393
Enolase	851	967
Pyruvate kinase	89	93
Misc.		
acetate kinase	914	1085
Fe-S cluster containing hydrogenase components 2	552	39
Periplasmic [Fe] hydrogenas large subunit	495	31

Expression data from Rosenthal *et al.* 2011.

[a]Belong to Cohort 3 (Table 3A-2).

[b]Belongs to Cohort 4 (Table 3A-2).

Table 3A-4: Genes unique to *Treponema* genomes relative to metagenome of *Nasutitermes* sp. P3 hindgut region

***Treponema primitia* str. ZAS-1**
DCMPdeaminase
UncharacterizedproteinTP_0813
SignalpeptidaseI(EC3.4.21.89)
Probablesignalpeptideprotein
UncharacterizedproteinTP_0181
Xyloseisomerase(EC5.3.1.5)
transcriptionalregulator,AraCfamily
oxidoreductase,aldo>ketoreductasefamily
anti-sigmaFfactorantagonist(spoIIAA-2);
HPrkinase>phosphorylase(EC2.7.1.-)(EC
Aspartyl-tRNA(Asn)amidotransferasesubunitB
RiboseABCtransportsystem,periplasmic
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,permeaseprotein
Biofilm-associatedprotein
HydroxymethylpyrimidineABCtransporter,ATPase
HydroxymethylpyrimidineABCtransporter,
CelldivisionproteinmraZ
S-adenosyl-methyltransferasemraW(EC2.1.1.-)
RossmannfoldnucleotideindingproteinSmf
Ribosome-bindingfactorA
sensoryboxhistidinekinase>response
PhnBprotein;putativeDNAbinding
Probableextracellularnuclease
Nitrogenaseironprotein(EC1.18.6.1)
Single-strandedDNA-bindingprotein
Threonyl-tRNAsynthetase(EC6.1.1.3)
ABCtransporter,ATP-bindingprotein
ABCtransporter,permeaseprotein
Two-componentresponseregulator
RiboseABCtransportsystem,periplasmic
RiboseABCtransportsystem,permeaseprotein
Riboseoperonrepressor
Esterase>lipase
Transcriptionalregulator,MarRfamily
Cellsurfaceprotein
Streptococcalhemagglutininprotein
Cystathioninebeta-lyase(EC4.4.1.8)
Sensoryboxhistidinekinase>response
Aspartatecarbamoyltransferase(EC2.1.3.2)
Putativesugartransporter
Nitrogenasemolybdenum-cofactorsynthesis
Nitrogenasevanadium-cofactorsynthesisprotein
ABC-typenitrate>sulfonate>bicarbonate
ABC-typenitrate>sulfonate>bicarbonate
TaurinetransportATP-bindingproteintauB
Cysteinesynthase(EC2.5.1.47)

Nitrogenasevanadium-cofactorsynthesisprotein
Nitrogenase(molybdenum-iron)alphachain(EC
Transketolase,C-terminalsection(EC2.2.1.1)
Transketolase,N-terminalsection(EC2.2.1.1)
L-idonate5-dehydrogenase(EC1.1.1.264)>
Melibioseoperonregulatoryprotein
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,permeaseprotein
Methyl-acceptingchemotaxisprotein
dTDP-rhamnosyltransferaserfbF(EC2.-.-.-)
O-antigenpolymerase
Mannosyltransferase(EC2.4.1.-)
UDP-glucose4-epimerase(EC5.1.3.2)
MaltoseO-acetyltransferase(EC2.3.1.79)
O-antigenflippaseWzx
Beta-1,3-glucosyltransferase
UDP-glucose4-epimerase(EC5.1.3.2)
dTDP-4-dehydrorhamnose3,5-epimerase(EC
Exopolysaccharideproductionprotein
capsularpolysaccharidebiosynthesisprotein
Glycosyltransferase
Beta-1,3-glucosyltransferase
Beta-1,3-glucosyltransferase
oxidoreductaseofaldo>ketoreductasefamily,
3-oxoacyl-acyl-carrier-protein]synthase,
3-oxoacyl-acyl-carrier-protein]reductase
3-oxoacyl-ACP]reductase
3-oxoacyl-acyl-carrier-protein]synthase,
3-oxoacyl-acyl-carrier-protein]synthase,
RelatedtoF420H2-dehydrogenase,betasubunit
dTDP-glucose4,6-dehydratase(EC4.2.1.46)
Cytosine>purine>uracil>thiamine>allantoin
PutativeDNAprimase>helicase
Tyrosine-proteinkinasewzc(EC2.7.1.112)
PolysialicacidtransportproteinkpsD
SugarABCtransporter,periplasmic
Maltose>maltodextrinABCtransporter,permease
Maltose>maltodextrinABCtransporter,permease
DNA-bindingtranscriptionalregulator
Kef-typeK+transportsystems(NAD-binding
HITfamilyprotein
NAD(P)-dependentglyceraldehyde3-phosphate
Phosphoglyceratekinase(EC2.7.2.3)
Inososedehydratase(EC4.2.1.44)
Epi-inositolhydrolase
5-keto-2-deoxygluconokinaseB(EC2.7.1.92)
SugarABCtransporter,periplasmic
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,permeaseprotein

IolIprotein
Branched-chainaminoacidABCtransporter,
High-affinitybranched-chainaminoacid
Branched-chainaminoacidtransportsystem
Branched-chainaminoacidtransportATP-binding
Branched-chainaminoacidtransportATP-binding
Branched-chainaminoacidtransportATP-binding
Branched-chainaminoacidtransportATP-binding
Branched-chainaminoacidtransportATP-binding
High-affinitybranched-chainaminoacid
Branched-chainaminoacidtransportsystem
Branched-chainaminoacidABCtransporter,
diguanylatecyclase(GGDEFdomain)withPAS>PAC
Membranespanningprotein
NitrogenregulatoryproteinP-II
Ornithinedecarboxylase(EC4.1.1.17)>
UPF0118membraneproteinBB_0006
methyl-acceptingchemotaxisprotein
methyl-acceptingchemotaxisprotein
OligopeptidetransportATP-bindingproteinoppF
OligopeptidetransportATP-bindingproteinoppD
Oligopeptidetransportsystempermeaseprotein
archaealATPase,fusedtoC-terminalDUF234
AnaerobicdimethylsulfoxidereductasechainA
O-acetylhomoserinesulfhydrylase(EC4.2.99.10)
Putativeamino-acidtransporterperiplasmic
MethionineABCtransporterATP-bindingprotein
D-galactose-bindingperiplasmicprotein
GalactosidetransportATP-bindingproteinmglA
Galactosidetransportsystempermeaseprotein
CircadianinputkinaseA
Two-componentsystemsensorprotein
Ureacarboxylase-relatedABCtransporter,
NitrogenregulatoryproteinP-II
NitrogenregulatoryproteinP-II
Ureacarboxylase(EC6.3.4.6)
sensoryboxhistidinekinase>response
Acylcarrierprotein
BchE>P-methylasefamilyprotein
Xaa-Proaminopeptidase(EC3.4.11.9)
Uncharacterizedproteininarchaea
TVG1219960protein
Cellsurfaceprotein
AntiholinlikeproteinLrgA
lrgA-associatedmembraneproteinLrgB
MultiplesugarABCtransporter,
MultiplesugarABCtransporter,
MultiplesugarABCtransporter,
MaltoseoperontranscriptionalrepressorMalR,

OmpAfamilyprotein
3-oxoacyl-(acyl-carrier-protein)synthase
Carbondioxideconcentratingmechanismprotein
Propanediolutilizationpolyhedralbodyprotein
Propanediolutilizationpolyhedralbodyprotein
classIIaldolase>adducindomainprotein
L-fuculokinase(EC2.7.1.51)
RiboseABCtransportsystem,periplasmic
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,permeaseprotein
Redox-sensitivetranscriptionalregulator
D-tagatose3-epimerase(EC5.3.1.-)
Putativeoxidoreductase
IolIprotein
MultiplesugarABCtransporter,
MultiplesugarABCtransporter,
Probablehemagglutinin>hemolysin-related
putativeDNAprimase>helicase
Regulatorofpolyketidesynthaseexpression
Hydantoinracemase(EC5.1.99.-)
Catalyzesthecleavageof
FlagellinproteinflaA
Two-componentsystemsensorprotein
CTP:Inositol-1-phosphatecytidylyltransferase
CapsularpolysaccharidesynthesisenzymecpsI,
Innermembraneprotein
Innermembraneprotein
putativebifunctionalpolymerase
Alpha-1,4-N-acetylgalactosaminetransferase
UDP-glucosedehydrogenase(EC1.1.1.22)
Bll3360protein
Beta-1,3-glucosyltransferase
Phospholipid-lipopolysaccharideABC
MoxR-likeATPases
Benzoyl-CoAreductasesubunitBadG(EC
FormatedehydrogenasechainD(EC1.2.1.2)
FormatedehydrogenaseH(EC1.2.1.2)
Tropomodulin1
DNA-damage-inducibleproteinJ,putative
Smallribosomalsubunit16SrRNA
anti-sigmaFfactorantagonist(spoIIAA-2);
PhageshockproteinA
FOG:CheY-likereceiver
Transcriptionalregulator,luxRfamily,
Glycerol-3-phosphateABCtransporter,
Methyl-acceptingchemotaxisprotein
TypeIIIrestriction-modificationenzyme
ATPasefamilyprotein
RadicalSAMdomainprotein

Mlr4706protein
Sensorytransductionhistidinekinase
Antibioticresistanceprotein
GMPsynthaseglutamine-hydrolyzing](EC
Nitrogenase(iron-iron)betachain(EC
Nitrogenase(iron-iron)deltachain(EC
Nitrogenase(iron-iron)alphachain(EC
NitrogenregulatoryproteinP-II
NitrogenregulatoryproteinP-II
Nitrogenaseironprotein(EC1.18.6.1)
L-threonine3-O-phosphatedecarboxylase(EC
NAD-dependentproteindeacetylaseofSIR2
SAPDNA-bindingdomain-containingprotein
Transcriptionalregulator,DeoRfamily
Ribose5-phosphateisomeraseB(EC5.3.1.6)
Tributyrinesterase
RiboseABCtransportsystem,permeaseprotein
Oxidoreductase,shortchain
PutativeROK-familytranscriptionalregulator
Nitroreductase
Propionatecatabolismoperonregulatoryprotein
Phosphopantothenoylcysteinesynthetase(EC
ENDO-TYPE6-AMINOHEXANOATEOLIGOMERHYDROLASE
HcptranscriptionalregulatorHcpR(Crp>Fnr
Sensorproteinofzincsigma-54-dependent
Type4fimbriaeexpressionregulatoryprotein
Long-chain-fatty-acid--CoAligase(EC6.2.1.3)
Peptidoglycan-bindingLysM
Signaltransductionhistidinekinase
Integrase
TnpY
TPRrepeatprecursor
CatabolitecontrolproteinA
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,permeaseprotein
RiboseABCtransportsystem,permeaseprotein
RiboseABCtransporter,periplasmic
Mll7147protein
RiboseABCtransportsystem,periplasmic
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,permeaseprotein
Uroporphyrinogen-IIIdecarboxylase
RiboseABCtransportsystem,permeaseprotein
RiboseABCtransporter,periplasmic
RiboseABCtransportsystem,ATP-binding
Putativeexportedproteinprecursor
Sodium-dependentphosphatetransporter
Xylulosekinase(EC2.7.1.17)
HemerythrinlikeproteinMJ0747

Putativelipoprotein
Outermembranelipoproteinomp16precursor
Sensoryboxhistidinekinase
Two-componentresponseregulator
LSUribosomalproteinL32p
Transcriptionalregulator,GntRfamily
Dihydroxy-aciddehydratase(EC4.2.1.9)
Nitroreductase
methyl-acceptingchemotaxisprotein
EnergyconservinghydrogenaseEhb
Hydrolase(HADsuperfamily)
Iron-sulfurclusterregulatorIscR
Cystathioninebeta-lyase(EC4.4.1.8)
Transcriptionalregulator,Cro>CIfamily
Phospholipid-lipopolysaccharideABC
GalactosideO-acetyltransferase(EC2.3.1.18)
Carbamoyl-phosphatesynthaselargechain(EC
Taurinetransportersubstrate-bindingprotein
ABC-typenitrate>sulfonate>bicarbonate
TaurinetransportsystempermeaseproteintauC
TaurinetransportATP-bindingproteintauB
UncharacterizedproteinRv1507c>MT1555
Glycosyltransferase
Beta-1,3-galactosyltransferase>
CDP-ribitol:poly(ribitolphosphate)ribitol
Glycosyltransferasedomainprotein
membraneprotein,putative
Choline-phosphatecytidylyltransferase
CholinepermeaseLicB
Beta-1,3-glucosyltransferase
Acetyltransferase,CYSE>LACA>LPXA>NODLfamily
O-antigenpolymerase
O-antigenpolymerase
Glycosyltransferase(EC2.4.1.-)
Succinoglycanbiosynthesisprotein
UDP-glucose4-epimerase(EC5.1.3.2)
Glycosyltransferase,group1familyprotein
Colanicacidbiosynthesisglycosyltransferase
Putativetelluriumresistanceprotein
Two-componentresponseregulator-likeprotein
serine>threoninekinase
serine>threoninekinase
Leaprotein-soybean
Probabletwo-componentsensor,nearpolyamine
sensoryboxhistidinekinase>response
cellwallsurfaceanchorfamilyprotein
twitchingmotilityproteinPilH
Transcriptionalregulator
ATP-dependenthelicaseHrpB

putativetransposase
Cellsurfaceprotein
InterPro:FibronectintypeIIIdomain
TonB-dependentreceptor
5'-nucleotidase(EC3.1.3.5)
Neopullulanase(EC3.2.1.135)
GlutathioneS-transferasedomainprotein
Cystathioninegamma-lyase(EC4.4.1.1)
ABCtransporter,substrate-bindingprotein
ABCtransporter,substrate-bindingprotein
Catalyzesthecleavageof
TetRfamilytranscriptionalregulatorprobably
GlutamatesynthaseNADPH]smallchain(EC
RegulatoryproteinrecX
Sensoryboxhistidinekinase>response
FOG:CheY-likereceiver
Helix-turn-helixprotein,CopG
Transcriptionalactivatorofmaltoseregulon,
Corrinoidmethyltransferaseprotein
Transposase,mutatorfamily
Sensoryboxhistidinekinase>response
Responseregulator>phosphatase
Na+drivenmultidrugefflux pump
Transcriptionalregulator,MarRfamily
Flagellarhook-lengthcontrolproteinfliK
MembranemucinMUC17
N-acyl-D-amino-acid deacylasefamilyprotein
Transcriptionalregulator,DeoRfamily
Proteininvolvedincatabolismofexternal DNA
DHHfamily>DHHA1domainprotein
Lactoylglutathionelyase(EC4.4.1.5)
Chloridechannelprotein
Mg-protoporphyrinIXmonomethylesteroxidative
OligoendopeptidaseF
CorecomponentFbpofpredictedfolateECF
ATPasecomponentofgeneralenergizingmodule
Transmembranecomponentofgeneralenergizing
bacterialseryl-tRNAsynthetaserelated
NADbindingoxidoreductase
2-hydroxy-6-oxohepta-2,4-dienoatehydrolase
Transcriptionalregulator,TetRfamily
TypeIrestriction-modificationsystem,
TypeIrestriction-modificationsystem,
putativetwo-componentsystemsensorhistidine
Adenosylcobinamide-phosphate
Nitrogenasemolybdenum-cofactorsynthesis
Nitrogenaseironprotein(EC1.18.6.1)
RibonucleasePproteincomponent(EC3.1.26.5)
DNA-bindingdomainofModE

D-tagatose3-epimerase(EC5.3.1.-)
Formiminotetrahydrofolatecyclodeaminase(EC
Lead,cadmium,zincandmercurytransporting
Acetylornithine
aminoacidABCtransporter,ATP-binding
aminoacidABCtransporter,aminoacid-binding
AminoacidABCtransporter,permeaseprotein
AminoacidABCtransporter,permeaseprotein
Methylcobalamin:coenzymeMmethyltransferase,
Beta-phosphoglucomutase(EC5.4.2.6)
Maltosephosphorylase(EC2.4.1.8)>Trehalose
MultiplesugarABCtransporter,
SOS-responsetranscriptionalrepressors
Fe-Soxidoreductase
Regulatorofpolyketidesynthaseexpression
VrlP
SuperfamilyIIDNA>RNAhelicases,SNF2family
ATP-dependentRNAhelicase
VrlK
VrlJ
ATPase
transposase,mutatorfamily
FOG:CheY-likereceiver
Sensoryboxhistidinekinase>response
MutTdomainprotein-like
AbortiveinfectionproteinAbiGII
abortiveinfectionproteinAbiGI
Uroporphyrinogen-IIIdecarboxylase
LysRfamilyregulatoryproteinCidR
Ferricsiderophoretransportsystem,
MolybdenumtransportATP-bindingproteinmodC
ABC-TYPEIRON(III)TRANSPORTSYSTEM
Iron(III)dicitratetransportsystempermease
IronABCtransporter,solute-bindingprotein
Methlytransferase,UbiE>COQ5family
Ni2+-bindingGTPaseinvolvedinregulationof
ABCtransporter,ATP-bindingprotein
Spermidine>putrescine-bindingprotein
Glucitoloperonrepressor
Transketolase,C-terminalsection(EC2.2.1.1)
Maltose>maltodextrinABCtransporter,substrate
MultiplesugarABCtransporter,
Glycerolkinase(EC2.7.1.30)
Serine>threonineproteinkinase
Beta-1,3-glucosyltransferase
Adenylatekinase(EC2.7.4.3)
MultiplesugarABCtransporter,
SN-glycerol-3-phosphatetransportsystem
MultiplesugarABCtransporter,

Ribose5-phosphateisomeraseB(EC5.3.1.6)
Helix-turn-helixmotif
Xylulosekinase(EC2.7.1.17)
MultiplesugarABCtransporter,
N-Acetyl-D-glucosamineABCtransportsystem,
L-fucosemutarotase
RiboseABCtransportsystem,permeaseprotein
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,periplasmic
Glucokinase(EC2.7.1.2)
PossibleD-erythrulose4-phosphate
Galactitol-1-phosphate5-dehydrogenase(EC
Ribulose-phosphate3-epimerase(EC5.1.3.1)
Mo>Fe-nitrogenase-specifictranscriptional
ATPase
Thiolperoxidase,Tpx-type(EC1.11.1.15)
PredictedmolybdateresponsiveregulatorYvgK
HuntingtoninteractingproteinHYPE
Probablemultipleantibioticresistanceprotein
Purinenucleotidesynthesisrepressor
RiboseABCtransportsystem,periplasmic
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,permeaseprotein
DNA-damage-inducibleproteinJ
Metaltransporter,ZIPfamily
L-rhamnosemutarotase
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,permeaseprotein
RiboseABCtransportsystem,permeaseprotein
Transcriptionalregulator,Cro>CIfamily
XyloserepressorXylR(ROKfamily)
RiboseABCtransportsystem,periplasmic
RiboseABCtransportsystem,ATP-binding
Sorbitoldehydrogenase(EC1.1.1.14)
L-xylulose>3-keto-L-gulonatekinase(EC
L-xylulose5-phosphate3-epimerase(EC5.1.3.-)
Ribosomallargesubunitpseudouridinesynthase
RiboseABCtransportsystem,periplasmic
RiboseABCtransportsystem,permeaseprotein
RiboseABCtransportsystem,ATP-binding
Leaderpeptidase(Prepilinpeptidase)(EC
glycosylhydrolase
Hydrolase(HADsuperfamily)
L-2-haloalkanoicaciddehalogenase
Putativehigh-affinityironpermease
Periplasmicproteinp19involvedin
probableintegralmembraneproteinCj1660
SimilartoABCtransporter:egYBJZ_ECOLI
Putativemembraneprotein

PutativeABCtransportsystemATP-binding
Putativepheromoneprecursorlipoprotein
Signaltransductionhistidinekinase
SpermidinePutrescinetransportATP-binding
LSUribosomalproteinL4p(L1e)
SSUribosomalproteinS14p(S29e)
Malatedehydrogenase(EC1.1.1.37)
Cellsurfaceprotein
Flagellarhook-lengthcontrolproteinfliK
ZincABCtransporter,periplasmic-binding
Oxidoreductase,aldo>ketoreductasefamily
Cellsurfaceprotein
transcriptionalregulator,GntRfamily
2-hydroxymuconicsemialdehydehydrolase(EC
4-oxalocrotonatedecarboxylase(EC4.1.1.77)
Acetaldehydedehydrogenase(acetylating)
4-hydroxy-2-oxovaleratealdolase(EC4.1.3.-)
2-polyprenylphenolhydroxylaseandrelated
Catechol2,3-dioxygenase(EC1.13.11.2)
Probableextracellularnuclease
Predictedmembraneprotein
Majorheadprotein(LateproteinGp8)
Putativeanti-terminatorregulatoryprotein
membraneprotein
HDfamilyhydrolase,diverged
Adenylatecyclase(EC4.6.1.1)
gnl
TPRrepeat
UncharacterizedproteinTM_0929
putativeABCtransporter,periplasmic
MultiplesugarABCtransporter,
ProbableABCtransporterpermeaseprotein
Transcriptionalrepressorofthearabinose
Alpha-mannosidase(EC3.2.1.24)
DNA-bindingresponseregulatorDegU
Flavodoxin
HeavymetaltranslocatingP-typeATPase
2-dehydropantoate2-reductase(EC1.1.1.169)
Responseregulator
CytochromeC-typebiogenesisproteinCcs1>ResB
Multiantimicrobialextrusionprotein
Metallo-beta-lactamasesuperfamilyhydrolase
HTHtranscriptionalregulatorTetRfamily
Diaminopimelateepimerasehomolog
Transporter
Histidineammonia-lyase(EC4.3.1.3)
GlycinebetainetransporterOpuD
Histidineammonia-lyase(EC4.3.1.3)
MethylaspartatemutaseSchain(EC5.4.99.1)

METHYLASPARTATEMUTASE(EC5.4.99.1)
PutativeglutamatemutasesubumitE
Methylaspartateammonia-lyase(EC4.3.1.2)
DNAfor3-methylaspartateammonia-lyase,
Transcriptionalregulator,GntRfamily>
TAP1protein
MutatormutTprotein
sensoryboxhistidinekinase>response
HDdomainprotein
Mll1436protein
Single-strandedDNA-bindingprotein
Dihydrolipoamideacetyltransferasecomponent
RiboseABCtransportsystem,permeaseprotein
D-xylosetransportATP-bindingproteinxylG
RiboseABCtransportsystem,periplasmic
Galactose>methylgalactosideABCtransport
SerinephosphataseRsbU,regulatorofsigma
SerinephosphataseRsbU,regulatorofsigma
Flagellarhook-lengthcontrolproteinfliK
Adenylatecyclase(EC4.6.1.1)
ATPsynthasedeltachain(EC3.6.3.14)
containsPfamdomainPF04685:Proteinof
Transcriptionalregulator,MerRfamily
Mll3043protein
Alanineracemase(EC5.1.1.1)
Pirin
Putativecytoplasmicprotein
Threoninesynthase(EC4.2.3.1)
D-ornithineaminomutaseScomponent
Methylaspartatemutase(EC5.4.99.1)
alpha-arabinosidesABCtransportsystem,
alpha-arabinosidesABCtransportsystem,
alpha-arabinosidesABCtransportsystem,
Sucrosephosphorylase(EC2.4.1.7)
sensoryboxhistidinekinase>response
Nucleotidyltransferase(EC2.7.7.-)
domainprotein
Branched-chainaminoacidtransportATP-binding
Branched-chainaminoacidtransportATP-binding
Branched-chainaminoacidtransportsystem
High-affinitybranched-chainaminoacid
Branched-chainaminoacidABCtransporter,
4-hydroxybenzoyl-CoAthioesterasefamilyactive
GDP-L-fucosesynthetase(EC1.1.1.271)
putativebacteriophageprotein
ADAregulatoryprotein
ATPase
Sensoryboxhistidinekinase>response
DNArecombinationandrepairproteinRecF

Ribokinase(EC2.7.1.15)
Sodium-dependentphosphatetransporter
ATPase
Probablepoly(beta-D-mannuronate)O-acetylase
RiboseABCtransportsystem,ATP-binding
TaurinetransportATP-bindingproteintauB
ABCtransporterpermeaseprotein
RNApolymerasesigmafactorRpoE
Transcriptionalregulator
MethionineABCtransporterATP-bindingprotein
MultiplesugarABCtransporter,
MultiplesugarABCtransporter,
Cytotoxictranslationalrepressorof
DNA-bindingprotein
GGDEF
Histoneprotein
Ureacarboxylase
Putativedeoxyribose-specificABCtransporter,
bmpfamilyprotein
Isochorismatasefamilyprotein
Isochorismatasefamilyprotein
Methioninegamma-lyase(EC4.4.1.11)
Phosphatebutyryltransferase(EC2.3.1.19)
Glutathioneperoxidase(EC1.11.1.9)
membraneprotein
ChaperoneproteinDnaK
ChaperoneproteinDnaJ
transcriptionalregulator,MarRfamily
Nitroreductase
FlagellarbiosynthesisproteinflhA
HuntingtininteractingproteinE-likeprotein
Two-componentsensorhistidinekinase
FlaAhomolog-1
DNA-bindingresponseregulator
Flagellarhook-lengthcontrolproteinfliK
Adenosylhomocysteinase(EC3.3.1.1)
Mo>Fe-nitrogenase-specifictranscriptional
Nitrogenaseironprotein(EC1.18.6.1)
2-isopropylmalatesynthase(EC2.3.3.13)
Xaa-Prodipeptidase(EC3.4.13.9)
ABCtransporterATP-bindingprotein-
AminoacidABCtransporter,permeaseprotein
ABCtransporter,substrate-bindingprotein
Cysteinesynthase(EC2.5.1.47)
PROBABLEACETYLTRANSFERASE(EC2.3.1.-)
putativetransferase
PutativeABC-typeamino-acidtransporter
Aminoacid(Glutamine)ABCtransporter
Amino-acidABCtransporterATP-bindingprotein

CRISPR-associatedproteinCas2
ProbableAAAfamilyATPase
Threonyl-tRNAsynthetase(EC6.1.1.3)
putativebacteriophageprotein
Single-strandedDNA-bindingprotein
probablemembraneproteinb2001
putativetransposase
Lactate-responsiveregulatorLldRin
Methioninebiosynthesisandtransportregulator
Transposase,mutatorfamily
Cellsurfaceprotein
serine>threoninekinase
serine>threoninekinase
Leaprotein-soybean
ReplicativeDNAhelicase(EC3.6.1.-)
ProbableAAAfamilyATPase
Putativekinaseprotein
Putativetelluriumresistanceprotein
Probableextracellularnuclease
Lysophospholipase(EC3.1.1.5);Monoglyceride
Aspartyl-tRNAsynthetase(EC6.1.1.12);
PTSsystem,mannitol-specificIICcomponent(EC
Mannitoloperonactivator,BglGfamily
PTSsystem,mannitol-specificIIAcomponent(EC
Putativeoxidoreductase
putativemembraneprotein
LactosetransportsystempermeaseproteinlacF
alpha-arabinosidesABCtransportsystem,
putativetwo-componentsystemsensorkinase
Methyl-acceptingchemotaxisprotein
Sugar>maltosefermentationstimulationprotein
Probableextracellularnuclease
Sensoryboxhistidinekinase>response
Putativepheromoneprecursorlipoprotein
PutativesugarABCtransportsystem,
MSM(multiplesugarmetabolism)operon
InositoltransportATP-bindingprotein
RiboseABCtransportsystem,permeaseprotein
RiboseABCtransportsystem,periplasmic
Pyruvateformate-lyase(EC2.3.1.54)
2-keto-3-deoxy-D-arabino-heptulosonate-7-
Two-componentsystemsensorhistidine
BatB
Lipase
Sensoryboxhistidinekinase>response
RiboseABCtransporter,periplasmic
Sensoryboxhistidinekinase>response
FOG:CheY-likereceiver
bmpfamilyprotein

FOG:CheY-likereceiver
Sensoryboxhistidinekinase>response
Transcriptionregulatorcontains
Lactoylglutathionelyase(EC4.4.1.5)
RNA-bindingprotein
Putativeinnermembraneprotein
Dihydroxyacetonekinase,ATP-dependent(EC
Heavy-metal-associateddomain(N-terminus)and
XyloserepressorXylR(ROKfamily)
RiboseABCtransportsystem,ATP-binding
RiboseABCtransportsystem,permeaseprotein
RiboseABCtransportsystem,periplasmic
Largerepetitiveprotein
RNDmultidrugefflux transporter;Acriflavin
ProbableCo>Zn>Cdeffluxsystemmembranefusion
RecAprotein
RNApolymerasesigmafactorRpoD
Molybdate-bindingdomainofModE
Possiblemembranetransportprotein
Transcriptionalregulator,GntRfamily
PutativegntRfamilyregulatoryprotein
PeroxidestressregulatorPerR,FURfamily
PutativedeoxyribonucleaseYcfH
DipeptidetransportATP-bindingproteindppF
DipeptidetransportATP-bindingproteindppD
Dipeptidetransportsystempermeaseprotein
Dipeptidetransportsystempermeaseprotein
Methyltransferase(EC2.1.1.-)
Ubiquinone>menaquinonebiosynthesis
ZincABCtransporter,innermembranepermease
ZincABCtransporter,ATP-bindingproteinZnuC
CationABCtransporter,periplasmicbinding
Putativehemolysin
TonB-dependentreceptor
***T. primitia* str. ZAS-2**
O-acyltransferase,putative
probablepoly(beta-D-mannuronate)O-acetylase(AlginatebiosynthesisproteinAlgI)
lipopolysaccharidebiosynthesis,putative
NAD-dependentepimerase-dehydratase
dTDP-rhamnosyltransferaseRfbG
glycosyltransferase,group
capsularpolysaccharidebiosynthesisprotein
membraneprotein,putative
membraneprotein,putative
transporter
glycosyltransferase
glycosyltransferasesugar-bindingregioncontainingDXDmotif
O-unitflippase,putative
putativepyruvate-formatelyase

tetratricopeptideTPR_
tetratricopeptiderepeatdomainprotein
lipoprotein,putative
probableextracellularnuclease
methyl-acceptingchemotaxisprotein
anti-anti-sigmafactor
phageshockproteinA,PspA
PotDprotein
spermidine-putrescineimportATP-bindingproteinPotA
PotBprotein
binding-protein-dependenttransportsystemsinnermembranecomponent
putativephosphodiesterase
lipoprotein,putative
lipoprotein,putative
lipoprotein,putative
serine-threonine-proteinkinasePkn
transcriptionalregulator,DeoRfamily
sorbosereductaseSOU
D-xylulosereductase(Xylitoldehydrogenase)(XDH)
dihydropteridinereductase
phosphopantothenoylcysteinesynthetase-decarboxylase
peptidasefamilyT
domainprotein
EF-handcalcium-bindingdomain-containingprotein
signaltransductionhistidinekinase,nitrogenspecific,NtrB,putative
radicalSAMdomainprotein,putative
ABCtransportercomponentA
taurinetransportsystempermeaseproteinTauC
nitratetransportATP-bindingproteinNrtC
multi-sensorhybridhistidinekinase
sulfatasedomainprotein
methyltransferasedomainfamily
ClpX,ATPaseregulatorysubunit,putative
lipopolysaccharidebiosynthesisprotein,putative
glycosyltransferase,group
WfbF
methyltransferaseFkbM
methyltransferasetype
methyltransferasetype
membraneprotein,putative
methyltransferaseFkbMfamily
glyoxalasefamilyprotein
riboseimportATP-bindingproteinRbsA
nitrogenaseironprotein
adenosine
transcriptionalregulator,LacIfamily,putative
ribosetransportsystempermeaseproteinRbsC
riboseimportATP-bindingproteinRbsA
multi-sensorhybridhistidinekinase

putativesignalingprotein
aldo-ketoreductase
VanZfamilyprotein
virulenceassociatedproteinC
bifunctionalacetyltransferase-isomerase(wxcM)
methyltransferasetype
methyltransferasetype
putativeacyltransferase,putative
-deoxy-d-xylulose-
pyrophosphatasePpaX
lipopolysaccharidebiosynthesisprotein
putativeglycosyltransferase
membraneprotein,putative
capsularpolysaccharidebiosynthesisprotein
lipopolysaccharidebiosynthesis,putative
VIpolysaccharidebiosynthesisproteinVipA-tviB
alpha-D-QuiNAcalpha-
UDP-glucose
undecaprenyl-phosphategalactosephosphotransferase
polysaccharidebiosynthesisproteinCapD
ATPase
polysaccharidebiosynthesis-exportprotein
DHHsuperfamilyprotein,subfamily
histidinetriad
glyceraldehyde-
phosphoglyceratekinase(pgk)
Myo-inositolcatabolismproteinIolE
probablemalonicsemialdehydeoxidativedecarboxylase
myo-inositolcatabolismprotein
xyloseisomerasedomainproteinTIMbarrel
D-ribose-bindingperiplasmicprotein
catabolitecontrolprotein
N-carbamyl-L-cysteineamidohydrolase
binding-protein-dependenttransportsystemsinnermembranecomponent
binding-protein-dependenttransportsystemsinnermembranecomponent
taurineimportATP-bindingproteinTauB
extracellularligand-bindingreceptor
inner-membranetranslocator
high-affinitybranchedchainaminoacidABCtransporter,permeaseprotein
highaffinitybranched-chainaminoacidABCtransporter,ATP-bindingprotein
high-affinitybranched-chainaminoacidtransportATP-bindingproteinlivF(LIV-IproteinF)
nitroreductase
branched-chainaminoacidtransportATP-bindingproteinLivG
highaffinitybranched-chainaminoacidABCtransporter,ATP-bindingprotein
inner-membranetranslocator,putative
inner-membranetranslocator,putative
receptorfamilyligandbindingregion
fattyacyl-CoAhydrolase,mediumchain(ThioesteraseB)
ggdefdomainprotein

D-serineammonia-lyase(dsdA)
phosphoenolpyruvate-proteinphosphotransferase(ptsP)
phosphocarrierproteinHPr(Histidine-containingprotein)
lichenan-specificphosphotransferaseenzymeiiacomponent(ptssystemlichenan-specificeiiacomponent)(eiia-lic)(eiii-lic)
lichenan-specificphosphotransferaseenzymeiibcomponent(ptssystemlichenan-specificeiibcomponent)(eiib-lic)
PTSsystem,IIccomponent
beta-glucosidase(Gentiobiase)(Cellobiase)(Beta-D-glucosideglucohydrolase)(Amygdalase)
anaerobicdimethylsulfoxidereductasechaina(dmsoreductase)
putativeXaa-Proaminopeptidase
acyl-coenzymeA:
uroporphyrinogendecarboxylase,putative
oligopeptidetransportATP-bindingproteinAppF
dipeptidetransportATP-bindingproteinDppD
oligopeptidetransportsystempermeaseproteinAppC
binding-protein-dependenttransportsystemsinnermembranecomponent
extracellularsolute-bindingprotein,family
methyl-acceptingchemotaxisprotein
methyl-acceptingchemotaxisprotein
ribonucleasePproteincomponent(rnpA)
NAD-dependentdeacetylase(RegulatoryproteinSIR
sirohydrochlorincobaltochelatase
TonB-dependentreceptor,putative
cobalaminbiosynthesisproteinCbiD(cbiD)
CbiG
precorrin-
threonine-phosphatedecarboxylase(cobD)
cobalt-precorrin-
HmuUprotein
ferricenterobactintransportATP-bindingproteinFepC
cobyricacidsynthaseCobQ(cobQ)
nitrogenfixationproteinNifH-NifE
oxidoreductase-nitrogenase,component
bifunctionaladenosylcobalaminbiosynthesisproteinCobP
AcrBfamilymembranetransportprotein,putative
TonB-dependentreceptorplugdomainprotein
effluxtransporter,RNDfamily,MFPsubunit
outermembraneeffluxprotein
lipolyticenzyme,gdsldomain
listeria-Bacteroidesrepeatdomain(List_Bact_rpt)family
ficfamilyprotein
theglugmotifdomainprotein
heavymetaltransport-detoxificationprotein
trapdicarboxylatetransporter,dctmsubunit
tripartiteATP-independentperiplasmictransporter,DctQcomponent,putative
transcriptionalregulator,GntRfamily
oxidoreductase
virulence-associatedprotein,putative
oligoendopeptidase,PepF-M
MutT-nudixfamilyprotein,putative

ferrousirontransportproteinB
cystathioninegamma-lyase(Gamma-cystathionase)(Probasin-relatedantigen)(PRB-RA)
RRF
glycosyltransferasefamily
OmpAfamilyprotein
transcriptionalRegulator,LacIfamily
sugarkinase,fggyfamily
ironABCtransportersubstrate-bindingprotein
spermidine-putrescineABCtransporterATP-bindingsubunit
inositol-
aspartyl-glutamyl-tRNA(Asn-Gln)amidotransferasesubunitB(Asp-Glu-ADTsubunitB),putative
ABCtransporterpeptide-bindingprotein,putative
glutathionetransportsystempermeaseproteinGsiD
oligopeptidetransportATP-bindingproteinOppD
oligopeptidetransportATP-bindingproteinAppF
radicalSAMdomainprotein
radicalSAMdomainprotein
radicalSAMdomainprotein
anti-anti-sigmafactor
PAS
innermembraneABCtransporterpermeaseproteinYddR
oligopeptidetransportATP-bindingproteinOppD
oligopeptidetransportATP-bindingproteinAppF
dipeptide-bindingprotein
creatininase
aldo-ketoreductase
-methylthioadenosine-S-adenosylhomocysteinedeaminase(MTA-SAHdeaminase),putative
aminoacidABCtransporter,periplasmicaminoacid-bindingprotein,putative
glutaminetransportATP-bindingproteinGlnQ
galactosidetransportsystempermeaseproteinMglC
mateeffluxfamilyprotein
RRF
nitroreductase
outermembraneautotransporterbarreldomain,putative
uroporphyrinogendecarboxylase
RNA-bindingprotein
multiphosphoryltransferprotein
ABCtransportercomponentA
cellsurfaceprotein
glyoxalasedomain-containingprotein
ATPaseassociatedwithvariouscellularactivities,AAA_
smallGTP-bindingprotein
typeIsite-specificdeoxyribonuclease
typeIrestriction-modificationsystemSsubunit
typeIrestrictionmodificationsystemMsubunit
thermostablecarboxypeptidase
D-methioninetransportsystempermeaseproteinMetI
methionineimportATP-bindingproteinMetN
D-methionine-bindinglipoproteinMetQ

cystathioninegamma-synthase(CGS)(O-succinylhomoserine(thiol)-lyase)
lipoprotein,putative
pyridoxamine
zinctransporter,ZIPfamily
sensorytransductionhistidinekinase
GH
trapdicarboxylatetransporter,dctmsubunit
filamentoushaemagglutininfamilyoutermembraneprotein,putative
surfaceantigenBspA,putative
transposase
acetyltransferase,gnatfamily
RNApolymerasebeta-subunit
sugar-phospahtenucleotidyltransferase
DNArepairproteinRadC
AAAATPase
transcriptionalregulator,putative
proteinofunknownfunction
DNApolymeraseIII,alphasubunit
plasmidpartitionprotein,putative
multi-sensorhybridhistidinekinase
oligopeptidetransportATP-bindingproteinAppF
oligopeptidetransportATP-bindingproteinAppD
probablepeptideABCtransporterpermeaseproteiny
glutathionetransportsystempermeaseproteinGsiD
bacterialextracellularsolute-bindingprotein,family
pyrrolidone-carboxylatepeptidase(
probableL-aspartatedehydrogenase
peptidaseM
leucineRichRepeatdomainprotein
glutaminetransportATP-bindingproteinGlnQ
glutamineABCtransporter,permease-substrate-bindingprotein
polaraminoacidABCtransporter,innermembranesubunit
phospho-
amidohydrolasefamilyprotein
domainofunknownfunction,putative
ggdef-eal-pas-pac-domaincontainingprotein
lipoprotein,putative
transposase,IS
transposase,Mutatorfamily
transcriptionalregulator,BadM-Rrf
putativeNADHdehydrogenase-nad(p)hnitroreductase
lipoprotein,putative
ThlR,HTHtranscriptionalregulatorTetR-AcrRfamily
thiolperoxidase
majormembraneimmunogen
sensoryboxhistidinekinase-responseregulator
membraneprotein,GPR
CrcBprotein,putative
outermembraneprotein

hypotheticalcytosolicprotein
crispr-associatedrampprotein,Csm
crispr-associatedprotein,family,putative
plasmidstabilityprotein,putative
TonBfamilyC-domainprotein
proteinofunknownfunction
Igdomainprotein,group
lipoprotein,putative
lipoprotein,putative
methyltransferasetype
cobalaminsynthesisprotein,P
ABCtransporter,ATP-bindingprotein
transcriptionalregulator,MarRfamily
addictionmoduleantidoteprotein,HigAfamily,putative
methioninegamma-lyase(L-methioninase)
transcriptionalregulator
lipoprotein,putative
aminoglycosideadenyltransferase
putativeNADHdehydrogenase-nad(p)hnitroreductase
mateeffluxfamilyprotein
membraneprotein,Bmpfamily
sensorproteinGacS
hybridsensorykinase
adenylatecyclase
sensorproteinGacS
sensoryboxhistidinekinase-responseregulator
adenylate-guanylatecyclasecatalyticdomainprotein
acetyl-CoAcarboxylase,carboxyltransferase,alphasubunit(accA)
acetyl-CoAcarboxylase,carboxyltransferase,betasubunit(accD)
acetyl-CoAcarboxylase,biotincarboxylcarrierprotein
membraneprotein,putative
membraneprotein,putative
aminobenzoyl-glutamateutilizationproteinB
citratelyase,alphasubunit(citF)
citrate(pro-
meioticexpressionup-regulatedprotein
vonWillebrandfactor,typeA,putative
lipoprotein,putative
aminobenzoyl-glutamateutilizationproteinB
methylaspartatemutase,Ssubunit(mamA)
MutLprotein
methylaspartatemutase,Esubunit
methylaspartateammonia-lyase
putativemethyltransferaseCmuC
tripartiteATP-independentperiplasmictransporter,DctQcomponent,putative
APendonuclease,family
SalB
catechol2,3-dioxygenase
binding-protein-dependenttransportsystemsinnermembranecomponent

ABCtransporter,ATP-bindingprotein
para-nitroBenzylesterase(pnbcarboxy-esterase)(intracellularesteraseb)(pnbce)
para-nitroBenzylesterase(pnbcarboxy-esterase)(intracellularesteraseb)(pnbce)
NAD+synthetase
probableextracellularnuclease
transcriptionalregulator
ABC-typeaminoacidtransport-signaltransductionsystem,periplasmiccomponent-domain
HAD-superfamilyhydrolase,subfamilyIIB
proteinUshA
MutLprotein
D-ornithineaminomutaseScomponent
pyridoxal-
dihydrodipicolinatereductase
alanineracemase,N-domainprotein
ribosomallargesubunitpseudouridinesynthaseD(rluD)
transglutaminasedomainprotein
ABC-typemultidrug-protein-lipidtransportsystem,ATPasecomponent
prepilin-typeN-cleavage-methylationdomainprotein
peptidase,A
putativeIgdomainfamily
Coffamilyprotein
probableextracellularnuclease,putative
addictionmoduleantidoteprotein,HigAfamily(higA)
PilTproteindomainprotein,putative
tetratricopeptiderepeatdomainprotein
tetratricopeptiderepeatdomainprotein
HAD-superfamilyhydrolase
macrolideexportATP-binding-permeaseproteinMacB
effluxtransporter,RNDfamily,MFPsubunit
helix-turn-helixdomainprotein
transposase,Mutatorfamily
ISPsy
lipoprotein,putative
bacterialextracellularsolute-bindingprotein,putative
sugarABCtransporterpermease
high-affinityironpermease
kDamembraneantigen(Pathogen-specificmembraneantigen)
membraneprotein,putative
ABCtransporterpermeaseprotein
permeasedomainprotein
macrolideexportATP-binding-permeaseproteinMacB
kDalipoprotein
parallelbeta-helixrepeat
outermembraneautotransporterbarreldomain
spermidine-putrescineABCtransporterATP-bindingsubunit
addictionmoduletoxin,RelE-StbEfamily,putative
ribosomalproteinS
phosphoesterasePHPdomainprotein,putative
ABCtransporterpermeaseprotein

transcriptionalregulator,putative
transcriptionalregulator,AraCfamilyprotein
riboseimportATP-bindingproteinRbsA
ribosetransportsystempermeaseproteinRbsC
probablesugarABCtransporter,substrate-bindingprotein,putative
hypotheticalABCtransporterperiplasmicsolute-bindingprotein
ABCtransporter,permeaseprotein
tobedomainfamily
proteinAdeh_
membraneprotein,putative
vanillate:corrinoidproteinmethyltransferase
transporter,majorfacilitatorfamily
veratrol:corrinoidproteinmetyltransferase,putative
transcriptionalregulator,LysRfamily,putative
cobalaminsynthesisprotein,P
putativeaminoacidtransporter
ferredoxin,putative
mateeffluxfamilyprotein
NADPH-dependentfmnreductase,putative
carboxylesterasefamily
para-nitroBenzylesterase(pnbcarboxy-esterase)(intracellularesteraseb)(pnbce)
dihydroxy-aciddehydratase(DAD)
oxidoreductaseYdhF
sugarABCtransporterpermease
beta-galactosidase
GntRdomainprotein
glycosyltransferasefamily
stageIIsporulationproteinE(SpoIIE)
stageIIsporulationproteinE(SpoIIE)
signaltransductionhistidinekinase
methyl-acceptingchemotaxisprotein
macrolideexportATP-binding-permeaseproteinMacB
ABCtransporter,permeaseprotein
heminimportATP-bindingproteinHmuV
HmuUprotein
trapdicarboxylatetransporter-dctpsubunit,putative
plasmidstabilizationsystemantitoxinprotein
ABCtransporter,ATP-bindingprotein
transcriptionalregulator,Crp-Fnrfamily
sulfitereductase,subunitA(asrA)
sulfitereductase,subunitC(asrC)
responseregulatorreceivermodulatedmetaldependentphosphohydrolase
methyl-acceptingchemotaxisprotein
bacterialextracellularsolute-bindingprotein,putative
transcriptionalregulator,LuxRfamilyprotein
hybridsensorykinase
proteintyrosine-serinephosphatase
lipoprotein,putative
tetratricopeptiderepeatdomainprotein

bacterialextracellularsolute-bindingproteins,family
putativeD-aminoacylase
X-Prodipeptidase
oligopeptidetransportATP-bindingproteinAppF
oligopeptidetransportATP-bindingproteinAppD
ATPase
putativesucrosephosphorylase(Sucroseglucosyltransferase)
hypotheticalcytosolicprotein
proteinofunknownfunction
UvrD-REPhelicase,putative
AAAATPase,putative
RNApolymerasesigmafactorRpoD(Sigma-A)(Sigma-
XylR
signaltransductionhistidinekinase,putative
surfaceantigenBspA
PAS
flagellarfilament
NADPH-dependentfmnreductase,putative
ABCtransporterpeptide-bindingprotein
glutathioneABCtransporter,permeaseprotein(gsiC)
glutathionetransportsystempermeaseproteinGsiD
oligopeptidetransportATP-bindingproteinOppD
oligopeptidetransportATP-bindingproteinAppF
cobalaminsynthesisprotein-P
veratrol:corrinoidproteinmetyltransferase
transcriptionalregulator,MerRfamily
putativeribokinase
hypotheticalABCtransporterperiplasmicsolute-bindingprotein
tobedomainfamily
binding-protein-dependenttransportsystemsinnermembranecomponent
HAD-superfamilyhydrolase,subfamilyIA,variant
hydrolase
PHPdomain,putative
sensoryboxsensorhistidinekinase-responseregulator
sensorproteinGacS
HADsuperfamilyhydrolase
glycosyltransferasegroup
membraneprotein,putative
sugarABCtransporter,putative
transcriptionalregulator,LuxRfamilyprotein
caudovirusproheadprotease,putative
endonucleaseandmethylaseLlaGI
peptidase,S
DNA-directedDNApolymerase
UvrD-REPhelicase,putative
binding-protein-dependenttransportsystemsinnermembranecomponent,putative
transposase,Mutatorfamily
mutt-nudixfamilyprotein
multi-sensorHybridHistidineKinase

multi-sensorhybridhistidinekinase,putative
DNApolymeraseIV(dinB)
MutT-nudixfamilyprotein(mutT)
lipoprotein,putative
phagecapsidfamily,putative
phageportalprotein
phageterminase,largesubunit,pbsxfamily
transcriptionalregulator,AlpAfamily
fibroin
OmpA
theglugmotifdomainprotein
nucleotidyltransferasesubstratebindingproteinlike
oligopeptidetransportATP-bindingproteinAppF
dipeptidetransportATP-bindingproteinDppD
oligopeptidetransportsystempermeaseproteinAppB
lipoprotein,putative
radicalSAMdomainprotein
radicalSAMdomainprotein
lipoprotein,putative
radicalSAMdomainprotein
lipoprotein,putative
transposase,Mutatorfamily
ISSod
transposase,Mutatorfamily
transposase,mutatorfamily
transcriptionalregulator,TetRfamily
iron-sulfurcluster-bindingprotein
SMI
probableNADPH:quinoneoxidoreductase
caenorhabditisproteinofunknownfunction
phagetranscriptionalregulator,AlpA
phage
glycoproteinG
transposase,Mutatorfamily
putativetranscriptionalregulator
proteinofunknownfunction
halomucin
multi-sensorHybridHistidineKinase
transcriptionalregulator,LuxRfamilyprotein
bacterialextracellularsolute-bindingprotein,putative
methylphosphotriester-DNAalkyltransferase
membraneprotein
nlpalipoprotein
D-methioninetransportsystempermeaseproteinMetI
lipoprotein,putative
MutT-nudixfamilyprotein(mutT)
dTDP-
paratosesynthase
glycosyltransferase,group

proteincontainingnucleotide-diphospho-sugartransferasedomain
UDP-glucuronicaciddecarboxylase
transposase,Mutatorfamily
glycosyltransferasegroup
innermembraneprotein
innermembraneprotein
two-componentsystemhybridsensorhistidinekinase-responseregulatorprotein
endoglucanaseM(EGM)(Endo-
oligopeptidetransportATP-bindingproteinAppF
dipeptidetransportATP-bindingproteinDppD
glutathionetransportsystempermeaseproteinGsiD
binding-protein-dependenttransportsystemsinnermembranecomponent
putativedipeptide-bindingabctransporterprotein
putativeABCtransportercomponent
adenylate-guanylatecyclasecatalyticdomainprotein
probableextracellularnuclease,putative
lipoprotein,putative
domainprotein
lipoprotein,putative
prolipoproteindiacylglyceryltransferase
membraneprotein,putative
lipoprotein,putative
YheOdomainprotein
glycosylhydrolase,family
TPRrepeat
glycosyltransferase,family
O-antigenexportsystempermeaseproteinRfbA
teichoicacidsexportATP-bindingproteinTagH(Teichoicacid-transportingATPase)
citratelyaseligaseC-domainprotein
citratelyaseligaseC-domainprotein
citratelyaseligaseC-domainprotein
citratelyaseligaseC-domainprotein
coiled-coildomaincontaining
transcriptionalRegulator,XREfamily
exopolygalacturonatelyase,putative
MDR-typeABCtransporter
ABC-typemultidrug-protein-lipidtransportsystem,ATPasecomponent
ATPasefamilyassociatedwithvariouscellularactivities(AAA)protein
transcriptionalregulator,MarRfamily
RRF
effluxtransporter,RNDfamily,MFPsubunit
acriflavinresistanceprotein,putative
outermembraneeffluxprotein
lipoprotein,putative
excisionase-Xis,DNA-binding
transposase,putative
putatIvetypeirestrictionenzymehindviipmprotein(m.hindviip)
divergentAAAATPase
restrictionmodificationsystemDNAspecificitydomain

putatIvetypeirestrictionenzymehindviiprprotein
HDfamilyhydrolase,diverged
NADPH-dependentfmnreductase
tetratricopeptiderepeatdomainprotein
tetratricopeptiderepeatdomainprotein
fibronectintypeIIIdomainprotein
probableextracellularnuclease,putative
fibronectintypeIIIdomainprotein
regulatoryprotein,ArsR
positiveregulatorofsigmaE,RseC-MucC,putative
mateeffluxfamilyprotein
lipoprotein,putative
transcriptionalregulator
hypotheticalcytosolicprotein
lipoprotein,putative
chondroitinsulfate-heparinutilizationregulationprotein
carboxymethylenebutenolidase
-isopropylmalatesynthase(Alpha-isopropylmalatesynthase)(Alpha-IPMsynthetase)
glutamineamidotransferaseclass-I
transcriptionalregulator,NifAsubfamily,FisFamily
cobalaminsynthesisproteinCobW
transporter,majorfacilitatorfamily
methioninesynthase
MmoS
methioninesynthase
domainprotein
PadR
aldo-ketoreductase
aldo-ketoreductase
lipoprotein,putative
antibioticbiosynthesismonooxygenasedomainprotein
phospholipase-carboxylesterasefamily
methyl-acceptingchemotaxisprotein
phosphoenolpyruvate-proteinphosphotransferase(ptsP)
'-Nucleotidasedomainprotein,putative
phosphoglyceratemutasefamilyprotein
YcsE
transcriptionalantiterminator,BglG
cyclicnucleotide-bindingprotein
ptssystemmannitol-specificeiicbacomponent(eiicba-mtl)(eii-mtl)
HTHdomainfamily
phosphocarrierproteinHPr(Histidine-containingprotein)
ptssystemmannitol-specificeiicbacomponent(eiicba-mtl)(eii-mtl)
Mcp-
binding-protein-dependenttransportsystemsinnermembranecomponent
binding-protein-dependenttransportsystemsinnermembranecomponent
PfkBdomainprotein
uroporphyrinogendecarboxylase(URO-D)superfamily
methyltransferasetype

Xaa-Hisdipeptidase
oligopeptideABCtransportersubstrate-bindingprotein
oligopeptideABCtransportersubstrate-bindingprotein
oligopeptidetransportsystempermeaseproteinAppB
dipeptidetransportATP-bindingproteinDppD
oligopeptidetransportATP-bindingproteinAppF
extracellularsolute-bindingprotein,family
-carboxyhexanoate--CoAligase(bioW)
Hptsensorhybridhistidinekinase
L-cystinetransportsystempermeaseproteinTcyB
probableamino-acidABCtransporter,substrate-bindingprotein
L-cystineimportATP-bindingproteinTcyN
activatorof(R)-
molybdopterin-guaninedinucleotidebiosynthesisproteinA
ironhydrogenase
proteinAegA
phenylacetylCoA
formatedehydrogenase,alphasubunit(fdhA)
proteinT
RecF-RecN-SMCNdomain,putative
ferredoxin
cobalttransportpermease,CbiQfamily,putative
cobaltimportATP-bindingproteinCbiO
flagellarmotorswitchproteinFliM(fliM)
lipoprotein,putative
PINdomaincontainingprotein
PAS-PACsensorhybridhistidinekinase
twocomponenttranscriptionalregulator,LuxRfamily
tetratricopeptideTPR_
lipoprotein,putative
lipoprotein,putative
lipoprotein,putative
probableextracellularnuclease,putative
lipoprotein,putative
multi-sensorhybridhistidinekinase
cyclasefamilyprotein
bacterialpre-peptidaseC-domainfamily
YbaK-ebsCprotein(ybaK)
mureinhydrolaseexporter
LrgBfamilyprotein
rubredoxin,putative
withinP.aerophilum
nickelimportATP-bindingproteinNikE
oligopeptide-dipeptideABCtransporter,ATP-bindingprotein
binding-protein-dependenttransportsystemsinnermembranecomponent
binding-protein-dependenttransportsystemsinnermembranecomponent
peptide-opine-nickelABCtransporter
ferrichrometransportATP-bindingproteinFhuC
transportsystempermeaseprotein

ABC-typeFe
FwdEfamilyprotein
iron(III)dicitratetransportATP-bindingproteinFecE
transportsystempermeaseprotein
periplasmicbindingprotein
ubiquinone-menaquinonebiosynthesismethyltransferaseUbiE
methyltransferasetype
transcriptionalregulator,putative
DNApolymeraseIIIsubunitalpha
glutathioneimportATP-bindingproteinGsiA
hydrogenase
NAD-dependentformatedehydrogenasebetasubunit
peroxiredoxinHyr
permease
RelA-SpoTdomainprotein
chaperoneproteinDnaK(Heatshockprotein
molecularchaperone,DnaJfamily
Zn-fingercontainingprotein
transcriptionalregulator,LysRfamily
aconitasefamily(aconitatehydratase)
citratetransportersuperfamily
bacterialcapsulesynthesisproteinPGA_cap
L-cystineimportATP-bindingproteinTcyN
ABC-typeaminoacidtransportsystem,permeasecomponent
bacterialextracellularsolute-bindingprotein,family
antibioticbiosynthesismonooxygenase
ABC-typeaminoacidtransportsystem,permeasecomponent
ABC-typeaminoacidtransportsystem,permeasecomponent
ExpA
two-componentsystemsensorkinase
peptidase,M
OmpAfamilyprotein
lipoprotein,putative
cobalaminB
transcriptionalregulator,AraCfamilyprotein
transcriptionalregulator,putative
sperm-activatingpeptidesfamily
virulencegenerepressorRsaL
PINdomainprotein
prevent-host-deathfamilyprotein,putative
pirindomainprotein
DNApolymerase,betadomainproteinregion
hydroxylaminereductase(hcp)
transcriptionalregulator
lipoprotein,putative
lipoprotein,putative
hybridsensorykinase
Nif-specificregulatoryprotein(nifA)
HicB

deoxyribonuclease,TatDfamily
heavymetaltranslocatingP-typeATPase
heavymetaltranslocatingP-typeATPase
nitrogenaseironprotein(nifH)
probablepeptideABCtransporterpermeaseproteiny
glutathioneimportATP-bindingproteinGsiA
high-affinitynickel-transporter
carbonmonoxidedehydrogenaseaccessoryproteinCooC
extracellularsolute-bindingprotein,family
putativehelix-turn-helixprotein
AsnCfamilytranscriptionalregulator
transcriptionalregulator,LysRfamily,putative
'-Nucleotidasedomainprotein,putative
transporter,majorfacilitatorfamily
vanillate:corrinoidproteinmethyltransferase
nitroreductase
lipoprotein,putative
probableextracellularnuclease,putative
two-componentresponseregulator
ferricuptakeregulator,Furfamily
nitroreductase
polymorphicoutermembraneprotein,putative
RNApolymerasesigmafactorRpoD(Sigma-
HmuUprotein
heminimportATP-bindingproteinHmuV
periplasmicbindingprotein
TfoXN-domainsuperfamily
carboxymuconolactonedecarboxylase
transposase,mutatorfamily
aspartyl-tRNAsynthetase(Aspartate--tRNAligase)(AspRS)
transportergatedomainprotein
transcriptionalregulator
lipoprotein,putative
oligopeptidetransportATP-bindingproteinOppD
oligopeptidetransportATP-bindingproteinAppF
isochorismatasefamilyprotein
transcriptionalregulator,MerRfamily
carboxymuconolactonedecarboxylase
serine-threonine-proteinkinase
PINdomainprotein
DnaJdomainprotein
transposase,Mutatorfamily
transposaseTpnA
peptidase,S
dnamethylase-typeIrestriction-modificationsystem
ATPaseoftheAAA+class
ATPaseoftheAAA+class
NADH-ubiquinoneoxidoreductasechain
transcriptionalregulator,TetRfamily

Xaa-Proaminopeptidase
thiamineSprotein
cellulose-binding,familyII
polysaccharidebiosynthesis-exportdomainprotein,putative
lipopolysaccharidebiosynthesisprotein,putative
thiogalactosideacetyltransferase
proteincontainingnucleotide-diphospho-sugartransferasedomain
putativesugartransferase
UDP-glucose
epimerase-dehydratase,putative
putativeLPSbiosynthesisrelatedglycosyltransferase
membraneprotein,putative
filamentationinducedbycAMPproteinFic
protein,withAweakD-galactaratedehydratase-altronatehydrolasedomain
YwbO,putative
indigoidinesynthaseAfamilyprotein
kinase,PfkBfamily
domainprotein
peptidase,M
hemerythrinfamilyprotein
mateeffluxfamilyprotein,putative
phageinfectionprotein
metallo-beta-lactamasesuperfamilyhydrolase
transcriptionalregulator,TetRfamily
flavodoxin
lipoprotein,putative
PINdomainprotein
lipoprotein,putative
amidohydrolase
membraneprotein,putative
membraneprotein,putative
TPRrepeat
TPRrepeat
TPRrepeat
TPRrepeat
flagellarbiosynthesisproteinFlhA
exonuclease,putative
Cl-channel,voltage-gatedfamilyprotein
radicalSAMdomainprotein
O-methyltransferase
acyltransferase
peptidaseS
***T. azotonutricium* str. ZAS-9**
responseregulator
biotincarboxylase
ribosetransportsystempermeaseproteinRbsC
riboseimportATP-bindingproteinRbsA
L-rhamnose
ferricuptakeregulator,Furfamily

activatorof
PilTproteindomainprotein
prevent-host-deathfamilyprotein
isochorismatasehydrolase
Smfprotein
domainprotein
transcriptionalregulator,LuxRfamilyprotein
domainprotein
transcriptionalregulator,LuxRfamilyprotein
sugardehydrogenase,putative
APsuperfamily
legionellavirregionprotein
transposaseIS
outermembraneautotransporterbarrel,putative
probableextracellularnuclease,putative
majormembraneimmunogen
Zn-dependentalcoholdehydrogenase
endo-arabinase
helix-turn-helix-domaincontainingprotein,AraCtype
extracellularsolute-bindingprotein,family
IstBdomainproteinATP-bindingprotein
IstBdomainproteinATP-bindingprotein
transcriptionalregulator,LacIfamily
L-fucoseisomerasedomainprotein
probablefructose-bisphosphatealdolase
inositol
probablemalonicsemialdehydeoxidativedecarboxylase
inososedehydratase(
adenylate-guanylatecyclase
transcriptionalregulator
probableinsertionsequencetransposaseprotein,IS
transposaseIS
transposase,Mutatorfamily
tetratricopeptiderepeatdomainprotein
glucose-
dsbaoxidoreductase,putative
tetratricopeptiderepeatdomainprotein
pirindomainprotein
integrasedomainprotein
proteinofunknownfunctionDUF
DNA-directedDNApolymerase,putative
DNApolymeraseIII,alphasubunit
mateeffluxfamilyprotein
APendonuclease,family
sporulationinitiationinhibitorproteinsoj
transposaseIS
TetR-familyregulator
adenylate-guanylatecyclasecatalyticdomainprotein
transcriptionalregulator,ArsRfamily

arabinoseoperonproteinAraM
cyclicnucleotide-bindingprotein
TrksystempotassiumuptakeproteinTrkH
transposase,IS
PINdomainprotein
transposase,Mutatorfamily
,putative
TonB-dependentreceptor
cationABCtransporter,periplasmicbindingprotein,putative
zincimportATP-bindingproteinZnuC
methyl-acceptingchemotaxisprotein
cationicoutermembraneprotein
transposase,is
integrasedomainprotein
glutamine-bindingperiplasmicprotein-glutaminetransportsystempermeaseprotein
sensorproteinGacS
ATP-dependenthelicaseHrpB(hrpB)
integrasedomainprotein
phagecapsidfamily
domainprotein
CTPpyrophosphohydrolase
PemKfamilyprotein
transposaseInsIforinsertionsequenceelementIS
glycosyltransferase
glycosyltransferase,group
CDP-Glycerol:Poly(glycerophosphate)glycerophosphotransferasefamily
innermembraneprotein
glycosyltransferase,group
CDP-alcoholphosphatidyltransferasefamily
nitroreductase
NADPH-dependentfmnreductase
NAD-dependentepimerase-dehydratase
adenylatekinase
transposase,is
tetratricopeptiderepeatdomainprotein
transposaseIS
ATPaseassociatedwithvariouscellularactivities,AAA_
swimzincfingerdomainprotein
ankyrin
tetratricopeptiderepeatdomainprotein
transcriptionalregulator,TetRfamily,putative
PAS
ribosetransportsystempermeaseproteinRbsC
riboseimportATP-bindingproteinRbsA
mateeffluxfamilyprotein
radicalSAMdomainprotein
sensorproteinGacS
responseregulatorreceiver:Metal-dependentphosphohydrolase,HDsubdomain
transcriptionalregulator,LuxRfamilyprotein

sensorproteinGacS
outermembraneprotein
PAS
lipoate-proteinligaseA
methyl-acceptingchemotaxisprotein(MCP)signalingdomain
responseregulatorreceiver:Metal-dependentphosphohydrolase,HDsubdomain
bacterialextracellularsolute-bindingprotein,putative
sugartransportsystem
sugarABCtransporter,permeaseprotein
sugarphosphateisomerases-epimerases
xyloseisomerasedomainproteinTIMbarrel,putative
dehydrogenase
binding-protein-dependenttransportsystemsinnermembranecomponent
L-arabinoseABCtransportpermeaseprotein
hexulose-
transcriptionalrepressoroftheriboseoperon
glyoxalase-bleomycinresistanceprotein-dioxygenase
alpha-galactosidase(Melibiase)
glucose--fructoseoxidoreductase(gfor),putative
ribosetransportsystempermeaseproteinRbsC
riboseimportATP-bindingproteinRbsA
haloaciddehalogenasedomainproteinhydrolase
PfkBdomainprotein
prolylaminopeptidase
L-fucoseisomeraserelatedprotein
chitinase
proteinofunknownfunction
typeIIIrestrictionenzyme,ressubunit
phageportalprotein,HK
site-specificrecombinase,phageintegrasefamilyprotein
aldo-ketoreductase
oxidoreductase,aldo-ketoreductasefamily
integrasedomainprotein
majorfacilitatorfamilytransporter,putative
transcriptionalregulator,TetRfamilyprotein
fkbp-type
aryl-alcoholdehydrogenase
mateeffluxfamilyprotein
aspartyl-tRNAsynthetase(aspS)
metallophosphoesterase
glycosyltransferases,putative
glycosyltransferase,group
fibronectintypeIIIdomainprotein
fibronectintypeIIIdomainprotein
TPRrepeat,putative
lipolyticproteinG-D-S-Lfamily
mucin-desulfatingsulfatase,putative
membraneprotein,putative
glucanendo-

AlgI
glycosyltransferase,involvedincellwallbiogenesis
CotHprotein
fibronectintypeIIIdomainprotein
sensorytransductionhistidinekinase
heminimportATP-bindingproteinHmuV
transportsystempermeaseprotein
HmuUprotein
periplasmicbindingprotein
LcoC
-hydroxythreonine-
extracellularsolute-bindingprotein,family
zincfingerprotein
acetyl-CoAcarboxylase,biotincarboxylcarrierprotein(accB)
helix-turn-helixdomainprotein
cobalttransportprotein
SfsA
tetratricopeptideTPR_
cystathioninegamma-synthase(CGS)(O-succinylhomoserine(thiol)-lyase)
transcriptionalregulator,BadM-Rrf
extracellularligand-bindingreceptor
HmgE
high-affinitybranchedchainaminoacidABCtransporter,permeaseprotein
highaffinitybranched-chainaminoacidABCtransporter,ATP-bindingprotein
abctransporter,hydrophobicaminoaciduptaketransporter(haat)family,ATP-bindingprotein
polaraminoacidABCtransporter,innermembranesubunit
polaraminoacidABCtransporter,innermembranesubunit
glutaminetransportATP-bindingproteinGlnQ
helicase,Snf
aspartyl-glutamyl-tRNA(Asn-Gln)amidotransferasesubunitB(Asp-Glu-ADTsubunitB),putative
zinctransporter,ZIPfamily
PAS
ROKdomaincontainingprotein
ClcA
voltage-gatedchloridechannel
radicalSAMdomainprotein
oxidoreductase
ferritin
OmpAfamilyprotein,putative
HmuUprotein
TonBfamilyprotein
proteinofunknownfunction
methyltransferase
cobalaminsynthesisprotein,P
ABCtransporter,ATP-bindingprotein
methylcobamide:CoMmethyltransferaseMtbA(Methylcobamide:CoMmethyltransferaseIIisozymeA)(MT
nitrogenaseironprotein(nifH)
transposase,Mutatorfamily
acetyltransferase,gnatfamily

RNApolymerasesigma-
surfaceantigenBspA
novelprotein
oxaloacetatedecarboxylase,gammasubunit
Rrf
cytoplasmicfilamentproteinA
ferricuptakeregulatoryprotein
transcriptionalregulator,AraCfamilyprotein
ADP-ribosylglycohydrolasesuperfamily
bacterialextracellularsolute-bindingprotein
binding-protein-dependenttransportsystemsinnermembranecomponent
binding-protein-dependenttransportsystemsinnermembranecomponent
regulatoryproteinLacI
glutamatesynthaseNADPH]smallchain(nadph-gogat)
methioninegamma-lyase(L-methioninase)
mureinhydrolaseexportregulator
LrgAfamilyprotein,putative
responseregulatorreceiver:Metal-dependentphosphohydrolase,HDsubdomain
Hptsensorhybridhistidinekinase
APendonuclease,family
bacterialextracellularsolute-bindingprotein,putative
transcriptionalregulator
membraneprotein,putative
twocomponenttranscriptionalregulator,LuxRfamily
nitroreductase
transcriptionalregulator,MarRfamily,putative
peroxiredoxinHyr
TPRDomaincontainingprotein
translocase
beta-lactamase
transcriptionalregulator,Cro-CIfamily,putative
transcriptionalregulator
NifK
oxidoreductase-nitrogenase,component
nitrogenaseironprotein
aliphaticsulfonatesimportATP-bindingproteinSsuB
radicalSAMdomainprotein
phosphoadenosinephosphosulfatereductase
GH
TPRrepeatprotein
trapdicarboxylatetransporter,dctmsubunit
cellwallsurfaceanchorfamilyprotein,putative
oxidoreductasedomainprotein
glucose-
transcriptionalregulator
L-ribulokinase(araB)
periplasmicbindingprotein-LacItranscriptionalregulator,putative
phosphateacetyltransferase(Phosphotransacetylase)
RNApolymerasesigmafactorRpoD(rpoD)

isochorismatasehydrolase
ferrousirontransportproteinA
signaltransductionhistidine-proteinkinaseBarA
extracellularsolute-bindingprotein,family
L-arabinosetransportsystempermeaseproteinAraP
antibioticbiosynthesismonooxygenasedomainprotein
PAS
two-componentsystemsensory-regulatoryprotein
mateefflux familyprotein,putative
transcriptionalregulator,TetRfamily
uroporphyrinogendecarboxylase(URO-D)superfamily
transcriptionalregulator
transcriptionalregulator,ArsRfamily
DNA-bindingprotein
phagemajorcapsidprotein,HK
transcriptionalregulator
transposaseandinactivatedderivatives
membraneprotein,putative
integrasedomainprotein
macrolideexportATP-binding-permeaseproteinMacB
ABCtransporter,permeaseprotein
-hexulose-
dihydropteridinereductase
fggyfamilyofcarbohydratekiNases,domainprotein
proteinofunknownfunction
deathoncuringprotein,putative
HDdomainprotein
metaldependentphosphohydrolase
D-mannonateoxidoreductase(Fructuronatereductase)
extracellularsolute-bindingprotein,family
binding-protein-dependenttransportsystemsinnermembranecomponent
oligopeptidetransportATP-bindingproteinOppD
riboseimportATP-bindingproteinRbsA
periplasmicsugar-bindingproteins
integrasedomainprotein
periplasmicbindingproteinsandsugarbindingdomainoftheLacIfamily,putative
extracellularsolute-bindingproteinfamily
sugarABCtransporterpermease
sugarABCtransporterpermease
methyl-acceptingchemotaxisprotein
multi-sensorhybridhistidinekinase
sensorytransductionhistidinekinase
-oxoacyl-acyl-carrier-protein]reductase(
integrasedomainprotein
ribose-
sugarABCtransporter,permeaseprotein
extracellularsolute-bindingprotein,family
alpha-L-rhamnosidase
regulatoryproteinLacI

ROKfamilyprotein
extracellularsolute-bindingproteinfamily
binding-protein-dependenttransportsystemsinnermembranecomponent
binding-protein-dependenttransportsystemsinnermembranecomponent
OmpAfamilyprotein
polymorphicoutermembraneprotein
bacterialextracellularsolute-bindingprotein,putative
ABCtransporter,carbohydrateuptaketransporter-
binding-protein-dependenttransportsystemsinnermembranecomponent
innermembraneABCtransporterpermeaseproteinYcjP
sensoryboxhistidinekinase-responseregulator
ribonucleasePproteincomponent(rnpA)
PAS
ATPase
bacterialpre-peptidaseC-domainfamily
ATPase,AAAfamily,putative
ABCsugartransporter,periplasmicligandbindingprotein
PAS-PACsensorhybridhistidinekinase
peptidase,M
aquaporin
diguanylatecyclase
sensorproteinGacS
membraneproteincontaining
putativesignalingprotein
beta-lactamase
integralmembraneprotein
HipAprotein,putative
xyloseisomerasedomainproteinTIMbarrel
rhomboidfamilyprotein
transcriptionalregulator,GntRfamily
tripartiteATP-independentperiplasmictransporter,DctQcomponent,putative
trapdicarboxylatetransporter,dctmsubunit
methyl-acceptingchemotaxisprotein
polymorphicoutermembraneprotein
nitrogenfixationproteinNifH-NifE
oxidoreductase-nitrogenase,component
bifunctionaladenosylcobalaminbiosynthesisproteinCobP
methyltransferaseMtaA-CmuA
methyltransferaseMtaA-CmuA
transposase,Mutatorfamily
putativeCelldivisionproteaseFtsHhomolog
myo-inositolcatabolismprotein
galactosidetransportsystempermeaseproteinMglC
two-componenthybridprotein
HAD-superfamilyhydrolase,subfamilyIA,variant
multiplesugar-bindingtransportsystemmultiplesugar-bindingprotein,putative
ribosetransportsystempermeaseproteinRbsC
xyloseimportATP-bindingproteinXylG
uroporphyrinogendecarboxylase(URO-D)superfamily

ATPase
helix-turn-helixdomainprotein
lactosetransportsystem
leucineRichRepeatdomainprotein
binding-protein-dependenttransportsystemsinnermembranecomponent
multi-sensorhybridhistidinekinase
PINdomainprotein
ATPase
integrasedomainprotein
domainprotein
phosphoglycolatephosphatase(PGPase)(PGP),putative
transcriptionalregulator,GntRfamily
Zn-dependentalcoholdehydrogenase
transposase,Mutatorfamily
transposase,Mutatorfamily
transposase,Mutatorfamily
transposaseforinsertionsequenceelementisrm
transposaseIS
transposaseIS
peptidaseM
transposase,IS
PINdomainprotein
multi-sensorhybridhistidinekinase
putativetranscriptionalregulator
multi-sensorhybridhistidinekinase
spfh-band
transposaseforinsertionsequenceelementisrm
transposase,Mutatorfamily
domainprotein
transcriptionalregulator,XREfamily
domainprotein
phageterminase,smallsubunit,putative
glutamineABCtransportersubstrate-bindingprotein
proteasePrsW(Proteaseresponsibleforactivatingsigma-W)
extracellularsolute-bindingproteinfamily
binding-protein-dependenttransportsystemsinnermembranecomponent
responseregulatorreceiverprotein
oxidoreductasedomainprotein
glucose-inhibiteddivisionproteinA
radicalSAMdomainprotein
periplasmicbindingprotein-LacItranscriptionalregulator,putative
insertionsequenceputativeATP-bindingprotein(ORF
binding-protein-dependenttransportsystemsinnermembranecomponent
responseregulatorreceiverprotein
xyloseisomerasedomainproteinTIMbarrel
uroporphyrinogendecarboxylase
PINdomain,putative
two-componentresponseregulatorYesN
bacterialextracellularsolute-bindingprotein

methyl-acceptingchemotaxisproteinDmcB
transcriptionalregulator,MarRfamily
transposase,Mutatorfamily
putativeDnaJhomologsubfamilyBmember
metallo-beta-lactamasesuperfamilyprotein
glycosidehydrolase,family
YbaC
transcriptionalregulator,TetRfamilyprotein
innermembraneprotein
innermembraneprotein
oxidoreductasedomainprotein,putative
sorbitoldehydrogenase(L-iditol
isochorismatasehydrolase
proteinofunknownfunction
extracellularsolute-bindingproteinfamily
binding-protein-dependenttransportsystemsinnermembranecomponent
binding-protein-dependenttransportsystemsinnermembranecomponent
ribosomalproteinS
transposase,Mutatorfamily
flavodoxin
proteincontainingandDnaJdomain,putative
phageshockproteinA,PspA
anti-anti-sigmafactor
chainN,ArchitectureOfMammalianFattyAcidSynthase
Igdomainprotein,group
repressor,putative
proteinrelatedtomifh-dopdproteinfamily,functioninbacteriaisunknown,putative
alpha-D-mannosidase
transcriptionalregulator,putative
sigma-
signaltransductionhistidinekinase,nitrogenspecific,NtrB
YkrA
radicalSAMdomainprotein,putative
transcriptionalregulator,DeoRfamily,putative
crispr-associatedproteinCas
crispr-associatedrampprotein,Cmr
hydrolaseoftheHDsuperfamily
crispr-associatedrampprotein,Cmr
crispr-associatedprotein,Cmr
crispr-associatedrampprotein,Cmr
crispr-associatedprotein,family,putative
transposaseIS
kDalipoprotein
permeasedomainprotein
membraneprotein,putative
kDamembraneantigen(Pathogen-specificmembraneantigen)
high-affinityironpermease
thymidylatesynthase(thyA)
ABCtransporter,ATP-bindingprotein,MsbAfamily

amidohydrolase
transcriptionalregulator,AraCfamilyprotein
ribulose-phosphate
sorbitoldehydrogenase(L-iditol
putativeD-erythrulose-
ROKdomaincontainingprotein
putativeoxidoreductase
YurM
sugarABCtransporterpermease
bacterialextracellularsolute-bindingprotein,putative
L-threonine
sensoryboxhistidinekinase-responseregulator
integrasedomainprotein
fumaratereductase-succinatedehydrogenaseflavoproteindomainprotein
FMN-bindingdomainprotein
binding-protein-dependenttransportsystemsinnermembranecomponent
nitratetransportATP-bindingproteinNrtD
dgqhrdomain,putative
proteinofunknownfunction
transcriptionalregulator,NifAsubfamily,FisFamily
cobaltimportATP-bindingproteinCbiO
transcriptionalregulator,TrmB
sugartransportersugarbindingprotein
ABC-typesugartransportsystems,permeasecomponents
sugarABCtransporterpermease
glutaminase
multi-sensorhybridhistidinekinase
proteinR
methyl-acceptingchemotaxisprotein
spermidine-putrescineimportATP-bindingproteinPotA
SoxRprotein
phospho-
deoxyribose-phosphatealdolase(deoC)
VanZlikeprotein
domainofunknownfunction
dTDP-
CDP-abequosesynthase
WbyH
putativeparatosetransferase
glycosyltransferase,putative
EpsT
glycosyltransferase,group
membraneprotein,putative
putativerhamnosyltransferase
EpsD
GDP-mannose
responseregulator
transcriptionalregulator,TrmB
alpha-N-arabinofuranosidase(Arabinosidase)

periplasmicbindingprotein
binding-protein-dependenttransportsystemsinnermembranecomponent
binding-protein-dependenttransportsysteminnermembranecomponent
oligopeptidetransportATP-bindingproteinAppF
anti-anti-sigmafactor
hemolysin-
peptidaseM
oxidoreductasedomainprotein
transcriptionalregulator,TetRfamilyprotein
long-chain-fatty-acid--CoAligase,putative
HipAN-domain,putative
domainprotein
NAD-dependentepimerase-dehydratase
cellsurfaceprotein
O-acetylhomoserine(thiol)-lyase(O-acetylhomoserinesulfhydrylase)(OAHsulfhydrylase)(Homocysteinesynthase)
cystathioninegamma-synthase(CGS)(O-succinylhomoserine(thiol)-lyase)
D-methionine-bindinglipoproteinMetQ
methionineimportATP-bindingproteinMetN
D-methioninetransportsystempermeaseproteinMetI
hydrogenase-
-isopropylmalatesynthase(Alpha-isopropylmalatesynthase)(Alpha-IPMsynthetase)
domainprotein
periplasmicmolybdate-bindingprotein-domain
oxidoreductasedomainprotein,putative
transcriptionalregulator,AraCfamily,putative
probableOxidoreductase,putative
sulfatase,putative
APendonuclease,family
Hptsensorhybridhistidinekinase
responseregulator,NarL-family
membraneprotein,putative
bacterialextracellularsolute-bindingprotein,putative
transcriptionalregulator,LuxRfamilyprotein
transcriptionalregulator,LuxRfamily
listeria-Bacteroidesrepeatdomain(List_Bact_rpt)family
Igdomainproteingroup
outermembraneautotransporterbarreldomain,putative
outermembraneautotransporterbarrel,putative
cellsurfaceprotein,putative
cellsurfaceprotein,putative
elongationfactorTu
ribosomalproteinL
lactoylglutathionelyase
sensorproteinGacS
transposaseInsIforinsertionsequenceelementIS
ABCtransporter,permeaseprotein
ABCtransporter,ATP-bindingprotein
L-serinedehydratase,iron-sulfur-dependent,betasubunit(sdaAB)
PAS

membraneprotein,putative
mateeffluxfamilyprotein
phosphonatesimportATP-bindingproteinPhnC
ROK,putative
cobalaminB
putativetranscriptionalregulator
methylphosphotriester-DNAalkyltransferase
methylated-dna--protein-cysteinemethyltransferase
sorbosereductaseSOU
mannosyltransferase
polymorphicoutermembraneprotein,putative
transcriptionalregulator,AraCfamily,putative
oligopeptidetransportATP-bindingproteinAppD
oligopeptidetransportsystempermeaseproteinAppC
oligopeptidetransportATP-bindingproteinAppF
alpha-galactosidase(Melibiase)
trimethylaminecorrinoidprotein
oxidoreductase
proteinofunknownfunction
two-componentsensorhistidinekinasewithresponseregulatorreceiverdomain
D-methionineABCtransporter,periplasmicD-methionine-bindingprotein(metQ)
purinenucleosidephosphorylase(deoD)
peptidaseU
sensorproteinGacS
responseregulatorreceiversensorhybridhistidinekinase
PASfoldfamily
multi-sensorhybridhistidinekinase
YbaK-ebsCprotein(ybaK)
transporter,majorfacilitatorfamily
nachtnucleosidetriphosphatase
tetratricopeptiderepeatdomainprotein
glutathioneperoxidase
bacterialextracellularsolute-bindingprotein,putative
binding-protein-dependenttransportsystemsinnermembranecomponent,putative
sugarABCtransporterpermease
penicillin-bindingprotein
insertionsequenceputativeATP-bindingprotein(ORF
ribose
dihydroxyacetonekinase,Lsubunit
PTS-dependentdihydroxyacetonekinase,dihydroxyacetone-bindingsubunitDhaK
glucose
Zn-dependentalcoholdehydrogenase
invasionproteinIbeA
invasionproteinIbeA
bacterialextracellularsolute-bindingprotein,putative
ABCtransporter,permeaseprotein
thiaminepyrophosphateenzyme,centraldomainfamily
iron-containingalcoholdehydrogenase,putative
methyl-acceptingchemotaxisprotein

tetratricopeptiderepeatdomainprotein
domainprotein
PilTproteindomainprotein,putative

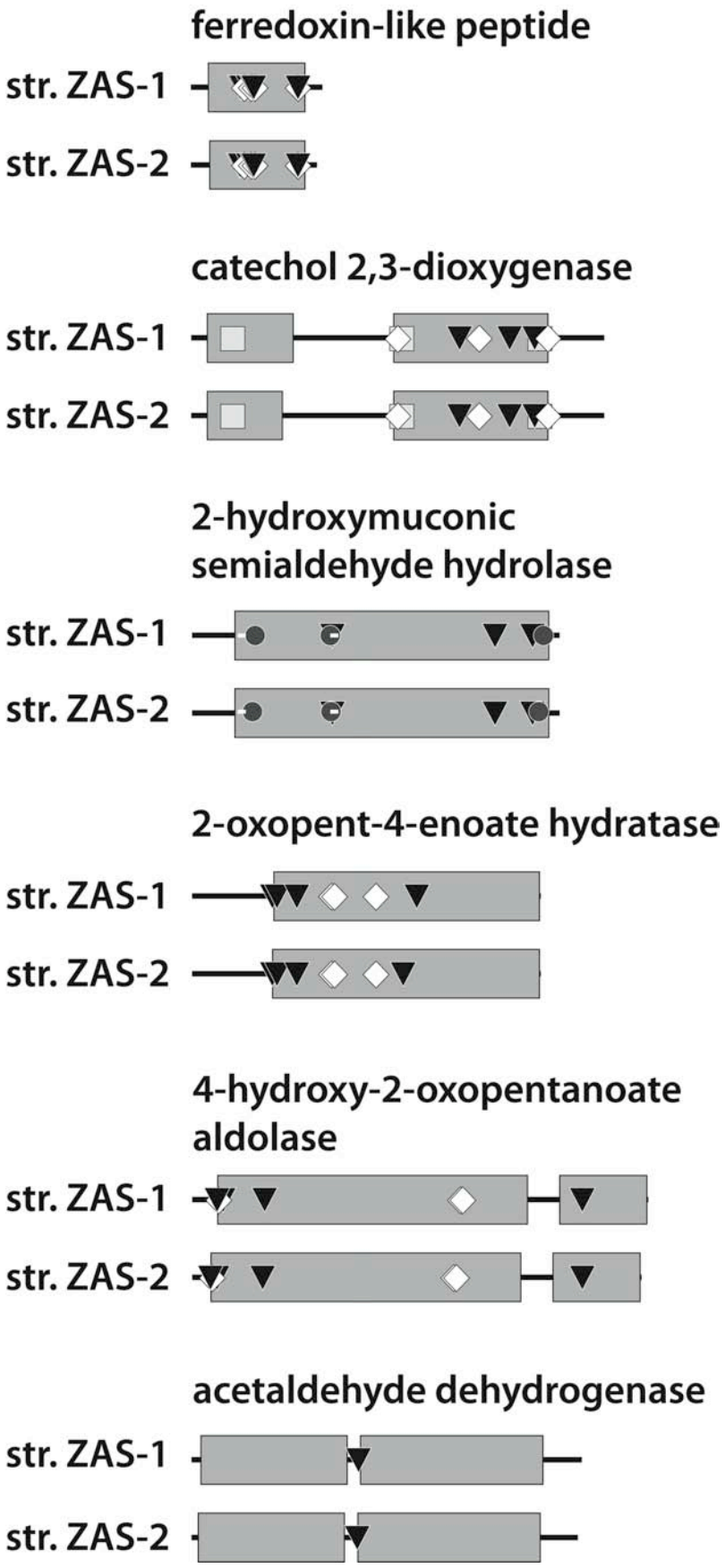

ferredoxin-like peptide
str. ZAS-1
str. ZAS-2
catechol 2,3-dioxygenase
str. ZAS-1
str. ZAS-2
2-hydroxymuconic
semialdehyde hydrolase
str. ZAS-1
str. ZAS-2
2-oxopent-4-enoate hydratase
str. ZAS-1
str. ZAS-2
4-hydroxy-2-oxopentanoate
aldolase
str. ZAS-1
str. ZAS-2
acetaldehyde dehydrogenase
str. ZAS-1
str. ZAS-2

Fig. 3A-1: PFAM domains and conserved functional residues and sequence motifs in *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-cleavage pathway proteins. The primary structure span of *meta*-cleavage pathway proteins is represented by black lines, and PFAM domains within the *meta*-cleavage pathway proteins are represented by grey rectangles. Symbols representing conserved functional residues are centered over the location of the functional residues. Active site residues are represented by black inverted triangles, metal-binding residues are represented by white diamonds, residues contributing to protein structure are represented by light grey squares, and substrate-binding residues are represented by grey circles. Conserved sequence motifs are represented by white bars.

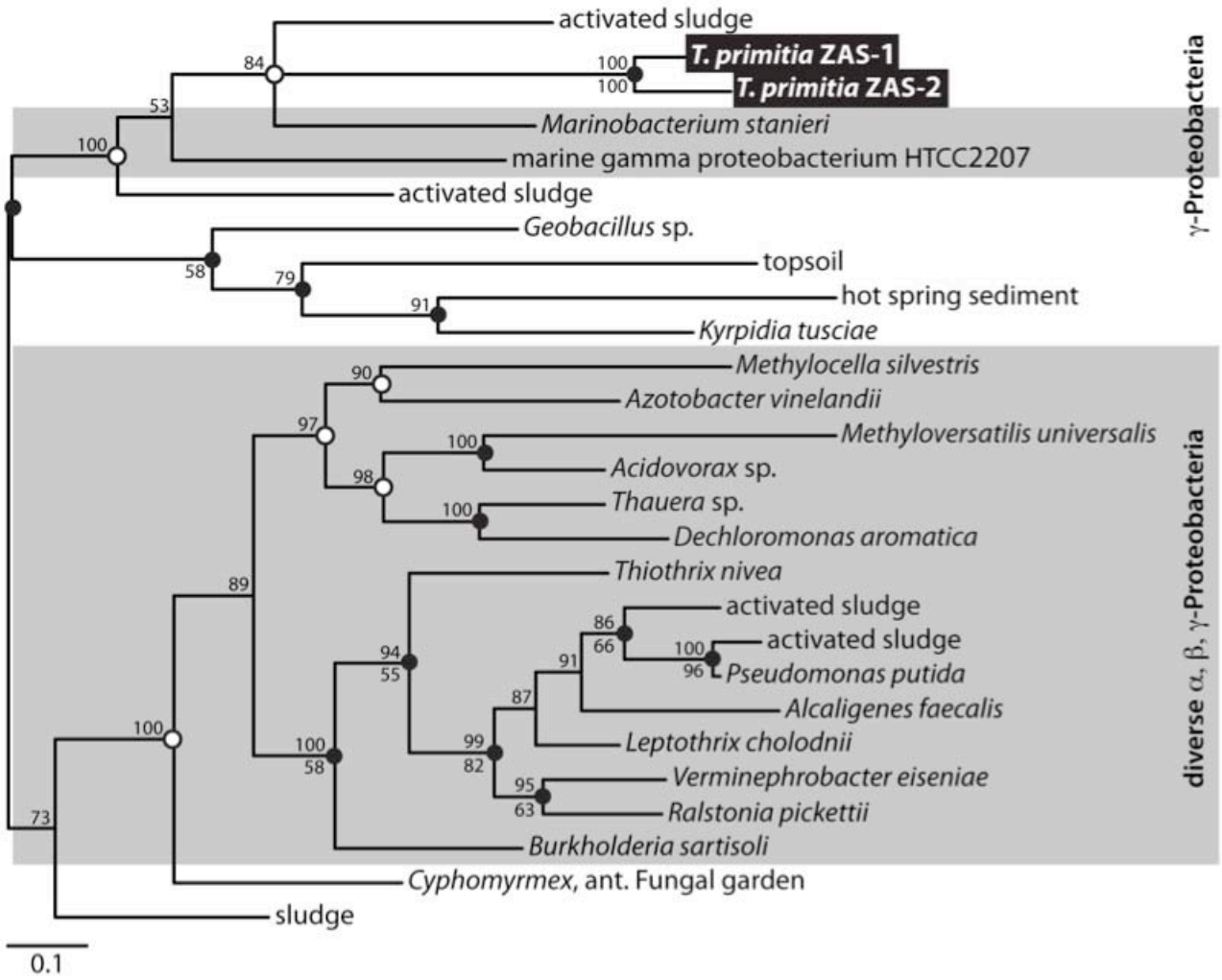


Fig. 3A-2: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 ferredoxin-like peptide (PF00111, Step 1'). Bayesian protein phylogenetic analysis (183 trees from 73,000 generations; PSRF = 1.001; average standard deviation of split frequencies = 0.011280) is based on 85 unambiguously aligned amino acid positions of a 98 amino acid-long protein. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

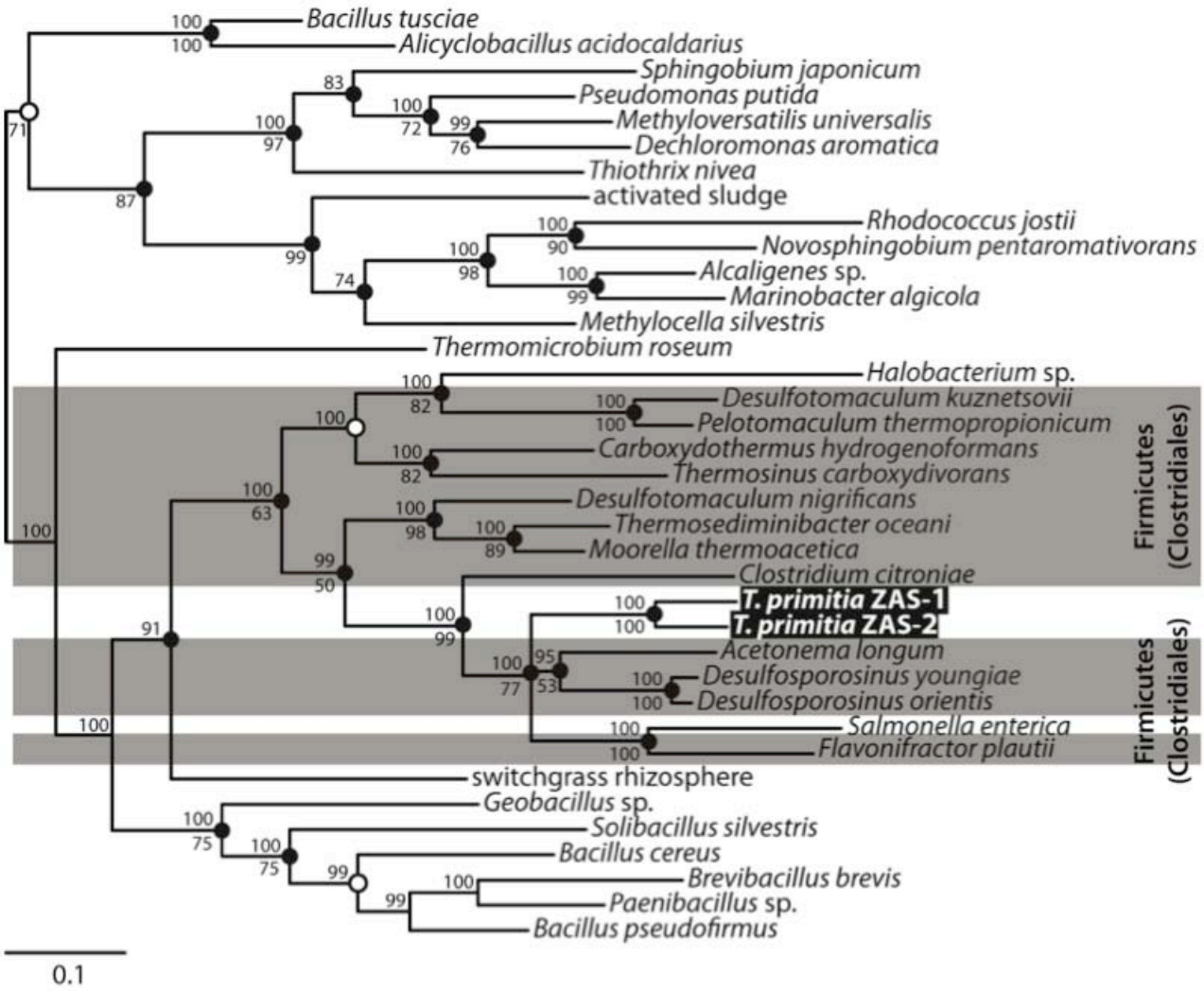


Fig. 3A-3: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 4-hydroxy-2-oxopentanoate aldolase (PF00682 and PF07836, Step 4). Bayesian protein phylogenetic analysis (375 trees from 150,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.002661) is based on 331 unambiguously aligned amino acid positions of a 340 amino acid-long protein. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

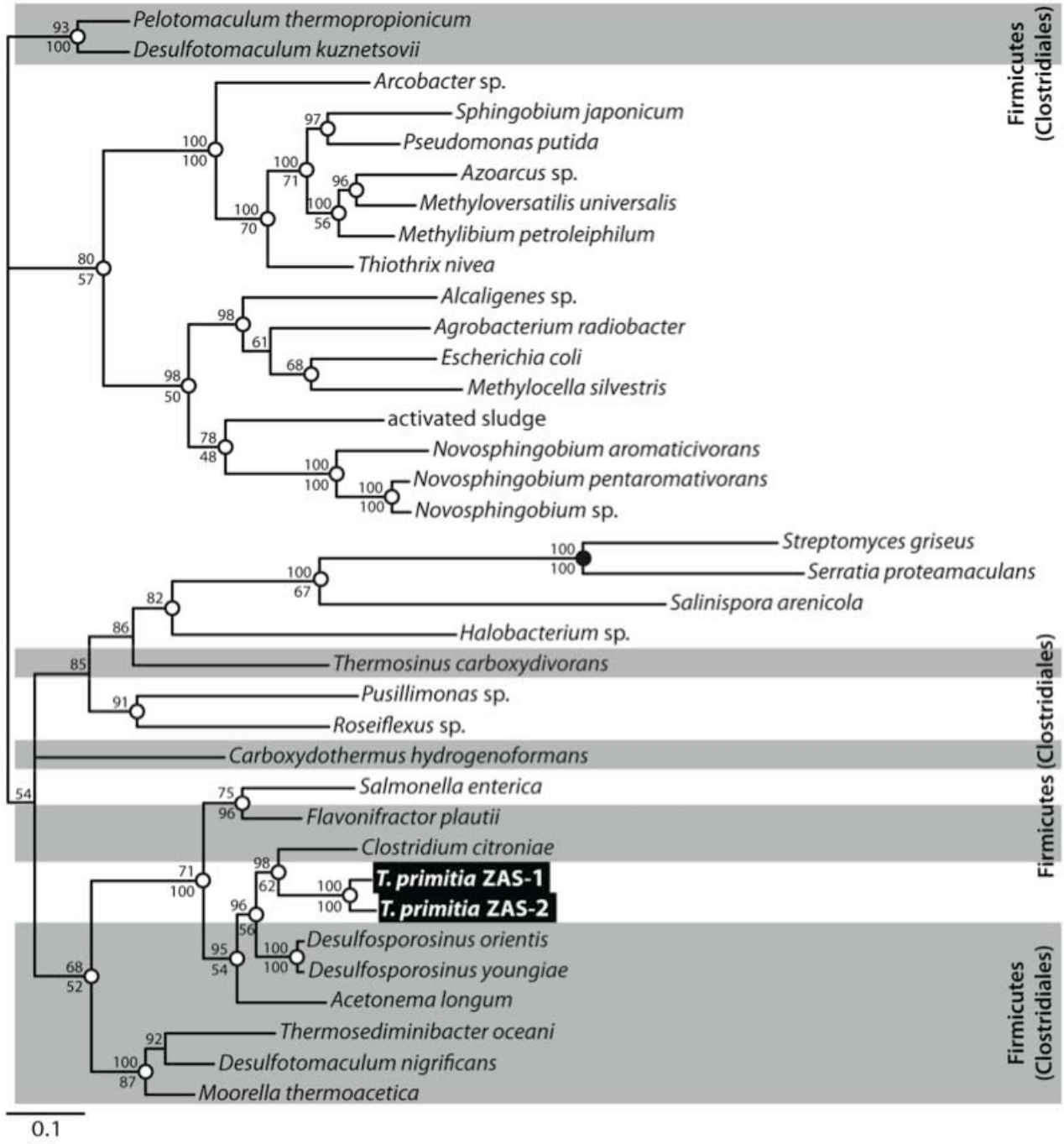


Fig. 3A-4: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 acetaldehyde dehydrogenase (PF01118 and PF09290, Step 5). Bayesian protein phylogenetic analysis (279 trees from 111,500 generations; PSRF = 1.004; average standard deviation of split frequencies = 0.018287) is based on 259 unambiguously aligned amino acid positions of a 291 amino acid-long protein. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum

parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

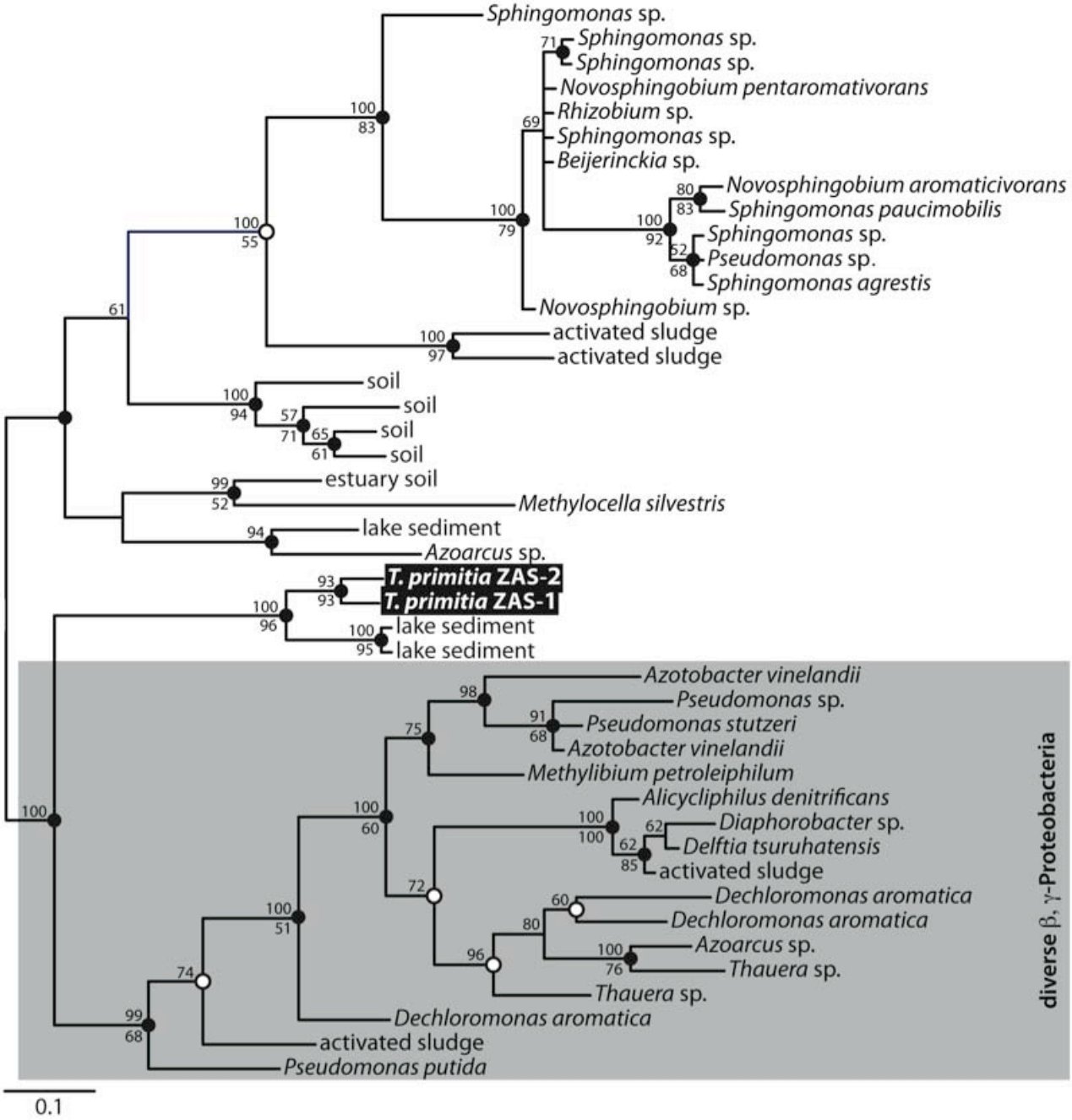


Fig. 3A-5: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 catechol 2,3-dioxygenase N-terminal Domain (PF00903). Bayesian protein phylogenetic analysis (220 trees from 88,000 generations; PSRF = 0.999; average standard deviation of split frequencies = 0.014790) is based on 56 unambiguously aligned amino acid positions of a 64 amino acid-long domain. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

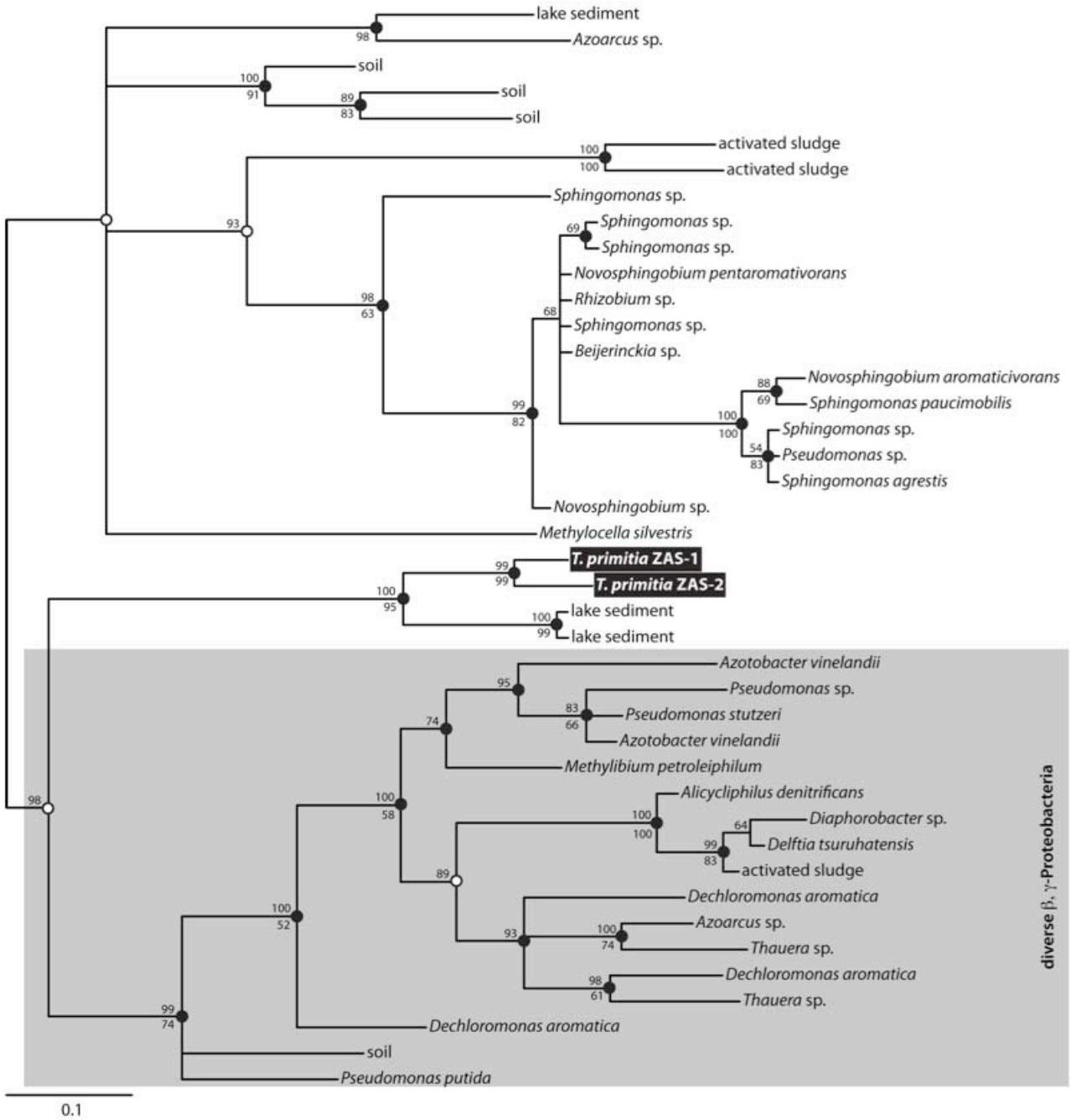


Fig. 3A-6: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 catechol 2,3-dioxygenase N-terminal Domain (PF00903) with extra-domain region. Bayesian protein phylogenetic analysis (600 trees from 240,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.010855) is based on 67 unambiguously aligned amino acid positions of a 68 amino acid-long region. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

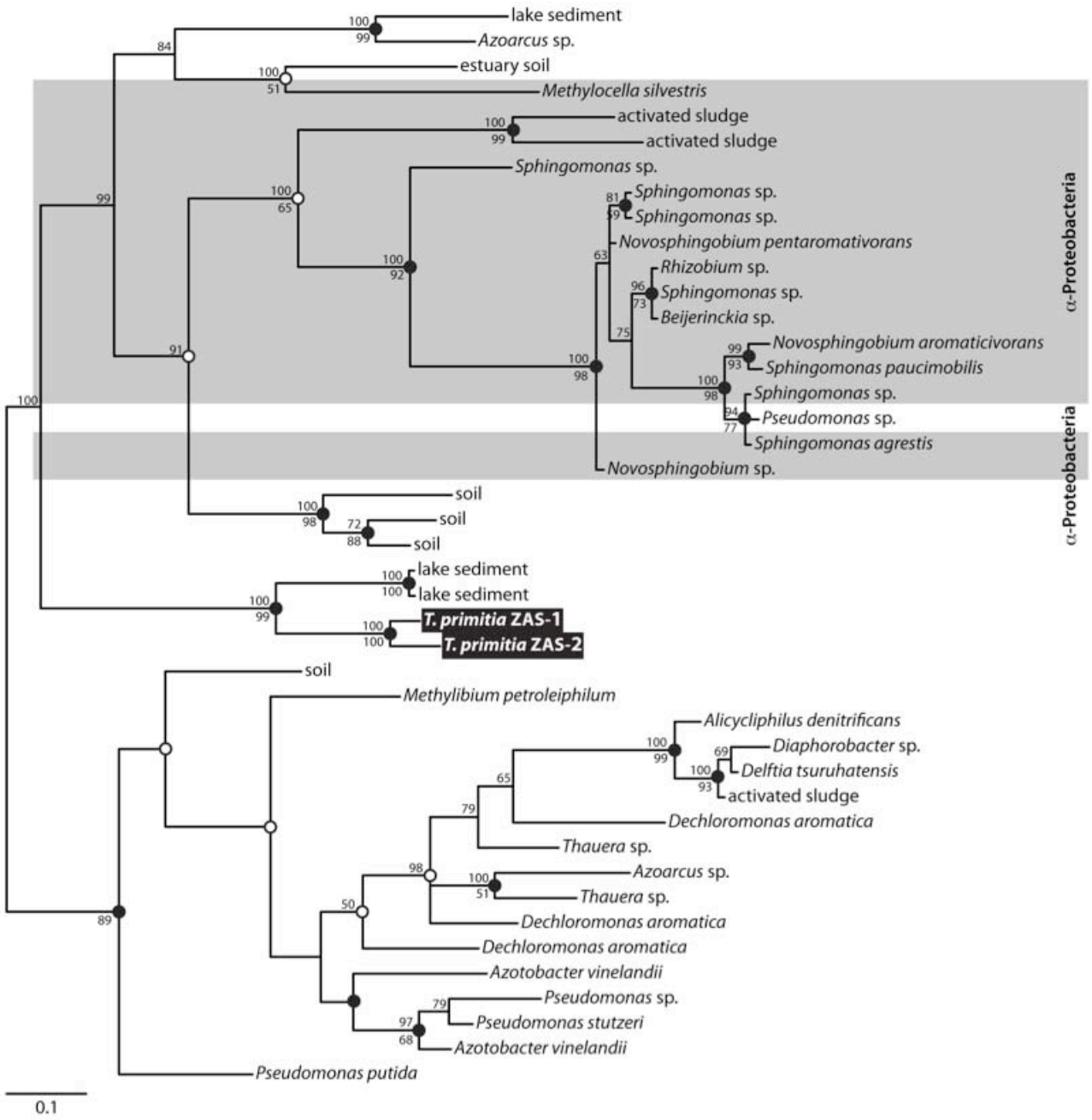


Fig. 3A-7: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 catechol 2,3-doxygenase N-terminal Domain (PF00903) with intra-domain region. Bayesian protein phylogenetic analysis (375 trees from 150,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.010443) is based on 88 unambiguously aligned amino acid positions of a 139 amino acid-long region. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍)

indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

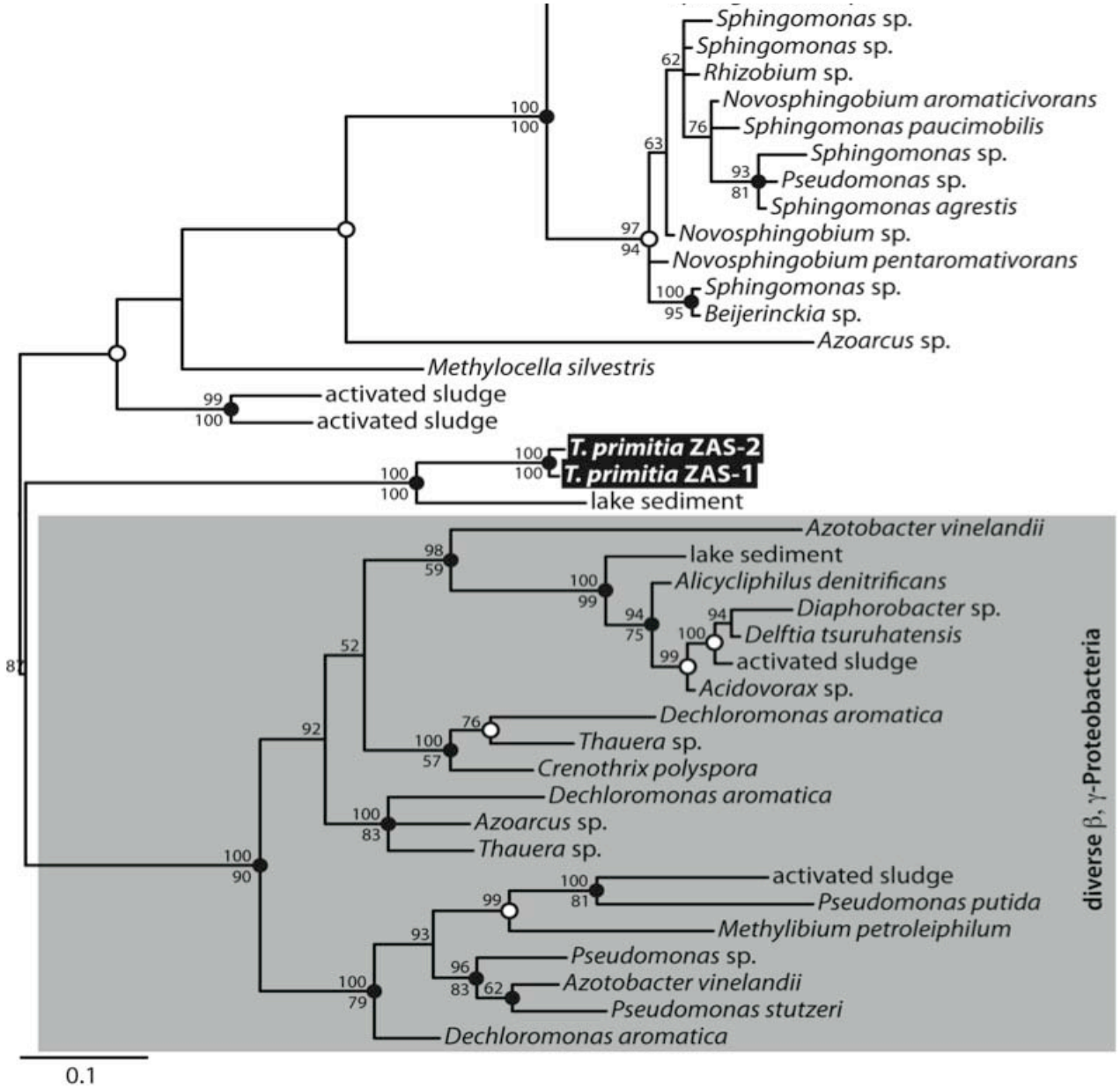


Fig. 3A-8: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 catechol 2,3-dioxygenase C-terminal Domain (PF00903). Bayesian protein phylogenetic analysis (350 trees from 140,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.014588) is based on 115 unambiguously aligned amino acid positions of a 116 amino acid-long domain. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

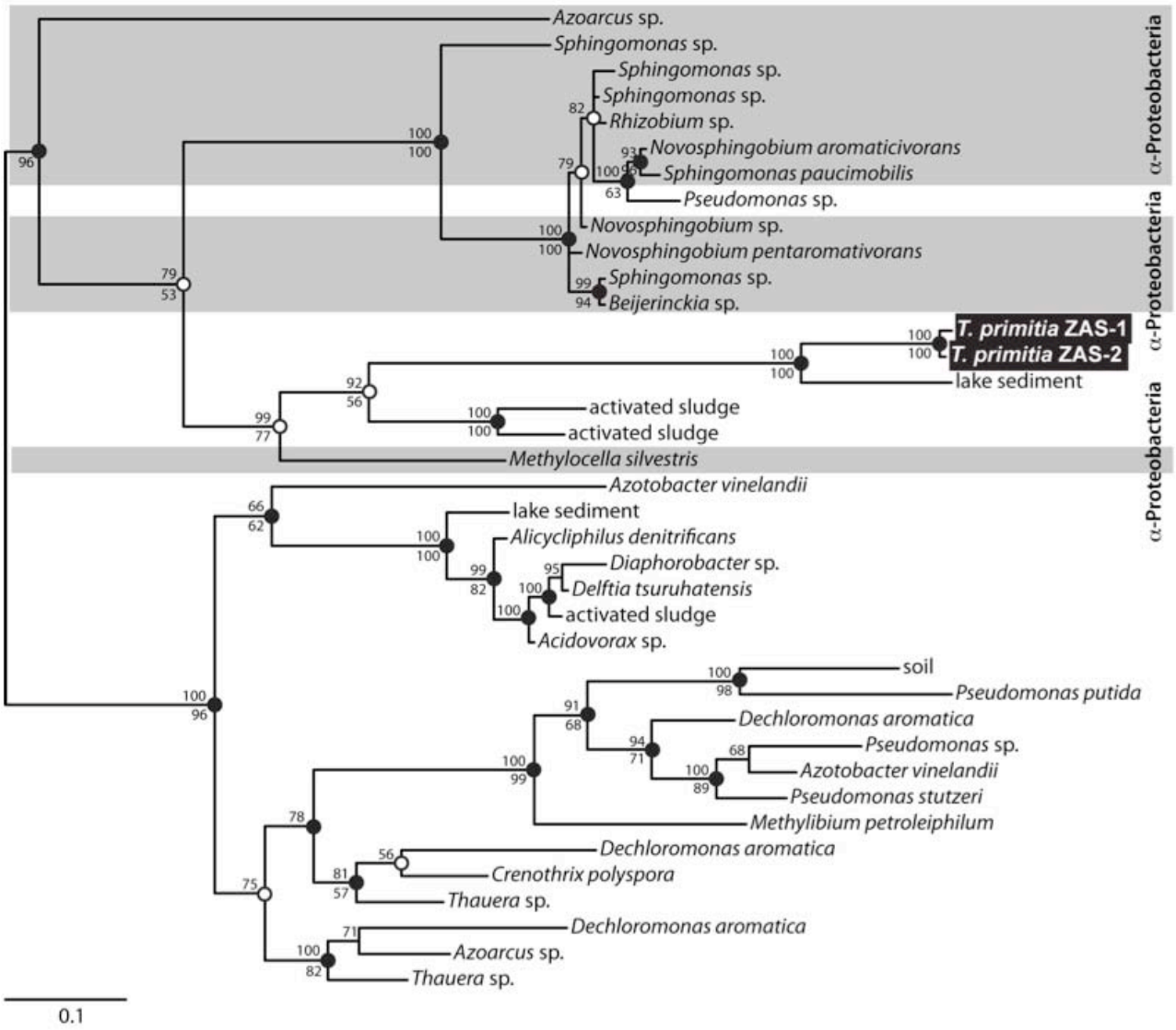


Fig. 3A-9: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 catechol 2,3-dioxygenase C-terminal Domain (PF00903) with extra-domain region. Bayesian protein phylogenetic analysis (450 trees from 180,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.010860) is based on 155 unambiguously aligned amino acid positions of a 158 amino acid-long region. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

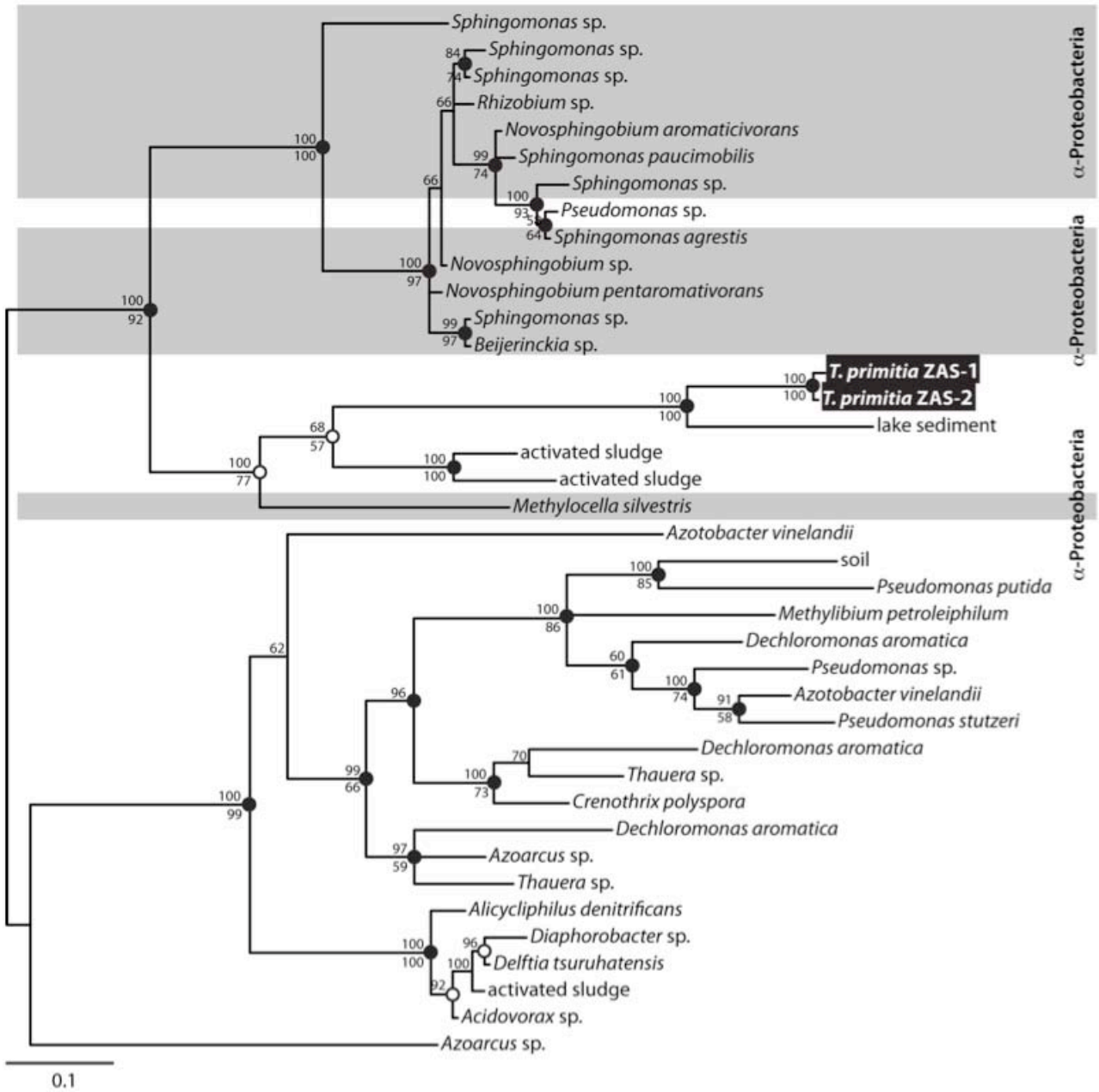


Fig. 3A-10: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 catechol 2,3-dioxygenase C-terminal Domain (PF00903) with intra-domain region. Bayesian protein phylogenetic analysis (500 trees from 200,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.010844) is based on 140 unambiguously aligned amino acid positions of a 141 amino acid-long region. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

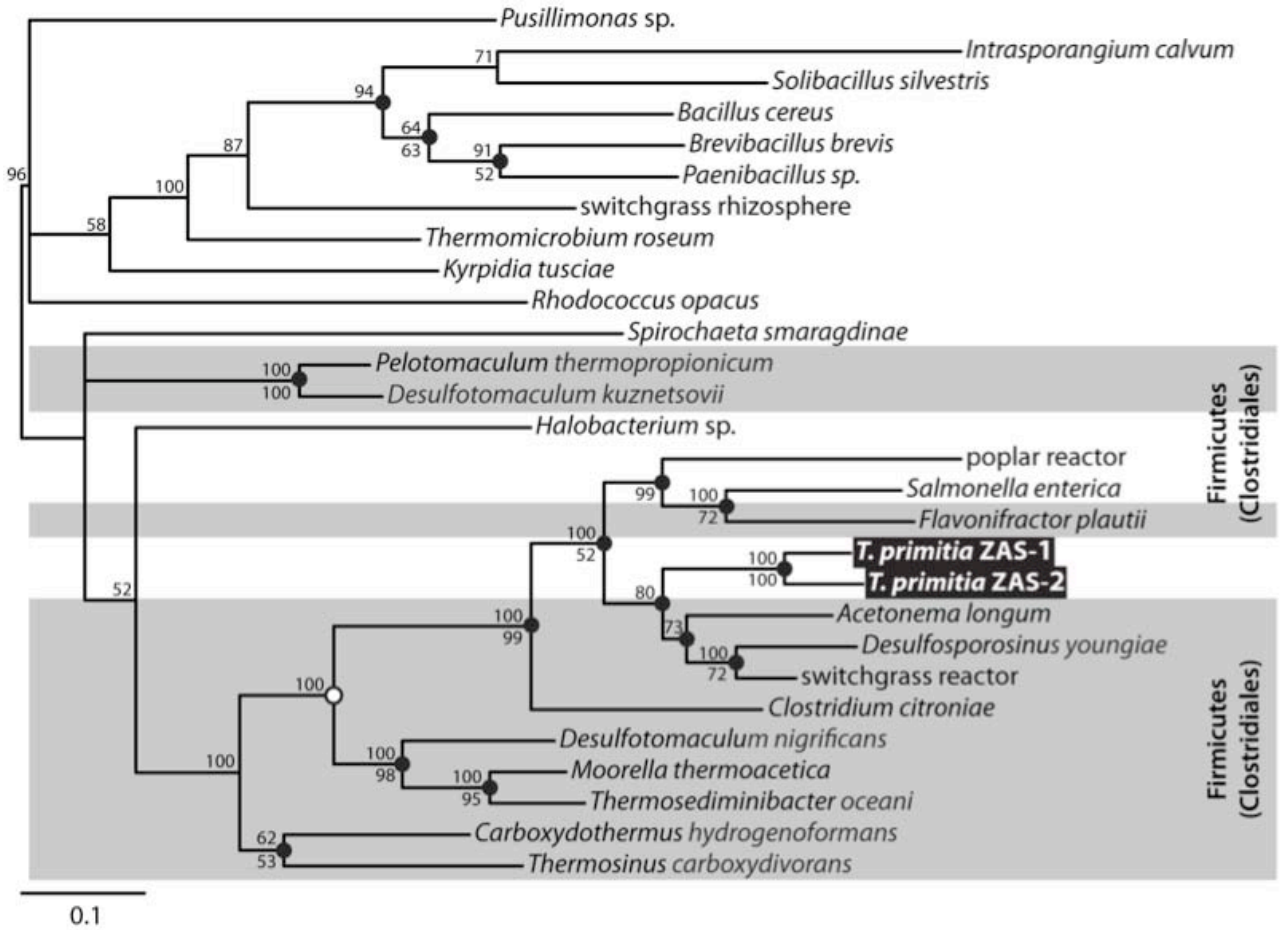


Fig. 3A-11: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 4-hydroxy-2-oxopentanoate aldolase HMGL-like domain (PF00682). Bayesian protein phylogenetic analysis (313 trees from 125,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.009037) is based on 243 unambiguously aligned amino acid positions of a 232 amino acid-long domain. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

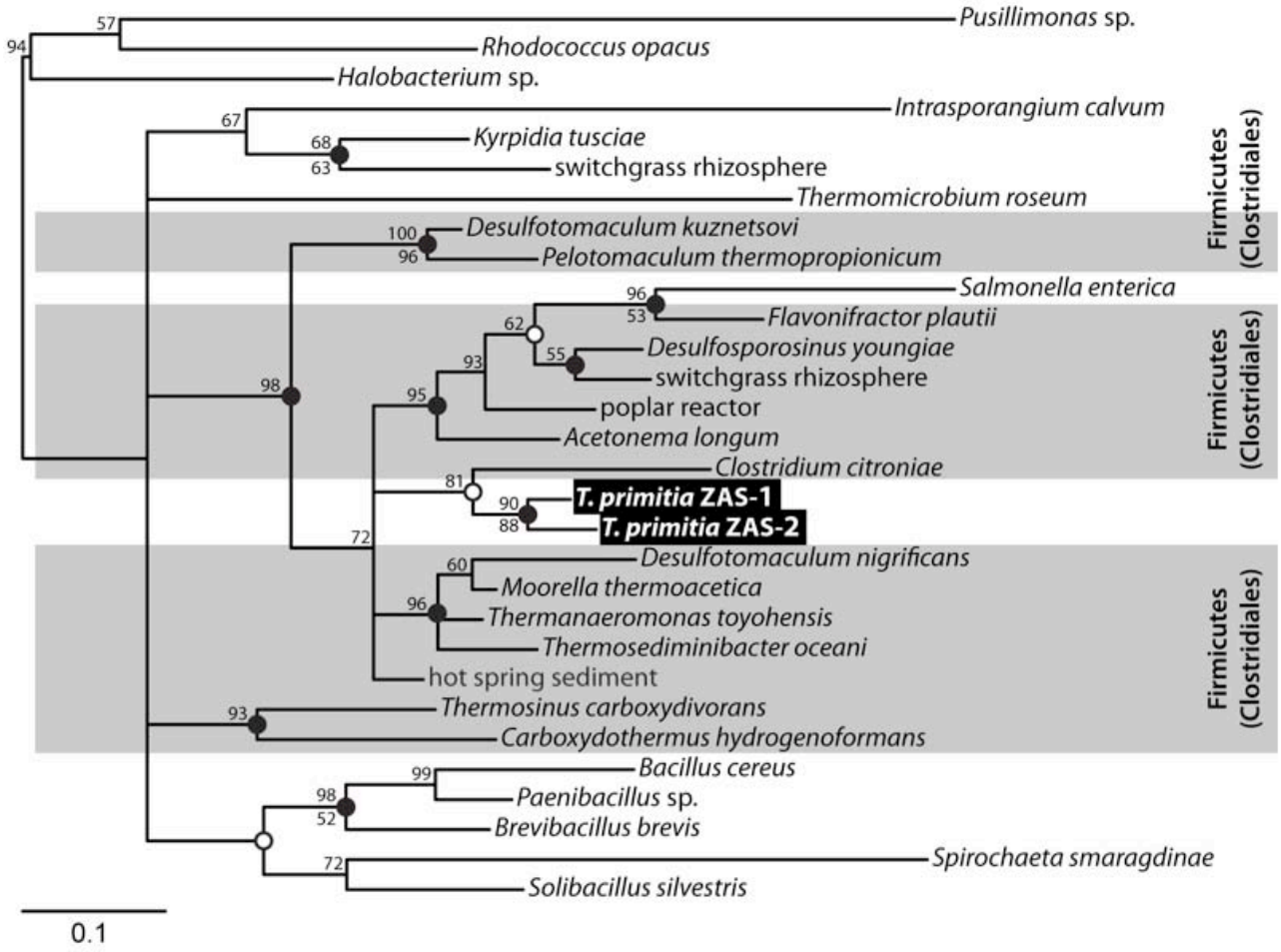


Fig. 3A-12: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 4-hydroxy-2-oxopentanoate aldolase DmpG-like communication domain (PF07836). Bayesian protein phylogenetic analysis (480 trees from 192,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.016899) is based on 65 unambiguously aligned amino acid positions of a 66 amino acid-long domain. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

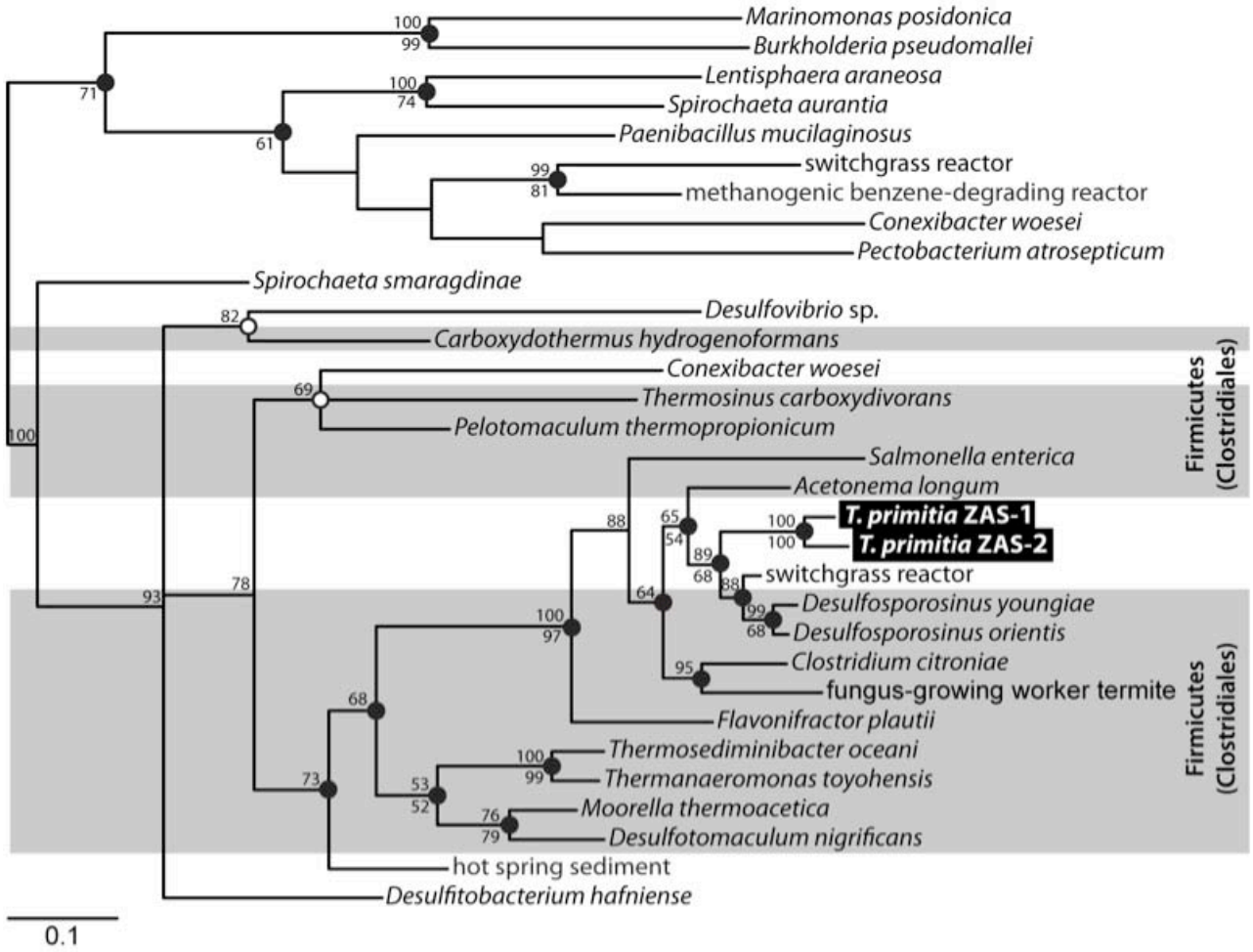


Fig. 3A-13: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 acetaldehyde dehydrogenase, NAD binding domain (PF01118). Bayesian protein phylogenetic analysis (123 trees from 205,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.010781) is based on 110 unambiguously aligned amino acid positions of a 110 amino acid-long domain. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

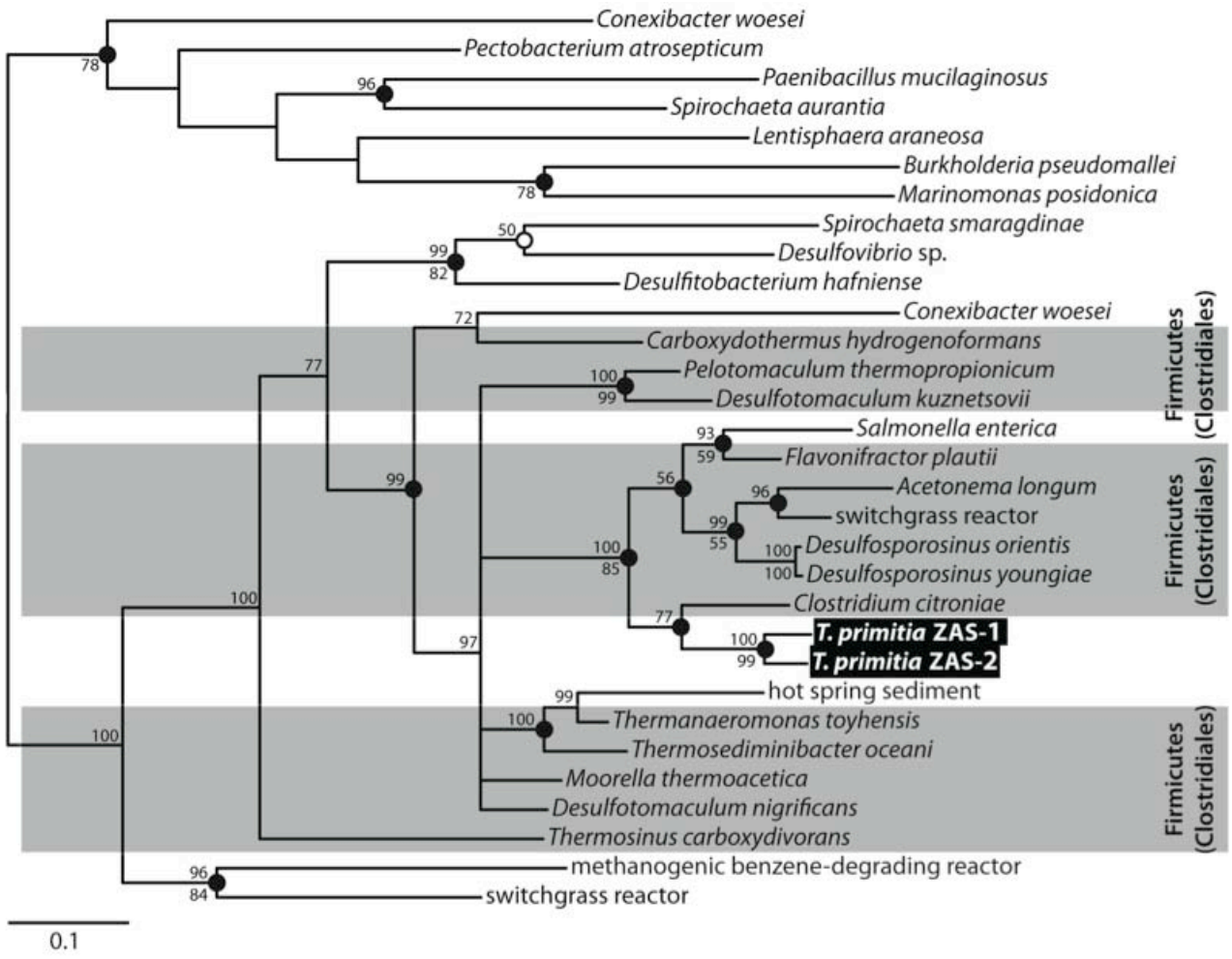


Fig. 3A-14: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 acetaldehyde dehydrogenase dimerisation domain (PF09290). Bayesian protein phylogenetic analysis (145 trees from 58,000 generations; PSRF = 0.999; average standard deviation of split frequencies = 0.013476) is based on 135 unambiguously aligned amino acid positions of a 137 amino acid-long domain. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

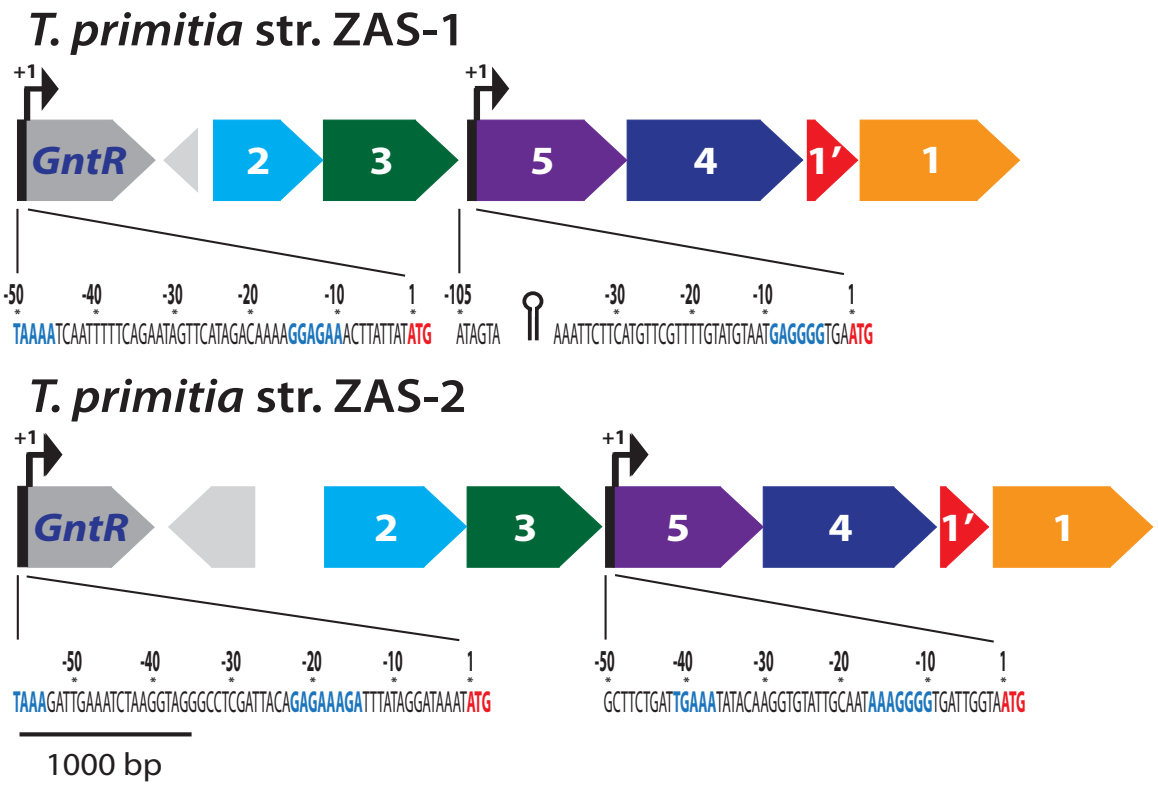


Fig. 3A-15: Transcriptional regulatory elements within the *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-cleavage pathway gene neighborhoods. Promoter regions are black rectangles and arrows indicate direction of transcription. Nucleotide bases of the promoter region are shown with conserved promoter elements in blue and transcriptional start sites in red. Stem loop structure is represented by a hairpin-like symbol. Ferredoxin-like peptide (step 1', red), catechol 2,3-dioxygenase (step 1, orange), 2-hydroxymuconic semialdehyde hydrolase (step 2, light blue), 2-oxopent-4-enoate hydratase (step 3, dark green), 4-hydroxy-2-oxopentanoate aldolase (step 4, dark blue), acetaldehyde dehydrogenase (step 5, purple), putative regulatory genes (dark grey), and hypothetical proteins (light grey) are depicted. Regulatory families are noted.

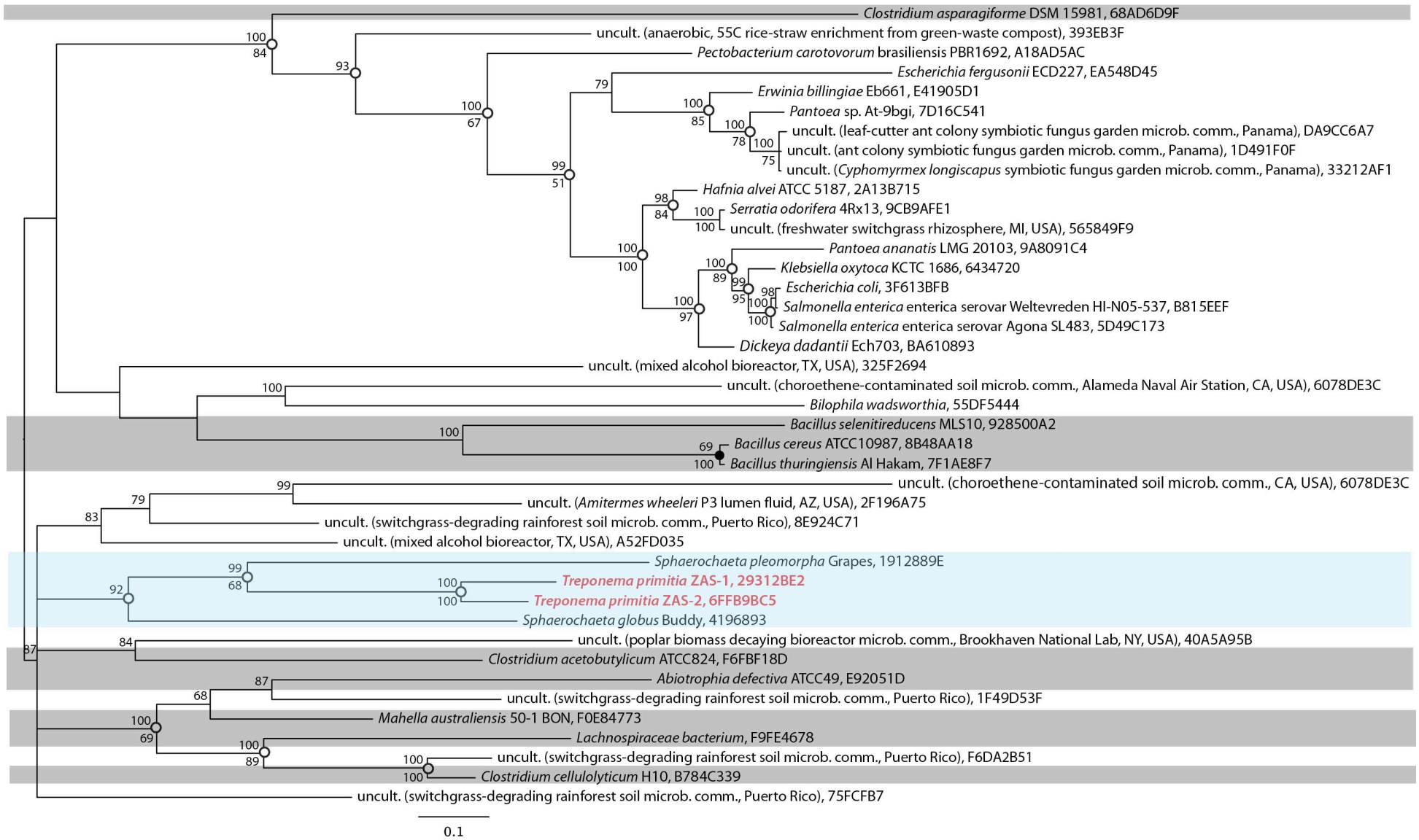

Clostridium asparagiforme DSM 15981, 68AD6D9F
uncult. (anaerobic, 55C rice-straw enrichment from green-waste compost), 393EB3F
Pectobacterium carotovorum brasiliensis PBR1692, A18AD5AC
Escherichia fergusonii ECD227, EA548D45
Erwinia billingiae Eb661, E41905D1
Pantoea sp. At-9bgi, 7D16C541
uncult. (leaf-cutter ant colony symbiotic fungus garden microb. comm., Panama), DA9CC6A7
uncult. (ant colony symbiotic fungus garden microb. comm., Panama), 1D491F0F
uncult. (Cyphomyrmex longiscapus symbiotic fungus garden microb. comm., Panama), 33212AF1
Hafnia alvei ATCC 5187, 2A13B715
Serratia odorifera 4Rx13, 9CB9AFE1
uncult. (freshwater switchgrass rhizosphere, MI, USA), 565849F9
Pantoea ananatis LMG 20103, 9A8091C4
Klebsiella oxytoca KCTC 1686, 6434720
Escherichia coli, 3F613BFB
Salmonella enterica enterica serovar Weltevreden HI-N05-537, B815EEF
Salmonella enterica enterica serovar Agona SL483, 5D49C173
Dickeya dadantii Ech703, BA610893
uncult. (mixed alcohol bioreactor, TX, USA), 325F2694
uncult. (choroethene-contaminated soil microb. comm., Alameda Naval Air Station, CA, USA), 6078DE3C
Bilophila wadsworthia, 55DF5444
Bacillus selenitireducens MLS10, 928500A2
Bacillus cereus ATCC10987, 8B48AA18
Bacillus thuringiensis Al Hakam, 7F1AE8F7
uncult. (choroethene-contaminated soil microb. comm., CA, USA), 6078DE3C
uncult. (Amitermes wheeleri P3 lumen fluid, AZ, USA), 2F196A75
uncult. (switchgrass-degrading rainforest soil microb. comm., Puerto Rico), 8E924C71
uncult. (mixed alcohol bioreactor, TX, USA), A52FD035
Sphaerochaeta pleomorpha Grapes, 1912889E
Treponema primitia ZAS-1, 29312BE2
Treponema primitia ZAS-2, 6FFB9BC5
Sphaerochaeta globus Buddy, 4196893
uncult. (poplar biomass decaying bioreactor microb. comm., Brookhaven National Lab, NY, USA), 40A5A95B
Clostridium acetobutylicum ATCC824, F6FBF18D
Abiotrophia defectiva ATCC49, E92051D
uncult. (switchgrass-degrading rainforest soil microb. comm., Puerto Rico), 1F49D53F
Mahella australiensis 50-1 BON, F0E84773
Lachnospiraceae bacterium, F9FE4678
uncult. (switchgrass-degrading rainforest soil microb. comm., Puerto Rico), F6DA2B51
Clostridium cellulolyticum H10, B784C339
uncult. (switchgrass-degrading rainforest soil microb. comm., Puerto Rico), 75FCFB7
0.1

Fig 3A-16: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-pathway associated GntR family transcriptional regulator (PF00392). Bayesian protein phylogenetic analysis (488 trees from 195,000 generations; PSRF = 1.001; average standard deviation of split frequencies = 0.015683) is based on 228 unambiguously aligned amino acid positions of a 249 amino acid-long protein. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

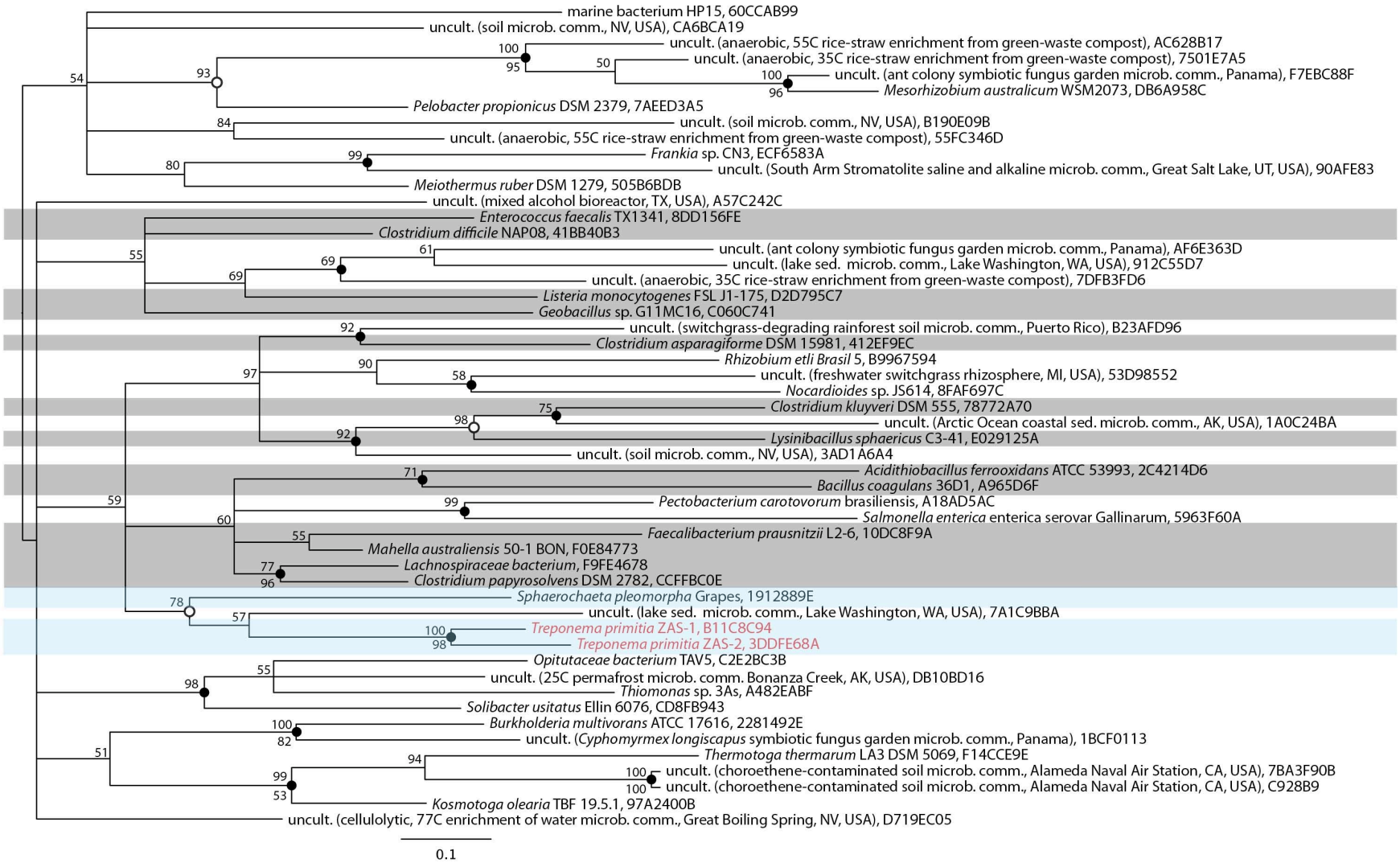

marine bacterium HP15, 60CCAB99
uncult. (soil microb. comm., NV, USA), CA6BCA19
uncult. (anaerobic, 55C rice-straw enrichment from green-waste compost), AC628B17
uncult. (anaerobic, 35C rice-straw enrichment from green-waste compost), 7501E7A5
uncult. (ant colony symbiotic fungus garden microb. comm., Panama), F7EBC88F
Mesorhizobium australicum WSM2073, DB6A958C
Pelobacter propionicus DSM 2379, 7AEED3A5
uncult. (soil microb. comm., NV, USA), B190E09B
uncult. (anaerobic, 55C rice-straw enrichment from green-waste compost), 55FC346D
Frankia sp. CN3, ECF6583A
uncult. (South Arm Stromatolite saline and alkaline microb. comm., Great Salt Lake, UT, USA), 90AFE83
Meiothermus ruber DSM 1279, 505B6BDB
uncult. (mixed alcohol bioreactor, TX, USA), A57C242C
Enterococcus faecalis TX1341, 8DD156FE
Clostridium difficile NAP08, 41BB40B3
uncult. (ant colony symbiotic fungus garden microb. comm., Panama), AF6E363D
uncult. (lake sed. microb. comm., Lake Washington, WA, USA), 912C55D7
uncult. (anaerobic, 35C rice-straw enrichment from green-waste compost), 7DFB3FD6
Listeria monocytogenes FSL J1-175, D2D795C7
Geobacillus sp. G11MC16, C060C741
uncult. (switchgrass-degrading rainforest soil microb. comm., Puerto Rico), B23AFD96
Clostridium asparagiforme DSM 15981, 412EF9EC
Rhizobium etli Brasil 5, B9967594
uncult. (freshwater switchgrass rhizosphere, MI, USA), 53D98552
Nocardioides sp. JS614, 8FAF697C
Clostridium kluyveri DSM 555, 78772A70
uncult. (Arctic Ocean coastal sed. microb. comm., AK, USA), 1A0C24BA
Lysinibacillus sphaericus C3-41, E029125A
uncult. (soil microb. comm., NV, USA), 3AD1A6A4
Acidithiobacillus ferrooxidans ATCC 53993, 2C4214D6
Bacillus coagulans 36D1, A965D6F
Pectobacterium carotovorum brasiliensis, A18AD5AC
Salmonella enterica enterica serovar Gallinarum, 5963F60A
Faecalibacterium prausnitzii L2-6, 10DC8F9A
Mahella australiensis 50-1 BON, F0E84773
Lachnospiraceae bacterium, F9FE4678
Clostridium papyrosolvens DSM 2782, CCFFBC0E
Sphaerochaeta pleomorpha Grapes, 1912889E
uncult. (lake sed. microb. comm., Lake Washington, WA, USA), 7A1C9BBA
Treponema primitia ZAS-1, B11C8C94
Treponema primitia ZAS-2, 3DDFE68A
Opitutaceae bacterium TAV5, C2E2BC3B
uncult. (25C permafrost microb. comm. Bonanza Creek, AK, USA), DB10BD16
Thiomonas sp. 3As, A482EABF
Solibacter usitatus Ellin 6076, CD8FB943
Burkholderia multivorans ATCC 17616, 2281492E
uncult. (Cyphomyrmex longiscapus symbiotic fungus garden microb. comm., Panama), 1BCF0113
Thermotoga thermarum LA3 DSM 5069, F14CCE9E
uncult. (choroethene-contaminated soil microb. comm., Alameda Naval Air Station, CA, USA), 7BA3F90B
uncult. (choroethene-contaminated soil microb. comm., Alameda Naval Air Station, CA, USA), C928B9
Kosmotoga olearia TBF 19.5.1, 97A2400B
uncult. (cellulolytic, 77C enrichment of water microb. comm., Great Boiling Spring, NV, USA), D719EC05
0.1

Fig. 3A-17: Phylogenetic position of *Treponema primitia* str. ZAS-1 and ZAS-2 *meta*-pathway associated GntR family transcriptional regulator, Bacterial regulatory proteins gntR family domain (PF00392). Bayesian protein phylogenetic analysis (950 trees from 380,000 generations; PSRF = 1.000; average standard deviation of split frequencies = 0.025667) is based on 61 unambiguously aligned amino acid positions of a 63 amino acid-long domain. Bayes values (when greater than 50) are reported above the nodes. Phylip PROTPARS maximum parsimony support after analysis of 1000 bootstraps (when greater than 50%) is reported below nodes. Shaded circles (●) indicate nodes supported by both maximum parsimony and Fitch distance matrix methods. Open circles (❍) indicate nodes supported by one of those methods. Scale bar indicates distance depicted as 0.1 amino acid changes per alignment position.

Fig. 3A-18: Yellow *Treponema primitia* str. ZAS-1 cultures. *T. primitia* str. ZAS-1 growing in 2YACo medium under an 80% N_2/20% CO_2 headspace (Leadbetter *et al*. 1999; Graber & Breznak 2004; Graber *et al*. 2004) reformulated to contain no dithiothreitol (DTT) or other reducing agent, no resazurin (a redox indicator). Maintaining sterile and anoxic cultures, 10mM maltose was added to the right culture. Maintaining sterile and anoxic cultures, 10mM maltose, 0.5mM catechol, and room air to achieve a final concentration of 0.5% O_2 vol/vol were added to the left culture. These cultures were incubated at 25°C in a horizontal position without agitation. After 5-7 days of growth fluids in the left culture turned a distinct lemon-yellow color. Analysis of culture fluids revealed that the catechol-dependent yellow material had an absorbance maximum at 375nm consistent with the presence of the expected dioxygenase-mediated ring cleavage product, 2-hydroxymuconic semialdehyde. UV/Vis absorbance spectra were obtained using a Cary WinUV Spectrophotometer.

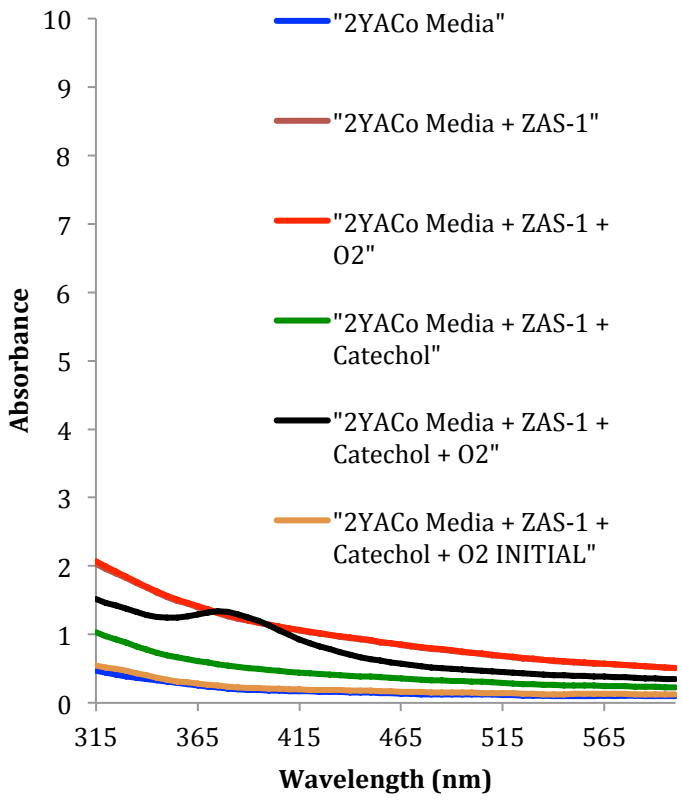
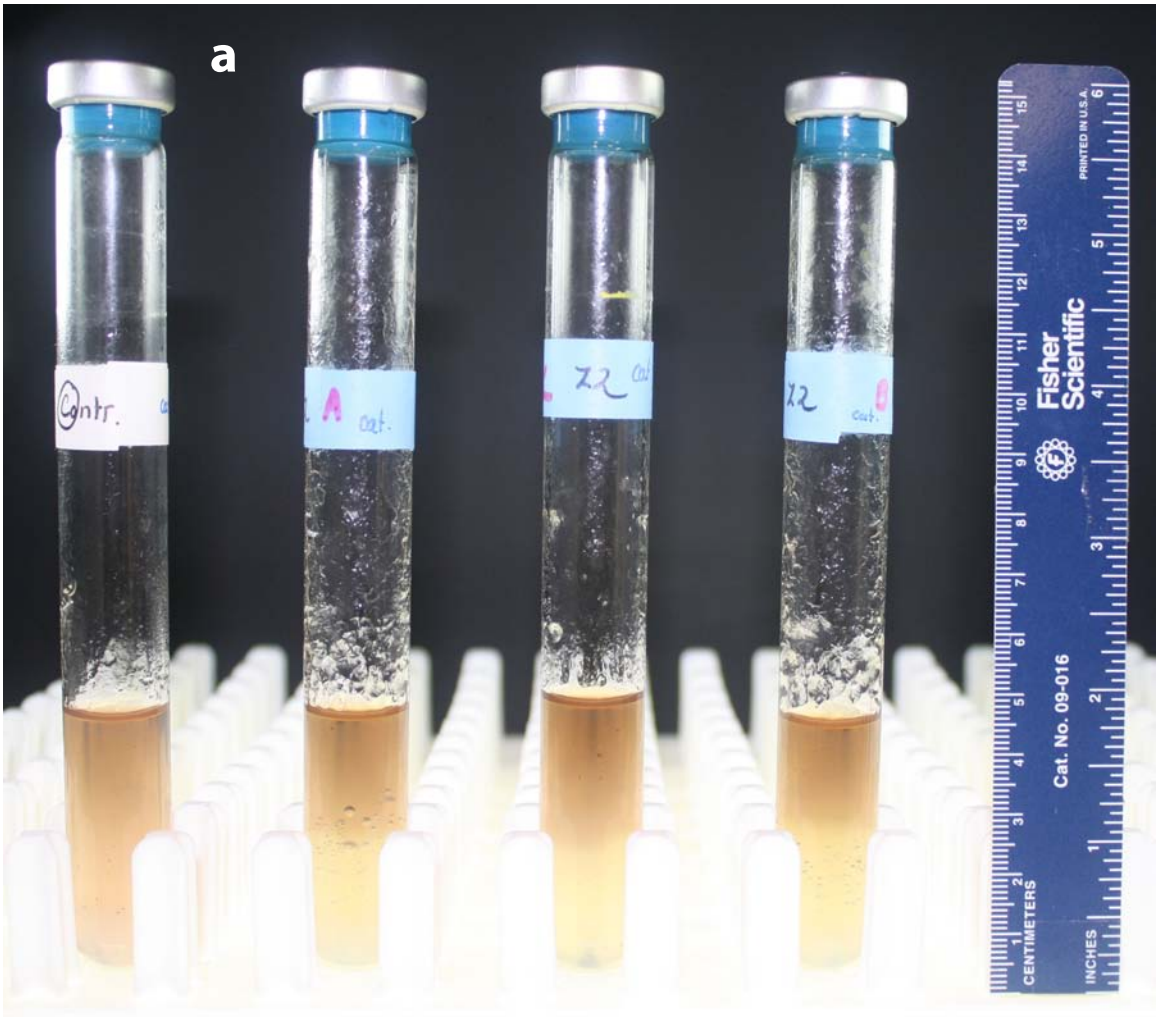


Fig. 3A-19: Absorbance spectra of *Treponema primitia* str. ZAS-1 cultures. *T. primitia* str. ZAS-1 was grown in 5mL of anaerobic, 2YACo, liquid media with an 80% N_2/20% CO_2 headspace in 25mL butyl rubber-stoppered Balch tubes. Where noted in the accompanying legend, cultures were supplemented with 0.5mM catechol and/or sterile air added to achieve 0.5% vol/vol O_2 in the headspace. For comparison, an initial spectra of *T. primitia* str. ZAS-1 with 0.5mM catechol and 0.5% vol/vol O_2 (upon inoculation and 4 days before culture fluid turned yellow) is shown. Media lacked both DTT and resazurin, and no other reducing agents were added. Cultures were incubated at 25°C in a horizontal position and still. Data taken 4 days after inoculation. UV/Vis absorbance spectra data was obtaind with a Cary WinUV Spectrophotometer.


a
Contr.
A
cat.
22
22
Fisher Scientific
Cat. No. 09-016
CENTIMETERS
INCHES
PRINTED IN U.S.A.

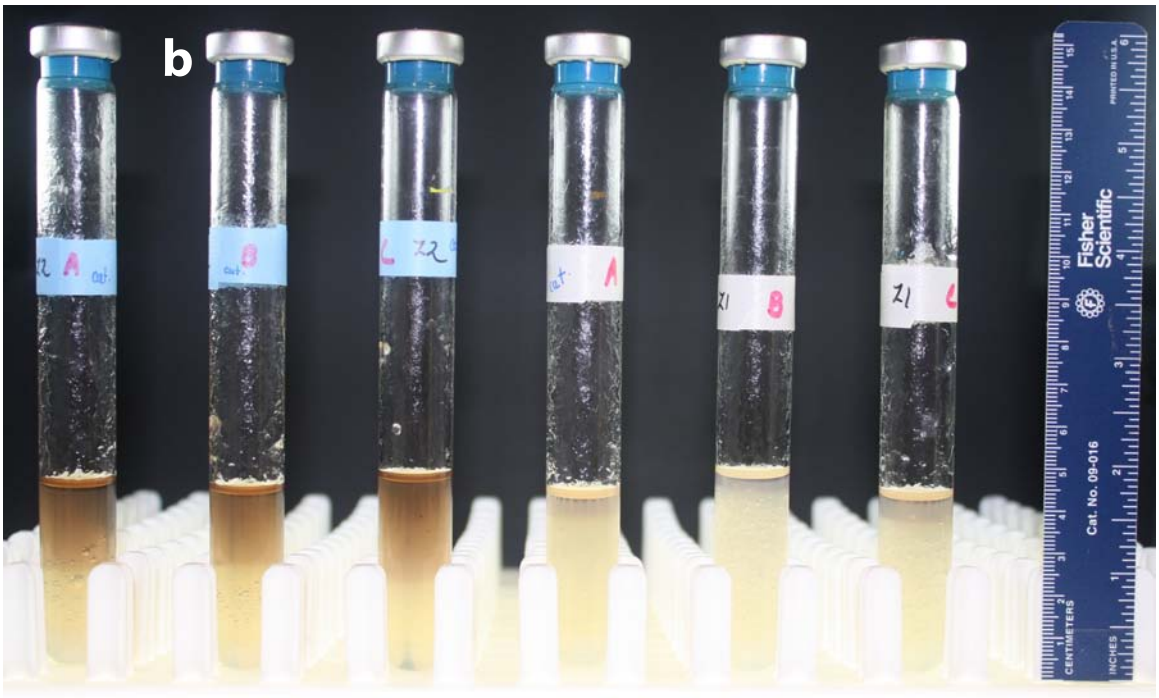
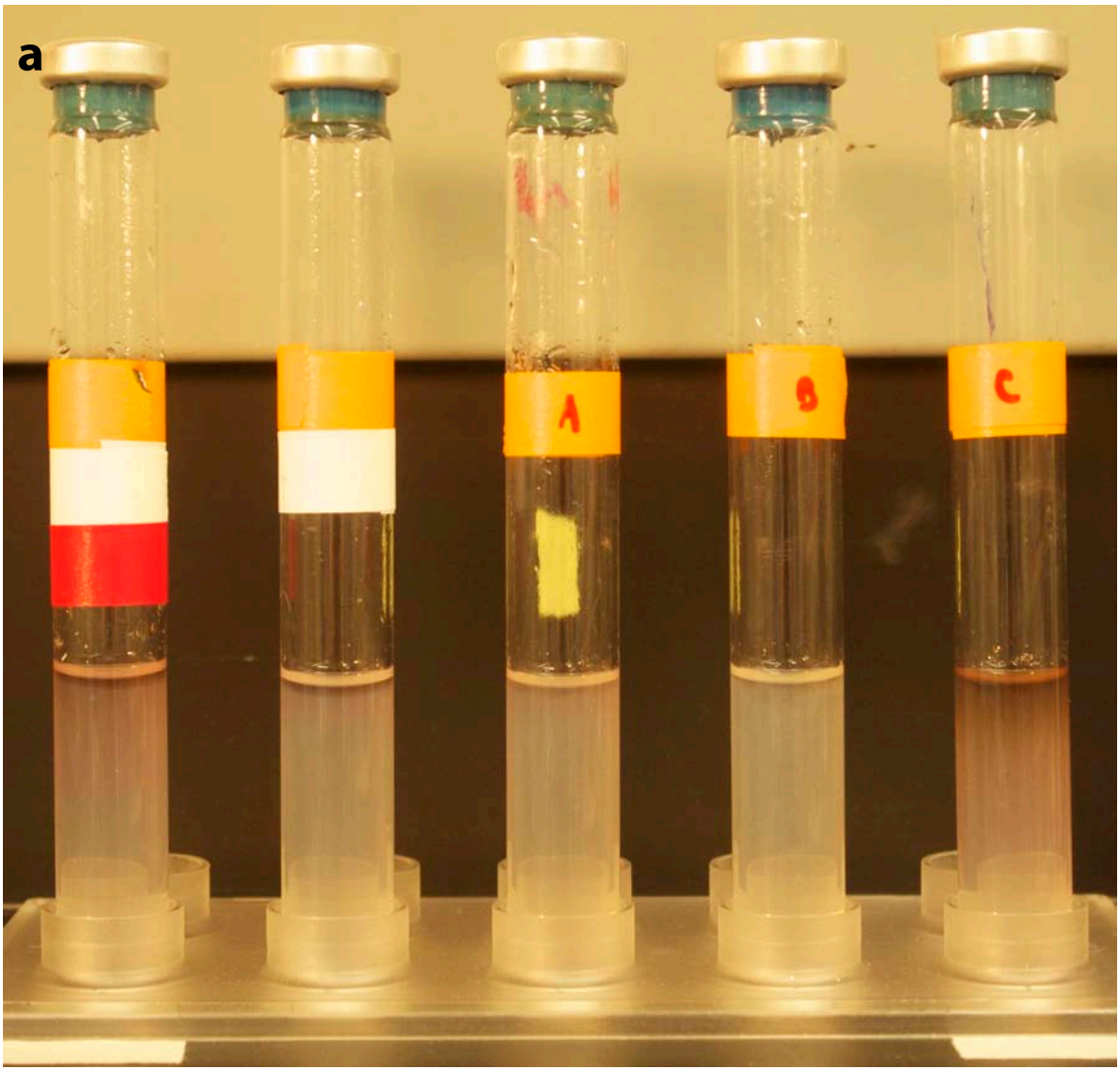


Fig. 3A-20: *Treponema primitia* str. ZAS-1 and ZAS-2 O_2/catechol gradient cultures and controls. Oxygen gradient tubes under a headspace of $N_2/CO_2/O_2$ (70:10:4) and depicted after 7 days of incubation. Gradient tubes prepared as described in Materials and Methods above but instead of adding *T. primitia* str. ZAS-1 and ZAS-2 to cultures along the vertical length of the agarose, *T. primitia* str. ZAS-1 and ZAS-2 was added homogenously **(a)** From left to right: uninoculated catechol (10mM plug at bottom of tube) and O_2 control, and triplicates of O_2/catechol gradient tubes inoculated with of *T. primitia* str. ZAS-2. **(b)** From left to right: triplicates of O_2/catechol gradient tubes inoculated with of *T. primitia* str. ZAS-2 as seen in **(a)** compared to triplicates of O_2/catechol gradient tubes inoculated with of *T. primitia* str. ZAS-1. (*for more thorough analyses of growth of T. primitia str. ZAS-1 in O_2/aromatic gradient tubes see Table 3-4. Fig. 3-7 and Appendix Fig. 3A-21a and b*).


a
A
B
C

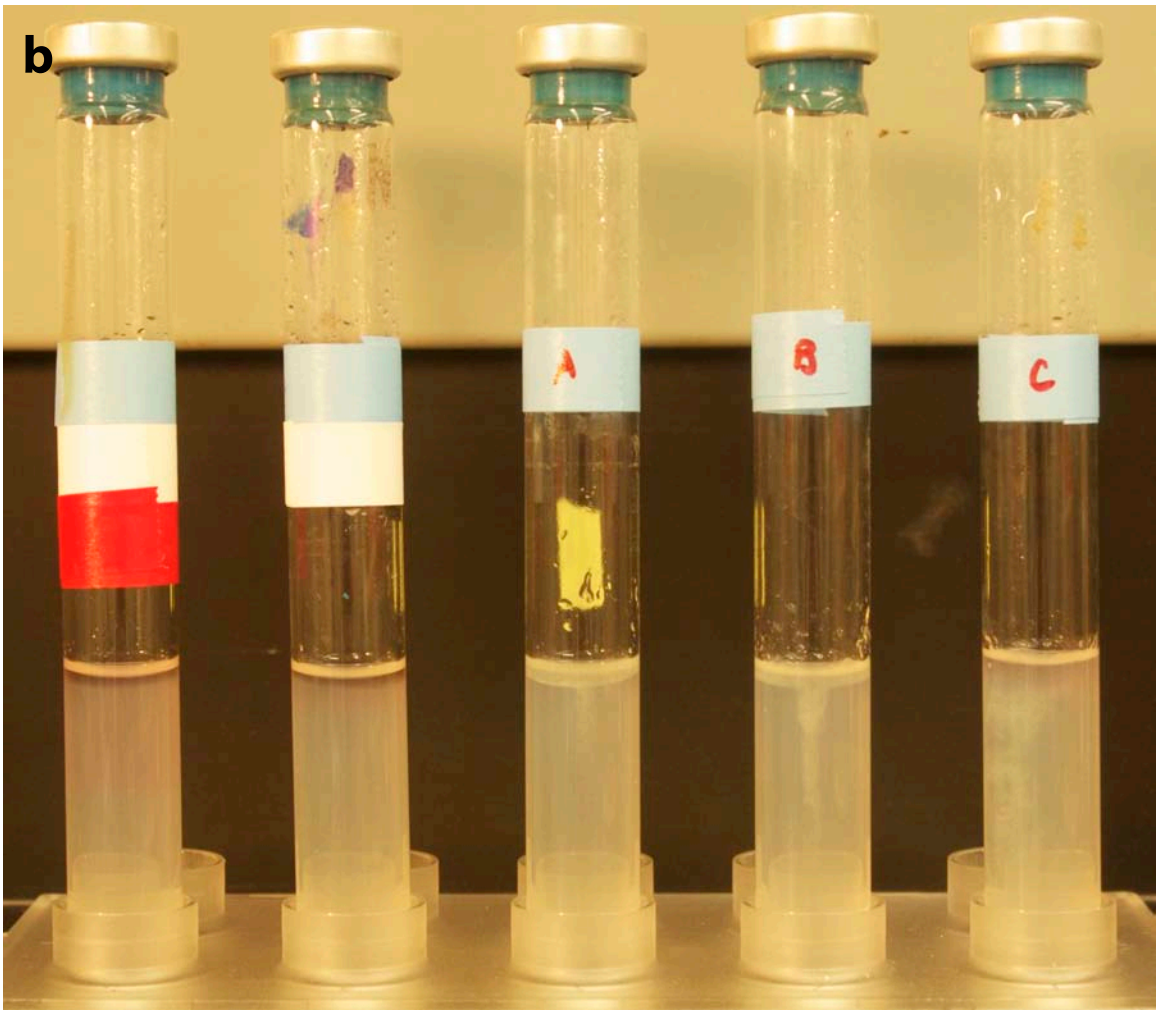
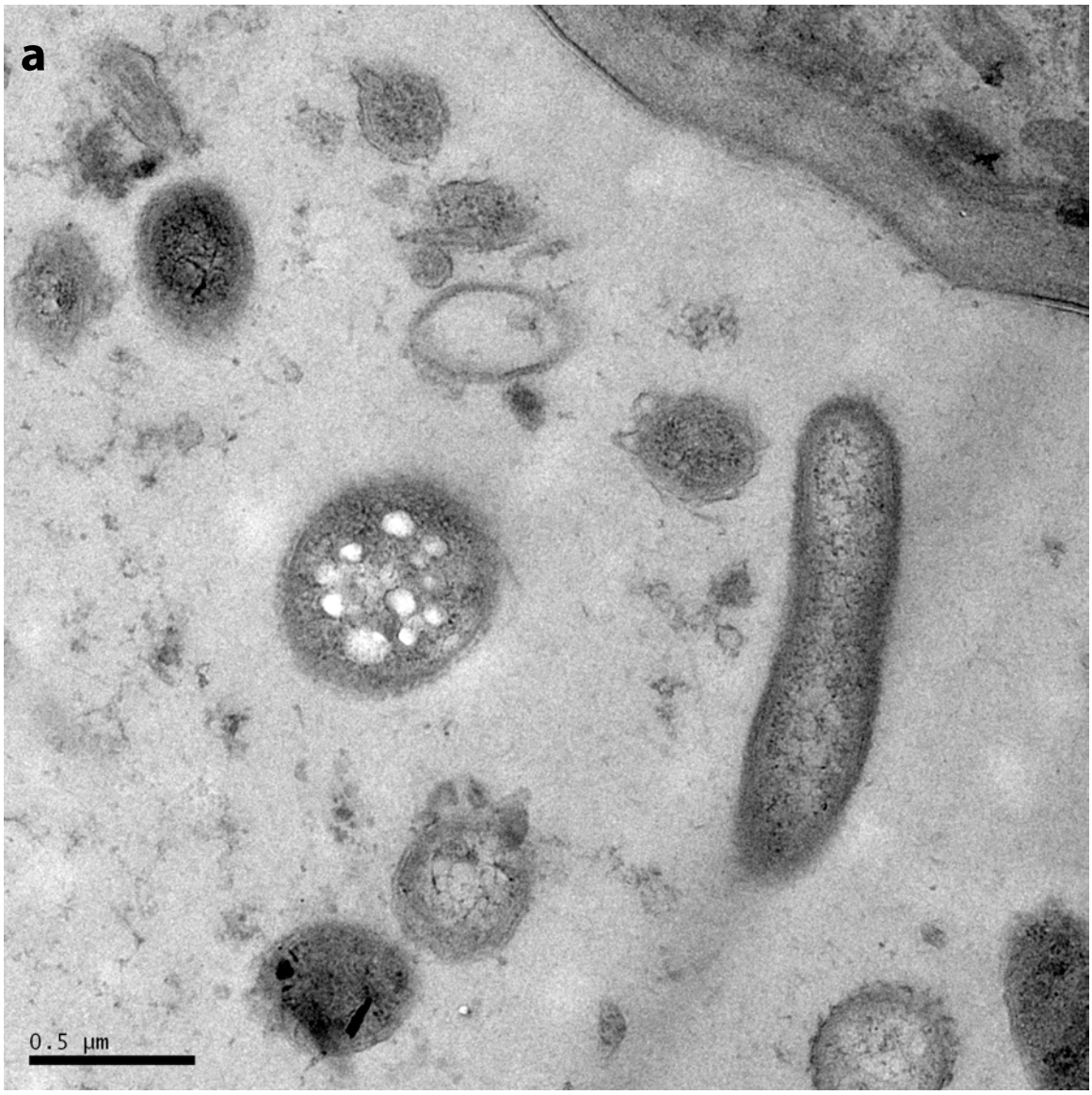
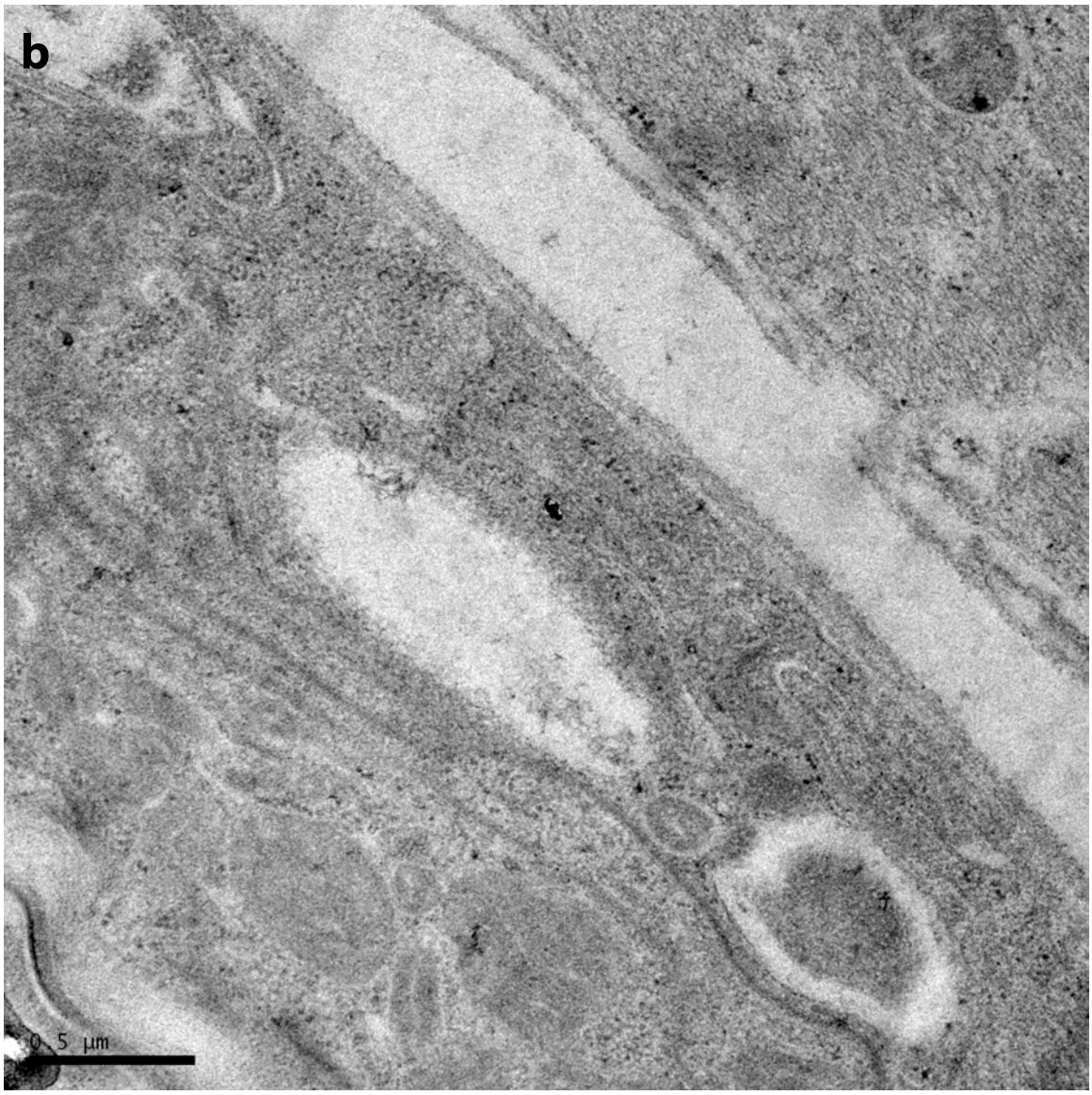
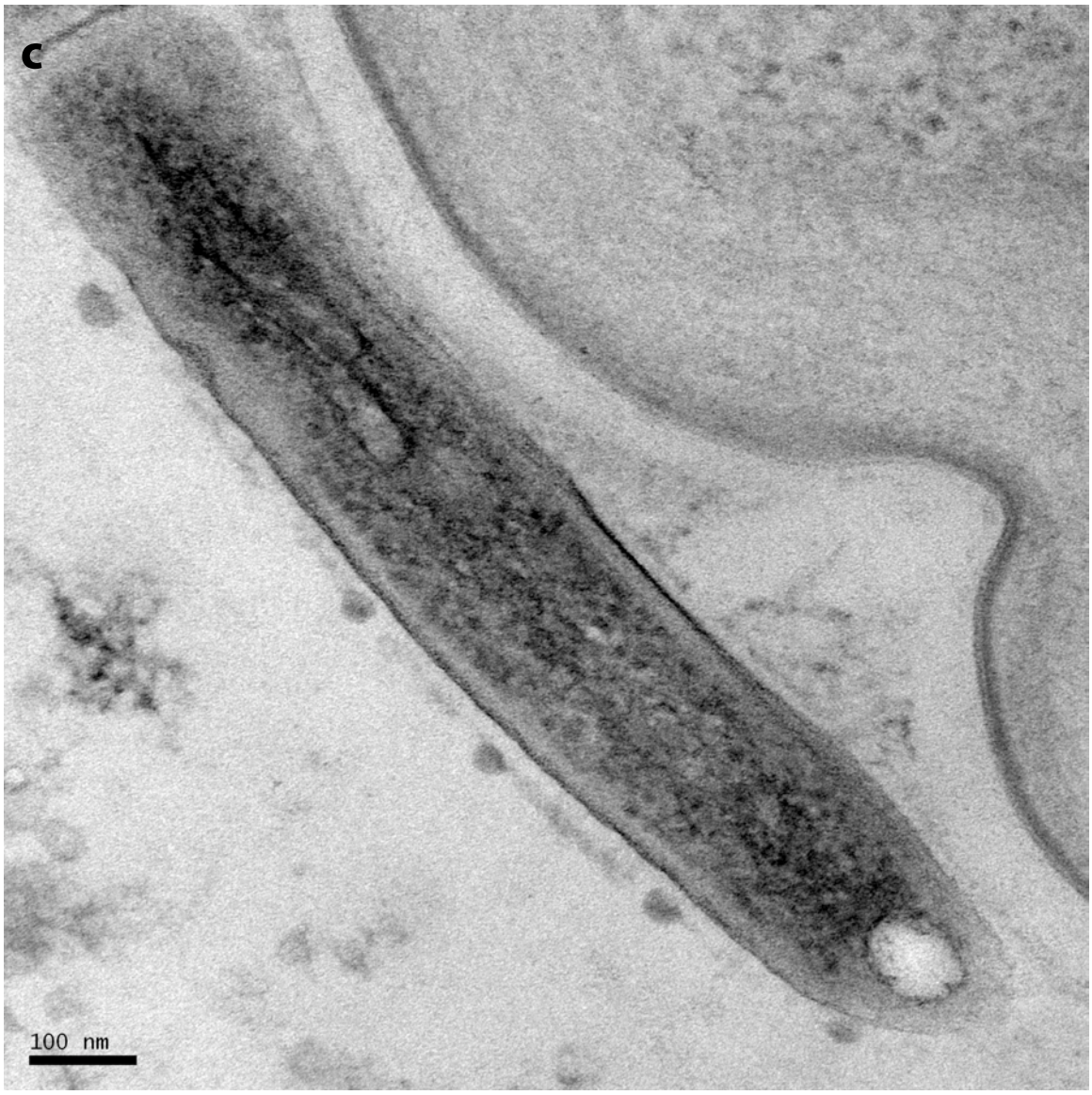


Fig. 3A-21: *Treponema primitia* str. ZAS-1 **(a)** catechol gradient cultures and controls and **(b)** water gradient cultures and controls. Gradient tubes under a headspace of 80% N_2/20% CO_2 (unless noted otherwise) depicted after 7 days of incubation. Gradient tubes prepared as described in Materials and Methods above. At the bottom of each tube there is a plug of **(a)** 10mM catechol or **(b)** water. **(a)** From left to right: uninoculated 4% vol/vol O_2 and catechol control; uninoculated catechol-only control; triplicates of *T. primitia* str. ZAS-1 inoculated in gradient tubes in which a 10mM catechol plug is located at the bottom but to which no O_2 was added. Orange = catechol, white = no cells, red = O_2. **(b)** From left to right: uninoculated 4% vol/vol O_2 and water control; uninoculated water-only control; triplicates of *T. primitia* str. ZAS-1 inoculated in gradient tubes in which a 10mM plug of water is located at the bottom but to which no O_2 was added. Blue = water, white = no cells, red = O_2.


a
0.5 µm


b
0.5 µm


c
100 nm

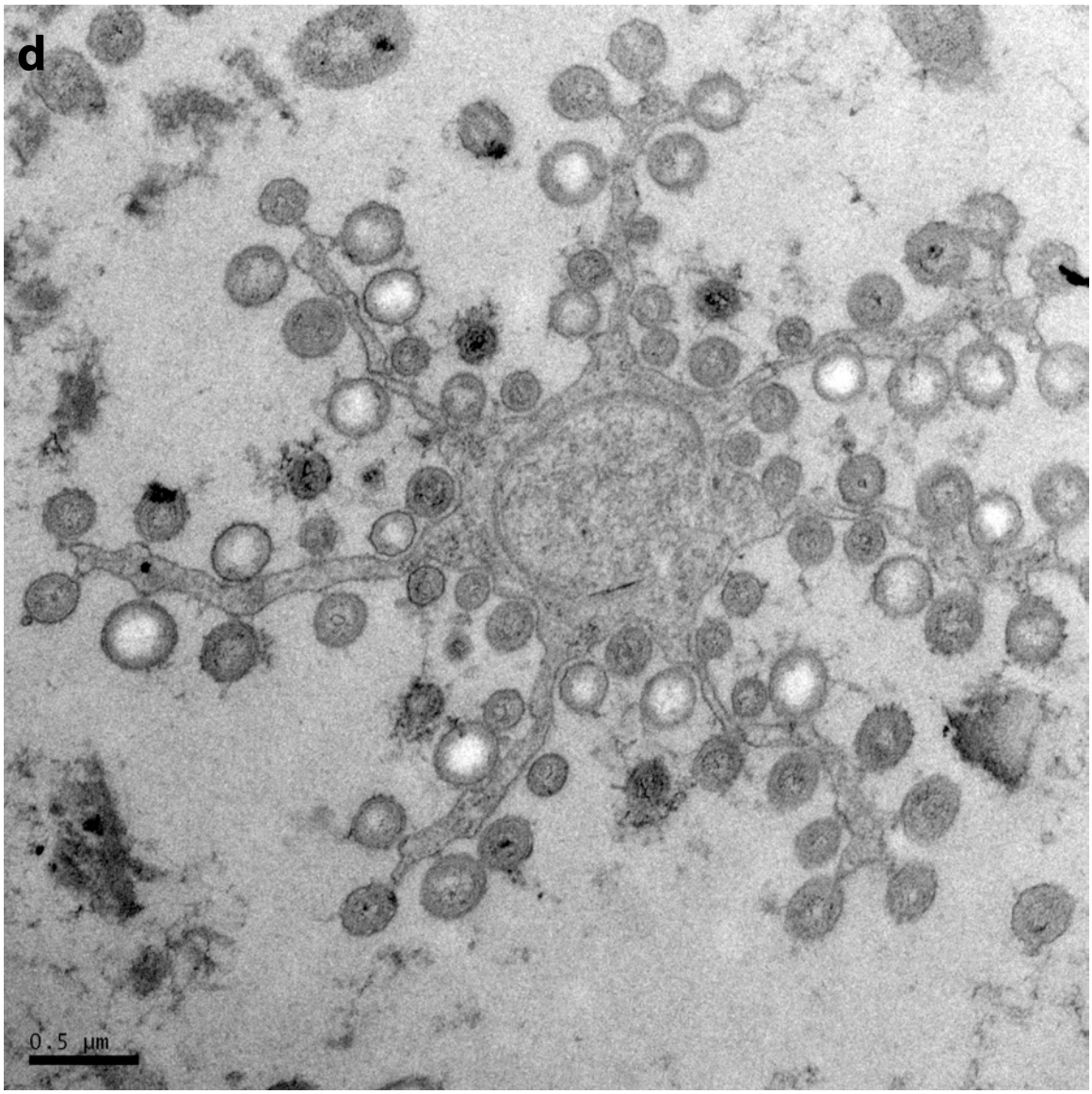


Fig. 3A-22: Transmission electron microscopy images of sections of *Zootermopsis nevadensis* hindgut **(a, b, c, and d)**. The hindgut of a medium-sized *Zootermopsis nevadensis* worker specimen was dissected out and immediately fixed in 2.5% glutaraldehyde in 100mM phosphate buffer at pH 7.0 for 24hrs. Post fixation, the hindgut was washed in 200mM phosphate buffer and treated in 1% osmium tetroxide in 100mM phosphate buffer at 4°C for 2hrs. Next, the hindgut was washed in distilled water and stained with 2% aquaeous uranyl acetate at 4°C for 2hrs in the dark. Then, the hindgut sample underwent a series of steps for dehydration that consisted of 30% acetone for 15min, 50% acetone for 15min, 70% acetone for 15min, 90% acetone for 15min, and three changes of 100% acetone for 30min each. Prior to resin embedding, the hindgut sample was exposed to propylene oxide twice for 15mins each. Then the hindgut sample underwent a resin infiltration regime that consisted of 2:1 mix of propylene oxide:resin for 1hr, 1:1 mix of propylene oxide:resin for 1hr, 1:2 mix of propylene oxide:resin for 1hr, 100% resin overnight, fresh resin 1hr, and polymerise at 60°C for 24hrs. Unless noted otherwise, sample

preparation occurred at room temperature under room lighting. An ultramicrotome was used to cut hidgut sample into <100nm cross sections. Images were obtained with a FEI Tecnai T12 transmission electron microscope. **(d)** Likely *Streblomastix strix* (*images couresy of Elitza Tocheva*).

Method 3A-1: Visualizing *Treponema primitia* in termite hindgut.

To prepare cross-sections of a termite hindgut first carefully dissected out gut as to not disrupt epithelium and contents. Then, immediately fix it in 2.5% glutaraldehyde in 100mM phosphate buffer at pH 7.0. Depending on size of sample this can take 2-24hrs. Fixation is typically begin at room (or physiological) temperature, and after 15-30min is continued at 4°C. 4°C slows autolytic processes and tissue shrinkage but is not appropriate for all samples.

Post fixation, wash gut in 200mM phosphate buffer. Then, treat gut in 1% osmium tetroxide in 100mM phosphate buffer at 4°C for 1-2hrs, again depending on sample size. Next, wash gut in distilled water at least five times to remove all excess phosphate ions and prevent uranyl acetate stain from precipitating.

Stain gut with 2% aquaeous uranyl acetate at 4°C for 2hrs in the dark. Then, the expose gut to a series of dehydration steps including 30% acetone for 15min, 50% acetone for 15min, 70% acetone for 15min, 90% acetone for 15min, and three changes of 100% acetone for 30min each.

Prior to resin embedding, expose the gut sample to propylene oxide twice for 15mins each. Then the gut is ready for a resin infiltration regime that includes 2:1 mix of propylene oxide:resin for 1hr, 1:1 mix of propylene oxide:resin for 1hr, 1:2 mix of propylene oxide:resin for 1hr, 100% resin overnight, fresh resin 1hr, and polymerise at 60-70°C for 12-24hrs.

Unless noted otherwise, sample preparation occurs at room temperature under room lighting. Use an ultramicrotome to cut gut sample into <100nm cross sections (Glauert & Lewis 1998). Hybridization Chain Reaction FISH could then be used to map mRNA expression to genes of interest (Choi *et al*. 2010).

Method 3A-2: qRT-PCR of each *meta*-cleavage pathway gene.

Although not conducted as part of the research presented here, utilization of qRT-PCR was explored as a means to gauge expression of each *meta*-cleavage pathway gene. As an initial step primers were designed based on *Treponema primitia meta*-pathway gene sequences. The following two sets of potentially useful primers for each *meta*-cleavage pathway gene have been evaluated sequence-wise (self- and hetero-dimer tools available in OligoAnalyzer 3.1 at www.idtdna.com) and await testing with qRT-PCR itself. A benefit of performing qRT-PCR for each *meta*-cleavage pathway gene is that a time course of expression can be obtained including information on under what conditions a particular gene is transcribed, how long does it take the be activated, and what substrates does it respond to, among other questions.

<u>Ferredoxin-like peptide</u>:
Set A:
FWD: 5' – CGG TGA GTT ATG GCT GCG CC – 3'
RVS: 5' – CCT CCG AAA TAT GGG CGG CG – 3'
Set B:
FWD: 5' – GTC CGG TGA GTT ATG GCT GC – 3'
RVS: 5' – CCG AAA TAT GGG CGG CGC TC – 3'

<u>Catechol 2,3-dioxygenase</u>:
Set A:
FWD: 5' – CGA CGC CAA GAC CCT GGA GG – 3'
RVS: 5' – GCG TAG CGT TTT CCA AAA CC – 3'
Set B:
FWD: 5' – CCG ACG GGA AAC GGC TTG CC – 3'
RVS: 5' – GGA AGA ATG CGA AGT GGT GC – 3'

<u>2-hydroxymuconic semialdehyde hydrolase</u>:
Set A:
FWD: 5' – GGA CTG GAC CAG GTT TGG GG – 3'
RVS: 5' – GCT TTT AAC AAG GTC CTT GG – 3'
Set B:

FWD: 5' – CTC TGG CGA TCA AAT ACC CC – 3'
RVS: 5' – GTT CGT AAC CCC AAA CCT GG – 3'

<u>2-oxopent-4-enoate hydratase</u>:
Set A:
FWD: 5' – CCA ACA AGG TTC GGC TCA GC – 3'
RVS: 5' – CTC CCA GGA CCG CTG CGC CG – 3'
Set B:
FWD: 5' – GGT TCG GCT CAG CGA TGT GG – 3'
RVS: 5' – CAA GGG TCT CCC AGG ACC GC – 3'

<u>4-hydroxy-2-oxovalerate aldolase</u>:
Set A:
FWD: 5' – GAT GTT CGC CGT GGG CTT CC – 3'
RVS: 5' – GCC AGA TTG ATA TAG TCG GC – 3'
Set B:
FWD: 5' CGG GAA TAC CCA GGG TGA GG – 3'
RVS: 5' GGC TCA ACC ACA TCC TCC GC – 3'

<u>Acetaldehyde dehydrogenase</u>:
Set A:
FWD: 5' – CCA TTA AAG GTG TGG ACG CC – 3'
RVS: 5' – GGG CAG CCT TGA GGA TAG CG – 3'
Set B:
FWD: 5' – CCA TAC CCA TTG CCT ACG CC – 3'
RVS: 5' – CGA TGT TCG CCC GTG TCC CG – 3'

CHAPTER 4

Conclusions

As we discover more and more about the complexity of microorganisms, especially the intricacies of microbial populations and communities in diverse ecosystems such as the termite hindgut and the human body, research in microbial ecology requires asking and answering ***who*, *what*, *where*, *when*, *how*, *how much*, *with whom*, and *why*** (Mackie *et al*. 1999; Elowitz *et al*. 2002; Grozdanov *et al*. 2004; Hejnova *et al.* 2005; Maharjan *et al*. 2006; Tenaillon *et al*. 2010; Lencastre-Fernandes *et al*. 2011; Rosenthal *et al*. 2011; Abraham *et al*. 2012; Lebret *et al*. 2012; Ackermann 2013; Freeman *et al*. 2013; Jami *et al*. 2013; Koeppel *et al.* 2013; Watrous *et al*. 2013; Wessel *et al*. 2013; Rosenthal *et al*. *in review*). To address these questions, microbial ecologists should strive to combine and apply a myriad of traditional and cutting-edge bioinformatic, culturing, physiological, and molecular skill sets to *in vitro*, *in vivo*, and *in situ* work. Microbial ecology research should also be approached with an overall ecological awareness.

Explore diverse approaches to answer research questions

With efficiency and resources in mind, ***research approaches that combine several diverse techniques are effective*** in obtaining as complete a picture as possible of how microorganisms are influencing, and being influenced by, their physical surroundings. A common theme of my PhD work has been to test genomic-, metagenomic-, and bioinformatic-inspired hypotheses at the bench with traditional culturing and physiological, as well as cutting-edge molecular, tools. Prime

examples are my contributions to work led by Adam Rosenthal and Xinning Zhang investigating "ZnD2Sec," the phylotype that is responsible for the majority (40%) of the formate dehydrogenase expression in the termite hindgut (*see Chapter 1*) (Rosenthal *et al*. *in review*). Specifically, I combined enrichment work with microfluidic digital PCR to successfully identify the termite hindgut microbial community member that encodes the "ZnD2Sec" phylotype as a deltaproteobacteria (*see Chapter 1*). In addition, my physiological experiments investigating potential acetogenic demethylation by *Treponema azotonutricium* str. ZAS-9 were premised upon genomic observations (*see Chapter 2*). Moreover, my bioinformatic and physiological work investigating the *meta*-cleavage catabolic pathway capabilities of *T. primitia* str. ZAS-1 and ZAS-2 were inspired by bioinformatic-based findings (*see Chapter 3*) (Lucey & Leadbetter *in review*).

Regarding the *T. azotonutricium* str. ZAS-9 acetogenic demethylation project, while physiological experiments were inconclusive in determining whether or not the strain can bypass the methyl branch of the Wood-Ljungdahl pathway and generate acetate with methyl groups it takes from wood, it is possible that additional diverse research approaches might achieve more definitive results. First, it would be worthwhile to examine the relevant *T. azotonutricium* str. ZAS-9 gene and protein sequences for Pfam domains and residues required for functionality or, alternatively, mutations in these key domains and residues (*see Chapter 2*). Likewise, analyses of promoter, operon, and other regulatory elements could also provide insight into function. This hypothesis could also be re-evaluated by "priming" *T. azotonutricium* str. ZAS-9 that has been sub-cultured away from its

original environment for over a decade with meth(ox)ylated compounds to "kickstart" this metabolic capability and/or decreasing the concentration of meth(ox)ylated substrate tested as the 5mM concentrations examined here may not inhibit growth but may indeed inhibit enzyme function. Directly measuring acetate production in *T. azotonutricium* str. ZAS-9 culture fluid over time using GC-MS, ICS, and LCMS, and/or tracing radio-labeled methyl groups into possible acetate generated by the strain, could alternatives to measuring growth yield changes (*see Chapters 2 and 3*). qRT-PCR of relevant genes and cloning/expression/ enzyme activity assay work could also measure activity, expression, or lack thereof of these relevant genes (*see Chapter 2*).

If with additional work the hypothesis that *T. azotonutricium* str. ZAS-9 is able to bypass the methyl branch of the Wood-Ljungdahl pathway and generate acetate with methyl groups it takes from wood is not supported, at least such experiments demonstrate that ***an organism, even with a complete set of genes representing a particular function, does not necessarily perform that specific function*** (*see Chapter 2*).

Regarding the *meta*-cleavage pathway project, activity assays for enzymes related to oxidative stress protection were performed with crude cell extracts of *T. primitia* str. ZAS-1 and ZAS-2 upon their isolation in pure culture (Graber & Breznak 2004). In so doing, activities for the hallmark oxidative stress defense enzymes, catalase and superoxide dismutase, were not detected and suggested that the *T. primitia* strains are especially sensitive to O_2 (Graber & Breznak 2004). Nevertheless, once

genomic datasets of the *T. primitia* strains were obtained, analyses revealed genes representing a complete *meta*-cleavage aromatic catabolic pathway (*see Chapter 3*) (Rosenthal *et al*. 2011; Ballor *et al*. 2012; Lucey & Leadbetter *in review*). In particular, the first step of the *meta*-pathway is performed by catechol 2,3-dioxygenase and uses O_2 as a co-substrate to cleave catechol-like aromatic rings (*see Chapter 3*) (Lucey & Leadbetter *in review*). While individual pathway steps are still being evaluated, physiological work premised upon genomic and metagenomic data have demonstrated the first evidence for aromatic ring cleavage in the phylum (division) *Spirochetes*. Catechol 2,3-dioxygenase and the *meta*-pathway might also represent a previously unknown O_2-sink in the termite system. This is an example where ***back-and-forth between diverse approaches*** allowed for a novel discovery (*see Chapter 3*) (Lucey & Leadbetter *in review*). Furthermore, GC-MS, ICS, and LCMS measurements, experiments with radio-labeled *meta*-pathway substrates and intermediates, promoter, operon, and regulatory element determination, as well as qRT-PCR of *meta*-pathway genes, and cloning/expression/ enzyme activity assay work are additional, diverse approaches that could be utilized to further evaluate complete *meta*-cleavage pathway functionality (*see Chapter 3*).

In vitro and *in situ* behaviors can be different

By approaching research questions with diverse tools, microbial ecologists have the best opportunity to obtain a near complete picture of microbial behavior and experience *in situ*. Thoughtfully and creatively combining various approaches with an ecological awareness can also reconcile and/or minimize differences in microbial behavior and experience observed *in vitro* compared to *in vivo* or *in situ*. This is

because typically ***in vitro conditions are at best approximations of microorganisms' natural environments***.

In vitro conditions are typically suboptimal

One of my initial projects in the Leadbetter Lab was to develop a yeast autolysate preparation that improved growth rates of the *Treponema* termite hindgut isolates on 4YACo media (*see Chapter 1*). Upon isolation, *Treponema* doubling times ranged from 22 to 35hrs (Graber & Breznak 2004; Graber *et al*. 2004). After approximately a decade of continuous passage in 4YACo media, however, *Treponema* doubling times increased to between 54 and 77hrs. In media prepared with my new autolysate from Fleischmann's brand active dry yeast, doubling times of *T. primitia* str. ZAS-2 were restored to approximately 47hrs (*see Chapter 1*). While this is an improvement over recent doubling times, the termite hindgut turns over every 24 hours (Bignell 1984) and if *in situ* the *Treponema* isolates were to double every 47hrs, presumably they would be removed from the termite system. These observations, therefore, hint that *Treponema* grows with a doubling time under 24hrs *in situ* and that current culture conditions are suboptimal.

In vitro vs. *in situ* ecology

In addition to thoughtfully and creatively combining diverse research approaches, ***when studying complex microbial communities such as the termite hindgut microbial community and human microbiome, it is important to also approach research with an overall ecological awareness***. For instance, physiological experiments to evaluate acetogenic demethylation by *Treponema azotonutricium*

str. ZAS-9 and putative *meta*-cleavage catabolic pathway capabilities of *T. primitia* str. ZAS-1 and ZAS-2 by observing increases in growth yield of experimental cultures, were inconclusive (*see Chapters 2 and 3*) (Lucey & Leadbetter *in review*). When considering these three *Treponema* isolates in the context of the termite hindgut environment and their relationships with both their termite host and the other microbial community members, however, that possible acetate production did not result in increase growth rates and yields is not surprising (Graber & Breznak 2004; Graber *et al*. 2004).

This result may reflect possible *in situ* behavior by which acetate is not used by the organism that generates it, but instead is utilized by other members of the hindgut microbial community. For instance, it has been suggested that related to their symbiotic relationships with the termite gut microbial community and the termite host itself, *T. primitia* str. ZAS-1 and ZAS-2 and *T. azotonutricium* str. ZAS-9 do not convert substrates to products for rapid growth and efficient generation of cell material for themselves (Graber & Breznak 2004; Graber *et al*. 2004). Rather, they give the acetate they produce to the rest of the hindgut microbial community and their host, and display limited production of biomass (Graber & Breznak 2004; Graber *et al*. 2004). Therefore, perhaps better ways to gauge acetate production consist of measuring acetate production in *Treponema* culture fluid over time using ion chromatography, "feeding" cell-free *Treponema* culture fluid to other cultures that would demonstrate a marked increase in growth rate and yield from acetate, and observing co-culture experiments with *Treponema* and these additional organisms (*see Chapters 2 and 3*).

Similarly, while many *meta*-cleavage pathway-containing organisms perform preparatory or "upper" pathway reactions before proceeding to cleave aromatic substrates via their "lower" or *meta*-pathways, there is no evidence of "upper" pathway functionality in the genomes of *T. primitia* str. ZAS-1 and ZAS-2 (*see Chapter 3*) (Lucey & Leadbetter *in review*). While both *T. primitia* strains do not appear to have the ability to perform these preparatory reactions, it is quite likely that other organisms in the complex termite hindgut microbial community do and that they generate substrates for *T. primitia* (Lucey & Leadbetter *in review*). A useful experiment would be to feed *T. primitia* str. ZAS-1 and ZAS-2 cell-free culture fluid in lieu of aromatic-like substrates from an organism known to perform the "upper" pathway such as *Azotobacter vinelandii*, *Novosphingobium aromaticivorans*, *Novosphingobium* sp., *Methylocella silvestris*, *Sphingobium japonicum*, *Azoarcus* sp., *Pseudomonas putida*, *Thauera* sp., *Dechloromonas aromatica*, and *Methylibium petroleiphilum* (*see Chapter 3*).

It is well established that microbial activity can affect conditions in the surrounding environment (Brune 1998; Brune & Friedrich 2000; Watrous *et al*. 2013). For example, the *T. primitia* strains' catechol 2,3-dioxygenase may function in part to help remove O_2 from the termite hindgut and, in turn, make that environment more amenable to the anaerobes residing there (*see Chapter 3*). This could be tested by examining the success of co-cultures with *T. primitia* and a known strict anaerobe, such as *T. azotonutricium* str. ZAS-9, to which O_2 and aromatic substrates are added (*see Chapter 3*) (Rosenthal *et al*. 2011). When working with isolates from a complex microbial community, therefore, one should keep in mind that ***in its natural***

environment this organism may have peers who enhance or inhibit its function through their own activities (Brune 1998; Brune & Friedrich 2000; Watrous *et al*. 2013).

Despite the fact that *in vitro* can be very different from *in situ* behavior, however, I still advocate for enriching for and isolating, if possible, microorganisms from the natural environment. ***With more and more enrichments and isolates in hand, relationships in the natural environment can be better mimicked and manipulated at the bench via co-culture and microcosm experiments*** (Rosenthal *et al*. 2011).

***In vitro* vs. *in situ* physical parameters**

In pure culture *in vitro* work, in addition to isolating microorganisms from their potentially symbiotic microbial peers, ***physical parameters of cultures are often disparate from those of the isolate's natural environment. This may influence function***. In complex communities, it is important to consider that microorganisms can be located in a myriad of different microenvironments (Wessel *et al*. 2013). Termite hindgut prokaryotes, for example, are either free-swimming, attached to the gut epithelium, or in association with the intestinal protozoa (Brune 2006). Liquid, semi-solid, or solid media with different incubation regimes are rough approximations for *in situ* physical experiences, and culture design and set-up should attempt to replicate *in situ* conditions as best possible. Therefore, thoughtful and creative culture design and set-up should be explored and should involve providing various surfaces for growth, spikes of metabolites, fluctuations

temperature, pH, etc… depending on what is known about the isolate's natural environment. I have experience with this already regarding the finding that the deltaproteobacteria that encodes the "ZnD2Sec" phylotype in the termite hindgut is attached to protozoa, instead of being associated with wood or free-living (Rosenthal *et al*. *in review*).

Microorganisms that are classified together do not necessarily demonstrate the same physiology

My work with "ZnD2Sec" has taught me to ***remain open to the unexpected*** (*see Chapter 1*) (Rosenthal *et al*. *in review*). That "ZnD2Sec," the phylotype that is responsible for the majority of the formate dehydrogenase expression in the termite hindgut, is encoded by a deltaproteobacteria and not a spirochete was unexpected, as FDH is a key enzyme in the Wood-Ljungdahl pathway of acetogenesis and spirochetes have long been theorized as prominent acetogens in the termite hindgut (*see Chapter 1*) (Leadbetter *et al*. 1999; Rosenthal *et al*. *in review*). It is important to realize that ***not all members of a genus, species, or strain within the same environment share the same physiology*** (Elowitz *et al*. 2002; Lencastre-Fernandes *et al*. 2011; Abraham *et al.* 2012; Lebret *et al*. 2012; Ackermann 2013).

In my enrichment and isolation work to obtain methanogens and spirochetes in pure culture, I isolated a spirochete, *Treponema primitia* str. "ZNS-1," from the hindgut of *Zootermopsis nevadensis* (*see Chapter 1*). There is 100% 16S rRNA sequence identity between the *T. primitia* str. "ZNS-1" and *T. primitia* str. ZAS-2, and 99% 16S rRNA sequence identity between *T. primitia* str. "ZNS-1" and *T. primitia* str. ZAS-2, which is logical given that str. "ZNS-1" was isolated using the same media

used for str. ZAS-1 and ZAS-2 growth (*see Chapter 1*). Despite their shared 16S rRNA identity, however, these organisms may not necessarily have the same physiology. Further testing should be undertaken to compare growth on various substrates and under various conditions between this new isolate and *T. primitia* str. ZAS-1 and ZAS-2 (Graber & Breznak 2004; Graber *et al*. 2004; Lencastre-Fernandes *et al*. 2011; Abraham *et al.* 2012; Lebret *et al*. 2012; Ackermann 2013).

While it would be useful to obtain the genome of *T. primitia* str. "ZNS-1" and compare it to the available genomes of str. ZAS-1 and ZAS-2 to determine the similarities and differences between a strain of the same species recently isolated from the environment to those transferred in pure culture for over a decade, it may be difficult to discern if differences are indeed artifacts of natural vs. test tube environments, or because of inherent strain deviations (Barrick *et al*. 2009). It would instead be more valuable to obtain the genome of a new isolate right when that isolate is obtained, examine physiological parameters in pure culture, and then re-evaluate the genome and those physical parameters some years later (Barrick *et al*. 2009).

Looking ahead

Although the following questions were not addressed in, and were not directly applicable to, my PhD research, by keeping up-to-date with a variety of microbial ecology literature and brainstorming with colleagues I have learned that it is important when studying complex microbial communities, especially host/microbe interactions, to consider the following:

- Which microorganisms in an ecosystem are permanent and which are transient?

- How do transient microorganisms influence the other microorganisms, physical parameters, and/or the host of that ecosystem?

- While there is a lot of focus on how the microbiota influences its host, how is the host influencing its microbiota?

- What makes a host-associated microorganism act as a commensal vs. a pathogen? (Grozdanov *et al*. 2004; Hejnova *et al.* 2005; Tenaillon *et al*. 2010)

- Are numerically less-abundant microbial community members significant contributors to the ecosystem? (Freeman *et al*. 2013)

- Are microbial populations splitting into ecologically diverse populations? (Maharjan *et al*. 2006; Koeppel *et al*. 2013)

- How is an antibiotic, prebiotic, probiotic, metabolite, etc… added to the system affecting the host vs. the microorganisms? Are those effects similar or different for the host and microorganisms? (Cabreiro *et al*. 2013)

- What are developmental milestones of the model host organism being researched? (Mackie *et al*. 1999; Jami *et al*. 2013)

- Do host developmental milestones reflect microbial colonization and/or composition? (Mackie *et al*. 1999; Jami *et al*. 2013)

REFERENCES:

Abraham S, Gordon DM, Chin J *et al*. (2012) Molecular characterization of commensal *Escherichia coli* adapted to different compartments of the porcine gastrointestinal tract. *Applied and Environmental Microbiology*, **78**, 6799-6803.

Ackermann M (2013) Microbial individuality in the natural environment. *The ISME Journal*, **7**, 465-467.

Ballor NR, Paulsen I, Leadbetter JR (2012) Genomic analysis reveals multiple [FeFe] hydrogenases and hydrogen sensors encoded by treponemes from the H_2-rich termite gut. *Microbial Ecology*, **63**, 282-294.

Barrick JE, Dong SY, Yoon SH *et al*. (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature*, **461**, 1243-1247.

Bignell DE (1984) Direct potentiometric determination of redox potentials of the gut contents in the termites *Zootermopsis nevadensis* and *Cubitermes severus* and in 3 other arthropods. *Journal of Insect Physiology*, **30**, 169-174.

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

Brune A (2006) Symbiotic Associations Between Termites and Prokaryotes. In: *Prokaryotes* (eds Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E), pp. 439-474. Springer, New York.

Brune A, Emerson D, Breznak JA (1995) The termite gut microflora as an oxygen sink: microelectrode determination of oxygen and pH gradients in guts of lower and higher termites. *Applied and Environmental Microbiology*, **61**, 2681-2687.

Brune A, Friedrich M (2000) Microecology of the termite gut: structure and function on a microscale. *Currently Opinion in Microbiology*, **3**, 263-269.

Cabreiro F, Au C, Leung K *et al*. (2013) Metformin retards aging in *C. elegans* by altering microbial folate and methionine metabolism. *Cell*, **153**, 228-239.

Elowitz MB, Levine AJ, Siggia ER, Swain PS (2002) Stochastic gene expression in a single cell. *Science*, **297**, 1183-1186.

Freeman CJ, Thacker RW, Baker DM, Fogel ML (2013) Quality or quantity: is nutrient transfer driven more by symbiont identity and productivity than by symbiont abundance? *The ISME Journal*, **6**, 1116-1125.

Graber JR, Breznak JA (2004) Physiology and nutrition of *Treponema primitia*, an H_2/CO_2-acetogenic spirochete from termite hindguts. *Applied and Environmental Microbiology*, **70**, 1307-1314.

Graber JR, Leadbetter JR, Breznak JA (2004) Description of *Treponema azotonutricium* sp. nov. and *Treponema primitia* sp. nov., the first spirochetes isolated from termite guts. *Applied and Environmental Microbiology*, **70**, 1315-1320.

Grozdanov L, Raasch C, Schulze J *et al*. (2004) Analysis of the genome structure of the non-pathogenic probiotic *Escherichia coli* strain NIssle 1917. *Journal of Bacteriology*, **186**, 5432-5441.

Hejnova J, Dobrindt U, Nemcova R *et al*. (2005) Characterization of the flexible genome complement of the commensal *Escherichia coli* strain A0 34/86 (O38: K24: H31). *Microbiology*, **151**, 385-398.

Jami E, Israel A, Kotser A, Mzrahi I (2013) Exploring the bovine rumen bacterial community from birth to adulthood. *The ISME Journal*, **7**, 1069-1079.

Koeppel AF, Wertheim JO, Barone L *et al*. (2013) Speedy speciation in a bacterial microcosm: new species can arise as frequently as adaptations within a species. *The ISME Journal*, **7**, 1080-1091.

Leadbetter JR, Schmidt TM, Graber JR, Breznak JA (1999) Acetogenesis from H_2 plus CO_2 by spirochetes from termite guts. *Science*, **283**, 686-689.

Lebret K, Kritzberg ES, Figueroa R, Rengefors K (2012) Genetic diversity within and genetic differentiation between blooms of a microalgal species. *Environmental Microbiology*, **14**, 2395-2404.

Lencastre-Fernandes R, Nierychlo M, Lundin L (2011) Experimental methods and modeling techniques for description of cell population heterogeneity. *Biotechnology Advances*, **29**, 575-599.

Lucey KS, Leadbetter JR (2013) Catechol 2,3-dioxygenase and other *meta*-cleavage catabolic pathway genes in the "anaerobic" termite gut spirochete *Treponema primitia*. *In review (Molecular Ecology)*

Mackie RI, Sghir A, Gaskins HR (1999) Developmental microbial ecology of the neonatal gastrointestinal tract. *The American Journal of Clinical Nutrition*, **69**, 1035S-1045S.

Maharjan R, Seeto S, Notley-McRobb L, Ferenci T (2006) Clonal adaptive radiation in a constant environment. *Science*, **313**, 514-517.

Rosenthal AZ, Matson EG, Eldar A, Leadbetter JR (2011) RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. *The ISME Journal*, **5**, 1133-1142.

Rosenthal AZ, Zhang X, Lucey KS *et al*. Localizing transcripts to single cells suggests an ecological role for the deltaproteobacteria living on termite gut protozoa. *In review (PNAS)*

Tenaillon O, Skurnik D, Picard B, Denamur E (2010) The population genetics of commensal *Escherichia coli*. *Nature Reviews Microbiology*, **8**, 207-217.

Watrous JD, Phelan VV, Hsu C-C *et al*. (2013) Microbial metabolic exchange in 3D. *The ISME Journal*, **7**, 770-780.

Wessel AK, Hmelo L, Parsek MR, Whiteley M (2013) Going local: technologies for exploring bacterial microenvironments. *Nature Reviews Microbiology*, **11**, 337-348.