#### ARTICLE





### Genome analyses of uncultured TG2/ZB3 bacteria in 'Margulisbacteria' specifically attached to ectosymbiotic spirochetes of protists in the termite gut

Yuniar Devi Utami<sup>1</sup> • Hirokazu Kuwahara<sup>1</sup> • Katsura Igai<sup>1</sup> • Takumi Murakami<sup>1</sup> • Kaito Sugaya<sup>1</sup> • Takahiro Morikawa<sup>1</sup> • Yuichi Nagura<sup>1</sup> • Masahiro Yuki<sup>2</sup> • Pinsurang Deevong<sup>3</sup> • Tetsushi Inoue<sup>4</sup> • Kumiko Kihara<sup>1</sup> • Nathan Lo<sup>5</sup> • Akinori Yamada<sup>1,4</sup> • Moriya Ohkuma<sup>2</sup> • Yuichi Hongoh<sup>1,2</sup>

Received: 23 June 2018 / Revised: 20 September 2018 / Accepted: 25 September 2018 / Published online: 4 October 2018 © International Society for Microbial Ecology 2018

#### Abstract

We investigated the phylogenetic diversity, localisation and metabolism of an uncultured bacterial clade, Termite Group 2 (TG2), or ZB3, in the termite gut, which belongs to the candidate phylum 'Margulisbacteria'. We performed 16S rRNA amplicon sequencing analysis and detected TG2/ZB3 sequences in 40 out of 72 termite and cockroach species, which exclusively constituted a monophyletic cluster in the TG2/ZB3 clade. Fluorescence in situ hybridisation analysis in lower termites revealed that these bacteria are specifically attached to ectosymbiotic spirochetes of oxymonad gut protists. Draft genomes of four TG2/ZB3 phylotypes from a small number of bacterial cells were reconstructed, and functional genome analysis suggested that these bacteria hydrolyse and ferment cellulose/cellobiose to H<sub>2</sub>, CO<sub>2</sub>, acetate and ethanol. We also assembled a draft genome for a partner *Treponema* spirochete and found that it encoded genes for reductive acetogenesis from H<sub>2</sub> and CO<sub>2</sub>. We hypothesise that the TG2/ZB3 bacteria we report here are commensal or mutualistic symbionts of the spirochetes as H<sub>2</sub> sinks. For these bacteria, we propose a novel genus, '*Candidatus* Termititenax', which represents a hitherto uncharacterised class-level clade in 'Margulisbacteria'. Our findings add another layer, i.e., cellular association between bacteria, to the multi-layered symbiotic system in the termite gut.

**Electronic supplementary material** The online version of this article (https://doi.org/10.1038/s41396-018-0297-4) contains supplementary material, which is available to authorized users.

☑ Yuichi Hongoh yhongo@bio.titech.ac.jp

- <sup>1</sup> Department of Biological Sciences, Tokyo Institute of Technology, Tokyo 152–8550, Japan
- <sup>2</sup> Japan Collection of Microorganisms, RIKEN BioResource Research Center, Tsukuba 305–0074, Japan
- <sup>3</sup> Department of Microbiology, Kasetsart University, Bangkok 10900, Thailand
- <sup>4</sup> Graduate School of Fisheries and Environmental Sciences, Nagasaki University, Nagasaki 852–8521, Japan
- <sup>5</sup> School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia

#### Introduction

Termite guts harbour abundant and diverse gut microbes, comprising protists, bacteria, archaea and viruses [1, 2]. Most of these microbes are unique to termites, and together they constitute a complex microbiota that supplements the nutritional requirements of their insect hosts [1, 3]. Due to their resistance to cultivation, culture-independent studies have extensively been conducted to investigate the biology of these microbes [3–6].

Among the diverse bacterial lineages detected in the termite gut, Hongoh et al. reported that 16S rRNA phylotypes Rs-A23 (AB089067) and Rs-J96 (AB089068) from the gut of the termite *Reticulitermes speratus* did not cluster with any known bacterial phyla at that time [7] and designated the clade as the candidate phylum Termite Group 2 (TG2) [8]. This TG2 clade also included sequences from a sulphide-rich spring, the cluster of which had been named ZB3 [9].

Numerous 16S rRNA sequences affiliated with this TG2/ ZB3 clade have now accumulated in public databases from various environments such as the deep Arctic ocean [10], lake sediment [11], and hypersaline microbial mats [12], although no cultured representative exists thus far. Anantharaman et al. reported draft genome sequences of a clade, RIF30, which were reconstructed from metagenomes of groundwater and the sediment, and the authors proposed the name 'Margulisbacteria' for the RIF30 clade [13]. In the Genome Taxonomy Database (GTDB: http://gtdb. ecogenomic.org), the candidate phylum 'Margulisbacteria' consists of three class-level clusters, namely 'WOR-1', 'GWF2-35-9' and 'ZB3' [14], the latter two belonging to the original 16S rRNA-based TG2/ZB3 clade [8, 9].

Draft genomes of 'WOR-1', recently named as 'Saganbacteria' [15], were first recovered from estuary sediments [16], those of 'GWF2-35-9' were from the groundwater sediments [13], and those of 'ZB3' from marine water samples [17]. Matheus-Carnevali et al. characterised representative draft genomes of these three bacterial groups and compared them with the genomes of their sister phylum *Cyanobacteria*, proposing the candidate names 'Riflemargulisbacteria' for 'GWF2-35-9' and 'Marinamargulisbacteria' for 'ZB3' [18].

In the present study, we surveyed the distribution of TG2/ZB3 bacteria among diverse termite and cockroach lineages, located their cells, and obtained draft genomes in order to predict their ecological and physiological functions in the termite gut. Our results indicate that several TG2/ZB3 phylotypes are specifically attached to ectosymbiotic spirochetes of protists in the termite gut, and that the obtained genomes represent another class-level cluster in 'Margulisbacteria'. Our study expands our knowledge on the uncultured candidate phylum 'Margulisbacteria' and provides novel insights into the complex, multi-layered symbiotic system in the termite gut.

#### Materials and methods

#### Sample collection and DNA extraction

Sixty-two termite and 10 cockroach species were collected from six different regions (Table S1). DNA extraction from the entire gut of termites and cockroaches was performed as described previously [19, 20]. Identification of the host insect species was based on both morphological characteristics and the mitochondrial cytochrome oxidase II (COII) gene sequence [21]. Most of these were the same DNA samples used in our previous studies [8, 20–24].

#### Amplicon sequencing analysis of 16S rRNA gene

16S rRNA genes were amplified by PCR using *Bacteria*-specific primers, 341F and 806R (Table S2) [21], and the

products were subjected to sequencing on an Illumina MiSeq platform [25]. Sequence reads were trimmed, quality-filtered and sorted into amplicon sequence variants (ASVs) using the DADA2 v1.4 program package [26, 27] as described previously [20], except for the setting: trim-Left = 5 bp for both forward and reverse sequence reads. The obtained ASVs were phylogenetically classified using SINA v1.2.11 [28] with database SILVA SSURef NR99 release 132 [29] and minimum similarity of  $\geq$  85%. ASVs affiliated with *Eukarya*, *Archaea*, mitochondria, chloroplast or *Blattabacterium* and those not alignable to database sequences were discarded from subsequent analyses.

#### Design of PCR primers specific to 16S rRNA of TG2/ ZB3 and clone sequencing analysis

We designed a novel set of primers to specifically amplify the 16S rRNA gene of TG2/ZB3 bacteria, based on previously obtained sequences [7–9, 30–32]. The forward primer TG2-T-FW-243 is specific to termite-derived TG2/ ZB3 sequences and the reverse primer TG2-RV-1340 is universal for the majority of the entire TG2/ZB3 clade (Table S2). Nested PCR was performed with the annealing temperature 58 °C using diluted products of PCR with the *Bacteria*-specific 27F and 1390R primer set as the template, as described previously [22]. Cloning and Sanger sequencing were conducted also as described previously [22].

This TG2/ZB3-specific primer set was also used for screening clones of 16S rRNA genes amplified by PCR using primers 27F and 1492R (Table S2) from single or multiple cells of oxymonad protists as described previously [33]. The near full-length 16S rRNA gene clones of TG2/ZB3 were sequenced using the Sanger method [33].

#### **Phylogenetic analysis**

The 16S rRNA ASVs and clone sequences obtained were aligned with reference sequences in the SILVA database, using the ARB software [34] with manual corrections. The deduced amino acid sequences of COII were aligned using MUSCLE [35] with ambiguously aligned sites being removed by Gblocks [36]. Phylogenetic trees were constructed using MEGA7 [37] or FastTree v.2.1.9 [38]. Nucleotide or amino acid substitution models were selected using the model test implemented in MEGA7.

#### Fluorescence in situ hybridisation

Oligonucleotide probes for fluorescence in situ hybridisation (FISH) used in this study are listed in Table S2. Probe TG2-688, specific to 16S rRNA of the majority of the termite-derived TG2/ZB3 sequences, was designed using ARB. 16S rRNA phylotype-specific probes, RsDinE6-011039, HsPyr-01-1028, NkOx7-1-1002, NkOx7-2-1002 and NkOx-clu11-1257, were also designed, and non-labelled helpers or competitors were designed and used when needed. Probes Spiro-36 and Bactd-937 were used to detect the majority of the orders *Spirochaetales* and *Bacteroidales*, respectively [22]. These probes were labelled at their 5' end with either of 6-carboxyfluorescein (6FAM), Texas Red, or Alexa Fluor 647. FISH was performed basically as described previously [20, 39]. Observation was conducted under an Olympus BX51 epifluorescence microscope or an Olympus FV1000D-IX81 confocal laser scanning microscope (Olympus, Tokyo, Japan).

#### Preparation of single-cell genome samples

The gut contents from a worker of the termite Hodotermopsis sjoestedti (family Archotermopsidae) collected in Kagoshima Prefecture, Japan, were suspended and homogenised in 1 mL of sterile solution U buffer [40]. The mixture was centrifuged at  $500 \times g$  for 3 min at 4 °C. The supernatant was discarded, and the pellet was resuspended in 750 µL sterile solution U. This process was repeated three times. Pulse sonication was performed on ice in order to break the protist cells. The mixture was then added to 5 mL sterile solution U and filtrated using 80, 40, 12, 5 and 3 µm membrane filters in this order (Merck Millipore, Burlington, MA, USA). The filtrate was combined with 5 µL of fluorescent dye FM1-43 (Molecular Probes, Eugene, OR, USA) and subjected to fluorescence-activated single-cell sorting on a BD FACSJazz (BD Biosciences, San Jose, CA, USA). Cell-sorting and subsequent whole genome amplification (WGA) using the Illustra GenomiPhi V2 Kit (GE Healthcare, Little Chalfont, UK) were carried out as described previously [20]. The single-cell WGA samples were taxonomically identified by Sanger sequencing of 16S rRNA genes amplified by PCR as described previously [20]. The degree of contamination was judged from the quality of 16S rRNA gene sequences as described by Yuki et al. [41].

### Preparation of metagenomic samples from single protist cells

Single protist cells were physically isolated by micromanipulation basically according to Kuwahara et al. [42]. The termites *R. speratus* (family Rhinotermitidae) sampled in Saitama or Ibaraki Prefecture, Japan, and *Neotermes koshunensis* (family Kalotermitidae) sampled in Okinawa Prefecture, Japan, were used for the experiments. The isolated protist cells were dissected using a TransferMan NK2 micromanipulator (Eppendorf, Hamburg, Germany) equipped with a microblade (FEATHER Safety Razor, Osaka, Japan) as described previously [33, 42], to remove the host nucleus and collect a portion with cells of TG2/ZB3 bacteria. Bacterial cells collected from a single protist cell were subjected to WGA using the Illustra GenomiPhi V2 Kit as described previously [33].

To check the presence of DNA derived from TG2/ZB3 bacteria in the WGA samples, PCR was performed using the TG2-T-FW-243 and TG2-RV-1340 primer set. For the samples containing TG2/ZB3 DNA, second WGA using the Illustra GenomiPhi HY Kit (GE Healthcare) was performed at 30 °C for 4 h, as described previously [43].

### Genome sequencing, assembly, binning and gene annotation

Sequencing libraries for WGA samples were prepared using the TruSeq DNA Sample Preparation Kit for low DNA input (Illumina). Sequencing was performed on the Illumina MiSeq platform with the Reagent Kit V3. For each of the WGA samples NkOx7 and RsDinE6, an additional library was prepared using the Nextera Mate Pair Library Preparation Kit (Illumina) and sequenced. Adaptor removal and quality trimming were performed using the cutadapt [44] and prinseq program [45], respectively. Qualified reads were assembled to construct contigs using SPAdes 3.10.0 [46].

Contigs were binned into respective bacterial assemblages contained in the sample, based on the tetranucleotide frequency and single-copy gene markers, using the MyCC program with default settings [47]. Bins with single-copy gene markers showing the highest sequence similarity to sequences of 'Margulisbacteria' by BLASTp searches against the NCBI non-redundant (nr) protein database were picked up and submitted to MiGAP (http://www.miga p.org/) for automatic gene finding and annotation. Additional gene annotation was conducted using PROKKA v1.12 [48].

#### Genome characterisation and phylogenomics

Genome completeness was estimated using CheckM v1.07 based on 104 single-copy gene markers conserved in *Bacteria* [49]. The automatic annotation using MiGAP was manually checked by BLAST searches against the NCBI nr protein, nucleotide database or Conserved Domain Database v.3.16 [50]. The overall nucleotide sequence similarity between genomes was calculated using the Genome-to-Genome Distance Calculator with formula 2 [51]. Metabolic pathways were inferred with the KEGG automatic annotation server (KAAS) [52]. Clustered regularly interspaced short palindromic repeat (CRISPR) loci were identified using CRISPRFinder [53]. Twenty-six single-copy gene markers listed in Table S3 were used for construction of a phylogenomic tree of 'Margulisbacteria'. Deduced amino acid sequences were aligned using hmmalign in HMMER 3.1b2 (http://hmmer.org/), concatenated, edited using Gblocks, and a maximum-likelihood tree was constructed using the FastTree program.

#### **Accession numbers**

The sequence data obtained in this study have been deposited in DDBJ under the BioProject PRJDB7102 for the phylotypes HsPyr-01 draft genomes of (BGZM01000001-22), RsDinE6-01 (BGZP01000001-32), NkOx7-01 (BGZN0100001-233) NkOx7-02 and (BGZO01000001-201). Representative 16S rRNA sequences of TG2/ZB3 will appear under accession numbers LC387877-953. The 16S rRNA gene and genes involved in reductive acetogenesis from the ectosymbiotic spirochete NkOx-clu11 will appear under accession numbers LC387957-71.

#### Results

### Phylogenetic diversity and abundance of TG2/ZB3 bacteria in termite guts

16S rRNA gene amplicons of TG2/ZB3 bacteria were detected using the MiSeq platform in 34 out of the 62 termite species and in none of the 10 cockroach species, with relative abundance ranging from 0.01 to 1.58% (Fig. 1). No sequence affiliated with the 'WOR-1' clade was detected. With such low relative abundance, we additionally performed PCR amplification of 16S rRNA genes using the TG2/ZB3-specific primer set for 17 samples (Table S1), cloned the amplicons, and sequenced them using the Sanger method, to further explore TG2/ZB3 sequence diversity and to obtain longer sequences for phylogenetic analysis with higher resolution. We obtained TG2/ZB3 sequences from 15 samples; these included five termite species and one cockroach species, Cryptocercus punctulatus, from which TG2/ZB3 sequences were not detected in the above amplicon analysis. In total, TG2/ZB3 sequences were recovered from 40 out of the 72 termite and cockroach species (Fig. 1).

The 16S rRNA genes of TG2/ZB3 bacteria derived from termites and *C. punctulatus* with  $\geq$ 1200 bp, obtained in this and previous studies, exclusively constituted a robust monophyletic cluster (Fig. 2). The sequence identity within this cluster is  $\geq$ 94.2%. All 32 TG2/ZB3 ASVs obtained in the amplicon analysis were also affiliated with this cluster. One to three ASVs were identified from each termite species, and most ASVs were detected only in one termite species. Multiple samples from one termite species tended to share the same ASVs; ASV5472 was detected in two *Nasutitermes takasagoensis* samples collected from distinct



Fig. 1 Maximum-likelihood tree of termites and cockroaches used in this study based on deduced amino acid sequences of the mitochondrial cytochrome oxidase II gene. The 16S rRNA genes of TG2/ZB3 bacteria were detected using the MiSeq platform from the insect species shown in red, and those detected using PCR with primers specific to TG2/ZB3 are shown in blue. The frequency of TG2/ ZB3 sequences is indicated for each sample. A total of 200 amino acid sites were used with the mtRev + G + I substitution model and 500 bootstrap resamplings in MEGA7. Bootstrap confidence values ≥ 50% (open circles) and ≥70% (closed circles) are indicated

localities, and ASV1770 was detected in three *Neotermes koshunensis* samples from different colonies in the same locality (Fig. S1). Five ASVs were detected in samples from two or more different but closely related termite species (Fig. S1).



Fig. 2 Phylogenetic positions of termite/cockroach gut-derived TG2/ ZB3 sequences in the candidate phylum 'Margulisbacteria' based on the 16S rRNA gene (≥1200 bp). A maximum-likelihood tree was constructed with the GTR + G substitution model and 100 bootstrap resamplings in FastTree. A total of 1248 nucleotide positions were used. Representatives of the phylum Cyanobacteria were used as outgroup taxa (Table S6). All sequences from termite and cockroach gut samples fell into one cluster. Clone ZB15 (\*), used for the original designation of the candidate phylum ZB3 [9], was affiliated with the indicated cluster, although the short sequence was not included in this tree. Clone ZB21 (similarly not included in this tree) and clone ML635J-21, also used by Elshahed et al. [9], were affiliated with the 'GWF2-35-9' clade. We consider that the 'ZB3' clade designated in GTDB [14] corresponds to the cluster 'marine water and sediment', because this robust monophyletic cluster contains the 16S rRNA sequences of several genomes of the 'ZB3' clade. Local support values based on SH test ≥50% are indicated on internal branches

#### Localisation of TG2/ZB3 bacteria in the termite gut

TG2/ZB3 bacterial cells were detected by FISH using probe TG2-688 in the guts of three lower termites, H. sjoestedti, R. speratus and Neotermes koshunensis. All TG2/ZB3 cells detected were attached to cells of ectosymbiotic spirochetes of protists belonging to the order Oxymonadida in the phylum Preaxostyla: Pyrsonympha sp. from H. sjoestedti, Dinenympha leidyi and Dinenympha porteri type III from R. speratus, and Oxymonas sp. from Neotermes koshunensis (Fig. 3). Their attachment to the ectosymbiotic spirochetes and not to ectosymbionts belonging to the order Bacteroidales [54-56] was confirmed in Oxymonas sp. (Fig. S2). Not all cells of the oxymonad host species harboured TG2/ ZB3 attached to the ectosymbiotic spirochetes: the frequency of the TG2/ZB3-associated oxymonad host cells was ~10% or below. No TG2/ZB3 cells free from oxymonad cells were found in these lower termite guts. We failed to detect TG2/ZB3 cells by FISH using probe TG2-688 in the guts of two higher termite species, *Microcerotermes* sp. collected in Bangkok, Thailand and *Nasutitermes takasagoensis* collected in Iriomote Island, Okinawa Prefecture, Japan.

# General genome features of TG2/ZB3 bacteria attached to ectosymbiotic spirochetes of oxymonad protists

We obtained draft genomes or genome fragments of four TG2/ZB3 phylotypes from the three lower termite species *H. sjoestedti*, *R. speratus* and *Neotermes koshunensis* (Table 1; Fig. S3). Genome fragments of phylotype HsPyr-01 from *H. sjoestedti* was obtained by single-cell genomics using FACS. Its 16S rRNA gene was coded in a large contig. The association of phylotype HsPyr-01 with *Pyrsonympha* sp. was confirmed by cloning and sequencing 16S rRNA genes amplified by PCR from a pool of *Pyrsonympha* sp. cells and also by FISH analysis using the specific probe HsPyr-01-1028 (Fig. S4).

Genomes of the other three TG2/ZB3 phylotypes were reconstructed by binning metagenomic contigs derived from oxymonad protist cells: phylotype RsDinE6-01 from a single D. leidyi cell in a R. speratus gut, and two phylotypes, NkOx7-01 and NkOx7-02, from a single Oxymonas sp. cell in a Neotermes koshunensis gut. The 16S rRNA gene of RsDinE6-01 was not found in the contigs, but recovered by PCR using the TG2/ZB3-specific primer set from the RsDinE6 sample. Only one 16S rRNA phylotype of TG2/ZB3 was detected, and only one TG2/ZB3 bin was generated; therefore, it seems safe to consider that the 16S rRNA gene was of the RsDinE6-01 genome. While a large contig of the NkOx7-02 bin contained a full-length 16S rRNA gene, only a truncated 16S rRNA gene was found on the edge of a large contig in the NkOx7-01 bin. We recovered a near full-length 16S rRNA gene of the latter phylotype by PCR using primer 1492R and a primer specific to a contig region adjacent to the 16S rRNA gene. The obtained 16S rRNA gene of NkOx7-01 showed 97.2% sequence similarity with that of NkOx7-02. The attachment of these three phylotypes to ectosymbiotic spirochetes of the respective oxymonad hosts was confirmed by FISH using the phylotype-specific probes (Fig. S5 and S6).

The estimated genome completeness based on the singlecopy gene markers was low in HsPyr-01 (4.9%) and RsDinE6-01 (6.9%) and high in NkOx7-01 (89.7%) and NkOx7-02 (87.1%) (Table 1). No contaminating contig was detected using CheckM. The genome size of these phylotypes was estimated to be 2.1 to 2.5 Mbp except that of RsDinE6-01 (Table 1). Their GC contents were 45.3 to 49.8%. Three and one CRISPR regions were identified in the NkOx7-01 and NkOx7-02 genomes, respectively



Fig. 3 Fluorescence in situ hybridisation (FISH) analysis of TG2/ZB3 bacteria in the gut of lower termites. Phase contrast (a-d) and epifluorescence (e-h) images are shown. *Pyrsonympha* sp. from the gut of *Hodotermopsis sjoestedti* (a, e), *Dinenympha leidyi* from the gut of *Reticulitermes speratus* (b, f), and *Oxymonas* sp. from the gut of *Neotermes koshunensis* (c, g) were observed. TG2/ZB3 cells were

Table 1 General features ofTG2/ZB3 genomes obtained in

this study

detected with probe TG2-688 (Texas red-labelled, red), and spirochetes were detected with probe Spiro-36 (6-carboxyflorescein, green) (Table S2). e Magnified FISH image of the area surrounded by a rectangle in a. d, h Magnified images of the area surrounded by a rectangle in c. Arrows indicate examples of detected TG2/ZB3 cells. Bars indicate 10 µm in b, d–f and h; 50 µm in a, c and g

| Phylotype                      | HsPyr-01                                    | RsDinE6-01                            | NkOx7-01                                 | NkOx7-02                              |
|--------------------------------|---|---------------------------------------|--|---------------------------------------|
| Proposed name                  | <i>'Ca.</i> Termititenax spirochaetophilus' | <i>'Ca.</i> Termititenax dinenymphae' | <i>'Ca.</i><br>Termititenax<br>aidoneus' | <i>'Ca.</i> Termititenax persephonae' |
| Termite host                   | Hodotermopsis<br>sjoestedti                 | Reticulitermes<br>speratus            | Neotermes<br>koshunensis                 | Neotermes<br>koshunensis              |
| Protist host                   | Pyrsonympha sp.                             | Dinenympha leidyi                     | Oxymonas sp.                             | Oxymonas sp.                          |
| Number of contigs              | 22  | 32                                    | 233                                      | 201                                   |
| Total length (bp)              | 122,812                                     | 359,801                               | 2,258,481                                | 1,834,680                             |
| Completeness (%) <sup>a</sup>  | 4.9   | 6.9                                   | 89.7                                     | 87.1                                  |
| Contamination (%) <sup>a</sup> | 0   | 0                                     | 0  | 0                                     |
| Predicted genome<br>size (Mbp) | 2.5   | 5.2 <sup>b</sup>                      | 2.5                                      | 2.1                                   |
| G+C (%)                        | 47.8  | 45.3                                  | 47.7                                     | 49.8                                  |
| CDS                            | 135   | 355                                   | 2309                                     | 1742                                  |
| rRNA                           | 16S   | 16S                                   | 58, 238, 168                             | 16S, 23S                              |

<sup>a</sup>The completeness and contamination rate were estimated based on single-copy marker genes using CheckM [49]

<sup>b</sup>This bin contained only 6 out of 104 single-copy gene markers used in the CheckM program, and therefore the predicted genome size is unreliable

(Table S4). The overall genetic distance in alignable genome regions between these phylotypes ranged from 22.2% to 24.4% (Table S5). Their phylogenetic positions within the termite/cockroach-derived cluster are shown based on the 16S rRNA gene in Fig. S1. The phylogenetic relationships of these four genomes with known draft genomes of



**Fig. 4** Phylogenomic tree of TG2/ZB3 bacteria based on concatenated amino acid sequences of single-copy marker genes. A maximum-likelihood tree was constructed with the LG + G substitution model and 100 bootstrap resamplings. A total of 4039 amino acid sites from 26 single-copy marker genes conserved within the candidate phylum 'Margulisbacteria' were used (Table S3). Local support values  $\geq$ 50% (open circles) and  $\geq$ 70% (closed circles) are indicated. Members of the phylum *Cyanobacteria* were used as outgroup (Table S6) and omitted from the tree

'Margulisbacteria' (listed in Table S6) are shown based on the concatenated amino acid sequences (Fig. 4).

Phylotypes HsPyr-01, RsDinE6-01 and NkOx7-01 were detected with low abundances by the 16S rRNA amplicon analysis of the entire gut microbiota from each host termite species, as ASV5974 (frequency: 0.36%), ASV8091 (0.14%) and ASV1770 (0.26, 0.51, 0.52% for three *Neotermes koshunensis* gut samples), respectively (Fig. S1). Phylotype NkOx7-02 was not detected in the amplicon sequence analysis, but detected by cloning analysis of 16S rRNA genes from a single *Oxymonas* sp. cell simultaneously with phylotype NkOx7-01. Multiple TG2/ZB3 phylotypes were also detected in a single *D. leidyi* cell from an *R. speratus* gut (Rs-Dine-2-A, -B, -C in Fig. S1).

#### Predicted metabolism of the TG2/ZB3 phylotypes

We predicted the metabolic pathways of these TG2/ZB3 bacteria based on the genome sequences mostly of NkOx7-01 and NkOx7-02, due to their high genome completeness. Most genes present in the genome fragments of HsPyr-01 and RsDinE6-01 were found in the genomes of NkOx7-01 and/or NkOx7-02. The genomes encoded pathways for glycolysis, gluconeogenesis and the non-oxidative pentose phosphate pathway (Fig. 5). The genes required for the synthesis of 2-oxoglutarate from pyruvate and fumarate hydratase were found, but other genes coding for cytochrome oxidase, catalase and superoxide dismutase were found, but these genomes encoded rubrerythrin and rubredoxin.

The NkOx7-02 genome encoded a gene for putative endo-\beta-1,4-glucanase CelB belonging to the glycoside hydrolase family (GHF) 9, the closest homologue of which is secreted extracellularly in Ruminoclostridium thermocellum [57] (Fig. S7a). The same contig also encoded a putative endo- $\beta$ -1,4-glucanase of GHF8, which showed the highest amino acid sequence similarities to those of other margulisbacteria in the NCBI nr protein database (Fig. S7b). The corresponding genome region was not recovered in the other three genomes. A gene coding for cellobiose phosphorylase CbpA (GHF94), which is possibly localised in the cytoplasm and cleaves cellobiose into  $\alpha$ -D-glucose 1phosphate and D-glucose [58], was found in the NkOx7-01 and RsDinE6-01 genomes. A gene for β-glucosidase (GHF116) adjacent to the uncharacterised sugar transporter genes *ycjNOPV* in the opposite direction was identified in both the NkOx7-01 and NkOx7-02 genomes. The cellular localisation of this  $\beta$ -glucosidase is unclear. The NkOx7-01 genome additionally encoded a putative periplasmic  $\beta$ -glucosidase BglX gene (GHF3). No genes coding for hemicellulase were found. Thus, it is likely that phylotype NkOx7-02 utilises cellulose and others also utilise at least cellobiose as a primary carbon and energy source.

The resulting glucose is probably fermented to  $H_2$ ,  $CO_2$ , ethanol and acetate (Fig. 5). The reduced ferredoxin generated during oxidation of pyruvate to acetyl-CoA is presumably re-oxidised by the action of an Rnf complex and a trimeric [FeFe] hydrogenase (HydABC). In the former, the oxidation of ferredoxin is coupled with the export of protons; the bacteria could also generate ATP using the proton motive force through the F-type or V-type ATPase (Fig. 5).

No genes encoding proteins for nitrogen fixation were found in any of the four genomes. However, genes encoding an ammonium transporter (Amt), an amino acid transporter for methionine (MetINQ) and a putative oligopeptide transporter (PotE) were identified. Biosynthetic pathways for most amino acids and several cofactors as well as nucleic acids were also found. Genes containing integrin alpha motif (smart00191), VCBS repeat (pfam13517) and FG-GAP repeat (pfam01839) were encoded in the genomes of RsDinE6-01, NkOx7-01 and NkOx7-02. These motifs are known as components of proteins involved in cell-cell adhesion [59]. NkOx7-02 possessed two sialic acid synthase genes, which may be involved in cell recognition [60]. Genes for flagellar components are present in both the NkOx7-01 and NkOx7-02 genomes; these bacteria are likely motile, and these flagella may assist the cell-cell adhesion [61]. The biosynthetic pathways for peptidoglycan, lipopolysaccharide (LPS) and S-layer-like component as well as exporters of LPS and lipidA were found. Genes for a type I secretion system, including the outer membrane protein TolC, were also found (Fig. 5). These suggest



Fig. 5 Predicted metabolic pathways of phylotypes NkOx7-01 and NkOx7-02. Pathways found only in one of the two draft genomes are highlighted with colours indicated below the pathway map. Incomplete pathways are indicated with dotted lines

that the TG2/ZB3 bacteria have a Gram-negative type cell wall.

#### Discussion

### Reductive acetogenesis from H<sub>2</sub> and CO<sub>2</sub> of an ectosymbiotic spirochete partner

During the binning of the metagenomic contigs from the single Oxymonas sp. cell, we also reconstructed the draft genome of an ectosymbiotic spirochete phylotype, designated here NkOx-clu11 (Fig. S3b), which was affiliated with the Treponema clade Ia [62] based on the 16S rRNA gene (data not shown). The genome completeness was estimated to be 63%, and the genome encoded genes responsible for the reductive acetogenesis from H<sub>2</sub> and CO<sub>2</sub> (Wood-Ljungdahl pathway), including an acsABCDEF (acetyl-CoA synthase complex) gene cluster and *fdhA* (formate dehydrogenase alpha subunit) (Table S7), although other components required for the pathway were not found possibly due to the incompleteness of the genome. The acsABCDEF and fdhA genes were phylogenetically closest to those of *Treponema primitia* (Fig. S8), which is an  $H_2/$ CO<sub>2</sub>-homoacetogen isolated from the gut of the termite Zootermopsis angusticollis [63, 64]. The attachment of the NkOx-clu11 treponeme with NkOx7-01 and NkOx7-02 cells was confirmed by FISH analysis using specific probes (Fig. S6).

SPRINGER NATURE

We detected the 16S rRNA genes of TG2/ZB3 bacteria from diverse termite species and a cockroach of the genus *Cryptocercus*, which is the sister group of termites. Because all TG2/ZB3 sequences derived from termites and the *Cryptocercus* cockroach exclusively formed a monophyletic cluster, it is likely that these bacteria are specific gut symbionts of termites and *Cryptocercus* cockroaches, although other cockroach lineages may harbour the bacteria with abundance below the detection limit of the present study. The fact that most ASVs were not shared by different termite species and that some ASVs were detected in distinct samples of single termite species, the relationship between the TG2/ZB3 bacteria and the termite hosts appears to be basically species specific.

Interestingly, the TG2/ZB3 cells in the gut of lower termites were found specifically from the surrounding area of oxymonad protist cells; they do not attach to the protist cell surface, but to the ectosymbiotic spirochete cells. Because we did not observe their association with other ectosymbiotic bacteria of oxymonads nor with ecto-symbiotic spirochetes of parabasalid protists, it appears that the TG2/ZB3 bacteria can recognise specific spirochete species as their partners. Since the TG2/ZB3 bacteria likely possess flagella, they may swim and find their specific

partner spirochetes. We still do not know the localisation of TG2/ZB3 cells in the guts of higher termites.

The NkOx7-01 and NkOx7-02 genomes encoded CRISPR–Cas systems. CRISPR loci have been found in the genomes of several other bacterial species from the termite gut [43, 65–67], and we also found CRISPR–Cas systems widely in other known genomes of 'Margulisbacteria' (data not shown). The absence of cytochrome oxidase, catalase and superoxide dismutase indicated that they are strict anaerobes, and the presence of rubrerythrin and rubredoxin may facilitate the survival of these bacteria during oxidative stress [68]. The TG2/ZB3 genomes encoded the biosynthetic pathways for a Gram-negative type cell wall; we found those pathways also in most genomes of 'Margulisbacteria' (data not shown).

The TG2/ZB3 genomes analysed in this study encoded genes involved in the utilisation of cellulose as a carbon and energy source. The potential to hydrolyse cellulose has been suggested in 'WOR-1' genomes [16], and one of the two putative endoglucanases encoded in the NkOx7-02 genome showed high sequence similarities to those of the 'WOR-1' genomes. The other endoglucanase CelB, on the other hand, may be unique to the termite/cockroach-derived TG2/ZB3 cluster. The participation of gut bacteria in lignocellulose digestion has been suggested in higher termites [32, 69–73] and in lower termites [41, 74]. It remains unknown whether the endoglucanases of the TG2/ZB3 bacteria are localised on their cell surface or released from the cells (Fig. 6). In any case, it appears that the TG2/ZB3 bacteria have a mutualistic relationship with the termite hosts via participation in cellulose digestion and also by upgrading nitrogenous compounds. Differences in ecological niche between phylotypes NkOx7-01 and NkOx7-02, both of which can attach to the same Treponema phylotype, were difficult to examine in this study only with the draft genome sequences.

The production of  $H_2$  during the fermentation of glucose may be a key aspect of the cellular symbiosis of TG2/ZB3 with ectosymbiotic spirochetes. At least one partner ectosymbiotic *Treponema* phylotype, NkOx-clu11, possesses key genes for reductive acetogenesis from  $H_2$  and CO<sub>2</sub>; the



Fig. 6 Hypothesised symbiotic relationship of the TG2/ZB3 bacteria and ectosymbiotic spirochetes of oxymonad protists

TG2/ZB3 bacteria have possibly exploited the spirochetes as  $H_2$  sinks to promote their fermentation process (Fig. 6). Reductive acetogenesis by spirochetes in the termite gut has been reported in many studies e.g., [75-78] and investigated in detail in T. primitia [63, 64, 79]. In co-cultures of T. primitia and H<sub>2</sub>-producing Treponema azotonutricium, interspecies transfer of H<sub>2</sub> was demonstrated [80]. In 'Candidatus Treponema intracellularis' and 'Candidatus Treponema teratonymphae', which are intracellular symbionts of parabasalid protists, it has been hypothesised that these endosymbiotic H<sub>2</sub>/CO<sub>2</sub>-homoacetogens decrease the high H<sub>2</sub> partial pressure inside their protist hosts, thereby promoting the fermentation process of host cells [65, 81]. However, to the best of our knowledge, the attachment of an H<sub>2</sub>-producing bacterium to an H<sub>2</sub>/CO<sub>2</sub>-homoacetogenic spirochete has hitherto never been reported. The direct attachment to an H2-oxidiser may benefit the metabolism of the TG2/ZB3 bacteria. This association also possibly helps the TG2/ZB3 bacteria to avoid washout from the gut.

On the other hand, assuming that the cellulolytic oxymonad protist hosts produce hydrogen as seen in cellulolytic parabasalid protists e.g., [82, 83], the ectosymbiotic spirochetes can take up hydrogen generated by their protist hosts [84]. Since the resulting hydrogen partial pressure in the termite hindgut varies among the termite lineages [85, 86], the impact of the H<sub>2</sub>-producing TG2/ZB3 bacteria on the metabolism of the ectosymbiotic spirochetes may also vary. For example, the association with the TG2/ZB3 bacteria in the hindgut of a *Neotermes* termite, where the H<sub>2</sub> partial pressure is relatively low [86], may be more beneficial to the ectosymbiotic spirochetes than in the hindgut of a Hodotermopsis termite, where the H<sub>2</sub> partial pressure would be much higher as predicted from the value for its related genus [85]. We therefore hypothesise that the TG2/ ZB3 bacteria described here are mutualistic or commensal symbionts of the ectosymbiotic spirochetes of the oxymonad protists. Interchanges of metabolites other than  $H_2$  are possible such as tetrahydrofolate [87], but we need the complete genome sequences of both partners to predict it.

In the candidate phylum 'Margulisbacteria', the four genomes obtained in this study constituted a clade distinct from the class-level clusters 'WOR-1', 'GWF2-35-9' or 'ZB3' in GTDB [14] (Fig. 4). Consistently, the 16S rRNA gene tree indicated that the termite/cockroach-derived clade is included in another class-level cluster (Fig. 2). On the basis of the phylogenetic, morphological and genomic data obtained in this study, we propose a novel genus, 'Candidatus Termititenax', for the termite/cockroach-derived monophyletic cluster sharing ≥94% 16S rRNA sequence identity, and tentatively name the class-level cluster 'Candidatus Termititenacia', the members of which in Fig. 2 identity. share  $\geq 80\%$ 16S rRNA Ca.sequence

Termititenacia' showed 66–75% 16S rRNA sequence similarities to members of 'WOR-1', 76–84% to members of 'GWF2-35-9' and 74–86% to members of 'ZB3' in Fig. 2. We also propose a novel order, '*Candidatus* Termititenacales', and a novel family, '*Candidatus* Termititenacaceae'. We propose four novel species of '*Ca*. Termititenax' as described below.

#### Description of 'Candidatus Termititenax' gen. nov

Termititenax [ter.mit.i.te'nax, L. masc. n. termes, termite; N. L. masc. substantive from L. adj. *tenax*, clinging, tenacious; N.L. masc. n. termititenax, (bacteria) clinging to termites]. The bacteria specifically attach to ectosymbiotic spirochetes of oxymonad protists in the gut of lower termites. Localisation in the guts of higher termites is unknown. The cells are short rods or curved rods 0.6-1.4 µm by 0.2-0.5 µm. Based on genome sequence data, the bacteria have a Gramnegative type cell wall and are motile with flagella. They are strict anaerobes, and hydrolyse and ferment sugars to H<sub>2</sub>, CO<sub>2</sub>, acetate and ethanol. Members of this genus have not been cultured thus far, and its taxonomic assignment is based on the 16S rRNA gene sequence and specific hybridisation with probe TG2-688 (Table S2), and the sequences of single-copy taxonomic marker genes. The type species is 'Candidatus Termititenax aidoneus' (corresponding to phylotype NkOx7-01) with its metagenome-assembled genome sequence as the type material, following the proposal for nomenclature of uncultured taxa by Chuvochina et al. [88].

### Description of 'Candidatus Termititenax aidoneus' sp. nov

*Termititenax aidoneus* (ai.do.neus, Gr. n. *Aidoneus*, an alternative name for Hades, meaning 'The Unseen One', referring to its inconspicuousness among ectosymbiotic bacteria). The bacterium attaches to ectosymbiotic spirochetes of *Oxymonas* sp. in the gut of *Neotermes koshunensis*. The cell dimensions are  $0.8-1.4 \,\mu$ m by  $0.3-0.4 \,\mu$ m. This assignment is based on the draft genome (BGZN01000001–233), 16S rRNA gene, and specific hybridisation with probe NkOx7-01-1002 (Table S2). This species corresponds to phylotype NkOx7-01.

## Description of 'Candidatus Termititenax persephonae' sp. nov

Termititenax persephonae (per.se.phonae, L. n. Persephone, the wife of Hades; L. gen. n. persephonae, of Persephone, referring to its simultaneous discovery with 'Ca. Termitenax aidoneus'). The bacterium attaches to ectosymbiotic spirochetes of Oxymonas sp. in the gut of *Neotermes koshunensis*. The cell dimensions are 0.8–1.1 µm by 0.3–0.4 µm. The bacterium possibly utilises cellulose. This assignment is based on the draft genome (BGZO01000001–201), 16S rRNA gene and specific hybridisation with probe NkOx7-02-1002 (Table S2). This species corresponds to phylotype NkOx7-02.

### Description of *'Candidatus* Termititenax spirochaetophilus' sp. nov

*Termititenax spirochaetophilus* (spi.ro.chae.to.phi.lus, N.L. fem. n. *Spirochaeta*, a genus name of bacteria, spirochete; N.L. masc. substantive *philus* from Gr. adj. *philos*, friend, loving; N.L. masc. n. *spirochaetophilus*, spirochete-loving). The bacterium specifically attaches to ectosymbiotic spirochetes of *Pyrsonympha* sp. in the gut of *Hodotermopsis sjoestedti*. The cell dimensions are 0.6–1.2 µm by 0.3–0.5 µm. This assignment is based on the genome fragments (BGZM01000001–22), 16S rRNA gene and specific hybridisation with probe HsPyr-01-1028 (Table S2). This species corresponds to phylotype HsPyr-01.

### Description of *'Candidatus* Termititenax dinenymphae' sp. nov

*Termititenax dinenymphae* (di.ne.nymphae, N.L. n. *Dinenympha*, a genus name of flagellated protists; N.L. gen. n. *dinenymphae*, of *Dinenympha*, referring to the host genus). The bacterium attaches to ectosymbiotic spirochetes of *Dinenympha leidyi* in the gut of *Reticulitermes speratus*. The cell dimensions are  $0.8-1.2 \,\mu\text{m}$  by  $0.2-0.3 \,\mu\text{m}$ . This assignment is based on the genome fragments (BGZP01000001–32), 16S rRNA gene and specific hybridisation with probe RsDinE6-01-1039 (Table S2). This species corresponds to phylotype RsDinE6-01.

Acknowledgements We thank Warin Boonriam, Sasitorn Hasin, Kazuki Izawa, Yukihiro Kinjo, Osamu Kitade, Tomonori Kume, Kiyoto Maekawa, Shigeharu Moriya, Christine Nalepa, Michael Neal, Yung Chul Park, Gaku Tokuda, Savitr Trakulnaleamsai and Charunee Vongkaluang for assisting with our termite and cockroach collection. We also thank Jun-ichi Inoue, Masato Kochi, Takako Mabuchi, Kei Matsushima, Yuto Mochizuki and Shu Onouchi for assisting with experiments. Sanger sequencing was performed at the Biomaterial Analysis Center of the Tokyo Institute of Technology or RIKEN BSI. This work was financially supported by a MEXT scholarship to YDU, and NEXT and KAKENHI grants from JSPS to YH (GS009, 23117003, 26650158, 16H04840, 17K19423), AY (25850104) and MO (17H01447). This work was also supported by a RIKEN Competitive Program for Creative Science and Technology to YH and MO, and JST-CREST (14532219) to YH.

#### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

#### References

- Brune A. Symbiotic digestion of lignocellulose in termite guts. Nat Rev Microbiol. 2014;12:168–80.
- Pramono AK, Kuwahara H, Itoh T, Toyoda A, Yamada A, Hongoh Y. Discovery and complete genome sequence of a bacteriophage from an obligate intracellular symbiont of a cellulolytic protist in the termite gut. Microbes Environ. 2017;32:112–7.
- Hongoh Y. Toward the functional analysis of uncultivable, symbiotic microorganisms in the termite gut. Cell Mol Life Sci. 2011;68:1311–25.
- Brune A, Dietrich C. The gut microbiota of termites: digesting the diversity in the light of ecology and evolution. Annu Rev Microbiol. 2015;69:145–66.
- Bourguignon T, Lo N, Dietrich C, Šobotník J, Sidek S, Roisin Y, et al. Rampant host switching shaped the termite gut microbiome. Curr Biol. 2018;28:649–54.
- Tikhe CV, Husseneder C. Metavirome sequencing of the termite gut reveals the presence of an unexplored bacteriophage community. Front Microbiol. 2018;8:e2548.
- Hongoh Y, Ohkuma M, Kudo T. Molecular analysis of bacterial microbiota in the gut of the termite *Reticulitermes speratus* (Isoptera; Rhinotermitidae). FEMS Microbiol Ecol. 2003;44:231–42.
- Hongoh Y, Deevong P, Inoue T, Moriya S, Trakulnaleamsai S, Ohkuma M, et al. Intra- and interspecific comparisons of bacterial diversity and community structure support coevolution of gut microbiota and termite host. Appl Environ Microbiol. 2005;71:6590–9.
- Elshahed MS, Senko JM, Najar FZ, Kenton SM, Roe BA, Dewers TA, et al. Bacterial diversity and sulfur cycling in a mesophilic sulfide-rich spring. Appl Environ Microbiol. 2003;69:5609–21.
- Galand PE, Potvin M, Casamayor EO, Lovejoy C. Hydrography shapes bacterial biogeography of the deep Arctic Ocean. ISME J. 2010;4:564–76.
- 11. Ferrer M, Guazzaroni ME, Richter M, García-Salamanca A, Yarza P, Suárez-Suárez A, et al. Taxonomic and functional metagenomic profiling of the microbial community in the anoxic sediment of a sub-saline shallow lake (Laguna de Carrizo, Central Spain). Microb Ecol. 2011;62:824–37.
- Harris JK, Caporaso JG, Walker JJ, Spear JR, Gold NJ, Robertson CE, et al. Phylogenetic stratigraphy in the Guerrero Negro hypersaline microbial mat. ISME J. 2013;7:50–60.
- Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Commun. 2016;7:e13219.
- Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 2018. https://doi.org/10.1038/nbt.4229
- Probst AJ, Ladd B, Jarett JK, Geller-Mcgrath DE, Sieber CMK, Emerson JB, et al. Differential depth distribution of microbial function and putative symbionts through sediment-hosted aquifers in the deep terrestrial subsurface. Nat Microbiol. 2018;3:328–36.
- Baker BJ, Lazar CS, Teske AP, Dick GJ. Genomic resolution of linkages in carbon, nitrogen, and sulfur cycling among widespread estuary sediment bacteria. Microbiome. 2015;3:e14.
- Karsenti E, Acinas SG, Bork P, Bowler C, de Vargas C, Raes J, et al. A holistic approach to marine eco-systems biology. PLoS Biol. 2011;9:7–11.
- Matheus-Carnevali PB, Schulz F, Castelle CJ, Kantor R, Shih P, Sharon I, et al. Hydrogen-based metabolism – an ancestral trait in lineages sibling to the *Cyanobacteria*. bioRxiv. 2018. https://doi. org/10.1101/328856

- Thongaram T, Kosono S, Ohkuma M, Hongoh Y, Kitada M, Yoshinaka T, et al. Gut of higher termites as a niche for alkaliphiles as shown by culture-based and culture-independent studies. Microbes Environ. 2003;18:152–9.
- Utami YD, Kuwahara H, Murakami T, Morikawa T, Sugaya K, Kihara K, 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.
- Izawa K, Kuwahara H, Sugaya K, Lo N, Ohkuma M, Hongoh Y. Discovery of ectosymbiotic *Endomicrobium* lineages associated with protists in the gut of stolotermitid termites. Environ Microbiol Rep. 2017;9:411–8.
- 22. Hongoh Y, Deevong P, Hattori S, Inoue T, Noda S, Noparatnaraporn N, et al. Phylogenetic diversity, localization, and cell morphologies of members of the candidate phylum TG3 and a subphylum in the phylum *Fibrobacteres*, recently discovered bacterial groups dominant in termite guts. Appl Environ Microbiol. 2006;72:6780–8.
- Hongoh Y, Ekpornprasit L, Inoue T, Moriya S, Trakulnaleamsai S, Ohkuma M, et al. Intracolony variation of bacterial gut microbiota among castes and ages in the fungus-growing termite *Macrotermes gilvus*. Mol Ecol. 2006;15:505–16.
- Hongoh Y, Sato T, Dolan MF, Noda S, Ui S, Kudo T, et al. The motility symbiont of the termite gut flagellate *Caduceia versatilis* is a member of the '*Synergistes*' group. Appl Environ Microbiol. 2007;73:6270–6.
- 25. Murakami T, Segawa T, Dial R, Takeuchi N, Kohshima S, Hongoh Y. Bacterial microbiota associated with the glacier ice worm is dominated by both worm-specific and glacier-derived facultative lineages. Microbes Environ. 2017;32:32–9.
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3.
- Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 2017;11:2639–43.
- Pruesse E, Peplies J, Glöckner FO. SINA: accurate highthroughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics. 2012;28:1823–9.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:590–6.
- 30. Nakajima H, Hongoh Y, Usami R, Kudo T, Ohkuma M. Spatial distribution of bacterial phylotypes in the gut of the termite *Reticulitermes speratus* and the bacterial community colonizing the gut epithelium. FEMS Microbiol Ecol. 2005;54:247–55.
- Thongaram T, Hongoh Y, Kosono S, Ohkuma M, Trakulnaleamsai S, Noparatnaraporn N, et al. Comparison of bacterial communities in the alkaline gut segment among various species of higher termites. Extremophiles. 2005;9:229–38.
- Warnecke F, Luginbühl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, et al. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature. 2007;450:560–5.
- 33. Sato T, Kuwahara H, Fujita K, Noda S, Kihara K, Yamada A, et al. Intranuclear verrucomicrobial symbionts and evidence of lateral gene transfer to the host protist in the termite gut. ISME J. 2014;8:1008–19.
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar A, et al. ARB: a software environment for sequence data. Nucleic Acids Res. 2004;32:1363–71.
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.

465

- Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007;56:564–77.
- Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33:1870–4.
- Price MN, Dehal PS, Arkin AP. FastTree 2 approximately maximum-likelihood trees for large alignments. PLoS ONE. 2010;5:e9490.
- Noda S, Ohkuma M, Yamada A, Hongoh Y, Kudo T. Phylogenetic position and in situ identification of ectosymbiotic spirochetes on protists in the termite gut. Appl Environ Microbiol. 2003;69:625–33.
- Trager W. The cultivation of a cellulose-digesting flagellate, *Trichomonas termopsidis*, and of certain other termite protozoa. Biol Bull. 1934;66:182–90.
- 41. Yuki M, Kuwahara H, Shintani M, Izawa K, Sato T, Starns D, et al. Dominant ectosymbiotic bacteria of cellulolytic protists in the termite gut also have the potential to digest lignocellulose. Environ Microbiol. 2015;17:4942–53.
- 42. Kuwahara H, Yuki M, Izawa K, Ohkuma M, Hongoh Y. Genome of 'Ca. Desulfovibrio trichonymphae', an H<sub>2</sub>-oxidizing bacterium in a tripartite symbiotic system within a protist cell in the termite gut. ISME J. 2017;11:766–76.
- 43. Hongoh Y, Sharma VK, Prakash T, Noda S, Taylor TD, Kudo 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.
- 44. Martin M. Cutadapt removes adapter sequences from highthroughput sequencing reads. EMBnet J. 2011;17:e10.
- 45. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics. 2011;27:863–4.
- 46. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77.
- Lin HH, Liao YC. Accurate binning of metagenomic contigs via automated clustering sequences using information of genomic signatures and marker genes. Sci Rep. 2016;6:12–9.
- Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–55.
- Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, et al. CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. Nucleic Acids Res. 2017; 45:D200–3.
- Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics. 2013;14:e60.
- Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: An automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 2007;35:182–5.
- Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 2007;35:52–7.
- Hongoh Y, Sato T, Noda S, Ui S, Kudo T, Ohkuma M. *Candidatus* Symbiothrix dinenymphae: bristle-like *Bacteroidales* ectosymbionts of termite gut protists. Environ Microbiol. 2007;9:2631–5.
- 55. Noda S, Kawai M, Nakajima H, Kudo T, Ohkuma M. Identification and in situ detection of two lineages of *Bacteroidales* ectosymbionts associated with a termite gut protist, *Oxymonas* sp. Microbes Environ. 2006;21:16–22.

- Desai MS, Strassert JFH, Meuser K, Hertel H, Ikeda-Ohtsubo W, Radek R, et al. Strict cospeciation of devescovinid flagellates and *Bacteroidales* ectosymbionts in the gut of dry-wood termites (Kalotermitidae). Environ Microbiol. 2010;12:2120–32.
- Kataeva I, Li XL, Chen H, Choi SK, Ljungdahl LG. Cloning and sequence analysis of a new cellulase gene encoding CelK, a major cellulosome component of *Clostridium thermocellum*: evidence for gene duplication and recombination. J Bacteriol. 1999;181:5288–95.
- Sekar R, Shin HD, Chen R. Engineering *Escherichia coli* cells for cellobiose assimilation through a phosphorolytic mechanism. Appl Environ Microbiol. 2012;78:1611–4.
- Springer TA. Folding of the N-terminal, ligand-binding region of integrin alpha-subunits into a beta-propeller domain. Proc Natl Acad Sci USA. 1997;94:65–72.
- Severi E, Hood DW, Thomas GH. Sialic acid utilization by bacterial pathogens. Microbiology. 2007;153:2817–22.
- Berne C, Ellison CK, Ducret A, Brun YV. Bacterial adhesion at the single-cell level. Nat Rev Microbiol. 2018;16:616–27.
- 62. Mikaelyan A, Köhler T, Lampert N, Rohland J, Boga H, Meuser K, et al. Classifying the bacterial gut microbiota of termites and cockroaches: a curated phylogenetic reference database (DictDb). Syst Appl Microbiol. 2015;38:472–82.
- Leadbetter JR, Schmidt TM, Graber JR, Breznak JA. Acetogenesis from H<sub>2</sub> plus CO<sub>2</sub> by spirochetes from termite guts. Science. 1999;283:686–9.
- Graber JR, Breznak JA. Physiology and nutrition of *Treponema* primitia, an H<sub>2</sub>/CO<sub>2</sub>-acetogenic spirochete from termite hindguts. Appl Environ Microbiol. 2004;70:1307–14.
- 65. Ohkuma M, Noda S, Hattori S, Iida T, Yuki M, Starns D, et al. Acetogenesis from H<sub>2</sub> plus CO<sub>2</sub> and nitrogen fixation by an endosymbiotic spirochete of a termite-gut cellulolytic protist. Proc Natl Acad Sci USA. 2015;112:10224–30.
- Zheng H, Dietrich C, Hongoh Y, Brune A. Restrictionmodification systems as mobile genetic elements in the evolution of an intracellular symbiont. Mol Biol Evol. 2016;33:721–5.
- 67. Izawa K, Kuwahara H, Kihara K, Yuki M, Lo N, Itoh T, 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.
- Lumppio HL, Shenvi NT, Summers AO, Voordouw G, Kurtz DM Jr. Rubrerythrin and rubredoxin oxidoreductase in *Desulfovibrio vulgaris*: a novel oxidative stress protection system. J Bacteriol. 2001;183:101–8.
- 69. Tokuda G, Watanabe H. Hidden cellulases in termites: revision of an old hypothesis. Biol Lett. 2007;3:336–9.
- He S, Ivanova N, Kirton E, Allgaier M, Bergin C, Scheffrahn RH, et al. Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. PLoS ONE. 2013;8:e61126.
- Poulsen M, Hu H, Li C, Chen Z, Xu L, Otani S, et al. Complementary symbiont contributions to plant decomposition in a fungus-farming termite. Proc Natl Acad Sci USA. 2014;111:14500–5.
- Mikaelyan A, Strassert JFH, Tokuda G, Brune A. The fibreassociated cellulolytic bacterial community in the hindgut of wood-feeding higher termites (*Nasutitermes* spp.). Environ Microbiol. 2014;16:2711–22.
- Abdul Rahman N, Parks DH, Vanwonterghem I, Morrison M, Tyson GW, Hugenholtz P. A phylogenomic analysis of the bacterial phylum *Fibrobacteres*. Front Microbiol. 2016;6:e1469.
- 74. Tokuda G, Tsuboi Y, Kihara K, Saitou S, Moriya S, Lo N, et al. Metabolomic profiling of <sup>13</sup>C-labelled cellulose digestion in a lower termite: insights into gut symbiont function. Proc Bio Sci. 2014;281:e20140990.

- Ottesen EA, Jong WH, Quake SR, Leadbetter JR. Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. Science. 2006;314:1464–7.
- Pester M, Brune A. Expression profiles of *fhs* (FTHFS) genes support the hypothesis that spirochaetes dominate reductive acetogenesis in the hindgut of lower termites. Environ Microbiol. 2006;8:1261–70.
- 77. Zhang X, Matson EG, Leadbetter JR. Genes for selenium dependent and independent formate dehydrogenase in the gut microbial communities of three lower, wood-feeding termites and a wood-feeding roach. Environ Microbiol. 2011;13:307–23.
- Matson EG, Gora KG, Leadbetter JR. Anaerobic carbon monoxide dehydrogenase diversity in the homoacetogenic hindgut microbial communities of lower termites and the wood roach. PLoS ONE. 2011;6:e19316.
- Matson EG, Rosenthal AZ, Zhang X, Leadbetter JR. Genomewide effects of selenium and translational uncoupling on transcription in the termite gut symbiont *Treponema primitia*. mBio. 2013;4:1–9.
- Rosenthal AZ, Matson EG, Eldar A, Leadbetter JR. RNA-seq reveals cooperative metabolic interactions between two termitegut spirochete species in co-culture. ISME J. 2011;5:1133–42.
- 81. Noda S, Shimizu D, Yuki M, Kitade O, Ohkuma M. Hostsymbiont cospeciation of termite-gut cellulolytic protists of the

genera *Teranympha* and *Eucomonympha* and their *Treponema* endosymbionts. Microbes Environ. 2018;33:26–33.

- Yamin MA. Cellulose metabolism by the flagellate *Trichonympha* from a termite is independent of endosymbiotic bacteria. Science . 1981;211:58–9.
- Inoue J, Saita K, Kudo T, Ui S, Ohkuma M. Hydrogen production by termite gut protists: characterization of iron hydrogenases of parabasalian symbionts of the termite *Coptotermes formosanus*. Eukaryot Cell. 2007;6:1925–32.
- Ohkuma M. Symbioses of flagellates and prokaryotes in the gut of lower termites. Trends Microbiol. 2008;16:345–52.
- Pester M, Brune A. Hydrogen is the central free intermediate during lignocellulose degradation by termite gut symbionts. ISME J. 2007;1:551–65.
- Desai MS, Brune A. *Bacteroidales* ectosymbionts of gut flagellates shape the nitrogen-fixing community in dry-wood termites. ISME J. 2012;6:1302–13.
- Graber JR, Breznak JA. Folate cross-feeding supports symbiotic homoacetogenic spirochetes. Appl Environ Microbiol. 2005;71:1883–9.
- Chuvochina M, Rinke C, Parks DH, Rappé MS, Tyson GW, Yilmaz P, et al. The importance of designating type material for uncultured taxa. Syst Appl Microbiol. 2018. https://doi.org/10. 1016/j.syapm.2018.07.003