BMC Microbiology

RESEARCH

Open Access

Comparison of microbial diversity and carbohydrate-active enzymes in the hindgut of two wood-feeding termites, *Globitermes sulphureus* (Blattaria: Termitidae) and *Coptotermes formosanus* (Blattaria: Rhinotermitidae)

Zhidong Zhang[1,2†], Kai Wang[1†], Chuanshan Zou[3], Ting Zhao[1], Wenbin Wu[1], Cai Wang[2*] and Yan Hua[1*]

Abstract

Background Wood-feeding termites have been employed as sources of novel and highly efficient lignocellulolytic enzymes due to their ability to degrade lignocellulose efficiently. As a higher wood-feeding termite, *Globitermes sulphureus* (Blattaria: Termitidae) plays a crucial role as a decomposer in regions such as Vietnam, Singapore, Myanmar, and Yunnan, China. However, the diversity of its gut microbiome and carbohydrate-active enzymes (CAZymes) remains unexplored. Here, we analyzed the diversity of hindgut microbial communities and CAZymes in a higher wood-feeding termite, *G. sulphureus*, and a lower wood-feeding termite, *Coptotermes formosanus* (Blattaria: Rhinotermitidae).

Results 16S rRNA sequencing revealed that Spirochaetota, Firmicutes, and Fibrobacterota were the dominant microbiota in the hindgut of the two termite species. At the phylum level, the relative abundances of Proteobacteria and Bacteroidota were significantly greater in the hindgut of *C. formosanus* than in *G. sulphureus*. At the genus level, the relative abundances of *Candidatus_Azobacteroides* and *Escherichia-Shigella* were significantly lower in the hindgut of *G. sulphureus* than in *C. formosanus*. Metagenomic analysis revealed that glycoside hydrolases (GHs) with cellulases and hemicellulases functions were not significantly different between *G. sulphureus* and *C. formosanus*. Interestingly, the cellulases in *G. sulphureus* were mainly GH5_2, GH5_4, GH6, GH9, and GH45, while the hemicellulases were mainly GH11, GH8, GH10, GH11, GH26, and GH53. In *C. formosanus*, the cellulases were mainly GH6 and GH9, and the hemicellulases were mainly GH5_7, GH5_21, GH10, GH12, and GH53. In addition, β-glucosidase, exo-β-1,4-glucanase,

†Zhidong Zhang and Kai Wang contributed equally to this work.

*Correspondence:
Cai Wang
wangcai@scau.edu.cn
Yan Hua
wildlife530@hotmail.com

Full list of author information is available at the end of the article

and endo-β-1,4-glucanase activities did not differ significantly between the two termite species, while xylanase activity was higher in *G. sulphureus* than in *C. formosanus*. The bacteria encoding GHs in *G. sulphureus* were mainly Firmicutes, Fibrobacterota, and Proteobacteria, whereas Bacteroidota and Spirochaetota were the main bacteria encoding GHs in *C. formosanus*.

Conclusions Our findings characterized the microbial composition and differences in the hindgut microbiota of *G. sulphureus* and *C. formosanus*. Compared to *C. formosanus*, *G. sulphureus* is enriched in genes encoding for hemicellulase and debranching enzymes. It also highlights the rich diversity of GHs in the hindgut microbiota of *G. sulphureus*, including the GH5 subfamily, GH6, and GH48, with the GH6 and GH48 not previously reported in other higher termites. These results strengthen the understanding of the diversity of termite gut microbiota and CAZymes.

Keywords Lignocellulose, Termite, Hindgut microbiota, Metagenomic sequencing, 16S rRNA sequencing, Carbohydrate-active enzymes (CAZymes)

Background

In nature, Lignocellulose is abundant and prevalent across forestry, agriculture, papermaking, and the food industry [1]. Lignocellulose has been proven to be a renewable resource that can effectively replace fossil fuels [2]. However, currently, there is a lack of highly efficient enzymes for the degradation of lignocellulose. Natural lignocellulosic materials are primarily composed of cellulose, hemicellulose, and lignin. The intricate and stable structure of lignocellulose constrains its efficient utilization. Under natural conditions, lignocellulose degradation involves a complex process mediated by various carbohydrate-activated enzymes (CAZymes), which can be categorized into five classes: glycoside hydrolases (GHs), glycosyltransferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), and auxiliary activities (AAs). Therefore, exploring novel and efficient lignocellulose-degrading enzymes is highly important for enhancing the utilization of lignocellulose.

Termites have evolved a specialized digestive system capable of degrading lignocellulose [3]. Despite their small size, termite guts possess the most efficient lignocellulose degradation system, capable of degrading 74-99% of cellulose and 65-78% of hemicellulose [3]. Moreover, they can produce methane and short-chain fatty acids during the degradation of lignocellulose [4]. The "termite-gut microbiome" serves as a crucial model for converting lignocellulose into renewable energy sources [5, 6]. Termites have different diets and could be classified as wood-feeding, fungus-feeding, grass-feeding, and soil-feeding termites [7]. Among them, wood-feeding termites have the most efficient lignocellulose utilization system, capable of digesting and utilizing lignocellulose within 24 h [8]. Lignocellulose degradation by termites mainly depends on the symbiotic gut microbiota, which includes bacteria, fungi, and protozoa (such as flagellates) [9]. Protozoa are unique to the gut of lower termites [10], while the gut of higher termites lacks such protozoa and is primarily composed of bacteria and fungi. Previous research has suggested that protozoa are the primary source of lignocellulolytic enzymes in lower termites, while many recent studies have shown that symbiotic bacteria also play a significant role in cellulose degradation [11, 12]. These symbiotic bacteria are also the main source of lignocellulolytic enzymes in higher termites [13]. Increasing evidence indicates that higher termites have richer symbiotic microorganisms for lignocellulose degradation than lower termites [14]. The rich bacterial diversity of higher termites may represent a potential reservoir for novel CAZymes.

Over the past few decades, many researches have been conducted on the function of termite gut microbiota in degrading lignocellulose using methods such as 16 S rRNA, metagenomic, and transcriptomic sequencing [15–19]. Previous research indicated that diet is a major driver of the termite gut microbiome [13]. He et al. (2019) found that both diet and phylogeny exert certain influences on the gut microbiome [20]. Moreover, recent studies have shown that the CAZymes in different species of termites are conserved. Grieco et al. (2019) analyzed the gut microbiota of seven termite species in Brazil and discovered that they share a conserved profile of CAZymes for the degradation of cellulose and chitin [21]. Marynowska et al. (2023) analyzed the functions associated with carbohydrate hydrolysis in the higher termites *Labiotermes labralis*, revealing a rich inventory of cellulases and hemicellulases. Compared with other dietary termites, the overall diversity of CAZymes was similar [22].

Globitermes sulphureus (Blattaria: Termitidae) is a higher wood-feeding termite species distributed in regions such as Vietnam, Singapore, Myanmar, and Yunnan Province in China [23–25]. It is an important economic pest that mainly affects coconut or oil palm plantations and wooden buildings [26–28]. Additionally, it plays a critical role as a decomposer in the Xishuangbanna Tropical Botanical Garden in Menglun, Yunnan, China [25]. However, its lignocellulose degradation ability remains largely unexplored. In this study, we used 16S rRNA sequencing and metagenomic sequencing

technologies to analyze the diversity of hindgut microbial communities and the bacteria encoding CAZymes of *G. sulphureus*. The results were compared with those for *C. formosanus*, a lower termite species that has received more attention for its hindgut bacterial communities and lignocellulose degradation abilities. Additionally, we measured β-glucosidase, exo-β-1,4-glucanase, endo-β-1,4-glucanase, and xylanase activities in the hindgut microbiota of both termite species. This study provides new information on the hindgut microbial communities of lower and higher wood-feeding termites and provides a research basis for identifying novel lignocellulose-degrading enzymes in the termite hindgut microbiota.

Materials and methods

Sample collection and preparation

The nests of *G. sulphureus* were excavated from Kunming, Yunnan, China, and placed in the Wildlife Rescue and Inspection Center in Guangdong, China. *C. formosanus* was collected from the Wildlife Rescue and Inspection Center in Guangdong Province, China using underground bait stations. The underground bait station is a rectangular plastic container (28.0×21.0×18.5 cm) with an array of circular holes (diameter: 2.5 mm) on the bottom and side walls, and the interior contains 10 pine wood blocks (24.0×8.0×2.5 cm). Briefly, 3–5 underground bait stations were buried in areas with *C. formosanus* activity. After collecting a large number of termites, the subterranean bait station was transported back to the laboratory and placed in a plastic container (49.5×35.0×27.5 cm). After being collected, termites were subsequently dissected immediately. Two termite species were identified for morphological analysis [29, 30]. Healthy and uniformly sized *G. sulphureus* and *C. formosanus* workers were selected for the following tests.

Dissect the termite hindgut with reference to Warnecke et al. [15]. In brief, the termites were surface disinfected with 70% alcohol for 1 min and then rinsed three times with phosphate-buffered saline. Termite heads were secured using forceps, while their posterior ends were gently compressed with a scalpel. The gut of the termites was then slowly removed. Based on the morphological characteristics of termite gut, the dilated portion of the hindgut, known as the 3rd proctodeal segment (P3), contains a large number of microbes in wood-feeding higher termites and lower termites, using a scalpel to amputate the P3 segment of the termite. The P3 segment was punctured with an insect needle and the contents were collected in a 1.5 mL centrifuge tube and subsequently placed on dry ice. All the samples were stored at -80 °C. To meet the demand for sample volume for the experiment, the hindgut contents of 800 termites were mixed into one sample, and five samples were collected from two termite species. Approximately 500 mg of hindgut content can be collected in one sample, which were used for 16 S rRNA and metagenomic sequencing. We employed the methodology described by Liu et al. (2011) for extracting microbial cells from the hindgut of termites [31]. Briefly, tissue debris from the hindgut contents of termites was digested using 0.25% trypsin and 0.02% EDTA. The sample was vortexed for 30 s to homogenize it, followed by centrifugation at 800 g for 10 min at 4 °C. The supernatant was collected and transferred to PBS buffer, and the centrifugation was repeated three times. Subsequently, the sample was centrifuged at 9000 g for 10 min at 4 °C. The supernatant was carefully removed, and the microbial cells were obtained for DNA extraction.

Assay of protein concentrations

Determination of protein concentration using the BCA protein assay kit (Suzhou Kemin Biotechnology Co., Ltd, Suzhou, China). Briefly, the sample (mg) to extraction solution (ml) ratio was set at 100:1 for ice bath homog-40enization, Subsequently, centrifugation was performed at 15,000 g and 4℃ for 10 min, and collecting the supernatant. The working solution was prepared by mixing reagent A (Na_2CO_3 solution, sodium tartrate solution, NaOH solution, $NaHCO_3$ solution) and reagent B ($CuSO_4$•$5H_2O$ solution) in a 50:1 ratio, and preheated in a 60℃ water bath for 30 min. Distilled water (4 μL) was added to the blank tube, standard (Bovine serum albumin, BSA) (4 μL) to the standard tube, and the supernatant (4 μL) to the assay tube. Subsequently, 200 μL of the working solution was added to the blank tube, standard tube, and assay tube respectively. After mixing, the solution was incubated in a 60℃ water bath for 30 min, and the absorbance values were measured at 562 nm using a microplate reader.

Assay of β-glucosidase activity

β-glucosidase activity was measured using a β-glucosidase activity assay kit (Suzhou Kemin Biotechnology Co., Ltd, Suzhou, China). β-glucosidase decomposition of p-Nitrophenyl-β-D-Glucopyranoside to produce p-Nitrophenol, which exhibits a maximum absorption peak at 400 nm, and β-glucosidase activity is calculated by measuring the rate of absorbance increase. Briefly, the sample (g) to citric acid-phosphate buffer (mL) ratio was set at 1:5–10 for ice bath homogenization. Subsequently, centrifugation was performed at 15,000 g and 4℃ for 10 min, and the supernatant. 120 μL of p-nitrophenyl-β-D-glucopyranoside solution was added to the assay tube, while 120 μL of distilled water was added to the control tube. Subsequently, 150 μL of citric acid-phosphate buffer and 30 μL of sample supernatant were added to both the assay and control tubes. The tubes were incubated in a 37℃ water bath for 30 min, followed by a 5 min

incubation in a 95℃ water bath. After the chill down, centrifugation was performed at 8,000 g and 4℃ for 5 min, and 70 μL of the supernatant was transferred to the assay and control tubes, respectively. Then, 130 μL of Na_2CO_3 solution was added to the assay and control tubes. The solutions were mixed and allowed to stand for 2 min. Finally, the absorbance values were measured by a microplate reader at 400 nm. At 37℃ and pH=5.0, one unit of β-glucosidase activity is defined as the amount that produces 1 nmol of p-nitrophenol per minute per milligram of protein.

Assay of exo-β-1,4-glucanase activity

Exo-β-1,4-glucanase activity was measured using an exo-β-1,4-glucanase activity assay kit (Suzhou Kemin Biotechnology Co., Ltd, Suzhou, China). Determination of reducing sugars produced by exo-β-1,4-glucanase catalyzed degradation of sodium carboxymethyl cellulose using the 3,5-dinitrosalicylic acid (DNS) method. Briefly, the sample (g) to acetic acid-anhydrous sodium acetate solution (mL) ratio is 1:5–10 for ice bath homogenization. Subsequently, centrifugation was performed at 15,000 g and 4℃ for 10 min, and the supernatant. 10 μL of the sample supernatant was added to both the assay tube and the control tube. 100 μL of sodium carboxymethyl cellulose solution was added to the assay tube, while 100 μL of distilled water was added to the control tube. Incubated in a 37℃ water bath for 2 h. 200 μL of DNS reagent (3,5-dinitro salicylic acid solution, sodium hydroxide solution, potassium sodium tartrate solution, redistilled phenol solution, and anhydrous sodium sulfite solution) was added to both the assay tube and the control tube. Mixed and incubated in a 90℃ water bath for 10 min. Finally, the absorbance values were measured by a microplate reader at 540 nm. At 37℃ and pH=6.5, one unit of exo-β-1,4-glucanase activity is defined as the amount that produces 1 μg of glucose per minute per milligram of protein.

Assay of endo-β-1,4-glucanase activity

Endo-β-1,4-glucanase activity was measured using an endo-β-1,4-glucanase activity assay kit (Suzhou Kemin Biotechnology Co., Ltd, Suzhou, China). Determination of reducing sugars produced by endo-β-1,4-glucanase catalyzed degradation of microcrystalline cellulose using the 3,5-dinitrosalicylic acid (DNS) method. Briefly, the sample (g) to acetic acid-anhydrous sodium acetate solution (mL) ratio is 1:5–10 for ice bath homogenization. Subsequently, centrifugation was performed at 15,000 g and 4℃ for 10 min, and the supernatant. 10 μL of the sample supernatant was added to both the assay tube and the control tube. 100 μL of microcrystalline cellulose solution was added to the assay tube, while 100 μL of distilled water was added to the control tube. Incubated in a 37℃ water bath for 2 h. 200 μL of DNS reagent (3,5-dinitro salicylic acid solution, sodium hydroxide solution, potassium sodium tartrate solution, redistilled phenol solution, and anhydrous sodium sulfite solution) was added to both the assay tube and the control tube. Mixed and incubated in a 90 °C water bath for 10 min. Finally, the absorbance values were measured by a microplate reader at 540 nm. At 37℃ and pH=6.5, one unit of endo-β-1,4-glucanase activity is defined as the amount that produces 1 μg of glucose per minute per milligram of protein.

Assay of xylanase activity

Xylanase activity was measured using a xylanase activity assay kit (Suzhou Kemin Biotechnology Co., Ltd, Suzhou, China). Xylanase can degrade xylan into reducing oligosaccharides and monosaccharides under acidic conditions, and further undergo a colorimetric reaction with 3,5-dinitrosalicylic acid in a boiling water bath, exhibiting a characteristic absorption peak at 540 nm. The depth of the reaction solution color is proportional to the amount of reducing sugars produced by enzymatic hydrolysis. The rate of increase in absorbance at 540 nm of the reaction solution is used to calculate the activity of xylanase. Briefly, the sample (g) to acetic acid - sodium acetate buffer (mL) ratio is 1:5–10 for ice bath homogenization. Subsequently, centrifugation was performed at 8,000 g and 4℃ for 10 min, and collecting the supernatant. 60 μL of sample supernatant and 90 μL of acetic acid - sodium acetate buffer were added to both the assay tube and the control tube, respectively. Subsequently, 60 μL of reagent 1 (xylan solution, NaOH solution, glacial acetic acid solution) was added to the assay tube. The mixture was then incubated at 50 °C for 30 min, followed by boiling for 10 min. 60 μL of reagent 1 was added to the control tube, and then 90 μL of reagent 2 (3,5-Dinitrosalicylic acid solution, NaOH solution, potassium sodium tartrate tetrahydrate solution, phenol solution, sodium bisulfite solution) was added to both the assay tube and the control tube. Mixed and incubated in a boiling water bath for 5 min. Finally, the absorbance values were measured by a microplate reader at 540 nm. At 50℃ and pH=4.8, one unit of xylanase activity is defined as the amount that produces of xylan to produce 1 nmol of reducing sugar per minute per milligram of protein.

DNA extraction and 16 S rRNA sequencing

The DNA of the total genome was extracted from the hindgut microbial cells of termites using the E.Z.N.A.® Stool DNA Kit (D4015-02, Omega, Inc., USA). The kit has been proven to be effective in preparing bacterial DNA from trace samples. DNA quality was detected by agarose gel electrophoresis, and DNA quality was quantified by an ultraviolet spectrophotometer. The primers used

for PCR amplification were 341 F (5'-CCTACGGGNGGCWGCAG-3') and 805R (5'-GACTACHVGGGTATCTAATCC-3'). PCR products were purified with AMPure XT beads (Beckman Coulter Genomics, Danvers, MA, USA) and quantified with a Qubit fluorometer (Invitrogen, USA). The purified PCR products were evaluated using an Agilent 2100 bioanalyzer (Agilent, USA) and a library quantification kit from Illumina (Kapa Biosciences, Woburn, MA, USA), and qualified samples were sequenced using a NovaSeq 6000 sequencer for 2×250 bp paired-end sequencing.

The double-end data obtained from 16 S rRNA sequencing were first subjected to data segmentation using Cutadapt (v1.9) to remove the raw street data and primer sequences [32]. Data splicing was completed using FLASH (v1.2.8) [33]. Sequences with lengths less than 100 bp and an N content greater than 5% were removed. Meanwhile, Vsearch (v2.3.4) was used to remove chimeric sequences to complete the filtering of the data [34]. DADA2 was used for denoising to obtain ASV (feature) feature sequences and abundance Table [35]. Species annotation was performed using the SILVA (Release 138) database with a confidence threshold of 0.7 [36]. Subsequently, the species abundance of each sample was determined according to the species abundance table. One sample was discarded from subsequent 16 S rRNA analysis due to insufficient sequencing depth.

Metagenomic sequencing

The DNA library was constructed using the TruSeq Nano DNA LT Library Preparation Kit (FC-121-4001). High-throughput sequencing was performed using the NovaSeq 6000 with the PE150 sequencing, and the sequencing kit used was the TruSeq Nano DNA LT Library Preparation Kit - Set A (FC-121-4001). The raw sequencing data were processed using Cutadapt (v1.9) [32] to remove adapters and fqtrim (v0.94) [37] to remove low-quality sequences. The data obtained from the previous step were assembled using IDBA-UD (v1.1.1) [38]. Subsequently, coding sequence prediction was performed on the assembled data using MetaGeneMark (v3.26) [39]. Clustering and deredundancy were performed using CD-HIT (v4.6.1) [40]. Using Bowtie2, the valid sequences were aligned to the Unigene sequences, and the number of reads mapped to each Unigene in each sample was calculated. Subsequently, the sequences were aligned to the CDS sequence library for TPM abundance calculation [41]. Subsequently, species annotation information was obtained by comparison in the NR_mate database.

Use the DIAMOND software to align the protein sequences of unigenes with the Kyoto Encyclopedia of Genes and Genomes (KEGG) and CAZy database and perform functional annotation [42, 43]. Subsequently, the abundance and number of unigene in each sample were counted, and statistical analysis of abundance and variance was performed at the functional and genetic levels. z.

The GHs were grouped according to the method of Allgaier et al. (2010) [44], including cellulases, hemicellulases, oligosaccharide enzymes, and debranching enzymes. To further determine the contribution of the hindgut microbiota to the degradation of different GHs (cellulases, hemicellulases, oligosaccharide enzymes, and debranching enzymes), the CAZyme annotation results of *G. sulphureus* and *C. formosanus* were compared with the non-redundant (NR) database to determine the taxonomic distribution of CAZymes at the phylum and genus levels.

Statistical analysis

Alpha diversity analysis was performed using QIIME2, employing various metrics including the Shannon, Simpson, Chao1, and Observed_species indices to assess the diversity within each sample [45]. To compare the differences in microbial community composition and structure between the two termite species, analysis of similarities (Anosim) was performed using the R vegan (v2.3) package. Principal component analysis (PCA) and nonmetric multidimensional scaling (NMDS) visualizations, based on the Bray-Curtis distance matrix, were performed using the R ggplot2 package. All statistical analyses were performed using IBM SPSS Statistics version 26. Data that conform to a normal distribution (such as xylanase activity) were analyzed using the Student's t test. Data that did not conform to a normal distribution (such as β-glucosidase, exo-β-1,4-glucanase, and endo-β-1,4-glucanase activities, alpha diversity, bacterial composition, and composition of GHs) were analyzed using the Mann-Whitney U test. All the statistical results are presented as mean±SEM. $P<0.05$ indicated a significant difference.

Results

Analysis of β-glucosidase, exo-β-1,4-glucanase, endo-β-1,4-glucanase, and xylanase activities

In the hindgut microbiota of *G. sulphureus*, the β-glucosidase activity was 4.34±0.126 nmol/min/mg prot, the exo-β-1,4-glucanase activity was 1.72±0.099 nmol/min/mg prot, the endo-β-1,4-glucanase activity was 4.10±0.133 nmol/min/mg prot, the xylanase activity was 6.98±0.353 nmol/min/mg prot. In the hindgut microbiota of *C. formosanus*, the β-glucosidase activity was 9.05±0.331 nmol/min/mg prot, the exo-β-1,4-glucanase activity was 4.85±0.147 nmol/min/mg prot, the endo-β-1,4-glucanase activity was 9.73±0.311 nmol/min/mg prot, the xylanase activity was 4.90±0.281 nmol/min/mg prot. β-glucosidase, exo-β-1,4-glucanase, and endo-β-1,4-glucanase activities did not differ significantly

between the two termite species (Mann-Whitney U test, Z = - 1.964, P=0.10). Xylanase activity was higher in *G. sulphureus* than *C. formosanus* (Student's t test, t=4.622, df=4, P<0.05) (Table 1).

16S rRNA sequencing data

Nine hindgut content samples from *G. sulphureus* and *C. formosanus* were used for 16 S rRNA sequencing, and 858,172 16 S rRNA gene sequences were generated. The average sequence lengths of *G. sulphureus* and *C. formosanus* were 41.94 M and 54.83 M, respectively. After filtering the low-quality sequences, 731,835 valid sequences were obtained for subsequent analysis (Table S1). Based on the feature abundance table, the number of shared and unique features was visualized by a Venn diagram. As shown in Figure S1, 9404 features were obtained from the two groups, 1619 of which features were shared.

The Shannon and Simpson indices indicated no significant differences in the diversity of the hindgut microbial communities between the two termite species (Shannon: Mann-Whitney U test, Z = - 1.960, P=0.06; Simpson: Mann-Whitney U test, Z = - 1.690, P=0.91). In contrast, the Chao1 index and Observed_species index revealed that *G. sulphureus* had a significantly higher number of hindgut bacterial communities than *C. formosanus* (Chao1: Mann-Whitney U test, Z = - 1.960, P=0.06; Observed_species: Mann-Whitney U test, Z = - 2.449, P<0.05) (Fig. 1A). PCA and NMDS based on the Bray-Curtis distance showed no significant differences in the structure of the hindgut microbial communities between the two termite species (Anosim, R=0.0187, P=0.364) (Fig. 1B, C).

Analysis of the hindgut microbiota composition of termites

In *G. sulphureus* and *C. formosanus*, we identified 30 and 35 bacterial phyla, respectively. In *G. sulphureus*, the dominant bacterial phyla were Spirochaetota (41.36%), Firmicutes (20.96%), Fibrobacterota (11.08%), Rs-K70_termite_group (3.76%), and Proteobacteria (2.31%). In *C. formosanus*, the dominant bacterial phyla were Spirochaetota (31.51%), Firmicutes (21.66%), Fibrobacterota (11.75%), Proteobacteria (7.81%), and Bacteroidota (6.10%). We found that the hindgut microbial communities of both *G. sulphureus* and *C. formosanus* were primarily composed of Spirochaetota (41.36% and 31.51%), followed by Firmicutes (20.96% and 21.66%) and Fibrobacterota (11.08% and 11.75%), which together accounted for 73.39% and 64.92% of the microbial community, respectively (Fig. 2A).

We also identified the top 20 most abundant bacterial genera in *G. sulphureus* and *C. sulphureus* (Fig. 2B). The five most abundant bacterial genera in their hindgut microbiota were *Termite_Treponema_cluster* (31.86% in *G. sulphureus* and 24.28% in *C. sulphureus*), *Clostridiales_unclassified* (11.33% in *G. sulphureus* and 8.83% in *C. sulphureus*), *Treponema* (8.54% in *G. sulphureus* and 6.06% in *C. sulphureus*), *Fibrobacterales_unclassified* (3.74% in *G. sulphureus* and 4.33% in *C. sulphureus*), and *Rs-K70_termite_group_unclassified* (3.76% in *G. sulphureus* and 3.41% in *C. sulphureus*). At the phylum level, the relative abundances of Proteobacteria and Bacteroidota were higher in the hindgut of *C. formosanus* than in that of *G. sulphureus* (Mann-Whitney U test, Z = - 2.449, P<0.05) (Fig. 2C, D). Compared to those of *G. sulphureus*, the relative abundances of *Candidatus_Azobacteroides* and *Escherichia-Shigella* were higher in the hindgut of *C. formosanus* (*Candidatus_Azobacteroides*: Mann-Whitney U test, Z = - 2.205, P<0.05; *Escherichia-Shigella*: Mann-Whitney U test, Z = - 2.183, P<0.05) (Fig. 2E, F). *Candidatus_Azobacteroides* was found only in the hindgut of *C. formosanus*.

Table 1 Analysis of β-glucosidase, exo-β-1,4-glucanase, endo-β-1,4-glucanase, and xylanase activities in the hindgut microbiota of *G. Sulphureus* and *C. formosanus*

Termite	*G. Sulphureus*	*C. formosanus*	Statistical results
β-glucosidase activity	4.34 ± 0.126	9.05 ± 0.331	Z = - 1.964, P=0.10
exo-β-1,4-glucanase activity	1.72 ± 0.099	4.85 ± 0.147	Z = - 1.964, P=0.10
endo-β-1,4-glucanase activity	4.10 ± 0.133	9.73 ± 0.311	Z = - 1.964, P=0.10
xylanase activity	6.98 ± 0.353	4.90 ± 0.281	t=4.622, df=4, P<0.05

Metagenomic data

In *G. sulphureus*, a total of 236,730,216 raw sequences were generated from the metagenomic sequencing. After quality filtering, a total of 194,913,396 reads were obtained. The assembly yielded a contig nucleotide length of 487,376 bp with a cumulative total length of 419,532,003 bp, with an average N50 value of 821 bp per sample. In *C. formosanus*, a total of 232,113,890 raw sequences were generated from the metagenomic sequencing. After quality filtering, a total of 185,891,986 reads were obtained. The assembly yielded a contig nucleotide length of 144,295 bp with a cumulative total length of 125,498,793 bp, with an average N50 value of 831 bp per sample. In both termite species, a total of 399,826 open reading frames (ORFs) were identified, with a combined length of 237.49 Mbp, and an average length of 593.99 bp per sample (Table S2).

Analysis of CAZymes in termites

In this study, 399,826 genes were annotated to the six primary functional classifications and 393 secondary functional classifications of CAZymes. In *G. sulphures*, GHs accounted for 38.08% of the primary functional

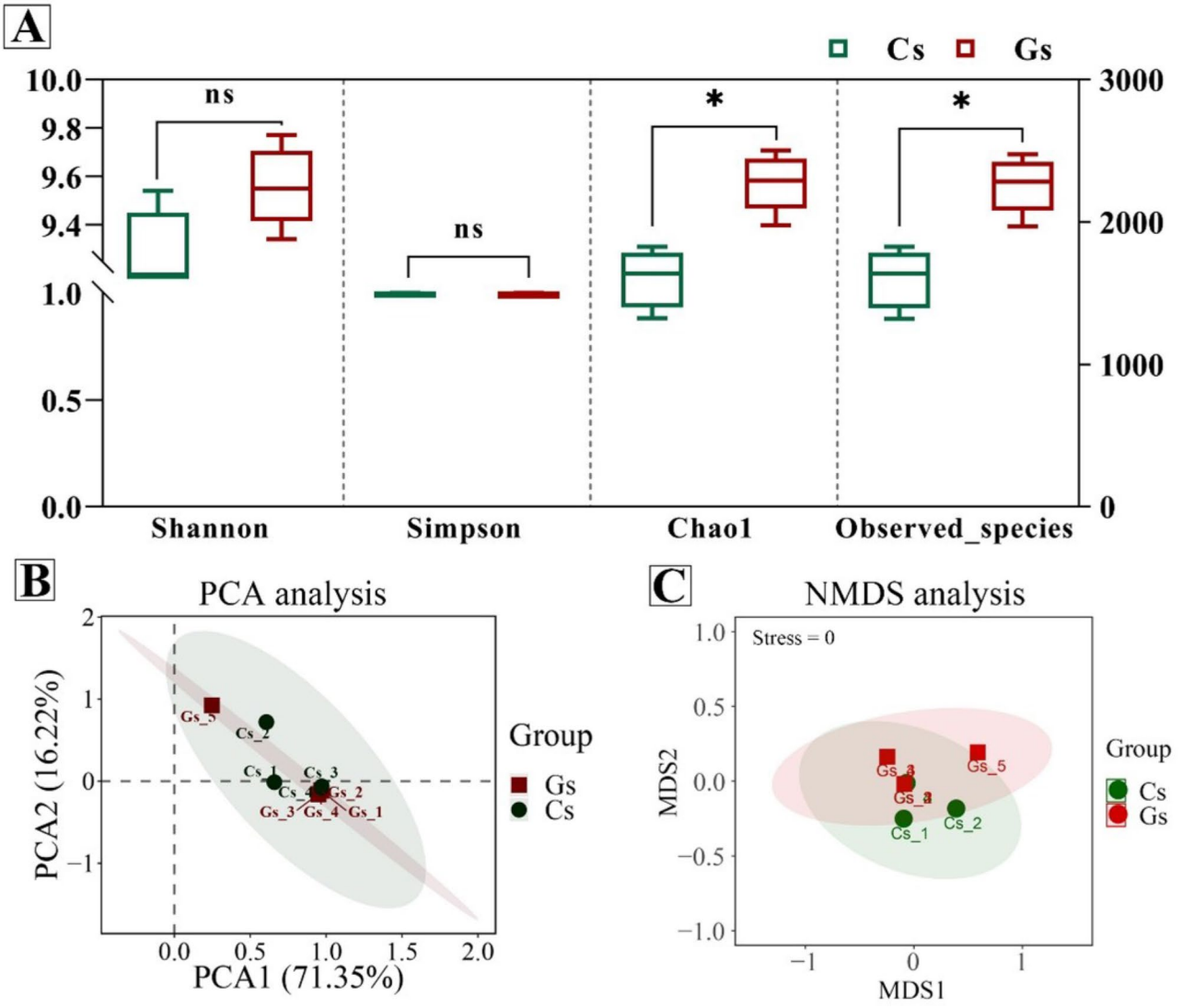

Fig. 1 Differences in microbial alpha diversity in the hindgut of *G. sulphureus* and *C. formosanus*. The alpha diversity estimated by Chao1, Shannon, Simpson, and Observed_species in *G. sulphureus* and *C. formosanus* (**A**). PCA (**B**) and NMDS (**C**) based on the Bray-Curtis distance matrix at ASVs level in the hindgut microbiota of *G. sulphureus* and *C. formosanus*. * $P < 0.05$. Gs, *G. sulphureus*. Cs, *C. formosanus*

classifications, GTs accounted for 36.69%, CBMs accounted for 20.09%, CEs accounted for 3.13%, PLs accounted for 1.07%, and AAs accounted for 0.94%. In *C. formosanus*, GHs accounted for 35.41% of the primary functional classifications, GTs accounted for 47.61%, CBMs accounted for 12.93%, CEs accounted for 2.96%, PLs accounted for 0.16%, and AAs accounted for 0.96% (Fig. 3A). At CAZymes level 2, PCA based on the Bray-Curtis distance revealed that the samples from *G. sulphureus* and *C. formosanus* clustered into distinct groups (Anosim, $R=1.0$, $P=0.006$) (Fig. 3B).

Since most lignocellulose degradation originates from GHs, further analysis of GHs was performed to gain insights into the carbohydrate degradation capabilities of *G. sulphureus* and *C. formosanus*. Genes were classified based on substrate targets of CAZymes, including cellulases, hemicellulases, oligosaccharide enzymes, and debranching enzymes. In both *G. sulphureus* and *C. formosanus*, there were no statistical differences in the relative abundances of the cellulases and debranching enzymes (cellulases: Mann-Whitney U test, Z = - 0.940, $P=0.35$; debranching enzymes: Mann-Whitney U test, Z = - 1.358, $P=0.18$). In *C. formosanus*, the relative abundances of oligosaccharide enzymes were higher than *G. sulphureus* (Mann-Whitney U test, Z = - 2.611, $P<0.05$), while the relative abundances of hemicellulases were lower than *G. sulphureus* (Mann-Whitney U test, Z = - 2.611, $P<0.05$). However, the relative abundances of oligosaccharide enzymes were higher in *C. formosanus* compared to *G. sulphureus* (Mann-Whitney U test, Z =

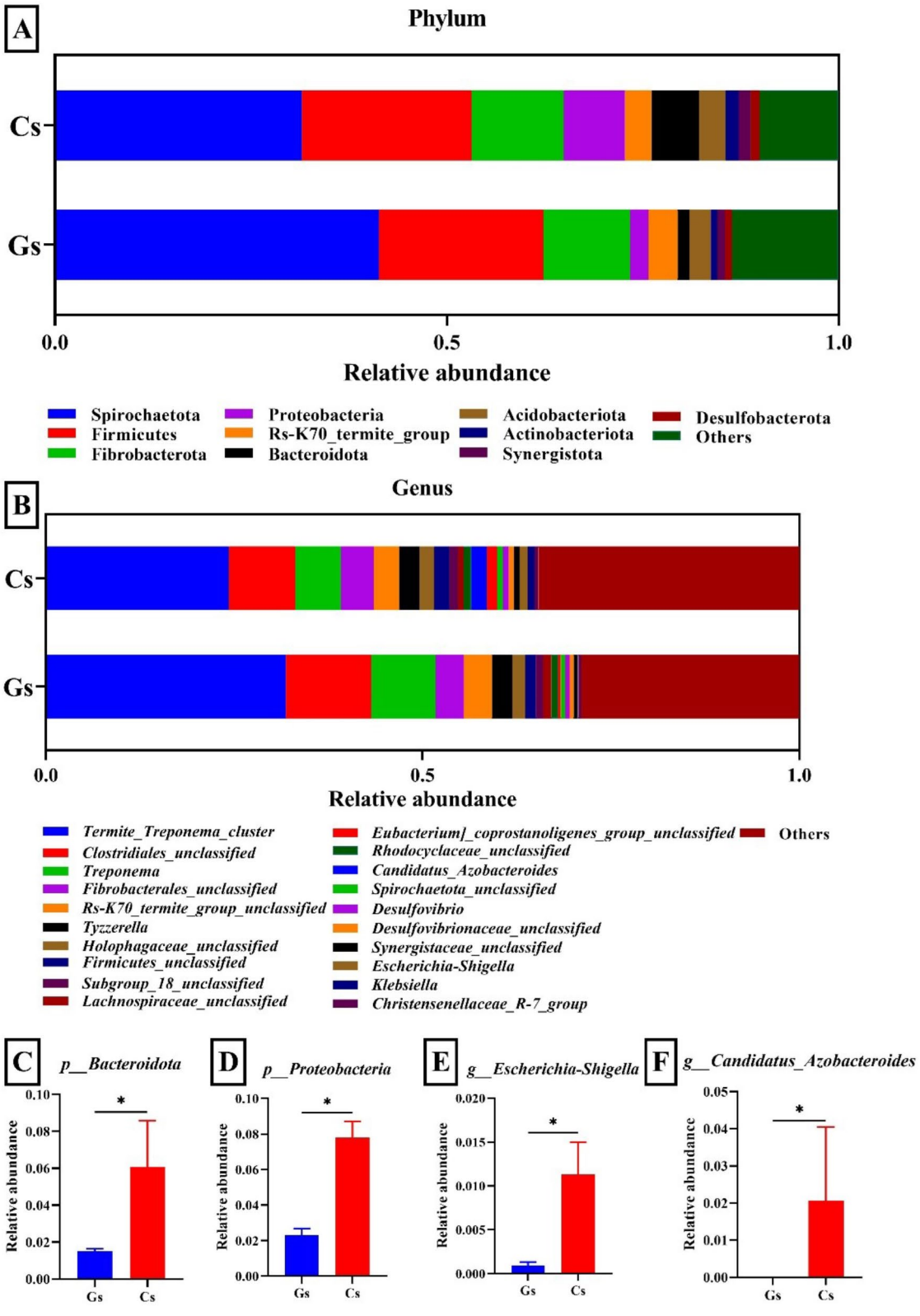

Fig. 2 Hindgut microbiota composition in *G. sulphureus* and *C. formosanus*. Relative abundances of the hindgut microbiota at (**A**) phylum and (**B**) genus levels of *G. sulphureus* and *C. formosanus*. The top 10 phyla and the top 20 genera were listed. Analysis of microbiota compositions at the phylum (**C**, **D**) and genus (**E**, **F**) levels. * $P < 0.05$. Gs, *G. sulphureus*. Cs, *C. formosanus*

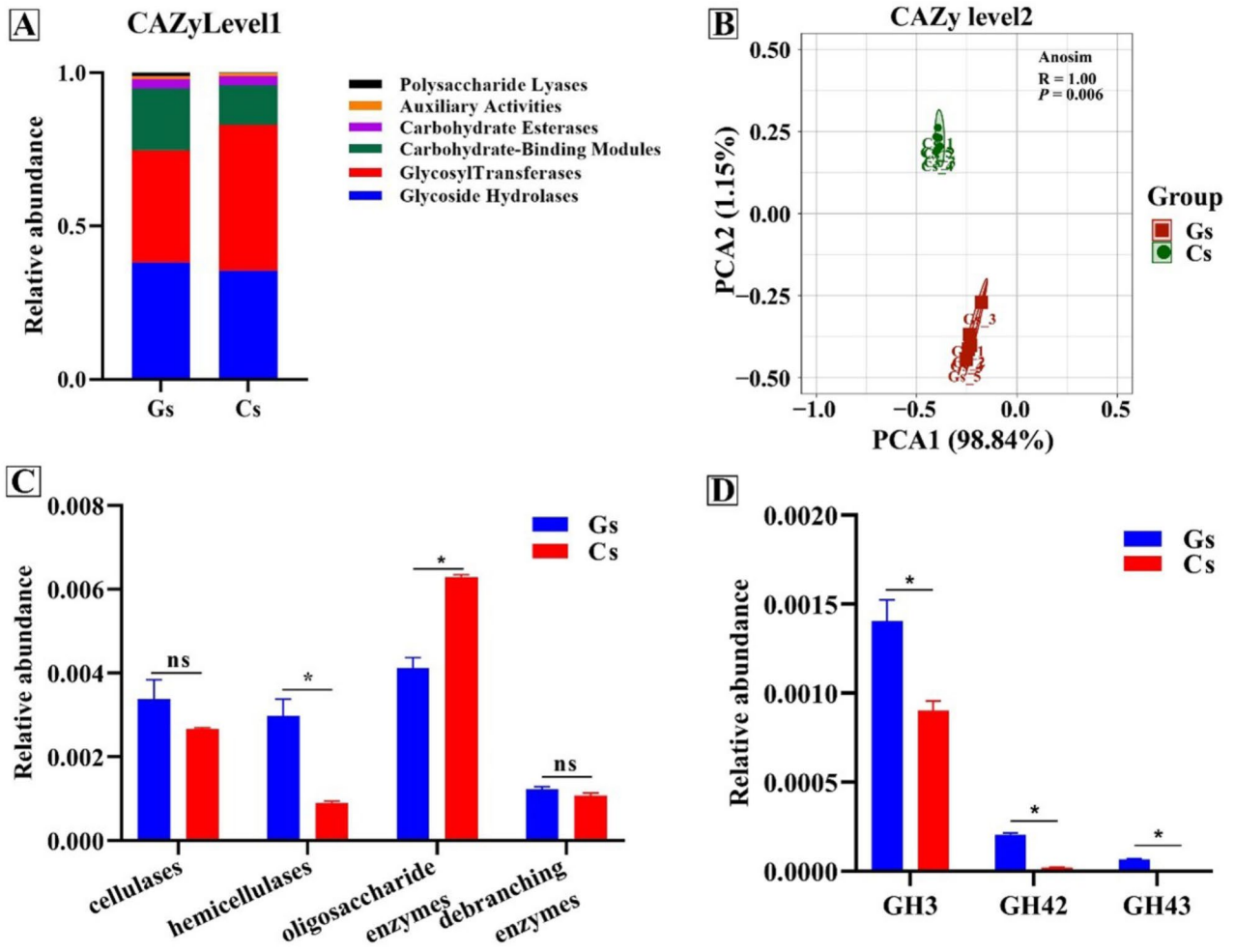

Fig. 3 Functional composition of *G. sulphureus* and *C. formosanus* at CAZymes level 1 (**A**). PCA analysis of *G. sulphureus* and *C. formosanus* at CAZymes level 2 (**B**). Differences in cellulases, hemicellulases, oligosaccharide enzymes, and debranching enzymes in the hindgut microbiota of *G. sulphureus* and *C. formosanus* (**C**). Differences in arabinoxylan degradation GHs in the hindgut microbiota of *G. sulphureus* and *C. formosanus* (**D**). * $P < 0.05$. Gs, *G. sulphureus*. Cs, *C. formosanus*

- 2.619, $P<0.05$) (Fig. 3C). Further analysis of oligosaccharides found that the relative abundance of GHs related to arabinoxylan degradation, including GH3, GH42, and GH43, which primarily catalyze the hydrolysis of α-L-arabinofuranosidic linkages, was significantly higher in *G. sulphureus* compared to *C. formosanus* (GH3: Mann-Whitney U test, Z = - 2.193, $P<0.05$; GH42: Mann-Whitney U test, Z = - 2.611, $P<0.05$; GH43: Mann-Whitney U test, Z = - 2.643, $P<0.05$) (Fig. 3D).

In *G. sulphureus*, the cellulases were mainly GH5_2, GH5_4, GH6, GH9, and GH45, while the hemicellulases were mainly GH11, GH8, GH10, GH11, GH26, and GH53. In *C. formosanus*, the cellulases were mainly GH6 and GH9, and the hemicellulases were mainly GH5_7, GH5_21, GH10, GH12 and GH53. GH5_52, GH5_1, GH5_25, GH5_55, GH5_9, GH_37, GH5_5, GH5_38, GH5_8, GH5_35, GH5_41, GH48, GH52, and GH54 were only found in the hindgut microbiota of *G. sulphureus*. In *G. sulphureus*, the diversity of CAZymes is higher than that in *C. formosanus*, particularly within the GH5 subfamily (Table 2).

Analysis of the functional contribution of microbiota to predicted CAZymes

In *G. sulphureus*, at the phylum level, the cellulases were mainly encoded by Firmicutes, Fibrobacterota, Proteobacteria, and Bacteroidota (Fig. 4A). The hemicellulases were mainly encoded by Firmicutes, Proteobacteria, Fibrobacterota, and Bacteroidota (Fig. 4C). The debranching enzymes were mainly encoded by Firmicutes, Bacteroidota, Proteobacteria, and Fibrobacterota (Fig. 4E). The oligosaccharide enzymes were mainly encoded by Firmicutes, Proteobacteria, Fibrobacterota, and Bacteroidota (Fig. 4G). At the genus level, the cellulases were mainly encoded by *Bacteria_noname*, *Fibrobacter*, *Treponema*, *Paenibacillus*, *Ruminococcus*, and

Table 2 Relative abundance (mean ± SEM) of carbohydrates in the gut metagenome of *G. sulphureus* and *C. formosanus*

CAZy Family	Main known activities (CAZy)	Relative abundance	
		G. Sulphureus	***C. formosanus***
Cellulases			
GH5_4	Glucomannan-specific endo-β-1,4-glucanase	605.47 ± 85.149	0.69 ± 0.208
GH5_2	Endo-β-1,4-glucanase	307.19 ± 51.223	0.81 ± 0.34
GH5_52	Endo-β-1,4-glucanase	70.29 ± 12.625	0
GH5_1	Endo-β-1,4-glucanase	53.05 ± 2.897	0
GH5_46	Endo-β-1,4-glucanase	22.66 ± 2.367	1.02 ± 0.527
GH5_25	Endo-β-1,4-glucanase	11.82 ± 1.309	0
GH5_55	Endo-β-1,4-glucanase	8.99 ± 0.967	0
GH5_9	β-glucosidase; exo-β-1,3-glucanase	6.38 ± 0.811	0
GH5_37	Endo-β-1,4-glucanase	0.9 ± 0.295	0
GH5_5	Endo-β-1,4-glucanase	0.85 ± 0.491	0
GH5_38	Endo-β-1,4-glucanase	0.71 ± 0.244	0
GH6	Endoglucanase, β-1,3-endoglucanase, β-1,4-endoglucanase	970.58 ± 103.943	2512.34 ± 19.797
GH9	Endoglucanase, cellobiohydrolase, β-glucosidase	864.28 ± 125.63	150.12 ± 12.902
GH44	Endoglucanase, xyloglucanase	23.06 ± 2.628	0.36 ± 0.223
GH45	Endoglucanase	406.02 ± 82.215	2.62 ± 0.576
GH48	Endo-processive celluloses	22.74 ± 3.166	0
Hemicellulases			
GH5_7	β-mannosidase	694.66 ± 83.52	594.42 ± 27.098
GH5_21	Endo-β-1,4-xylanase	0.58 ± 0.3	84.23 ± 5.993
GH5_36	Endo-β-1,4-mannanase	62.84 ± 1.165	6.96 ± 1.585
GH5_10	Xylan β-1,4-xylosidase	47.9 ± 7.457	0.74 ± 0.356
GH5_8	Endo-β-1,4-mannanase	8.34 ± 2.149	0
GH5_18	β-mannosidase	6.8 ± 2.113	0.03 ± 0.028
GH5_22	Xylan β-1,4-xylosidase	2.79 ± 0.337	0.5 ± 0.504
GH5_35	Endo-β-1,4-xylanase	2.26 ± 0.51	0
GH5_41	Endo-β-1,4-mannanase	1.73 ± 0.479	0
GH10	β-1,3-endoxylanase, β-1,4-endoxylanase	745.72 ± 87.480	72.40 ± 10.070
GH8	β-1,4-endo xylanase, β-1,4-endoglucanase, others	445.56 ± 83.924	1.76 ± 0.705
GH11	β-1,4-xylanase, β-1,3-xylanase	330.18 ± 52.903	1.98 ± 0.558
GH26	β-1,3-xylanase, mannanase	154.56 ± 15.334	40.38 ± 1.233
GH12	Endoglucanase, xyloglucan hydrolysis	41.20 ± 4.515	49.90 ± 2.799
GH52	Xylan β-1,4-xylosidase	38.44 ± 10.465	0
GH53	β-1,4-endogalactanase	392.36 ± 53.871	46.38 ± 6.074
Oligosaccharide enzymes			
GH3	β-glucosidase, xylan 1,4-β-xylosidase, α-L-arabinofuranosidase, others	1406.14 ± 117.934	901.90 ± 53.408
GH1	β-glucosidase, β-galactosidase, β-mannosidase, β-glucuronidase	431.12 ± 45.682	1256.46 ± 36.438
GH2	β-galactosidase, β-mannosidase, β-glucuronidase	1433.40 ± 78.199	3433.80 ± 45.433
GH29	α-L-fucosidase	90.00 ± 7.186	292.10 ± 19.288
GH35	β-Galactosidase	91.38 ± 24.848	240.76 ± 12.635
GH38	α-mannosidase	186.90 ± 11.424	133.12 ± 7.112
GH39	α-L-iduronidase, β-xylosidase	205.88 ± 8.899	9.72 ± 2.629
GH42	β-galactosidase, α-L-arabinopyranosidase	205.16 ± 10.005	20.62 ± 2.061
GH43	Xylanase, β-xylosidase, α-L arabinofuranosidase, arabinanase, others	66.64 ± 3.334	0.34 ± 0.236
Debranching enzymes			
GH51	α-L-arabinofuranosidase, β-xylosidase, endo-β-1,4-xylanase	290.72 ± 13.507	218.34 ± 14.496
GH54	α-L-arabinofuranosidase	0.34 ± 0.125	0
GH67	α-glucuronidase, xylan α-1,2-glucuronidase	106.40 ± 4.148	42.70 ± 4.079
GH78	α-L-rhamnosidase, α-L-rhamnohydrolase, rhamnogalacturonan	408.78 ± 16.162	514.32 ± 24.841
GH106	α-L-rhamnosidase	213.00 ± 11.109	217.98 ± 18.94
GH115	Xylan α-1,2-glucuronidase	207.00 ± 25.080	80.54 ± 4.265

Relative abundances are scaled by multiplication with 10^6

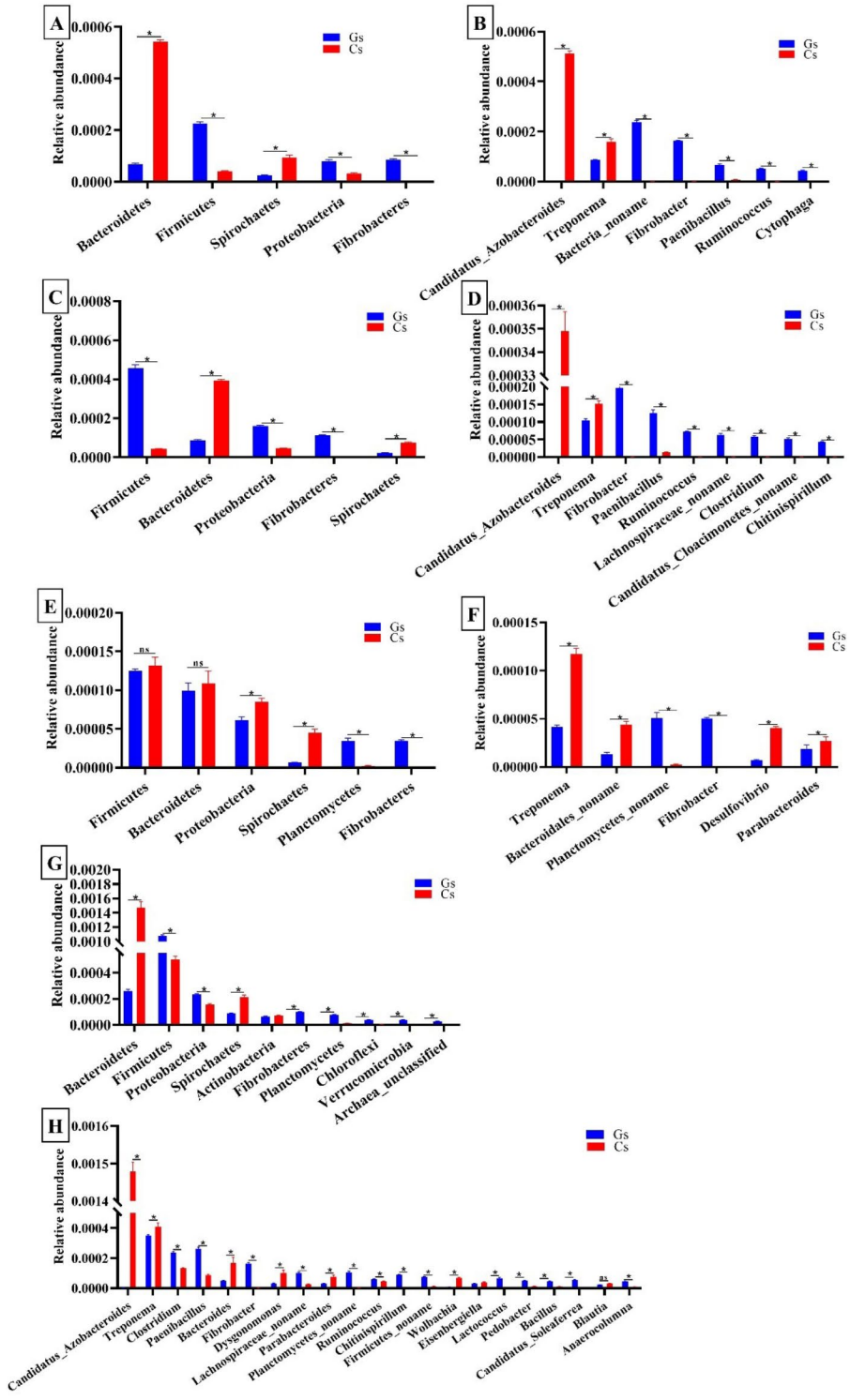

Fig. 4 Predicted phylum-level and genus-level taxonomic origin of cellulases (**A**, **B**), hemicellulases (**C**, **D**), oligosaccharide enzymes (**E**, **F**), and debranching enzymes (**G**, **H**). * $P < 0.05$. Gs, *G. sulphureus*. Cs, *C. formosanus*

Cytophaga (Fig. 4B). The hemicellulases were mainly encoded by *Fibrobacter*, *Paenibacillus*, *Treponema*, *Ruminococcus*, *Lachnospiraceae_noname*, and *Clostridium* (Fig. 4D). The debranching enzymes were mainly encoded by *Fibrobacter*, *Planctomycetes_momame*, *Treponema*, and *Parabacyeroides* (Fig. 4F). The debranching enzymes were mainly encoded by *Treponema*, *Paenibacillus*, *Clostridium*, *Fibrobacter*, *Planctomycetes_momame*, and *Lachnospiraceae_noname* (Fig. 4H).

In *C. formosanus*, at the phylum level, the cellulases and hemicellulases were mainly encoded by Bacteroidota, Spirochaetota, and Firmicutes (Fig. 4A, C). The hemicellulases were mainly encoded by Firmicutes, Bacteroidota, Proteobacteria, and Spirochaetota (Fig. 4E). The oligosaccharide enzymes were mainly encoded by Bacteroidota, Firmicutes, Spirochaetota, and Proteobacteria (Fig. 4G). At the genus level, the cellulases and hemicellulases were mainly encoded by *Candidatus_Azobacteroides* and *Treponema* (Fig. 4B, D). The debranching enzymes were mainly encoded by *Fibrobacter*, *Bacteroidales_noname*, *Desulfovibrio*, *and Parabacterioides* (Fig. 4F). The debranching enzymes were mainly encoded by *Candidatus_Azobacteroides*, *Treponema*, *Clostridium*, *Bacteroides*, *Paenibacillus*, *Dyssgonomonas*, and *Parabacteroides* (Fig. 4H). It is noteworthy that the abundance and diversity of the microbial community encoding oligosaccharide enzymes were significantly higher than those encoding cellulases, hemicellulases, and debranching enzymes.

Discussion

The conversion of lignocellulosic biomass into biofuels is a promising yet challenging avenue due to the complexities of plant cell wall polysaccharides. There is an urgent need to discover and develop more efficient and novel lignocellulolytic enzymes to enhance the saccharification of biomass. In nature, termites can utilize symbiotic microbiota, particularly wood-feeding species, to degrade plant biomass. *G. sulphureus* is a higher wood-feeding termite, but its ability to degrade lignocellulose has not been explored. Therefore, this study used 16 S rRNA and metagenomic sequencing to analyze the diversity and functional GHs of hindgut microbial communities in the higher termite *G. sulphureus*. The results were compared with *C. formosanus*, a lower termite species that has received more attention for its hindgut bacterial communities and lignocellulose degradation abilities. At the same time, the β-glucosidase, exo-β-1,4-glucanase, endo-β-1,4-glucanase, and xylanase activities in the hindgut microbiota of both termite species were measured. The results showed that Spirochaetota, Firmicutes, and Fibrobacterota were the dominant microbiota in the hindgut of two termite species. At the phylum level, the relative abundances of Proteobacteria and Bacteroidota were significantly higher in the hindgut of *C. formosanus* than in that of *G. sulphureus* (Mann-Whitney U test, $Z = -2.449$, $P<0.05$). At the genus level, the relative abundances of *Candidatus_Azobacteroides* and *Escherichia-Shigella* were significantly lower in the hindgut of *G. sulphureus* than in that of *C. formosanus* (*Candidatus_Azobacteroides*: Mann-Whitney U test, $Z = -2.183$, $P<0.05$; *Escherichia-Shigella*: Mann-Whitney U test, $Z = -2.205$, $P<0.05$). The hindgut microbiota of *G. sulphureus* exhibits a rich diversity of GHs involved in lignocellulose degradation. The classification and distribution of CAZymes indicated that the bacteria encoding GHs in *G. sulphureus* and *C. formosanus* were different. This study laid the foundation for the exploration of novel and efficient lignocellulose-degrading genes.

Many studies have shown that the alpha diversity of hindgut microbial communities in higher termites is significantly higher than that in lower termites [46–48], which is similar to our results. In this study, the Observed_species index and Chao1 index indicated that the number of hindgut microbiota in *G. sulphureus* (a higher termite) was significantly higher than that in *C. formosanus* (Chao1: Mann-Whitney U test, $Z = -1.960$, $P=0.06$; Observed_species: Mann-Whitney U test, $Z = -2.449$, $P<0.05$). Previous studies have shown that the ability of lower termite gut microbiota to degrade lignocellulose mainly comes from protozoa, while higher termites lack such symbiotic protozoa [49]. This may be one of the reasons why the bacterial microbial community is more abundant in higher termites, which need a greater abundance of bacteria to degrade lignocellulose [50]. Previous studies have reported that protists occupy most of the space in the gut of lower termites and that the surface of these protists is colonized by a large number of commensal bacteria [51]. Most of these bacteria have been identified as belonging to the Bacteroidales order, accounting for approximately 70% of the bacterial cells in termites [52]. Previous studies have reported that the dominant bacterial genus in the gut microbiota of *C. formosanus* is "*Candidatus Zobacteroides*", which belongs to Bacteroidota [52], and is also the main nitrogen-fixing bacteria in *C. formosanus* [53]. Su et al.. (2016) reported that the relative abundance of "*Candidatus Azobacteroides*" in the gut of lower wood-feeding termites was significantly higher than that in higher wood-feeding termites [3]. Our results are consistent with this finding, *Candidatus_Azobacteroides* was found only in the hindgut of *C. formosanus.*

The gut microbial diversity of termites is not only influenced by phylogeny but is also closely related to feeding habits [14, 54, 55]. Increasing evidence has revealed that termites with different feeding habits have dominant phyla. For example, Spirochaetota is the dominant phylum in wood-feeding termites [56], Bacteroidota and

Firmicutes are the dominant phyla in fungus-feeding termites [49], and Actinobacteria is the dominant phylum in soil-feeding termites [57]. Su et al. (2016) reported that the gut microbiota of lower and higher wood-feeding termites is highly similar, with Spirochaetota being the most dominant phylum [14]. This is similar to our study, which revealed that *G. sulphureus* and *C. formosanus* share common dominant bacterial phyla: Spirochaetota, Firmicutes, and Fibrobacterota, as well as common dominant bacterial genera: *Lachnospiraceae_unclassified*, *Termite_Treponema_cluster*, and *Treponema*. Furthermore, PCA and NMDS analysis revealed no significant differences in the structure of the hindgut microbial communities of the two termite species. Metagenomic sequencing results showed that GHs accounted for 38.08% and 35.41% of the CAZymes annotated in the hindgut microbiota of *G. sulphureus* and *C. formosanus*, respectively. Further analysis revealed that the relative abundance of cellulase was high in the hindgut microbiota of *C. formosanus*, while the relative abundance of hemicellulase was high in the hindgut microbiota of *G. sulphureus*. Similar to the results of the relative abundance of GHs, the enzyme activity assay showed that the cellulase (β-glucosidase, exo-β-1,4-glucanase, and endo-β-1,4-glucanase) activity was high in the hindgut microbiota of *C. formosanus*, while the hemicellulase (xylanase) activity was high in the hindgut microbiota of *G. sulphureus*. These findings suggested that *G. sulphureus* might have high hemicellulose degradation ability. Previous studies have shown that *G. sulphureus* mainly damages coconut trees and palm trees, and palm leaves have been proven to be a major potential source of xylose [58]. Our results showed that the relative abundances of GH3 (Mann-Whitney U test, Z = - 2.193, $P<0.05$), GH42 (Mann-Whitney U test, Z = - 2.611, $P<0.05$), and GH43 (Mann-Whitney U test, Z = - 2.643, $P<0.05$) were significantly higher in *G. sulphureus* compared to *C. formosanus*. These three GHs are mainly involved in the degradation of oligosaccharides and catalyze α-L-arabinofuranosidase activity. Additionally, the relative abundance of GH51, which is responsible for the debranching of arabinose side chains, was significantly higher in *G. sulphureus* than in *C. formosanus*. These findings support the notion that dietary preferences drive the composition of the termite gut microbiota.

Wood-feeding termites have abundant GHs for lignocellulose digestion [16, 21, 59]. Arora et al. (2022) demonstrated that a significant abundance of CAZymes is present in the majority of termites, including GH2, GH3, GH10, GH31, and GH77. Additionally, the GH8, GH26, GH45, GH5_2, and GH53 are found in high quantities specifically within the wood-feeding termites [54]. Victorica et al. (2020) analyzed the microbial community composition and lignocellulose metabolic potential between *C. fulviceps* and *Nasutitermes aquilinus*, revealing that GHs constitute 40.3% and 37.6% of the total CAZymes in these termites, respectively. The GHs with the highest abundance included GH1, GH3, GH5, GH9, GH10, GH11, and GH43 [18]. Liu et al. (2019) identified GH families with higher abundances in *G. brachycerastes*, including GH1, GH3, GH5, GH10, and GH11 [16]. This is similar to our results, both *G. sulphureus* and *C. formosanus* have abundant GHs, with higher abundances of GH2, GH3, GH5 subfamily, GH6, GH9, and GH10. We further analyzed the taxonomic distribution of CAZymes at the phylum and genus levels. The cellulases and hemicellulases in *G. sulphureus* were mainly encoded by the most abundant Spirochaetota, Fibrobacterota, Firmicutes, and Proteobacteria. In *C. formosanus*, the cellulases and hemicellulases were mainly encoded by Bacteroidota and Spirochaetota. Our results are similar to previous studies. In the wood-feeding termites *Microcerotermes parvus* and *Neocapritermes taracua*, the cellulase and hemicellulases are predominantly encoded by the most abundant *Spirochaetales*, *Fibrobacterales*, *Clostridiales*, and *Bacteroidates* [17]. CAZymes in *Nasutitermes* spp. are mainly encoded by the most abundant Spirochaetota and Fibrobacterota [15]. These results are consistent with previous studies, which found that the gut microbiota of termites have similar CAZymes, and are encoded by the most abundant bacterial groups [19, 21, 60].

GH6 and GH48 are key components of the lignocellulolytic enzyme system [61, 62]. However, these two GHs are absent in higher termites such as *Nasutitermes* spp. and *Macrotermes annandalei* [15, 63]. In difference to these findings, this study found the presence of GH6 in both *G. sulphureus* and *C. formosanus*, while GH48 was found exclusively in the hindgut of *G. sulphureus*. In *G. sulphureus*, the cellulases were mainly GH5_2, GH5_4, GH6, GH9, and GH45, while the hemicellulases were mainly GH11, GH8, GH10, GH11, GH26, and GH53. In *C. formosanus*, the cellulases were mainly GH6 and GH9, and the hemicellulases were mainly GH5_7, GH5_21, GH10, GH12, and GH53. Interestingly, the diversity of the GH5 subfamily is higher in *C. formosanus* compared to *G. sulphureus*. Previous studies have indicated that the gene abundance of cellulase GH5 is highest in certain higher termites, such as *C. fulviceps*, *N. aquilinus*, and *G. brachycerastes* [16, 19]. Calusinska et al. (2020) indicated that the abundance of GH5 is highest in the gut microbiota of Nasutitermitinae, predominantly consisting of GH5_2 and GH5_4. This finding is consistent with our results. In *G. sulphureus*, GH5_2 and GH5_4 are the predominant cellulases within the GH5 family [64]. The results suggested that the hindgut of *G. sulphureus* has diverse GHs for degrading cellulose and hemicellulose, which could be a potential source of novel lignocellulose-degrading enzymes. Our findings are limited by the lack

of expression information, which could restrict the effectiveness of GH screening efforts. Future research needs to conduct a thorough expression analysis of the identified GHs in *G. sulphureus* to fully assess their potential for biotechnological applications.

Conclusion

In summary, the study showed that Spirochaetota, Firmicutes, and Fibrobacterota dominated the hindgut microbiota of *C. formosanus* and *G. sulphureus*. Compared to *C. formosanus*, *G. sulphureus* is enriched in genes encoding hemicellulase and debranching enzymes. Additionally, it highlights the rich diversity of GHs in the hindgut microbiota of *G. sulphureus*, including the GH5 subfamily, GH6, and GH48, with the GH6 and GH48 not previously reported in other higher termites. These results strengthen the understanding of the diversity of termite gut microbiota and CAZymes.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12866-024-03623-8.

Supplementary Material 1

Acknowledgements
All authors thank the staff and students of the Wildlife Conservation Research Team of the Guangdong Academy of Forestry for the collection of samples.

Author contributions
HY, WC, WK, and ZCS designed the experiments; ZZD, ZT, and WWB performed the sample collection; ZZD and ZCS carried out the data analyses; ZZD and WK were responsible for the original manuscript; HY, WK, and WC provided critical revisions to this manuscript.

Funding
This study was funded by the Guangdong Basic and Applied Basic Research Foundation (2024A1515012617).

Data availability
The raw data from the 16S rRNA sequencing and metagenomic sequencing have been deposited in the NCBI database with the accession number PRJNA940600.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Guangdong Provincial Key Laboratory of Silviculture, Protection and Utilization, Guangdong Academy of Forestry, Guangzhou 510520, China
[2]College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou 510642, China
[3]College of Forestry, Northeast Forestry University, Harbin 150040, China

Received: 23 June 2024 / Accepted: 4 November 2024
Published online: 12 November 2024

References

1. Paul S, Dutta A. Challenges and opportunities of lignocellulosic biomass for anaerobic digestion. Resour Conserv Recycl. 2018;130(5):164–74. https://doi.org/10.1016/j.resconrec.2017.12.005.
2. Mujtaba M, Fraceto LF, Fazeli M, Mukherjee S, Savassa SM, de Medeiros GA, et al. Lignocellulosic biomass from agricultural waste to the circular economy: a review with focus on biofuels, biocomposites and bioplastics. J Clean Prod. 2023;402(18):136815. https://doi.org/10.1016/j.jclepro.2023.136815.
3. Brune A. Symbiotic digestion of lignocellulose in termite guts. Nat Rev Microbiol. 2014;12(3):168–80. https://doi.org/10.1038/nrmicro3182.
4. Tokuda G, Lo N, Watanabe H. Marked variations in patterns of cellulase activity against crystalline- vs. carboxymethyl-cellulose in the digestive systems of diverse, wood-feeding termites. Physiol Entomol. 2005;30(4):372–80. https://doi.org/10.1111/j.1365-3032.2005.00473.x.
5. Ohkuma M. Termite symbiotic systems: efficient bio-recycling of lignocellulose. Appl Microbiol biot. 2003;61(1):1–9. https://doi.org/10.1007/s00253-002-1189-z.
6. Scharf ME. Termites as targets and models for biotechnology. Annu Rev Entomol. 2015;60(1):77–102. https://doi.org/10.1146/annurev-ento-010814-020902.
7. Brauman A, Majeed MZ, Buatois B, Robert A, Pablo A-L, Miambi E. Nitrous oxide (N_2O) emissions by termites: does the feeding guild matter? PLoS ONE. 2015;10(12):e0144340. https://doi.org/10.1371/journal.pone.0144340.
8. Breznak JA, Brune A. Role of microorganisms in the digestion of lignocellulose by termites. Annu Rev Entomol. 1994;39(1):453–. https://doi.org/10.1146/annurev.en.39.010194.002321. 87.
9. Wong LJ, H'ng PS, Wong SY, Lee SH, Lum WC, Chai EW, et al. Termite digestomes as a potential source of symbiotic microbiota for lignocelluloses degradation: a review. PJBS. 2014;17(8):956–63. https://doi.org/10.3923/pjbs.2014.956.963.
10. Brennan Y, Callen WN, Christoffersen L, Dupree P, Goubet F, Healey S, et al. Unusual microbial xylanases from insect guts. Appl Environ Microb. 2004;70(6):3609–17. https://doi.org/10.1128/AEM.70.6.3609-3617.2004.
11. 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(12):4942–53. https://doi.org/10.1111/1462-2920.12945.
12. Do TH, Dao TK, Nguyen HD, Truong NH. Understanding the role of free-living bacteria in the gut of the lower termite *Coptotermes gestroi* based on metagenomic DNA analysis. Insects. 2023;14(11):832. https://doi.org/10.3390/insects14110832.
13. Mikaelyan A, Strassert JF, Tokuda G, Brune A. The fibre-associated cellulolytic bacterial community in the hindgut of wood-feeding higher termites (Nasutitermes spp). Environ Microbiol. 2014;16(9):2711–22. https://doi.org/10.1111/1462-2920.12425.
14. Su L, Yang L, Huang S, Su X, Li Y, Wang F, et al. Comparative gut microbiomes of four species representing the higher and the lower termites. J Insect Sci. 2016;16(1):97. https://doi.org/10.1093/jisesa/iew081.
15. 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(7169):560–5. https://doi.org/10.1038/nature06269.
16. Liu N, Li H, Chevrette MG, Zhang L, Cao L, Zhou H, et al. Functional metagenomics reveals abundant polysaccharide-degrading gene clusters and cellobiose utilization pathways within gut microbiota of a wood-feeding higher termite. ISME J. 2019;13(1):104–17. https://doi.org/10.1038/s41396-018-0255-1.
17. Hu H, da Costa RR, Pilgaard B, Schiøtt M, Lange L, Poulsen M. Fungiculture in termites is associated with a mycolytic gut bacterial community. Msphere. 2019;4(3):e00165–19. https://doi.org/10.1128/msphere.00165-19.
18. Romero Victorica M, Soria MA, Batista-García RA, Ceja-Navarro JA, Vikram S, Ortiz M, et al. Neotropical termite microbiomes as sources of novel plant cell wall degrading enzymes. Sci Rep. 2020;10(1):3864. https://doi.org/10.1038/s41598-020-60850-5.
19. Tokuda G, Mikaelyan A, Fukui C, Matsuura Y, Watanabe H, Fujishima M et al. Fiber-associated spirochetes are major agents of hemicellulose degradation in the hindgut of wood-feeding higher termites. PNAS. 2018;115(51):E11996-E2004. https://doi.org/10.1073/pnas.1810550115

20. 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(4):e61126. https://doi.org/10.1371/journal.pone.0061126.
21. Grieco MB, Lopes FA, Oliveira LS, Tschoeke DA, Popov CC, Thompson CC, et al. Metagenomic analysis of the whole gut microbiota in Brazilian Termitidae termites Cornitermes cumulans, Cyrilliotermes Strictinasus, Syntermes Dirus, Nasutitermes Jaraguae, Nasutitermes Aquilinus, Grigiotermes bequaerti, and Orthognathotermes Mirim. Curr Microbiol. 2019;76(6):687–97. https://doi.org/10.1007/s00284-019-01662-3.
22. Marynowska M, Sillam-Dussès D, Untereiner B, Klimek D, Goux X, Gawron P, et al. A holobiont approach towards polysaccharide degradation by the highly compartmentalised gut system of the soil-feeding higher termite Labiotermes labralis. BMC Genom. 2023;24(1):115. https://doi.org/10.1186/s12864-023-09224-5.
23. Kuswanto E, Ahmad I, Dungani R. Threat of subterranean termites attack in the Asian countries and their control: a review. AJAS. 2015;8(4):227–39. https://doi.org/10.3923/ajaps.2015.227.239.
24. Rust MK, Su N-Y. Managing social insects of urban importance. Annu Rev Entomol. 2012;57(1):355–. https://doi.org/10.1146/annurev-ento-120710-100634. 75.
25. Liu G, Cornwell WK, Cao K, Hu Y, Van Logtestijn RS, Yang S, et al. Termites amplify the effects of wood traits on decomposition rates among multiple bamboo and dicot woody species. J Ecol. 2015;103(5):1214–23. https://doi.org/10.1111/1365-2745.12427.
26. Majid AHA, Ahmad AH, Rashid M, Rawi CSM. Preliminary field efficacy of imidacloprid on *Globitermes sulphureus* (Isoptera: Termitidae)(subterranean termite) in Penang. Trop Life Sci Res. 2007;18(2):109–14.
27. Neoh K-B, Jalaludin NA, Lee C-Y. Elimination of field colonies of a mound-building termite *globitermes sulphureus* (Isoptera: Termitidae) by bistrifluron bait. J Econ Entomol. 2011;104(2):607–13. https://doi.org/10.1603/EC10161.
28. Aiman Hanis J, Abu Hassan A, Nurita A, Che Salmah M. Community structure of termites in a hill dipterocarp forest of Belum–Temengor Forest Complex, Malaysia: emergence of pest species. Raffles Bull Zool. 2014;62:3–11. https://doi.org/10.5281/zenodo.5353138.
29. Sornnuwat Y, Vongkaluang C, Takematsu Y. A systematic key to termites of Thailand. Agric Nat Resour. 2004;38(3):349–68. http://creativecommons.org/licenses/by-nc-nd/4.0/.
30. Huang FS, Zhu SM, Ping ZM, He XS, Li GX, Gao DR. Fauna Sinica, Insecta, volume 17, Isoptera. Beijing: Science; 2000.
31. Liu N, Yan X, Zhang M, Xie L, Wang Q, Huang Y, et al. Microbiome of fungus-growing termites: a new reservoir for lignocellulase genes. Appl Environ Microb. 2011;77(1):48–56. https://doi.org/10.1128/AEM.01521-10.
32. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335–6. https://doi.org/10.1038/nmeth.f.303.
33. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011;27(21):2957–63. https://doi.org/10.1093/bioinformatics/btr507.
34. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4:e2584. https://doi.org/10.7717/peerj.2584.
35. 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(7):581–3. https://doi.org/10.1038/nmeth.3869.
36. Jiao S, Xu Y, Zhang J, Lu Y. Environmental filtering drives distinct continental atlases of soil archaea between dryland and wetland agricultural ecosystems. Microbiome. 2019;7(1):1–13. https://doi.org/10.1186/s40168-019-0630-9.
37. Pertea G. Fqtrim: v0.9.4 release. CS. 2015. https://doi.org/10.5281/ZENODO.20552.
38. Peng Y, Leung HC, Yiu S-M, Chin FY. Bioinformatics. 2012;28(11):1420–8. https://doi.org/10.1093/bioinformatics/bts174. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth.
39. Zhu W, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences. Nucleic Acids Res. 2010;38(12):e132. https://doi.org/10.1093/nar/gkq275.
40. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28(23):3150–2. https://doi.org/10.1093/bioinformatics/bts565.
41. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3):R25. https://doi.org/10.1186/gb-2009-10-3-r25.
42. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30. https://doi.org/10.1093/nar/28.1.27.
43. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The carbohydrate-active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. 2009;37(Database issue):D233–8. https://doi.org/10.1093/nar/gkn663.
44. Allgaier M, Reddy A, Park JI, Ivanova N, D'haeseleer P, Lowry S, et al. Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. PLoS ONE. 2010;5(1):e8812. https://doi.org/10.1371/journal.pone.0008812.
45. Hall M, Beiko RG. 16S rRNA gene analysis with QIIME2. Methods Mol Biol. 2018;1849:113–29. https://doi.org/10.1007/978-1-4939-8728-3_8.
46. Huang X-F, Bakker MG, Judd TM, Reardon KF, Vivanco JM. Variations in diversity and richness of gut bacterial communities of termites (*Reticulitermes flavipes*) fed with grassy and woody plant substrates. Microb Ecol. 2013;65(3):531–6. https://doi.org/10.1007/s00248-013-0219-y.
47. Vikram S, Arneodo JD, Calcagno J, Ortiz M, Mon ML, Etcheverry C, et al. Diversity structure of the microbial communities in the guts of four neotropical termite species. PeerJ. 2021;9:e10959. https://doi.org/10.7717/peerj.10959.
48. Dietrich C, Köhler T, Brune A. The cockroach origin of the termite gut microbiota: patterns in bacterial community structure reflect major evolutionary events. Appl Environ Microb. 2014;80(7):2261–9. https://doi.org/10.1128/AEM.04206-13.
49. Otani S, Mikaelyan A, Nobre T, Hansen LH, Koné NGA, Sørensen SJ, et al. Identifying the core microbial community in the gut of fungus-growing termites. Mol Ecol. 2014;23(18):4631–44. https://doi.org/10.1111/mec.12874.
50. Ohkuma M. Symbioses of flagellates and prokaryotes in the gut of lower termites. Trends Microbiol. 2008;16(7):345–52. https://doi.org/10.1016/j.tim.2008.04.004.
51. Ohkuma M, Brune A. Diversity, structure, and evolution of the termite gut microbial community. In: Bignell D, Roisin Y, Lo N, editors. Biology of termites: a modern synthesis. Dordrecht: Springer; 2011. pp. 413–38. https://doi.org/10.1007/978-90-481-3977-4_15.
52. Noda S, Iida T, Kitade O, Nakajima H, Kudo T, Ohkuma M. Endosymbiotic Bacteroidales bacteria of the flagellated protist *Pseudotrichonympha Grassii* in the gut of the termite *Coptotermes formosanus*. Appl Environ Microb. 2005;71(12):8811–7. https://doi.org/10.1128/AEM.71.12.8811-8817.2005.
53. Hongoh Y, Sharma VK, Prakash T, Noda S, Toh H, Taylor TD, et al. Genome of an endosymbiont coupling N_2 fixation to cellulolysis within protist cells in termite gut. Science. 2008;322(5904):1108–9. https://doi.org/10.1126/science.1165578.
54. Arora J, Kinjo Y, Šobotník J, Buček A, Clitheroe C, Stiblik P, et al. The functional evolution of termite gut microbiota. Microbiome. 2022;10(1):78. https://doi.org/10.1186/s40168-022-01258-3.
55. 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(4):649–54. https://doi.org/10.1016/j.cub.2018.01.035. e2.
56. Mikaelyan A, Dietrich C, Köhler T, Poulsen M, Sillam-Dussès D, Brune A. Diet is the primary determinant of bacterial community structure in the guts of higher termites. Mol Ecol. 2015;24(20):5284–95. https://doi.org/10.1111/mec.13376.
57. Fall S, Hamelin J, Ndiaye F, Assigbetse K, Aragno M, Chotte JL, et al. Differences between bacterial communities in the gut of a soil-feeding termite (*Cubitermes Niokoloensis*) and its mounds. Appl Environ Microb. 2007;73(16):5199–208. https://doi.org/10.1128/AEM.02616-06.
58. Manaf SFA, Jahim JM, Harun S, Luthfi AAI. Fractionation of oil palm fronds (OPF) hemicellulose using dilute nitric acid for fermentative production of xylitol. Ind Crop Prod. 2018;115:6–15. https://doi.org/10.1016/j.indcrop.2018.01.067.
59. Lynd LR, Weimer PJ, Van Zyl WH, Pretorius IS. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol R. 2002;66(3):506–77. https://doi.org/10.1128/MMBR.66.3.506-577.2002.
60. Marynowska M, Goux X, Sillam-Dussès D, Rouland-Lefèvre C, Halder R, Wilmes P, et al. Compositional and functional characterisation of biomass-degrading microbial communities in guts of plant fibre-and soil-feeding higher termites. Microbiome. 2020;8:1–18. https://doi.org/10.1186/s40168-020-00872-3.
61. Taylor LE, Henrissat B, Coutinho PM, Ekborg NA, Hutcheson SW, Weiner RM. Complete cellulase system in the marine bacterium *Saccharophagus*

degradans strain 2-40^{T}. J Bacteriol. 2006;188(11):3849–61. https://doi.org/10.1128/JB.01348-05.
62. Xie G, Bruce DC, Challacombe JF, Chertkov O, Detter JC, Gilna P, et al. Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*. Appl Environ Microb. 2007;73(11):3536–46. https://doi.org/10.1128/AEM.00225-07.
63. Cui Y, Wang W, Lu Q, Wu HX, Ling X, Liu J, et al. Metagenomic analysis suggests unique gut-microbiota composition and GH family constitution of *Macrotermes Annandalei*. Entomol. 2022;52(3):127–34. https://doi.org/10.1111/1748-5967.12572.
64. Calusinska M, Marynowska M, Bertucci M, Untereiner B, Klimek D, Goux X, et al. Integrative omics analysis of the termite gut system adaptation to Miscanthus diet identifies lignocellulose degradation enzymes. Commun Biol. 2020;3(1):275. https://doi.org/10.1038/s42003-020-1004-3.

Publisher's note