Diversity of Formyltetrahydrofolate Synthetases in the Guts of the Wood-Feeding Cockroach Cryptocercus punctulatus and the Omnivorous Cockroach Periplaneta americana

Elizabeth A. Ottesen†‡ and Jared R. Leadbetter*§

Division of Biology and Environmental Science & Engineering, W. M. Keck Laboratories, M/C 138-78, California Institute of Technology, Pasadena, California 91125

Received 3 February 2010/Accepted 10 May 2010

We examined the diversity of a marker gene for homoacetogens in two cockroach gut microbial communities. Formyltetrahydrofolate synthetase (FTHFS or fhs) libraries prepared from a wood-feeding cockroach, Cryptocercus punctulatus, were dominated by sequences that affiliated with termite gut treponemes. No spirochete-like sequences were recovered from the omnivorous roach Periplaneta americana, which was dominated by Firmicutes-like sequences.

The guts of wood-feeding termites and Cryptocercus punctulatus cockroaches share an unusual pattern of electron flow, as high rates of CO2-reductive acetogenesis typically supplant methanogenesis as the terminal electron sink (2, 3). Past studies have shown that from 10 to 30% of gut acetate produced in environments of termites and wood-feeding cockroaches is microbiologically generated from CO2 (3, 28), ultimately powering 18% to 26% of the host insect’s own respiratory energy metabolism (25). Nevertheless, most termites emit methane (2), and termite emissions constitute approximately 4% of the global methane budget (27). Cockroaches have been proposed to represent an additional source of note (9). Interestingly, methanogenic termites and cockroaches exhibit increased acetogenesis following addition of exogenous H2 (3, 29). This suggests that these insects are host to a robust population of bacteria that are capable of homoacetogenesis but may be primarily using alternative electron donors (and other substrates and pathways) in vivo.

Acetogenic bacteria belonging to two bacterial phyla, Firmicutes and Spirochaetes, have been isolated from the guts of termites (1, 4, 11, 12, 14). Several surveys have targeted and used the gene for formyltetrahydrofolate synthetase (FTHFS), a key gene in the Wood-Ljungdahl pathway of acetogenesis (16), as a potential marker for the pathway (15, 18). For the wood-feeding termites that have been examined, the studies have revealed an abundance of FTHFS sequences that form a coherent phylogenetic cluster, together with genes from homoacetogenic termite gut spirochetes belonging to the genus Treponema (24, 26, 30). This suggests that treponemes may be among the more abundant of the homoacetogens active in these environments.

Little is known about the population structure and biology of CO2-reducing, acetogenic bacteria in the guts of either omnivorous or wood-feeding cockroaches. The wood-feeding cockroach Cryptocercus hosts an abundance of flagellate protozoa closely related to those believed to dominate polysaccharide fermentation in the guts of termites (5, 6, 22), suggesting that at least one key environmental niche is filled by similar microbes in both termites and Cryptocercidae. Additionally, Cryptocercidae cockroaches, like termites, house diverse spirochetes and are the site of intense CO2 reduction into acetate (3, 7). Possibly, spirochetes capable of CO2 reduction into acetate are present in the hindguts of cockroaches. However, no evidence has yet been presented for the existence of homoacetogenic treponemes in environments other than the guts of termites, and FTHFS surveys of human (21) or horse (15) fecal matter and bovine rumen samples (20) revealed only Firmicutes-like and other FTHFS alleles that are distinct from those comprising the termite treponeme cluster.

Here, by examining FTHFS gene diversity in Cryptocercus punctulatus and Periplaneta americana guts, we endeavored to learn more about the distribution and origins of homoacetogenic treponemes (and their genes) that are found in wood-feeding termites. In particular, we wished to ascertain whether FTHFS genes present in either of the two cockroaches are termite treponeme-like and, if so, whether analysis reveals any obvious signal indicating recent or ancient lateral community transfer events between insect lineages.

Analysis of FTHFS diversity in the guts of cockroaches. We catalogued FTHFS gene sequence diversity present in the guts of a wood-feeding cockroach of the family Cryptocercidae (Cryptocercus punctulatus) and an omnivorous cockroach of the family Blattidae (Periplaneta americana) in an effort to shed light on the evolutionary origins of the homoacetogenic spirochetes thought to dominate acetogenesis in the guts of wood-feeding termites.

Library preparation and sequence analysis. FTHFS gene diversity was examined in DNA extracted from pooled whole guts of two P. americana adults, from a single whole-gut sample
TABLE 1. Abundance of FTHFS sequence clades in libraries constructed from \textit{P. americana} and \textit{C. punctulatus} gut extracts\textsuperscript{a}

<table>
<thead>
<tr>
<th>Host insect</th>
<th>Abundance of FTHFS sequence clades</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treponeme affiliated</td>
</tr>
<tr>
<td>\textit{P. americana}</td>
<td>0</td>
</tr>
<tr>
<td>\textit{C. punctulatus} adult</td>
<td>88</td>
</tr>
<tr>
<td>\textit{C. punctulatus} nymph</td>
<td>50</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Clades defined as marked in Fig. 1 and 2. Abundance given as percentage of full-length clones.

Termite treponemas and the “\textit{Treponema-affiliated}” FTHFS cluster. Gut bacteria carrying termite treponeme-like FTHFS genes are commonly encountered in termite gut communities. In the four termite species (representing four of the six termite families) examined previously, sequences that affiliated with the termite treponeme group were more frequently encountered than were those from putative homoacetogenic \textit{Firmicutes} (24, 26, 30). Here, this pattern is shown to extend to the flagellate- and spirochete-rich hindgut communities of the wood-feeding cockroach, \textit{C. punctulatus} (Fig. 1). In contrast, treponeme-like FTHFS sequences were not identified in libraries prepared from \textit{P. americana}.

The termite treponeme cluster of FTHFS gene sequences is defined as encompassing FTHFS alleles from termite \textit{Treponema} isolates and termite gut environmental sequences that phylogenetically affiliate with them and share a conserved hexapeptide insert (26). Two major groups of FTHFS alleles (cockroach groups I and II) from \textit{C. punctulatus} fell within the termite treponeme cluster (Fig. 1). These cockroach-derived sequences were distinct from the FTHFS variants that originated from termite hindgut communities, with less than 93% amino acid similarity to the most closely related termite-derived sequences. Cockroach group II comprised three alleles affiliated with a set of FTHFS gene sequences (from \textit{Zootermopsis nevadensis} and dry-wood termites) lying at the base of the termite treponeme cluster. The FTHFS alleles that comprise “cockroach group I” are more numerous than those representing cockroach group II and fell well within the termite treponeme FTHFS radiation. The robustness with which the cockroach group I sequences cluster to the exclusion of termite community-derived sequences, and the long branch length separating this group from the others, is inconsistent with it having arisen after a recent lateral acquisition from termites. However, the short branch lengths between the alleles comprising this group, compared to most others on the tree, including those of cockroach group III, suggest that a breakout radiation of diversity has occurred in cockroach group I after either a lateral acquisition event occurred or an evolutionary bottleneck was resolved. As a result, lateral transfer of cockroach group I genes, or the organisms that encode them, from a termite hindgut community cannot be ruled out.

A cluster of \textit{C. punctulatus} FTHFS sequences, “cockroach group III,” represents a sister branch to the termite \textit{Treponema} clade. This group fell at the base of the termite treponeme cluster (to the exclusion of all other deduced FTHFS proteins available for analysis), using all three of the treeing algorithms employed. This group lacked the hexapeptide insert (Fig. 1, right panel) previously used to define the termite treponeme cluster (26) but was classified as treponeme-affiliated on the strength of its phylogenetic affiliation with the termite treponeme cluster. This cluster had a high level of within-group diversity, as evidenced by branch lengths. If FTHFS variants belonging to this cluster of sequences were present in the gut communities of the last common ancestor of termites and cockroaches, they appear to have subsequently gone extinct prior to the radiation of the four (of the six) lower termite families that have been examined. The position of this clade basal to the termite treponeme assemblage and its absence from all of the phylogenetic breadth of termites thus far examined suggest that the organisms encoding the cockroach
group III alleles should be targets for future study. In particular, it would be of interest to confirm whether these organisms are indeed homoacetogens and/or spirochetes and to examine any similarity that they have to homoacetogenic termite treponemes.

**Firmicutes homoacetogens and “Lovell cluster A.”** A number of sequences that fell outside the treponeme-affiliated cluster but within the “Lovell cluster A” of potential acetogens (18) were identified in the three cockroach-derived libraries (Fig. 2; see Fig. S2 in the supplemental material). This cluster incorporates all currently sequenced Firmicutes acetogens. However, with the addition to the analysis of FTHFS genes from newly sequenced microbial genomes, Lovell cluster A can no longer be considered to represent a single, monophyletic “Firmicutes acetogens” cluster. That notwithstanding, within this cluster, three groups with high representation in our clone libraries were identified.

A total of 41% of all clones recovered from the omnivorous cockroach *P. americana* and ca. 2% of those identified in the *C. punctulatus* nymphs clustered phylogenetically within the Lovell cluster A, with an FTHFS sequence from *Clostridium* sp. M62/1, a butyrate-producing Firmicutes strain originally isolated from human feces (17). This group represents the most abundant (and most OTU-rich) cluster of FTHFS sequences identified in the gut of the omnivorous cockroach *P. americana*. To our knowledge, *Clostridium* sp. M62/1 has not yet been demonstrated experimentally to be capable of homoacetogenic growth. The FTHFS gene in this organism is located in close proximity to a putative glutamate formiminotransferase and formimidoyltetrahydrofolate cyclodeaminase, which suggests that may be playing a role in other pathways, e.g., in histidine metabolism. However, as this cluster of FTHFS variants was found to be the most highly diverse and abundant sequence cluster in gene libraries of *P. americana*, it may represent the best candidate for the resident population of acetogenic bacteria in that insect’s gut environment. CO2-reductive acetogenesis has previously been demonstrated to occur at low yet significant rates in this insect (3, 13).

**FIG. 1.** Phylogenetic analysis of termite treponeme and treponeme-affiliated FTHFS sequences. (Left) Tree constructed using PHYML algorithm and 343 unambiguous, aligned amino acid positions. Open circles mark groupings also supported by either Phylip PROTPARS parsimony or Fitch distance algorithms. Closed circles identify clusters grouped by all three treeing methods. Scale bar represents 0.1 amino acid changes per alignment position. The tree was rooted using FTHFS sequences from *Ruminococcus* and *Clostridium* sp. M62/1 clades in Fig. 2. (Right) A highly variable region of the protein sequence, corresponding to residues 229 to 234 in *Moorella thermoacetica* (ABC18448.1). Each line of the alignment corresponds to a sequence in the tree to the left.
Two alleles representing 31% of the \textit{P. americana} library clustered with “Clone F” and “Clone Y” FTHFS sequences previously identified in a study on the dampwood termite \textit{Z. nevadensis} and also affiliated more remotely with FTHFS sequences from two \textit{Actinobacteria} species (\textit{Eggerthella lenta} and \textit{Cryptobacterium curtum}). These alleles are borne by organisms of uncertain phylogenetic position and appear to be present at various abundances in the guts of diverse termites and cockroaches.

One final FTHFS clade identified in this study is of potential interest, as it encompasses a diverse group of alleles that group with \textit{Z. nevadensis} clone \textit{E} and \textit{Reticulitermes santonensis} clone \textit{Rs10} (24, 26). Sequences affiliated with this group were identified in both \textit{C. punctulatus} libraries (Table 2; see Fig. S2 in the supplemental material) but were far more abundant in the library prepared from the nymph specimens (41%) than in that prepared from the adult (2%) specimens. The significance of this difference is hard to infer at this juncture, as these were insects from different colonies, geographical locations, and developmental stages. However, the magnitude of this difference suggests that a comprehensive exploration of the variability of FTHFS community composition in \textit{Cryptocercidae} may present a fruitful line of inquiry.

**Implications.** The hindgut microbial communities of both the omnivorous cockroach \textit{Periplaneta americana} and the wood-feeding cockroach \textit{Cryptocercus} bear a variety of FTHFS genes, however, the phylogenetic profiles of the genes from these two cockroaches are entirely distinct. The FTHFS alleles from the omnivorous cockroach appear to share more commonalities with those from vertebrate gut tracts, whereas those from \textit{C. punctulatus} appear to share more commonalities with those from termite gut communities. In particular, many of the FTHFS alleles recovered from \textit{C. punctulatus} are postulated to be encoded by homoacetogenic spirochetes based on phylogenetic inference. Similar alleles were not recovered from \textit{P. americana}, suggesting that homoacetogenic spirochetes are absent from the gut community of omnivorous cockroaches. This suggests that the emergence of a microbial community rich with homoacetogenic spirochetes may have been coincident with the evolution of the wood-feeding lifestyle and the widespread success of that host and its subsequent lineages. Given that the majority of cockroach species examined have been
methanogenic (9), it seems likely that the common ancestor of Blattidae and Cryptocercidae had a gut fermentation dominated by methanogenesis as the dominant electron sink. Since methane essentially represents calories lost to the nutrition of the host, the acquisition and/or evolution of the first homoaceticogenic treponemes may have helped drive the subsequent success of the progenitor of the Cryptocercidae cockroaches and the termites, during the transition from an omnivorous to lignocellulose-based diet.

Nucleotide sequence accession numbers. Sequences generated in this study have been deposited into GenBank under accession numbers GU552320 through GU552435.

This research was supported by the NIH (Caitech subcontract of R01-HG002644 and the NIH/NRSA Training Grant for Biology 5T32GM07616).

We are grateful to Christine Nalepa of the North Carolina Department of Agriculture for collecting and supplying us with viable C. punctulatus specimens. We thank Jennifer Ottesen and our laboratory colleagues for their helpful comments and suggestions during the development of the manuscript.

REFERENCES