# Discovery and genomics of $H_2$-oxidizing/$O_2$-reducing *Deferribacterota* ectosymbiotic with protists in the guts of termites and a *Cryptocercus* cockroach

**Naoya Maruoka[1], Rinpei Kudo[1], Katsura Igai[1], Michiru Shimizu[2], Masahiro Yuki[2], Moriya Ohkuma[2], Yuichi Hongoh[1,2*]**

[1]School of Life Science and Technology, Institute of Science Tokyo, Tokyo, Japan

[2]Japan Collection of Microorganisms/Microbe Division, RIKEN BioResource Research Center, Ibaraki, Japan

**Running headline: *Deferribacterota* attached to protists**

* Corresponding author: yhongo@life.isct.ac.jp

E-mail: maruoka.n.aa@m.titech.ac.jp; bananariyasu@gmail.com; kigai@life.isct.ac.jp; michiru.tahara@riken.jp; masahiro.yuki@a.riken.jp; mohkuma@riken.jp

**Keywords: *Deferribacterota*, protists, termites, *Cryptocercus*, symbiosis, genomics; gut bacteria**

**Abstract**

Members of the phylum *Deferribacterota* inhabit diverse environments, but their cellular symbiosis with protists has never been reported. We discovered an ectosymbiotic clade of *Deferribacterota* specifically associated with spirotrichonymphid protists in the guts of the termites *Reticulitermes speratus* and *Hodotermopsis sjostedti* and trichonymphid protists in the gut of the wood-feeding cockroach *Cryptocercus punctulatus*. These ectosymbionts were spiral rods, attached to 16–91% of the host protist cells. These formed a monophyletic cluster within an uncultured family-level clade exclusively comprising insect gut-associated lineages. The complete genome of an ectosymbiotic *Deferribacterota*, CpT32-4, was obtained from a *Trichonympha acuta* cell in a *C. punctulatus* gut and analyzed with a single-cell amplified genome of another ectosymbiotic *Deferribacterota*, RsTz2-092, associated with *Holomastigotes* sp. in the gut of *R. speratus*. The genome analyses suggested that these *Deferribacterota* ferment monosaccharides and conduct fumarate and oxidative respiration with $H_2$ as an electron donor. Thus, these bacteria possibly contribute to the removal of hydrogen and oxygen to protect the fermentative activity of the protist hosts. The ectosymbionts exhibited reduction in the number of genes related to signal transduction, implying that the association has provided a relatively stable environment for these bacteria. Notably, these ectosymbionts likely possess flagella with an unusually expanded number of flagellin variants up to 40, which may reflect an adaptation to their ectosymbiotic lifestyle. We propose the novel genus "*Candidatus* Termitispirillum" for these ectosymbionts and the novel family "*Candidatus* Termitispirillaceae" for the insect-gut clade. Our findings provide new insights into the ecology and evolution of *Deferribacterota*.

## Introduction

*Deferribacterota* is a phylum of Gram-negative, curved to spiral rod-shaped bacteria [1, 2], currently comprising one class (*Deferribacteres*), one order (*Deferribacterales*), and seven families (*Calditerrivibrionaceae*, *Deferribacteraceae*, *Deferrivibrionaceae*, *Flexistipitaceae*, *Geovibrionaceae*, *Mucispirillaceae*, and "*Candidatus* Microvillispirillaceae") [3–5]. The first five families have been isolated from a wide variety of environments, including deep-sea hydrothermal vents [6], activated sludges [7], terrestrial hot springs [8], and subsurface oil reservoirs [9], whereas the latter two families have exclusively been discovered in animal intestines.

*Mucispirillaceae* inhabit the guts of vertebrates, including rodents [10], chickens [11], and turkeys [12]. *Mucispirillum schaedleri*, a sole cultured species in the family *Mucispirillaceae*, inhabits the guts of rodents and is reportedly involved in host inflammation [10, 13]. "*Candidatus* Microvillispirillaceae" is an uncultured family represented by a species, "*Candidatus* Rimicarispirillum atlantis", which occupies the space among microvilli of the ectoperitrophic space in the midgut of the deep-sea hydrothermal vent shrimp genus *Rimicaris* [5, 14, 15]. "*Candidatus* Rimicarispirillum atlantis" can degrade chitin and presumably contributes to the turnover of chitin released during shrimp molt [5, 16]. Members of *Deferribacterota* have also been detected by 16S rRNA amplicon sequencing and metagenomic analyses as minor residents in the guts of insects, including the Auckland tree weta [17], cockroaches [18, 19], and termites [20–22].

Termites thrive solely on dead plant matter with the aid of their hindgut microbiota [23]. In "lower" termites (i.e., termites except for the family Termitidae) and their closest relative, the cockroach genus *Cryptocercus*, the digestion of lignocellulose is largely attributable to the gut protists belonging to either the phylum *Parabasalia* or *Preaxostyla* [24, 25]. Most of these protists harbor symbiotic prokaryotes in their cytoplasm, nucleoplasm, and/or on the cell

surface [26]. It has been suggested that these prokaryotic symbionts contribute to the provision of nitrogenous compounds [27–34], and/or hydrogen removal to promote fermentative activities in the gut [29, 35–38].

During our previous single-cell genomic survey of the bacterial microbiota in the gut of the termite *Reticulitermes speratus* [39], a single-cell amplified genome (SAG), designated MS499-41, was obtained and phylogenetically assigned to *Deferribacterota*. Our preliminary fluorescence in situ hybridization (FISH) analysis implied that the bacterium is an ectosymbiont of a gut protist, whereas cellular symbiosis between *Deferribacterota* and protists has never been reported thus far. Thus, in the present study, we aimed to clarify the localization and phylogenetic diversity of *Deferribacterota* in the guts of termites and cockroaches as well as their metabolic capacities and symbiotic roles based on their genome sequences. This study provides evolutionary and functional perspectives on the novel, uncultured *Deferribacterota* clade associated with cellulolytic gut protists.

## Materials and Methods

### Termite and cockroach collection

The Japanese subterranean termite *R. speratus* (family Rhinotermitidae) was collected in Tsukuba, Ibaraki Prefecture, and Ogose, Saitama Prefecture in Japan. The damp-wood termite *Hodotermopsis sjostedti* (family Archotermopsidae) was collected in Yakushima, Kagoshima Prefecture in Japan. The wood-feeding cockroach *Cryptocercus punctulatus* was collected in the Black Rock Mountain State Park, Georgia, U.S.A. The termites and cockroaches were kept with fragments of their nest logs in the laboratory.

### 16S rRNA gene amplicon sequencing and phylogenetic analysis

DNA was extracted from the entire guts of two *C. punctulatus* adults individually using the Power Soil DNA Isolation Kit (MoBio Laboratories). The V3–V4 regions of 16S rRNA (ca. 450 bp) genes were amplified by PCR using the prokaryote-universal primer set Pro341F and Pro805R (Table S1). The library preparation, sequencing on an Illumina MiSeq platform, and amplicon data processing were conducted as previously described [40]. The amplicon sequence variants (ASVs) obtained in this study and our previous study using 60 termite and eight cockroach species [40] were phylogenetically classified using SINA v1.7.2 [41] with the database SILVA SSURef NR99 release 138.2. The ASVs assigned to *Deferribacterota* were aligned with reference sequences using MAFFT v7 [42]. Ambiguously aligned regions were trimmed using trimAL v1.2 [43]. Maximum likelihood trees were constructed with 1,000 ultrafast bootstrap resampling using IQ-TREE v1.6.12 [44]. The best-fit nucleotide substitution models were selected using ModelFinder [45].

**FISH**

The entire guts were removed from worker termites and adult cockroaches, and the gut contents were suspended in Trager's solution U [46]. The samples were fixed in 4% paraformaldehyde for over 8 hrs at 4°C, rinsed twice with double-distilled water, and placed on MAS-GP slide glasses (Matsunami, Japan). FISH targeting 16S rRNA was performed as previously described [47, 48] using oligonucleotide probes labeled with Texas Red or 6-carboxyfluorescein (6FAM). Specific oligonucleotide probes (Table S2) were designed using ARB [49]. The specimens were observed under an Olympus BX51 epifluorescence microscope.

**Collection of protist cells, Sanger sequencing of small subunit (SSU) rRNA genes, and phylogenetic analysis**

Termites and cockroaches were dissected, and the gut contents were suspended in sterile Trager's solution U containing 0.01% bovine serum albumin (BSA). Single protist cells were collected in PCR tubes, respectively, using a TransferMan NK2 micromanipulator (Eppendorf) equipped with a hand-made glass capillary. The collected protist cells were subjected to whole genome amplification (WGA) using a GenomiPhi V2 DNA Amplification Kit (GE Healthcare Life Science). Near full-length 16S and 18S rRNA genes were amplified from the WGA products by PCR using Phusion High-Fidelity DNA polymerase (New England Biolabs), respectively. PCR was performed with initial denaturation at 98°C for 3 min, followed by 25 cycles of denaturation at 98□ for 30 sec, annealing at 50–70°C for 30 sec, and extension at 72□ for 1 min with a final extension at 72□ for 10 min. For *Holomastigotes* samples, nested PCR was performed for the amplification of the 18S rRNA genes using primer sets listed in Table S1 and the above program. The PCR products were cloned using a Zero Blunt TOPO PCR Cloning Kit (Thermo Fisher Scientific) after purification using a MonoFas DNA Purification Kit I (ANIMOS, Japan). Sanger sequencing of randomly chosen clones was conducted using a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) on an ABI 3730 Genetic Analyzer. The SSU rRNA gene trees of *Deferribacterota* and their host protists were constructed as described above, respectively.

**Genome sequencing and assembly**

A SAG of *Deferribacterota* (MS499-41: DRR253682) was obtained in our previous study from the gut content of *R. speratus* using the MDA-in-AGM method and an Illumina MiSeq platform with a MiSeq Reagent Kit V3 (600 cycles) [39]. In the present study, the MiSeq raw reads were subjected to quality trimming using Cutadapt 3.7 [50] and Prinseq-lite 0.20.4 [51] and were assembled using SPAdes v3.14.0 [52] with the --sc option. Contaminating contigs

were identified and removed using the Contig Annotation Tool 5.2.3 [53] and BLASTn searches [54]. Only contigs >500 bp were retained.

To acquire the genome sequence of *Deferribacterota* associated with *Trichonympha acuta*, the hindgut content of *C. punctulatus* was suspended in sterile Trager's solution U containing 0.01% BSA, and the membrane portion of a single *T. acuta* cell was separated using a Micro Feather Blade K-715 attached to the micromanipulator and collected in a PCR tube. The sample was subjected to WGA using EquiPhi29 DNA Polymerase (Thermo Fisher Scientific) as previously described [55]. The WGA product was purified using a Zymo Genomic DNA Clean & Concentrator kit (ZYMO Research). Sequencing libraries were prepared using a QIAseq FX DNA Library Kit (QIAGEN) for the MiSeq platform and a Native Barcoding Kit 24 V14 for a MinION platform (Oxford Nanopore Technologies). Prior to the latter library preparation, debranching, single-strand DNA digestion, purification, and size selection that retained fragments of >2 kbp were conducted as previously described [55]. Sequencing was performed on the MiSeq platform as above and on the MinION platform with a FLO-MIN114 (R10.4.1) flow cell. After base calling using dorado and quality-filtering using nanoq [56], the MinION reads were assembled using Raven [57]. The resulting circular contig was polished one time using Pilon v1.24 [58] with the quality-filtered MiSeq reads.

The genome quality was assessed using CheckM2 [59]. Circular genomes were visualized using Proksee [60]. The average nucleotide identity (ANI) and average amino acid identity (AAI) were calculated using FastANI [61] and EzAAI [62], respectively. The Skew Index (SkewI), which indicates the degree of GC skew, was calculated using SkewIT [63].

**Phylogenomic analysis**

The phylogenomic analysis was conducted using GToTree [64] as the following process. Seventy-four single-copy marker genes conserved among *Bacteria* were extracted from

*Deferribacterota* genomes using Prodigal [65] and HMMER [66]. Only genomes that contained ≥50% of the marker genes were retained as references. The deduced amino acid sequences were aligned using MUSCLE [67] and trimmed using trimAL v1.2. A maximum likelihood tree was constructed using IQ-TREE as above. The phylogenetic classification was also confirmed using the Genome Taxonomy Database (GTDB) [68].

**Functional annotation**

The reconstructed genomes were annotated using DFAST [69], KAAS [70], and DIAMOND [71]. Operon structures were predicted using Operon-mapper [72]. CRISPR-Cas systems were identified using CRISPRCasFinder (https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index). Protein-coding sequences (CDSs) were assigned to clusters of orthologous groups (COGs) using RPS-BLAST v2.13.0+ [73]. Principal component analysis (PCA) was performed using the sklearn.decomposition module of the Scikit-learn library [74]. Orthologous genes among the *Deferribacterota* genomes were identified using OrthoFinder [75]. Principal coordinate analysis (PCoA) based on orthogroup composition was performed using Phyloseq [76]. The functional annotation of each orthogroup was performed using GhostKOALA [77]. The orthologous genes characteristically associated with a phylogenetic group were identified using Scoary2 [78]. A maximum likelihood tree of the flagellin genes was constructed based on the deduced amino acid sequences with >350 aa as described above.

**Results**

**Phylogenetic diversity and relative abundance of *Deferribacterota* in termite and cockroach guts**

The phylogenetic diversity and relative abundance of *Deferribacterota* in the guts of termites and cockroaches were examined based on the 16S rRNA gene amplicon data derived from the guts of diverse termite and cockroach species. We identified 72 ASVs assigned to *Deferribacterota* from 47 out of the 68 termite and cockroach species (Fig. S1). Their relative abundance was consistently low (0.00–1.29%). These ASVs formed a monophyletic cluster with sequences derived from other insect guts (Fig. S1). The corresponding 16S rRNA region of the SAG MS499-41 was identical to the sequence of ASV13241 detected in the guts of *R. speratus* (relative abundance: 0.05%) and *Hodotermopsis* sp. (0.06%) (Fig. S1). A near full-length 16S rRNA sequence, RsTz2-092 (TACW01000013), previously obtained from the same *R. speratus* gut sample, showed 99.9% similarity to that of MS499-41. The bacterial species represented by the SAG MS499-41 and ASV13241 was designated as RsTz2-092 in the present study.

**Localization, morphology, and infection rate**

We designed probe RsTz2-092-190 (Table S2), which is specific to the 16S rRNA of RsTz2-092 and potentially its close relatives (Fig. S1). FISH using this probe detected curved or spiral-shaped long rods, which were specifically attached to the cell surface of the parabasalid protist *Holomastigotes* sp. (family *Spirotrichonymphidae*) in the gut of *R. speratus* (Fig. 1A, B). In addition, morphologically similar bacterial cells were detected with this probe specifically on the cell surface of *Holomastigotes* sp. (Fig. 1C, D) and *Brugerollina cincta* (*Spirotrichonymphidae*) (Fig. 1E, F) in the gut of *Hodotermopsis sjostedti*. These ectosymbionts appeared to be randomly distributed across the host cell surface (Fig. 1B, D, F) together with spirochaetes, and their cell dimensions were ca. 3–10 µm in length and 0.4 µm in width. Among the cells of the protist host species, 15.8–91.4% harbored the ectosymbionts detected with this probe (Table 1).

We then designed probe Deferri-term-661 (Table S2), which is specific to most members of the insect-gut clade, except for RsTz2-092 (Fig. S1). When FISH with this probe was performed for the gut content of *C. punctulatus*, spiral cells with dimensions 3–13 µm in length and 0.4 µm in width were detected, together with spirochaetes, specifically on the anterior surface area of the parabasalids *Trichonympha lata* (Fig. 1G, H) and *Trichonympha acuta* (Fig. S2A, B) (family *Trichonymphidae*). Of the *T. lata* and *T. acuta* cells in a *C. punctulatus* gut, 67.6% harbored the ectosymbiotic *Deferribacterota* (Table 1). No specific signals were detected in the gut contents of *R. speratus* and *H. sjostedti* with this probe.

**Phylogenetic analysis based on near full-length 16S rRNA genes**

We collected single cells of these host protist species, i.e., *Holomastigotes* spp., *B. cincta*, *T. acuta*, and *T. lata*, based on their morphological characteristics and performed WGA. Near full-length 18S rRNA and 16S rRNA genes were amplified by PCR, respectively, and the products were cloned and subjected to Sanger sequencing. We confirmed the taxonomic assignment of the host protist species based on the 18S rRNA genes (Fig. S3A, B). We identified the 16S rRNA genes of *Deferribacterota* among sequenced clones, which formed a monophyletic cluster within the insect-gut clade (Fig. 2A). The clone RsHol6-1, obtained from *Holomastigotes* sp. in an *R. speratus* gut, was nearly identical to the sequence of RsTz2-092. The clones HsHol1-2 from *Holomastigotes* sp. and HsBr1-3 from *B. cincta* in the gut of *Hodotermopsis sjostedti* were identical and showed 98.7% similarity to RsHol6-1 and RsTz2-092 (Fig. 2A). The clones CpT15-2 and CpT06-9 were obtained from *T. acuta* and *T. lata*, respectively, in the gut of *C. punctulatus*. These two sequences shared 97.9% identity with each other (Fig. 2A) and showed ≥99.3% similarities to ASV149, which was detected in one *C. punctulatus* gut (relative abundance: 0.21%) (Fig. S1).

**General features of the RsTz2-092 (MS499-41) and CpT32-4 genomes and phylogenomics**

In addition to the SAG MS499-41 (1,578,178 bp, 79% estimated completeness, 44% GC content) of RsTz2-092 from the gut of *R. speratus*, a circular chromosome (2,151,577 bp, 40% GC content) was reconstructed from a single *T. acuta* cell sample (Table 2). The 16S rRNA gene sequence of this genome shared 98.0% and 98.4% identity with CpT15-2 and CpT06-9, respectively. The bacterium represented by this genome is referred to as CpT32-4 hereafter (Fig. 2A). The total sequence length and GC content of the RsTz2-092 and CpT32-4 genomes were similar to those of other insect-gut clade species (Fig. S4). The GC skew pattern of the CpT32-4 genome is distinct from other genomes of *Deferribacterota*; the SkewI of CpT32-4 is only 0.176, whereas the others range 0.733–0.914 (Fig. S5). The CpT32-4 genome encodes CRISPR and associated genes (Table S3), whereas they were not found in the RsTz2-092 genome.

The genomes of RsTz2-092 and CpT32-4 were most closely related to metagenome-assembled genomes (MAGs) from the guts of "higher" termites (i.e., termites belonging to Termitidae) [79, 80], which generally do not harbor cellulolytic gut protists. These and two additional MAGs from lower termites (Pcl387_bin.162 and Ly174_bin.14) [79] formed a monophyletic cluster (Fig. 2B). This clade corresponds to the uncultured family "JAISNW01" in the GTDB and is phylogenetically sister to the vertebrate-gut inhabitants *Mucispirillaceae* (Fig. 2B). The ANI and AAI values between the RsTz2-092 and CpT32-4 genomes were <70% and 62.0%, respectively (Fig. S6A, B). Thus, these two bacteria are distinct species but in the same genus according to the proposed criterion of genus-level similarity (ANI ≤ 95%; AAI ≥ 60–85%) [61, 81].

**Metabolic capacities of CpT32-4 and RsTz2-092**

*Energy conservation and carbon metabolism*

The CpT32-4 genome contains a complete set of genes for glycolysis and gluconeogenesis (Fig. 3A, B). CpT32-4 likely imports monosaccharides and sugar acids, such as fructose, xylose, mannose, N-acetylglucosamine, and uronate, via phosphotransferase systems (PTS), ABC transporters, or transporters of major facilitator superfamilies (MFS). It also likely imports and utilizes glycerol, lactate, citrate, and unidentified carboxylates. Citrate can be metabolized to acetate and oxaloacetate by the action of citrate lyase (CitDEF) (Fig. 3A). It is predicted that pyruvate is fermented to acetate, ethanol, lactate, and possibly $H_2$ using monomeric iron hydrogenase B1/B3 and electron-bifurcating iron hydrogenase HydABC (Fig. 3A, B). The RsTz2-092 genome (MS499-41) lacks two genes in the glycolysis/gluconeogenesis pathway and all genes encoding the transporters for sugars and other carbon sources, except for those transporting uronates and carboxylates (Figs 3A, B, and S7), possibly due to its low completeness (Table 2). Thus, the main carbon source of RsTz2-092 remains unclear. No genes encoding hydrolytic enzymes for extracellular polysaccharides were identified in either genome.

In addition to substrate-level phosphorylation through glycolysis and oxidation to acetate, CpT32-4 has the potential to generate ATP via the agumatine deiminase pathway with the agmatine/putrescine antiporter AguD [82] (Fig. 3A). The CpT32-4 genome is equipped with several systems to generate $H^+/Na^+$ membrane potential to produce ATP via $FoF_1$-type ATPase. The genome encodes the membrane-bound [NiFe]-hydrogenase HybOABC, the Rnf complex, $Na^+$-transporting oxaloacetate/methylmalonyl-CoA decarboxylase (OadABC), and the $H^+/Na^+$ antiporter NhaB. It also encodes fumarate reductase (FrdABCD) with the fumarate/succinate antiporter DcuB and the succinate exporter DcuC, and in addition, *cbb*$_3$-type cytochrome *c* oxidase (Fig. 3A). All these components except for NhaB and FrdABCD were identified in the RsTz2-092 genome. Thus, these bacteria likely use $H_2$ as an electron

donor and fumarate and oxygen as electron acceptors for respiration. Considering the very low concentration of oxygen in the hindguts of termites and cockroaches, except for their peripheral regions [83], *cbb*$_3$-type cytochrome *c* oxidase may primarily function in oxygen removal to protect anaerobic reactions. The RsTz2-092 genome encodes a gene for periplasmic octaheme nitrite reductase (Onr) (Figs. 3A and S7), which possibly contributes to the detoxification of nitrite by dissimilatory reduction to ammonium, as suggested in *Desulfovibrio* species [84]. The genomes of these ectosymbionts contain several other genes for oxygen tolerance (Fig. S7).

*Nitrogen metabolism*

No genes responsible for nitrogen fixation were identified in the CpT32-4 and RsTz2-092 genomes. Both genomes encode several transporters for amino acids and di/oligo-peptides, which are likely the main nitrogen sources (Fig. 3A). The CpT32-4 genome also encodes transporters for nucleosides and uracil although both genomes apparently can de novo synthesize nucleotides. Both genomes lack the biosynthetic pathways for branched-chain amino acids, and only the *aroE, aroK, aroA, and aroC* genes were identified in the pathway for aromatic amino acids. Although an operon for the tryptophan synthase genes *trpA* and *trpB* was found in the CpT32-4 genome, both are pseudogenized. CpT32-4 also cannot synthesize histidine, whereas RsTz2-092 possesses its complete biosynthetic pathway. These biosynthetic pathways, absent in CpT32-4 and/or RsTz2-092, are typically retained by most other *Deferribacterota* (Fig. 3B). In the CpT32-4 genome, the biosynthetic pathways for riboflavin, tetrahydrofolate, cobalamin, and heme are complete, whereas the biosynthesis of pantothenate/coenzyme A (CoA), nicotinamide adenine dinucleotide (NAD), biotin, pyridoxin 5'-phosphate, and menaquinone requires precursors (Fig. S7).

*Sulfur metabolism*

CpT32-4 potentially takes up sulfate, which is subsequently reduced to sulfite via either adenosine 5'-phosphosulfate (APS) or 3'-phosphoadenosine 5'-phosphosulfate (PAPS) (Fig. 3A). It is predicted that sulfite is reduced to sulfide by the action of assimilatory sulfite reductase (Sir) in cooperation with the heterodisulfide reductase complex HdrABC-FlxABCD (Fig. S8A, B) [85]. Sulfides can be assimilated into methionine and cysteine. No genes involved in sulfur assimilation were detected in the RsTz2-092 genome (Fig. 3A, B).

*Cell wall, secretion systems, and motility*

Both genomes encode the biosynthetic pathways for peptidoglycan and lipopolysaccharides and also contain the gene sets encoding proteins for flagellar assembly and chemotaxis, similar to the cases in other *Deferribacterota* [3, 5, 6, 13] (Fig. 3A, B). Genes for the secretory (Sec) system and the twin arginine translocation (Tat) pathway, which facilitate protein transport, were present, whereas no other secretion systems were identified (Fig. 3A).

*Comparison of genomic compositions with other members of* Deferribacterota

PCA based on the relative abundance of COG categories exhibited the clustering of the gut-inhabiting clades, i.e., the insect-gut clade, *Mucispirillaceae*, and "*Ca*. Microvillispirillaceae" (Fig. 4). Among the COG categories, a decrease in the ratio and number of genes in category [T] (signal transduction mechanisms) was prominent in CpT32-4 and RsTz2-092 (Figs 4 and S9A, B). In PCoA based on the composition of all orthologs (Fig. S10A, B), the genomes of the insect-gut clade formed a cluster and were separated from those of other animal-gut inhabitants.

Most genomes of the insect-gut clade and *Mucispirillaceae* lack components for nitrogen fixation and NADH dehydrogenase complex I (Fig. 3B). Among the *Deferribacterota*,

CpT32-4 exceptionally has capacity to use diverse carbon and energy sources (Fig. S7) and to generate $H_2$ as a fermentation product (Fig. 3B). In addition, the agmatine deiminase pathway is almost unique to CpT32-4 in *Deferribacterota*, and its ability to assimilate sulfate is rare among the gut inhabitants. The biosynthetic capacity for amino acids and cofactors is reduced in the insect-gut clade and especially in the ectosymbionts CpT32-4 and RsTz2-092. This tendency is also conspicuous in the shrimp-gut symbiont "*Ca.* Rimicarispirillum atlantis" in "*Ca*. Microvillispirillaceae" (Figs 3B and S7). Notably, CpT32-4 and RsTz2-092 possess an unusually high copy number of flagellin genes compared with other *Deferribacterota* species (CpT32-4: 40 copies, RsTz2-092: 15 copies, others: 0–7 copies) (Figs 3B and S11).

## Discussion

In this study, we discovered the first examples of ectosymbiosis between *Deferribacterota* and protists. The ectosymbiotic *Deferribacterota* are facultatively associated with specific parabasalid protist species in the guts of termites and *Cryptocercus* cockroaches. The monophyly of these bacteria implies that their ability to establish ectosymbiosis with protist cells has evolved in the gut microbiota of the common ancestor of termites and *Cryptocercus*, i.e., before ~150 million years [86, 87].

The *Deferribacterota* detected in the guts of higher termites lacking cellulolytic gut protists are most likely free-swimming bacteria with flagella (Fig. 3B). The ectosymbionts CpT32-4 and RsTz-092 possess genes encoding proteins for flagellar assembly and chemotaxis; therefore, these bacteria also swim, possibly targeting host protists, and attach to the cell surface. The expansion of the number of flagellin gene variants in the genomes of CpT32-4 and RsTz-092, up to 40, is noteworthy. In many bacteria, flagella act not only for swimming but also for other roles such as adhesion and colonization [88]. It has been suggested that such multiple functions of flagella are associated with the structural variations of flagellin [89]. In

addition, in pathogenic bacteria, shifts in the expression of flagellin variants to avoid or confuse recognition of their flagella by the host cell were previously reported [88]. We thus hypothesize that the large repertoire of flagellin variants may be used for multiple functions, such as swimming and adhesion, and possibly also to avoid recognition and elimination by the protist host.

Most protist cells in the guts of termites and *Cryptocercus* cockroaches harbor ectosymbiotic bacteria of diverse lineages, which likely play various important roles in the multilayered symbiotic system [26]. Our genome analyses indicated that CpT32-4 and RsTz-092 are not able to fix dinitrogen. In addition, they have limited capacity to synthesize amino acids (Fig. 3B), whereas their ability to synthesize several cofactors, such as riboflavin, cobalamin, and tetrahydrofolate, may contribute to the nutrition of the protist and termite hosts (Fig. S7). Although CpT32-4 has the potential to assimilate diverse carbon sources, including plant-derived monosaccharides and uronate, it does not possess secretory enzymes for the hydrolysis of polysaccharides (Fig. 3A).

As CpT32-4 and RsTz-092 possess the ability to consume $H_2$ as an energy source and use $O_2$ as an electron acceptor, these ectosymbionts possibly play a role in the removal of these molecules to promote the fermentative activities of their anaerobic protist hosts. This is in concordance with the localization of CpT32-4 confined to the anterior part of *Trichonympha* cells, where hydrogenosomes are densely localized [90]. This localization pattern was also previously observed in $H_2$-oxidizing ectosymbiotic *Desulfovibrio* associated with *Trichonympha* in termite guts [36, 37]. However, unlike the cases in these obligate *Desulfovibrio* ectosymbionts, the ectosymbioses by CpT32-4 and RsTz-092 are facultative; their association as ectosymbionts is not essential for the protist host survival.

The reduction in the number of signal transduction-related genes, with no prominent genome size reduction, was previously found in several obligate ectosymbionts of protists in

the guts of termites and *Cryptocercus* cockroaches [37]. This trait is shared by CpT32-4 and RsTz-092; it is implied that these bacteria almost entirely live as ectosymbionts except for their infection phases. The unusual GC skew pattern of CpT32-4 suggests a non-canonical chromosome replication system and/or massive genome rearrangements, as previously found in endosymbiotic bacteria of termite-gut protists [27, 55, 91] (Fig. S5). However, as no complete genomes except for CpT32-4 were obtained, it remains unclear whether this GC skew pattern is a common trait in the insect-gut clade or attributable to the ectosymbiosis. Additional genome sequences of ectosymbiotic *Deferribacterota* and ecological characterization of other members of the insect-gut clade will further provide insights into the evolution of ectosymbiosis from free-swimming gut bacteria.

On the basis of the phylogenetic, morphological, and genomic data, we propose a novel genus, "*Candidatus* Termitispirillum" for the ectosymbionts CpT32-4 and RsTz2-092 and a novel family, "*Candidatus* Termitispirillaceae", for the insect-gut clade.

**Description of "*Candidatus* Termitispirillaceae" fam. nov.**

*Termitispirillaceae* (Ter.mi.ti.spi.ril.la'ce.ae. N.L. fem. pl. n. *Termitispirillum*, the type genus of the family; *-aceae*, suffix denoting a family; N.L. fem. pl. n. *Termitispirillaceae*, the family of *Termitispirillum*). Members of this family are uncultured and inhabit the guts of insects, including termites and *Cryptocercus* cockroaches. This family is a sister clade of the family *Mucispirillaceae* in the phylum *Deferribacterota*. The type genus of this family is "*Candidatus* Termitispirillum".

**Description of "*Candidatus* Termitispirillum" gen. nov.**

*Termitispirillum* (Ter.mi.ti.spi'ril.lum. L. masc. n. *Termes*, termite; L. neut. n. *spirillum*, a small spiral; N.L. neut. n. *Termitispirillum*, a spiral bacterium from termites). Members of this

genus are uncultured and spiral or curved long rods that specifically attach to parabasalid protists in the guts of termites and *Cryptocercus* cockroaches. Cells measure 3–13 µm in length and 0.4 µm in width. Genomic analyses indicated that these bacteria possess a Gram-negative-type cell wall and chemoheterotrophic metabolism with fermentation and respiratory pathways. Motility was inferred from the presence of genes involved in flagellar assembly and chemotaxis. Taxonomic assignment is based on the 16S rRNA gene sequence and specific detection by FISH using the oligonucleotide probe RsTz2-092-190 or Deferri-term-661 (Table S2). The type species is "*Candidatus* Termitispirillum reticulotermitis" with its genome sequence as the type material.

**Description of "*Candidatus* Termitispirillum reticulotermitis" sp. nov.**

*Termitispirillum reticulotermitis* (re.ti.cu.lo.ter'mi.tis. N.L. gen. masc. n. *reticulotermitis*, of *Reticulitermes*, referring to the host termite genus). The bacterium specifically attaches to the cell surface of *Holomastigotes* sp. in the gut of *Reticulitermes speratus*. The cell dimensions are 3–10 µm by 0.4 µm. This assignment is based on the 16S rRNA gene sequence (LC866559) and specific detection by FISH using the oligonucleotide probe RsTz2-092-190 (Table S2). The single-cell genome MS499-41 (GCA_025997275) was obtained. This species corresponds to the 16S rRNA phylotypes RsTz2-092 and RsHol6-1.

**Description of "*Candidatus* Termitispirillum hodotermopsidis" sp. nov.**

*Termitispirillum hodotermopsidis* (ho.do.ter.mop'si.dis. N.L. fem. n. *hodotermopsidis*, of *Hodotermopsis*, referring to the host termite genus). The bacterium specifically attaches to the cell surfaces of *Holomastigotes* sp. and *Brugerollina cincta* in the gut of *Hodotermopsis sjostedti*. The cell dimensions are 3–10 µm by 0.4 µm. This assignment is based on the 16S rRNA gene sequence (LC866557, LC866558) and specific detection by FISH using the

oligonucleotide probe RsTz2-092-190 (Table S2). This species corresponds to the 16S rRNA phylotypes HsHol1-2 and HsBr1-3.

**Description of "*Candidatus* Termitispirillum cryptocercusi" sp. nov.**

*Termitispirillum cryptocercusi* (cryp.to.cer'cu.si. N.L. gen. n. *cryptocercusi*, of *Cryptocercus*, referring to the host cockroach genus). The bacterium specifically attaches to the cell surfaces of *Trichonympha acuta* and *T. lata* in the gut of *Cryptocercus punctulatus*. The cell dimensions are 3–13 µm by 0.4 µm. This assignment is based on the 16S rRNA gene sequence (LC866555, LC866555) and detection by FISH using the oligonucleotide probe Deferri-term-661 (Table S2). A complete genome sequence (PRJDB20420) was obtained. This species includes phylotypes CpT06-9, CpT15-2, and CpT32-4.

**Acknowledgements**

We are grateful to Gaku Tokuda, Aram Mikaelyan, and Melbert Schwartz for cooperation in cockroach collection. We also thank Shu Onouchi, Hirokazu Kuwahara, and Kazuki Takahashi for assisting with experiments and Takumi Murakami for handling sequence data. Sanger sequencing was performed in the RIKEN CBS and the Biomaterial Analysis Center in the Tokyo Institute of Technology. This work was financially supported by the JST (Japan Science and Technology Agency) SPRING for PhD course students to N.M. (JPMJSP2106), by NEXT and KAKENHI grants from JSPS (Japan Society for the Promotion of Science) to Y.H. (GS009, 22241046, 16H04840, 20H02897, 20H05584, 22K19342, and 23H02553) and to M.O. (17H01447 and 19H05689), and also by JST-CREST (14532219) to Y.H. and by JST-GteX to M.O. (JPMJGX23B0).

**Author contributions**

Naoya Maruoka: conceptualization, methodology, investigation, formal analysis, data curation, visualization, resource, writing – original draft, funding acquisition. Rinpei Kudo: investigation, data curation. Katsura Igai: investigation, data curation. Masahiro Yuki: methodology, investigation, data curation, resource. Michiru Shimizu: methodology, investigation, data curation, resource. Moriya Ohkuma: resource, writing – review and editing, funding acquisition. Yuichi Hongoh: conceptualization, methodology, resource, formal analysis, writing – original draft, writing – review and editing, funding acquisition, project administration, supervision.

**Conflicts of interest**

The authors declare no conflict of interest.

**Data availability**

The genome sequences obtained in this study are available under the BioProject PRJDB20420 in DDBJ. The accession numbers of the SSU rRNA genes appears under DDBJ accession numbers LC866555–64 and LC875758.

**References**

1. Jumas-Bilak E, Roudière L, Marchandin, H. Despcription of "*Synergistetes*" phyl. nov. and emended description of the phylum "*Deferribacteres*" and of the family *Syntrophomonadaceae*, phylum "*Firmicutes*". *Int J Syst Evol Microbiol* 2009;**59**:1028–35. https://doi.org/10.1099/ijs.0.006718-0
2. Oren A, Garrity GM. Valid publication of the names of forty-two phyla of prokaryotes. *Int J Syst Evol Microbiol* 2021;**71**:e005056. https://doi.org/10.1099/ijsem.0.005056

3. Spring S, Rohde M, Bunk B *et al*. New insights into the energy metabolism and taxonomy of *Deferribacteres* revealed by the characterization of a new isolate from a hypersaline microbial mat. *Environ Microbiol* 2022;**24**:2543–75. https://doi.org/10.1111/1462-2920.15999

4. Zavarzina DG, Prokofeva MI, Pikhtereva VA *et al*. *Deferrivibrio essentukiensis* sp. nov., gen. nov., a representative of *Deferrivibrionaceae* fam. nov., isolated from the subsurface aquifer of caucasian mineral drinking waters. *Microbiology (Russian Fed)* 2022;**91**:143–59. https://doi.org/10.1134/S0026261722020114

5. Aubé J, Cambon-Bonavita MA, Velo-Suárez L *et al*. A novel and dual digestive symbiosis scales up the nutrition and immune system of the holobiont *Rimicaris exoculata*. *Microbiome* 2022;**10**:e189. https://doi.org/10.1186/s40168-022-01380-2

6. Takaki Y, Shimamura S, Nakagawa S *et al*. Bacterial lifestyle in a deep-sea hydrothermal vent chimney revealed by the genome sequence of the thermophilic bacterium *Deferribacter desulfuricans* SSM1. *DNA Res* 2010**;17**:123–37. https://doi.org/10.1093/dnares/dsq005

7. Rauschenbach I, Posternak V, Cantarella P *et al*. *Seleniivibrio woodruffii* gen. nov., sp. nov., a selenate- and arsenate-respiring bacterium in the *Deferribacteraceae*. *Int J Syst Evol Microbiol* 2013;**63**:3659–3665. https://doi.org/10.1099/ijs.0.043547-0

8. Iino T, Nakagawa T, Mori K *et al*. *Calditerrivibrio nitroreducens* gen. nov., sp. nov., a thermophilic, nitrate-reducing bacterium isolated from a terrestrial hot spring in Japan. *Int J Syst Evol Microbiol* 2008;**58**:1675–9. https://doi.org/10.1099/ijs.0.65714-0

9. Tamazawa S, Mayumi D, Mochimaru H *et al*. *Petrothermobacter organivorans* gen. nov., sp. nov., a thermophilic, strictly anaerobic bacterium of the phylum *Deferribacteres* isolated from a deep subsurface oil reservoir. *Int J Syst Evol Microbiol* 2017;**67**:3982–6. https://doi.org/10.1099/ijsem.0.002234

10. Robertson BR, O'Rourke JL, Neilan BA *et al*. *Mucispirillum schaedleri* gen. nov., sp. nov., a spiral-shaped bacterium colonizing the mucus layer of the gastrointestinal tract of laboratory rodents. *Int J Syst Evol Microbiol* 2005;**55**:1199–204. https://doi.org/10.1099/ijs.0.63472-0

11. Gilroy R, Ravi A, Getino M *et al*. Extensive microbial diversity within the chicken gut microbiome revealed by metagenomics and culture. *PeerJ* 2021;**9**:e10941. https://doi.org/10.7717/peerj.10941

12. Scupham AJ, Patton TG, Bent E *et al*. Comparison of the cecal microbiota of domestic and wild turkeys. *Microb Ecol* 2008;**56**:322–31. https://doi.org/10.1111/1462-2920.70010

13. Loy A, Pfann C, Steinberger M *et al*. Lifestyle and horizontal gene transfer-mediated evolution of *Mucispirillum schaedleri*, a core member of the murine gut microbiota. *mSystems* 2017;**2**:e00171-16. https://doi.org/10.1128/msystems.00171-16

14. Durand L, Zbinden M, Cueff-Gauchard V *et al.* Microbial diversity associated with the hydrothermal shrimp *Rimicaris exoculata* gut and occurrence of a resident microbial community. *FEMS Microbiol. Ecol* 2010;**71**:291–303. https://doi.org/10.1111/j.1574-6941.2009.00806.x

15. Guéganton M, Rouxel O, Durand L *et al.* Anatomy and symbiosis of the digestive system of the vent shrimp *Rimicaris exoculata* and *Rimicaris chacei* revealed through imaging approaches. *Front Mar Sci* 2022;**9**:e903748. https://doi.org/10.3389/fmars.2022.903748

16. Qi L, Shi M, Zhu FC *et al.* Genomic evidence for the first symbiotic Deferribacterota, a novel gut symbiont from the deep-sea hydrothermal vent shrimp *Rimicaris kairei*. *Front Microbiol* 2023;**14**:e1179935. https://doi.org/10.3389/fmicb.2023.1179935

17. Waite DW, Dsouza M, Biswas K *et al.* Microbial community structure in the gut of the New Zealand insect Auckland tree weta (*Hemideina thoracica*). *Arch Microbiol* 2015;**197**:603–12. https://doi.org/10.1007/s00203-015-1094-3

18. Schauer C, Thompson CL, Brune A. The bacterial community in the gut of the cockroach *Shelfordella lateralis* reflects the close evolutionary relatedness of cockroaches and termites. *Appl Environ Microbiol* 2012;**78**:2758–67. https://doi.org/10.1128/AEM.07788-11

19. Berlanga M, Llorens C, Comas J *et al.* Gut bacterial community of the xylophagous cockroaches *Cryptocercus punctulatus* and *Parasphaeria boleiriana*. *PLOS ONE* 2016;**11**:e0152400. https://doi.org/10.1371/journal.pone.0152400

20. Warnecke F, Luginbühl P, Ivanova N *et al.* Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 2007;**450**:560–5. https://doi.org/10.1038/nature06269

21. Liu N, Zhang L, Zhou H *et al.* Metagenomic insights into metabolic capacities of the gut microbiota in a fungus-cultivating termite (*Odontotermes yunnanensis*). *PLOS ONE* 2013;**8**:e69184. https://doi.org/10.1371/journal.pone.0069184

22. King JHP, Mahadi NM, Bong CFJ *et al*. Bacterial microbiome of *Coptotermes curvignathus* (Isoptera: Rhinotermitidae) reflects the coevolution of species and dietary pattern. *Insect Sci.* 2014;**21**:584–96. https://doi.org/10.1111/1744-7917.12061

23. Brune A. Symbiotic digestion of lignocellulose in termite guts. *Nat. Rev. Microbiol.* 2014;**12**:168–80. https://doi.org/10.1038/nrmicro3182

24. Cleveland LR, Hall SR, Sanders EP *et al.* The wood-feeding roach *Cryptocercus*, its protozoa, and the symbiosis between protozoa and roach. *Mem Am Acad Arts Sci* 1934;**17**:185–342. https://doi.org/10.2307/25058198

25. Ohkuma M, Brune A. Diversity, structure, and evolution of the termite gut microbial community. In: Bignell D, Roisin Y, Lo N (ed.), *Biology of Termites: a Modern Synthesis*. Switzerland: Springer Nature, 2011, 413–38. https://doi.org/10.1007/978-90-481-3977-4

26. Hongoh Y, Ohkuma M. Termite gut flagellates and their methanogenic and eubacterial symbionts. In: Hackstein JHP (ed.), *(Endo)symbiotic Methanogenic Archaea*. Switzerland: Springer Nature, 2018, 55–80. https://doi.org/10.1007/978-3-319-98836-8_5

27. Hongoh Y, Sharma VK, Prakash T *et al*. Complete genome of the uncultured Termite Group 1 bacteria in a single host protist cell. *Proc Natl Acad Sci USA* 2008;**105**:5555–60. https://doi.org/10.1073/pnas.0801389105

28. Hongoh Y, Sharma VK, Prakash T *et al*. Genome of an endosymbiont coupling $N_2$ fixation to cellulolysis within protist cells in termite gut. *Science* 2008;**322**:1108–9. https://doi.org/10.1126/science.1165578

29. Ohkuma M, Noda S, Hattori S *et al*. Acetogenesis from $H_2$ plus $CO_2$ and nitrogen fixation by an endosymbiotic spirochete of a termite-gut cellulolytic protist. *Proc Natl Acad Sci USA* 2015;**112:**10224–30. https://doi.org/10.1073/pnas.1423979112

30. Tai V, Carpenter KJ, Weber PK *et al.* Genome evolution and nitrogen fixation in bacterial ectosymbionts of a protist inhabiting wood-feeding cockroaches. *Appl Environ Microbiol* 2016;**82**:4682–95. https://doi.org/10.1128/AEM.00611-16

31. Strassert JFH, Mikaelyan A, Woyke T *et al*. Genome analysis of '*Candidatus* Ancillula trichonymphae', first representative of a deep-branching clade of bifidobacteriales, strengthens evidence for convergent evolution in flagellate endosymbionts. *Environ Microbiol Rep* 2016;**8**:865–73. https://doi.org/10.1111/1758-2229.12451

32. Treitli SC, Kolisko M, Husník F *et al.* Revealing the metabolic capacity of *Streblomastix strix* and its bacterial symbionts using single-cell metagenomics. *Proc Natl Acad Sci USA* 2019;**116**:19675–84. https://doi.org/10.1073/pnas.1910793116

33. Mies US, Hervé V, Kropp T *et al.* Genome reduction and horizontal gene transfer in the evolution of Endomicrobia—rise and fall of an intracellular symbiosis with termite gut flagellates. *mBio* 2024;**15**:e00826-24. https://doi.org/10.1128/mbio.00826-24

34. Silva JK, Hervé V, Mies US *et al.* A novel lineage of endosymbiotic *Actinomycetales*: genome reduction and acquisition of new functions in *Bifidobacteriaceae* associated with termite gut flagellates. *Environ Microbiol* 2025;**27**:e70010. https://doi.org/10.1111/1462-2920.70010

35. Ikeda-Ohtsubo W, Strassert JFH, Köhler T *et al.* "*Candidatus* Adiutrix intracellularis", an endosymbiont of termite gut flagellates, is the first representative of a deep-branching clade of *Deltaproteobacteria* and a putative homoacetogen. *Environ Microbiol* 2016;**18**:2548–64. https://doi.org/10.1111/1462-2920.13234

36. Kuwahara H, Yuki M, Izawa K *et al.* Genome of '*Ca.* Desulfovibrio trichonymphae', an $H_2$-oxidizing bacterium in a tripartite symbiotic system within a protist cell in the termite gut. *ISME J* 2017;**11**:766–76. https://doi.org/10.1038/ismej.2016.143

37. Takeuchi M, Kuwahara H, Murakami T *et al.* Parallel reductive genome evolution in *Desulfovibrio* ectosymbionts independently acquired by *Trichonympha* protists in the termite gut. *ISME J* 2020;**14**:2288–301. https://doi.org/10.1038/s41396-020-0688-1

38. Kaneko M, Omori T, Igai K *et al.* Facultative endosymbiosis between cellulolytic protists and methanogenic archaea in the gut of the Formosan termite *Coptotermes formosanus*. *ISME Commun* 2024;**4**:ycae097. https://doi.org/10.1093/ismeco/ycae097

39. Aoki H, Yuki M, Shimizu M *et al.* Agarose gel microcapsules enable easy-to-prepare, picolitre-scale, single-cell genomics, yielding high-coverage genome sequences. *Sci Rep* 2022;**12**:e17014 (2022). https://doi.org/10.1038/s41598-022-20923-z

40. Utami YD, Kuwahara H, Murakami T *et al.* Phylogenetic diversity and single-cell genome analysis of "*Melainabacteria*", a non-photosynthetic cyanobacterial group, in the termite gut. *Microbes Environ* 2018;**33**:50–7. https://doi.org/10.1264/jsme2.ME17137

41. Pruesse E, Peplies J, Glöckner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;**28**:1823–9. https://doi.org/10.1093/bioinformatics/bts252

42. Katoh K, Misawa K, Kuma KI *et al.* MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002;**30**:3059–66. https://doi.org/10.1093/nar/gkf436

43. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;**25**:1972–3. https://doi.org/10.1093/bioinformatics/btp348

44. Nguyen LT, Schmidt HA, von Haeseler A *et al.* IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;**32**:268–74. https://doi.org/10.1093/molbev/msu300

45. Kalyaanamoorthy S, Minh BQ, Wong TKF *et al*. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;**14**:587–9. https://doi.org/10.1038/nmeth.4285

46. Trager W. The cultivation of a cellulose-digesting flagellate, *Trichomonas termopsidis*, and of certain other termite protozoa. *Biol Bull* 1934;**66**:182–90.

47. Noda S, Ohkuma M, Yamada A *et al.* Phylogenetic position and in situ identification of ectosymbioticspirochetes onprotists in the termite gut. *Appl Environ Microbiol* 2003;**69**:625–33. https://doi.org/10.1128/AEM.69.1.625-633.2003

48. Murakami T, Onouchi S, Igai K *et al.* Ectosymbiotic bacterial microbiota densely colonize the surface of thelastomatid nematodes in the gut of the wood-feeding cockroach *Panesthia angustipennis*. *FEMS Microbiol Ecol* 2019;**95**:fiy238. https://doi.org/10.1093/femsec/fiy238

49. Ludwig W, Strunk O, Westram R *et al.* ARB: a software environment for sequence data. *Nucleic Acids Res* 2004;**32**:1363–71. https://doi.org/10.1093/nar/gkh293

50. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 2011;**17**:10–12. https://doi.org/10.14806/ej.17.1.200

51. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 2011;**27**:863–4. https://doi.org/10.1093/bioinformatics/btr026

52. Bankevich A, Nurk S, Antipov D *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;**19**:455–77. https://doi.org/10.1089/cmb.2012.0021

53. von Meijenfeldt FAB, Arkhipova K, Cambuy DD *et al.* Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biol* 2019;**20**:e217. https://doi.org/10.1186/s13059-019-1817-x

54. Camacho C, Coulouris G, Avagyan V *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* 2009;**10**:e421. https://doi.org/10.1186/1471-2105-10-421

55. Takahashi K, Kuwahara H, Horikawa Y *et al.* Emergence of putative energy parasites within *Clostridia* revealed by genome analysis of a novel endosymbiotic clade. *ISME J* 2023;**17**:1895–1906. https://doi.org/10.1038/s41396-023-01502-0

56. Steinig E, Coin L. Nanoq: ultra-fast quality control for nanopore reads. *J Open Source Softw* 2022;**7**:e2991. https://doi.org/10.21105/joss.02991

57. Vaser R, Šikić M. Time- and memory-efficient genome assembly with Raven. *Nat Comput Sci* 2021;**1**:332–6. https://doi.org/10.1038/s43588-021-00073-4

58. Walker BJ, Abeel T, Shea T *et al.* Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLOS ONE* 2014;**9**:e112963. https://doi.org/10.1371/journal.pone.0112963

59. Chklovski A, Parks DH, Woodcroft BJ *et al.* CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods* 2023;**20**:1203–12. https://doi.org/10.1038/s41592-023-01940-w

60. Grant JR, Enns E, Marinier E *et al.* Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res* 2023;**51**:W484–92. https://doi.org/10.1093/nar/gkad326

61. Jain C, Rodriguez-R LM, Phillippy AM *et al.* High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;**9**:1–8. https://doi.org/10.1038/s41467-018-07641-9

62. Kim D, Park S, Chun J. Introducing EzAAI: a pipeline for high throughput calculations of prokaryotic average amino acid identity. *J Microbiol* 2021;**59**:476–80. https://doi.org/10.1007/s12275-021-1154-0

63. Lu J, Salzberg SL. SkewIT: the Skew Index Test for large-scale GC Skew analysis of bacterial genomes. *PLOS Comput Biol* 2020;**16**:e1008439. https://doi.org/10.1371/journal.pcbi.1008439

64. Lee MD. GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics* 2019;**35**:4162–4. https://doi.org/10.1093/bioinformatics/btz188

65. Hyatt D, Chen G, LoCascio PF *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;**11**:e119. https://doi.org/10.1186/1471-2105-11-119

66. Eddy SR. Accelerated profile HMM searches. *PLOS Comput. Biol.* 2011;**7**:e1002195. https://doi.org/10.1371/journal.pcbi.1002195

67. Edgar RC. Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat Commun* 2022;**13**:e6989. https://doi.org/10.1038/s41467-022-34630-w

68. Parks DH, Chuvochina M, Rinke C *et al.* GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 2022;**50**:D785–94. https://doi.org/10.1093/nar/gkab776

69. Tanizawa Y, Fujisawa T, Nakamura Y. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 2018;**34**:1037–9. https://doi.org/10.1093/bioinformatics/btx713

70. Moriya Y, Itoh M, Okuda S *et al.* KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 2007;**35**:182–5. https://doi.org/10.1093/nar/gkm321

71. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015;**12**:59–60. https://doi.org/10.1038/nmeth.3176

72. Taboada B, Estrada K, Ciria R *et al.* Operon-mapper: a web server for precise operon identification in bacterial and archaeal genomes. *Bioinformatics* 2018;**34**:4118–20. https://doi.org/10.1093/bioinformatics/bty496

73. Marchler-Bauer A, Panchenko AR, Shoemarker BA *et al.* CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res* 2002;**30**:281–3. https://doi.org/10.1093/nar/30.1.281

74. Pedregosa F, Varoquaux G, Gramfort A *et al.* Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;**12**:2825–30.

75. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019;**20**:e238. https://doi.org/10.1186/s13059-019-1832-y

76. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE* 2013;**8**:e61217. https://doi.org/10.1371/journal.pone.0061217

77. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 2016;4**28**:726–31. https://doi.org/10.1016/j.jmb.2015.11.006

78. Roder T, Pimentel G, Fuchsmann P *et al.* Scoary2: rapid association of phenotypic multi-omics data with microbial pan-genomes. *Genome Biol* 2024;**25**:1–19. https://doi.org/10.1186/s13059-024-03233-7

79. Hervé V, Liu P, Dietrich C *et al.* Phylogenomic analysis of 589 metagenome-assembled genomes encompassing all major prokaryotic lineages from the gut of higher termites. *PeerJ* 2020;**8**:e8614. https://doi.org/10.7717/peerj.8614

80. Arora J, Kinjo Y, Šobotník J *et al.* The functional evolution of termite gut microbiota. *Microbiome* 2022;**10**:1–22. https://doi.org/10.1186/s40168-022-01258-3

81. Riesco R, Trujillo ME. Update on the proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2024;**74**:1–12. https://doi.org/10.1099/ijsem.0.006300

82. del Rio B, Linares DM, Ladero V *et al.* Putrescine production via the agmatine deiminase pathway increases the growth of *Lactococcus lactis* and causes the alkalinization of the culture medium. *Appl Microbiol Biotechnol* 2015;**99**:897–905. https://doi.org/10.1007/s00253-014-6130-8

83. Brune A. Termite guts: the world's smallest bioreactors. *Trends Biotechnol.* 1998;**16**: 16–21. https://doi.org/10.1016/S0167-7799(97)01151-7

84. Besson S, Almeida MG, Silveira CM. Nitrite reduction in bacteria: a comprehensive view of nitrite reductases. *Coord Chem Rev* 2022;**464**:e214560. https://doi.org/10.1016/j.ccr.2022.214560

85. Ramos AR, Grein F, Oliveira GP *et al.* The FlxABCD-HdrABC proteins correspond to a novel NADH dehydrogenase/heterodisulfide reductase widespread in anaerobic bacteria and involved in ethanol metabolism in *Desulfovibrio vulgaris* Hildenborough. *Environ Microbiol* 2015;**17**:2288–305. https://doi.org/10.1111/1462-2920.12689

86. Bourguignon T, Lo N, Cameron SL *et al.* The evolutionary history of termites as inferred from 66 mitochondrial genomes. *Mol. Biol. Evol.* 2015;**32**:406–21. https://doi.org/10.1093/molbev/msu308

87. Evangelista DA, Wipfler B, Béthoux O *et al.* An integrative phylogenomic approach illuminates the evolutionary history of cockroaches and termites (Blattodea). *Proc R Soc B Biol Sci* 2019;**286**:e20182076. https://doi.org/10.1098/rspb.2018.2076

88. Chaban B, Hughes HV, Beeby M. The flagellum in bacterial pathogens: for motility and a whole lot more. *Semin Cell Dev Biol* 2015;**46**:91–103. https://doi.org/10.1016/j.semcdb.2015.10.032

89. Fields JL, Zhang H, Bellis NF *et al.* Structural diversity and clustering of bacterial flagellar outer domains. *Nat Commun* 2024;**15**:e9500. https://doi.org/10.1038/s41467-024-53923-w

90. Carpenter KJ, Chow L, Keeling PJ. Morphology, phylogeny, and diversity of *Trichonympha* (Parabasalia: Hypermastigida) of the wood-feeding cockroach *Cryptocercus punctulatus*. *J Eukaryot Microbiol* 2009;**56**:305–13. https://doi.org/10.1111/j.1550-7408.2009.00406.x

91. Izawa K, Kuwahara H, Kihara K *et al.* Comparison of intracellular "*Ca.* Endomicrobium trichonymphae" genomovars illuminates the requirement and decay of defense systems against foreign DNA. *Genome Biol Evol* 2016;**8**:3099–107. https://doi.org/10.1093/gbe/evw227

## Legends to Figures

**Figure 1.** Detection of *Deferribacterota* associated with protists in the guts of termites and a *Cryptocercus* cockroach. Phase-contrast images of (A) *Holomastigotes* sp. from *Reticulitermes speratus*, (C) *Holomastigotes* sp. from *Hodotermopsis sjostedti*, (E) *Brugerollina cincta* from *H. sjostedti*, and (G) *Trichonympha lata* in *Cryptocercus punctulatus* are shown. *Deferribacterota* were detected by FISH using probe RsTz2-092-190 labeled with 6FAM (green) (B, D, F) and probe Deferri-term-661 labeled with 6FAM (green) (H), respectively. Spirochaetes were detected using probe Spiro-36 labeled with Texas red (red), and the images were overlayed (B, D, F, H). Bars indicate 10 µm and 5 µm (inset for magnification).

**Figure 2.** Phylogenetic positions of the ectosymbiotic *Deferribacterota*. (A) 16S rRNA gene tree constructed using 993 nucleotide sites with the GTR+F+I+G4 substitution model and an outgroup (AB729138). (B) Phylogenomic tree constructed using 12,478 amino acid sites of concatenated, partitioned 74 marker genes, with substitution models selected for each part and an outgroup (GCF_000469585 and GCF_000177635). Family-level taxonomic codes according to GTDB are indicated in parentheses. Filled and open circles indicate nodes with ultrafast bootstrap value ≥95% and ≥80% in panel A, and 100% and ≥90% in panel B, respectively. Asterisks indicate sequences obtained by Arora et al. (2022).

**Figure 3.** Predicted metabolic capacities of CpT32-4 and RsTz2-092. (A) Outline of predicted metabolic pathways based on genome sequences. Pathways detected exclusively in CpT32-4 or RsTz2-092 are indicated in blue and red, respectively. (B) Comparisons of repertoires of genes and pathways among members of *Deferribacterota*. The completeness of pathways is

indicated in a heatmap. Estimated genome completeness, total sequence length, and the number of flagellin genes are also shown.

**Figure 4.** Comparation of genomic contents among members of *Deferribacterota*. Principal component analysis based on the relative abundance of the clusters of orthologous groups (COGs) was performed. The COG categories: [C] Energy production and conversion, [D] Cell cycle control, cell division, chromosome partitioning, [E] Amino acid transport and metabolism, [F] Nucleotide transport and metabolism, [G] Carbohydrate transport and metabolism, [H] Coenzyme transport and metabolism, [I] Lipid transport and metabolism, [J] Translation, ribosomal structure and biogenesis, [K] Transcription, [L] Replication, recombination and repair, [M] Cell wall/membrane/envelope biogenesis, [N] Cell motility, [O] Posttranslational modification, protein turnover, chaperones, [P] Inorganic ion transport and metabolism, [Q] Secondary metabolites biosynthesis, transport and catabolism, [R] General function prediction only, [S] Function unknown, [T] Signal transduction mechanisms, [U] Intracellular trafficking, secretion, and vesicular transport, [V] Defense mechanisms, [W] Extracellular structures, [X] Mobilome: prophages, transposons.

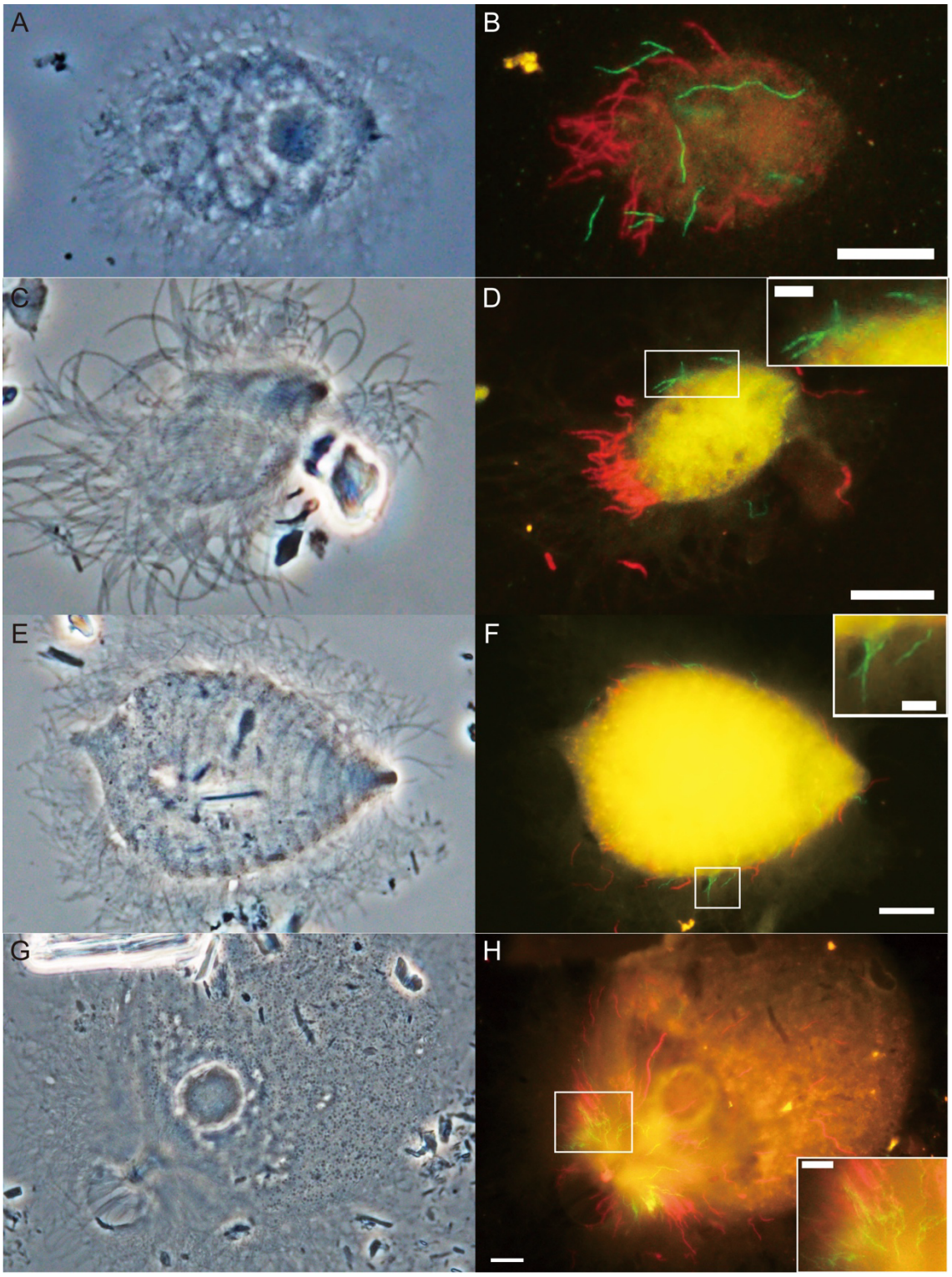


Figure 1.

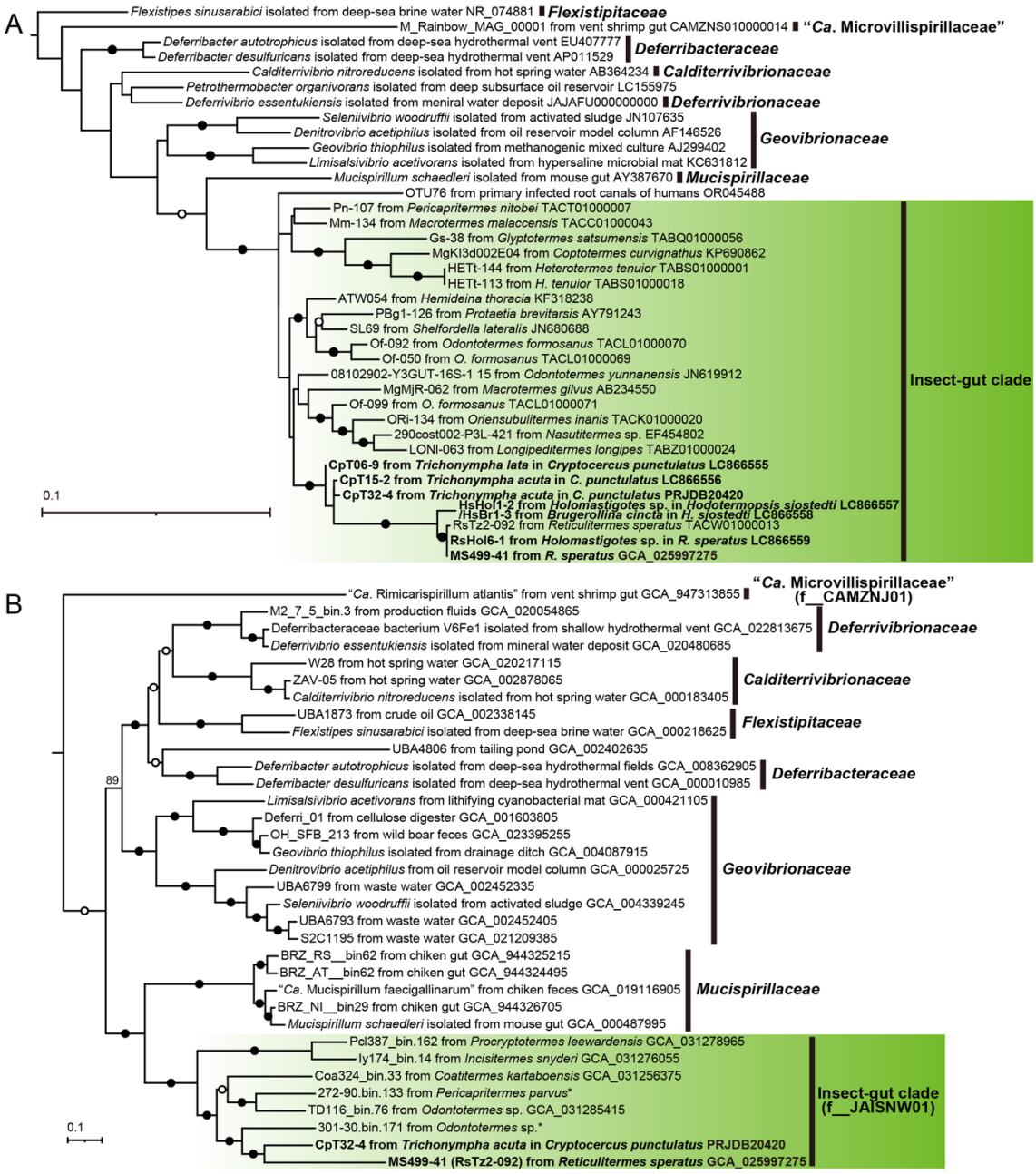
A
Flexistipes sinusarabici isolated from deep-sea brine water NR_074881
Flexistipitaceae
M_Rainbow_MAG_00001 from vent shrimp gut CAMZNS010000014
"Ca. Microvillispirillaceae"
Deferribacter autotrophicus isolated from deep-sea hydrothermal vent EU407777
Deferribacter desulfuricans isolated from deep-sea hydrothermal vent AP011529
Deferribacteraceae
Calditerrivibrio nitroreducens isolated from hot spring water AB364234
Calditerrivibrionaceae
Petrothermobacter organivorans isolated from deep subsurface oil reservoir LC155975
Deferrivibrio essentukiensis isolated from meniral water deposit JAJAFU000000000
Deferrivibrionaceae
Seleniivibrio woodruffii isolated from activated sludge JN107635
Denitrovibrio acetiphilus isolated from oil reservoir model column AF146526
Geovibrio thiophilus isolated from methanogenic mixed culture AJ299402
Limisalsivibrio acetivorans isolated from hypersaline microbial mat KC631812
Geovibrionaceae
Mucispirillum schaedleri isolated from mouse gut AY387670
Mucispirillaceae
OTU76 from primary infected root canals of humans OR045488
Pn-107 from Pericapritermes nitobei TACT01000007
Mm-134 from Macrotermes malaccensis TACC01000043
Gs-38 from Glyptotermes satsumensis TABQ01000056
MgKI3d002E04 from Coptotermes curvignathus KP690862
HETt-144 from Heterotermes tenuior TABS01000001
HETt-113 from H. tenuior TABS01000018
ATW054 from Hemideina thoracia KF318238
PBg1-126 from Protaetia brevitarsis AY791243
SL69 from Shelfordella lateralis JN680688
Of-092 from Odontotermes formosanus TACL01000070
Of-050 from O. formosanus TACL01000069
08102902-Y3GUT-16S-1 15 from Odontotermes yunnanensis JN619912
MgMjR-062 from Macrotermes gilvus AB234550
Of-099 from O. formosanus TACL01000071
ORi-134 from Oriensubulitermes inanis TACK01000020
290cost002-P3L-421 from Nasutitermes sp. EF454802
LONI-063 from Longipeditermes longipes TABZ01000024
CpT06-9 from Trichonympha lata in Cryptocercus punctulatus LC866555
CpT15-2 from Trichonympha acuta in C. punctulatus LC866556
CpT32-4 from Trichonympha acuta in C. punctulatus PRJDB20420
HsHol1-2 from Holomastigotes sp. in Hodotermopsis sjostedti LC866557
HsBr1-3 from Brugerolliina cincta in H. sjostedti LC866558
RsTz2-092 from Reticulitermes speratus TACW01000013
RsHol6-1 from Holomastigotes sp. in R. speratus LC866559
MS499-41 from R. speratus GCA_025997275
Insect-gut clade
0.1
B
"Ca. Rimicarispirillum atlantis" from vent shrimp gut GCA_947313855
"Ca. Microvillispirillaceae" (f__CAMZNJ01)
M2_7_5_bin.3 from production fluids GCA_020054865
Deferribacteraceae bacterium V6Fe1 isolated from shallow hydrothermal vent GCA_022813675
Deferrivibrio essentukiensis isolated from mineral water deposit GCA_020480685
Deferrivibrionaceae
W28 from hot spring water GCA_020217115
ZAV-05 from hot spring water GCA_002878065
Calditerrivibrio nitroreducens isolated from hot spring water GCA_000183405
Calditerrivibrionaceae
UBA1873 from crude oil GCA_002338145
Flexistipes sinusarabici isolated from deep-sea brine water GCA_000218625
Flexistipitaceae
UBA4806 from tailing pond GCA_002402635
Deferribacter autotrophicus isolated from deep-sea hydrothermal fields GCA_008362905
Deferribacter desulfuricans isolated from deep-sea hydrothermal vent GCA_000010985
Deferribacteraceae
89
Limisalsivibrio acetivorans from lithifying cyanobacterial mat GCA_000421105
Deferri_01 from cellulose digester GCA_001603805
OH_SFB_213 from wild boar feces GCA_023395255
Geovibrio thiophilus isolated from drainage ditch GCA_004087915
Denitrovibrio acetiphilus from oil reservoir model column GCA_000025725
UBA6799 from waste water GCA_002452335
Seleniivibrio woodruffii isolated from activated sludge GCA_004339245
UBA6793 from waste water GCA_002452405
S2C1195 from waste water GCA_021209385
Geovibrionaceae
BRZ_RS__bin62 from chiken gut GCA_944325215
BRZ_AT__bin62 from chiken gut GCA_944324495
"Ca. Mucispirillum faecigallinarum" from chiken feces GCA_019116905
BRZ_NI__bin29 from chiken gut GCA_944326705
Mucispirillum schaedleri isolated from mouse gut GCA_000487995
Mucispirillaceae
Pcl387_bin.162 from Procryptotermes leewardensis GCA_031278965
Iy174_bin.14 from Incisitermes snyderi GCA_031276055
Coa324_bin.33 from Coatitermes kartaboensis GCA_031256375
272-90.bin.133 from Pericapritermes parvus*
TD116_bin.76 from Odontotermes sp. GCA_031285415
301-30.bin.171 from Odontotermes sp.*
CpT32-4 from Trichonympha acuta in Cryptocercus punctulatus PRJDB20420
MS499-41 (RsTz2-092) from Reticulitermes speratus GCA_025997275
Insect-gut clade (f__JAISNW01)
0.1

Figure 2

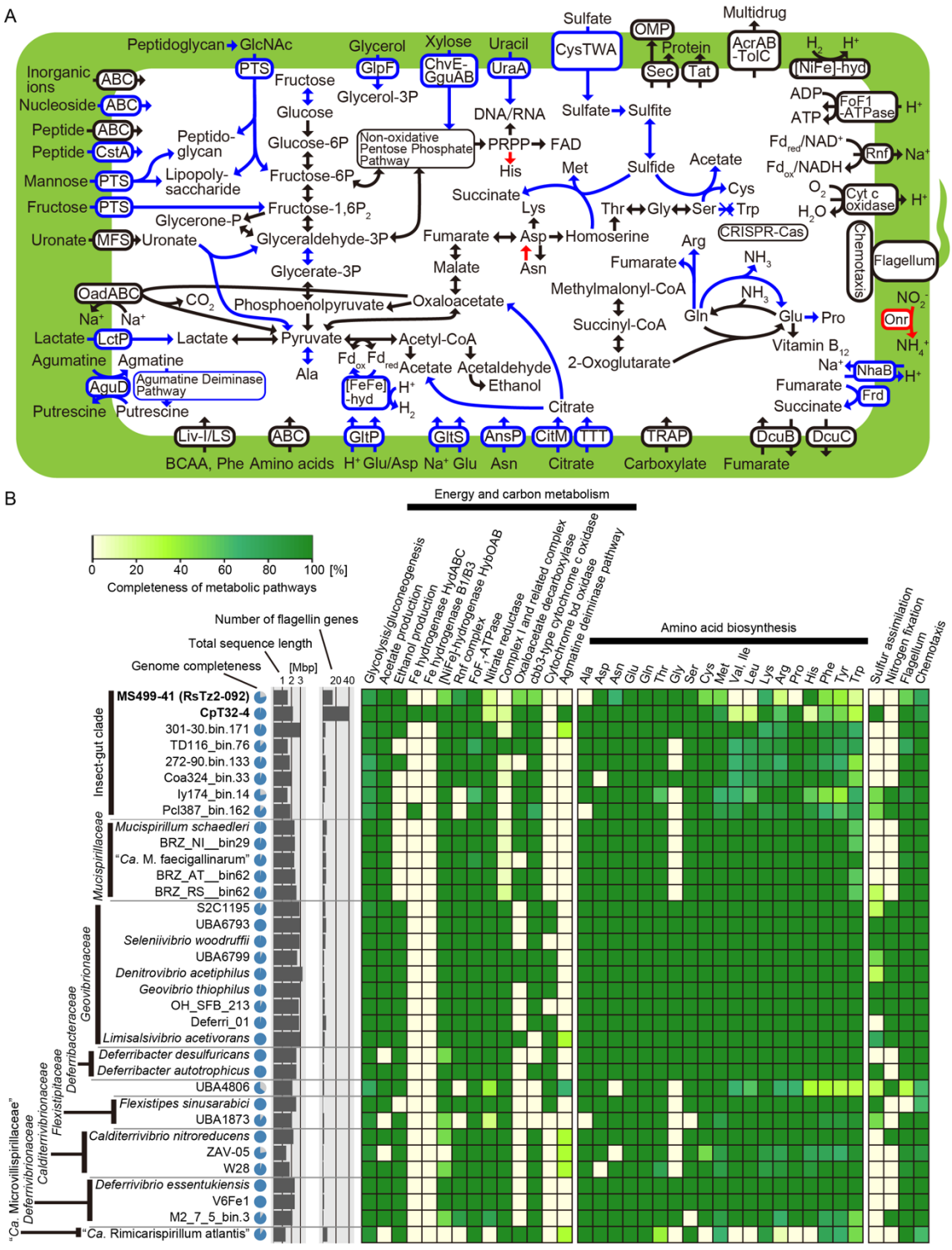
A
Peptidoglycan
GlcNAc
Glycerol
Xylose
Uracil
Sulfate
OMP
Multidrug
Inorganic ions
ABC
PTS
GlpF
ChvE-GguAB
UraA
CysTWA
Protein
Sec
Tat
AcrAB-TolC
[NiFe]-hyd
Nucleoside
Fructose
Glycerol-3P
Glucose
DNA/RNA
Sulfate
Sulfite
ADP
ATP
FoF1-ATPase
Peptide
Peptido-glycan
Glucose-6P
Non-oxidative Pentose Phosphate Pathway
PRPP
FAD
His
Met
Sulfide
Acetate
Fd_red/NAD+
Fd_ox/NADH
Rnf
Na+
Mannose
Lipopoly-saccharide
Fructose-6P
Succinate
Cys
O2
H2O
Cyt c oxidase
Fructose
Fructose-1,6P2
Lys
Thr
Gly
Ser
Trp
Glycerone-P
Uronate
Glyceraldehyde-3P
Fumarate
Asp
Homoserine
CRISPR-Cas
Chemotaxis
Flagellum
Glycerate-3P
Malate
Asn
Arg
Fumarate
NH3
OadABC
CO2
Phosphoenolpyruvate
Oxaloacetate
Methylmalonyl-CoA
NO2-
Na+
Lactate
LctP
Pyruvate
Acetyl-CoA
Succinyl-CoA
Gln
Glu
Pro
Onr
Agumatine
Ala
Fd_ox
Fd_red
Acetate
Acetaldehyde
2-Oxoglutarate
Vitamin B12
NH4+
AguD
Agumatine Deiminase Pathway
[FeFe]-hyd
Ethanol
NhaB
Fumarate
Frd
Putrescine
Citrate
Succinate
Liv-I/LS
ABC
GltP
GltS
AnsP
CitM
TTT
TRAP
DcuB
DcuC
BCAA, Phe
Amino acids
H+ Glu/Asp
Na+ Glu
Asn
Citrate
Carboxylate
Fumarate
B
Energy and carbon metabolism
0 20 40 60 80 100 [%]
Completeness of metabolic pathways
Number of flagellin genes
Total sequence length
Genome completeness
[Mbp]
1 2 3
20 40
Glycolysis/gluconeogenesis
Acetate production
Ethanol production
Fe hydrogenase HydABC
Fe hydrogenase B1/B3
[NiFe]-hydrogenase HybOAB
Rnf complex
FoF1-ATPase
Nitrate reductase
Complex I and related complex
Oxaloacetate decarboxylase
cbb3-type cytochrome c oxidase
Cytochrome bd oxidase
Agmatine deiminase pathway
Amino acid biosynthesis
Ala Asp Asn Glu Gln Thr Gly Ser Cys Met Val, Ile Leu Lys Arg Pro His Phe Tyr Trp
Sulfur assimilation
Nitrogen fixation
Flagellum
Chemotaxis
Insect-gut clade
MS499-41 (RsTz2-092)
CpT32-4
301-30.bin.171
TD116_bin.76
272-90.bin.133
Coa324_bin.33
Iy174_bin.14
Pcl387_bin.162
Mucispirillaceae
Mucispirillum schaedleri
BRZ_NI__bin29
"Ca. M. faecigallinarum"
BRZ_AT__bin62
BRZ_RS__bin62
Geovibrionaceae
S2C1195
UBA6793
Seleniivibrio woodruffii
UBA6799
Denitrovibrio acetiphilus
Geovibrio thiophilus
OH_SFB_213
Deferri_01
Limisalsivibrio acetivorans
Deferribacteraceae
Deferribacter desulfuricans
Deferribacter autotrophicus
Flexistipitaceae
UBA4806
Flexistipes sinusarabici
UBA1873
Calditerrivibrionaceae
Calditerrivibrio nitroreducens
ZAV-05
W28
Deferrivibrionaceae
Deferrivibrio essentukiensis
V6Fe1
M2_7_5_bin.3
"Ca. Microvillispirillaceae"
"Ca. Rimicarispirillum atlantis"

Figure 3

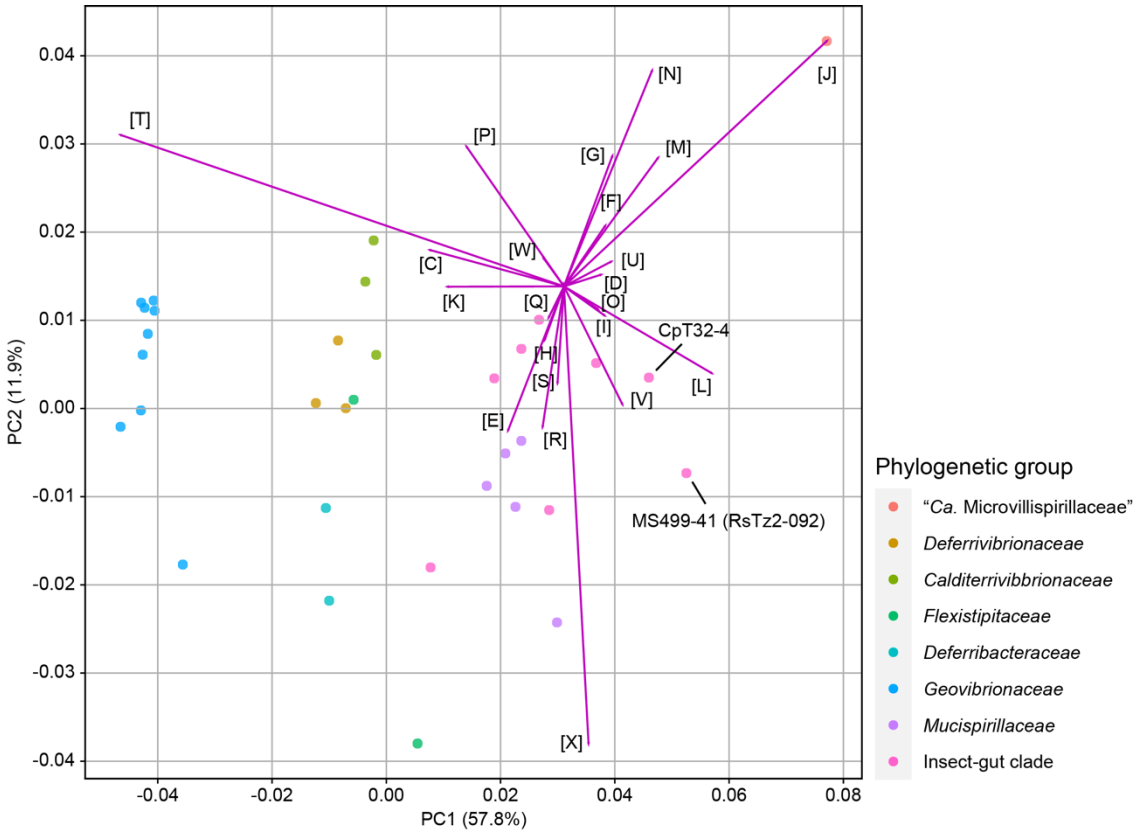
0.04
0.03
0.02
0.01
0.00
-0.01
-0.02
-0.03
-0.04
PC2 (11.9%)
-0.04
-0.02
0.00
0.02
0.04
0.06
0.08
PC1 (57.8%)
[T]
[P]
[N]
[J]
[G]
[M]
[F]
[W]
[C]
[U]
[D]
[K]
[Q]
[O]
[I]
CpT32-4
[H]
[S]
[L]
[V]
[E]
[R]
MS499-41 (RsTz2-092)
[X]
Phylogenetic group
"Ca. Microvillispirillaceae"
Deferrivibrionaceae
Calditerrivibbrionaceae
Flexistipitaceae
Deferribacteraceae
Geovibrionaceae
Mucispirillaceae
Insect-gut clade

Figure 4.