



Metagenomic insights into mangrove lignocellulolytic bacteria and functional analysis of a glucose-tolerant GH 1 β -glucosidase

Kheng Loong Chong¹ · Kok Jun Liew¹ · Faezah Mohd Salleh¹ · Chun Shiong Chong¹

Received: 11 September 2025 / Accepted: 26 March 2026 / Published online: 15 April 2026
 © The Author(s) 2026

Abstract

Mangrove ecosystems contain abundant lignocellulosic biomass and mangrove microorganisms that are capable of degrading plant polymers. In this study, a shotgun metagenomic approach was employed to explore the bacterial communities from Tanjung Piai National Park, Malaysia and their genes involved in lignocellulosic biomass degradation. A total of 148 of carbohydrate active enzymes (CAZy) genes spanning GH, CE, and AA families were identified with lignocellulolytic abilities. These enzymes included 20 cellulases, 46 hemicellulases, and 82 lignin-modifying enzymes. Approximately 89.19% of these genes were found from underexplored bacterial lineages. A set of lignocellulolytic genes derived from diverse bacterial taxa highlighted the synergistic action of mangrove bacteria in lignocellulose degradation. To validate the functionality of these genetic resources, one of the genes (*BGL3_GH1*) encoding a β -glucosidase was selected for expression and characterisation. The recombinant enzyme showed optimal activity at 60 °C and pH 7, retained up to 75% activity at 10% (w/v) NaCl. The enzyme exhibited a 1.6 to 2.1-fold in enzyme activity with glucose concentration up to 2 M. In a two-step saccharification assay using sugarcane bagasse, supplementation with recombinant BGL3_GH1 enhanced the saccharification yield (0.0674 g g⁻¹ biomass) compared with treatments using commercial cellulase or recombinant BGL3_GH1 alone. These findings reveal the functional diversity of lignocellulose-degrading genes in mangrove bacteria and identify recombinant BGL3_GH1 as a potential enzyme candidate for biomass conversion application.

Keywords Bacterial communities · Lignocellulose biodegradation · Mangrove · Metagenomics · Taxonomy and functional profiling

Abbreviations

AA	Auxiliary activity
CAZy	Carbohydrate-active enzyme database
CAZymes	Carbohydrate-active enzymes
CBM	Carbohydrate-binding module
CE	Carbohydrate esterase
CMC	Carboxymethylcellulose

EDTA	Ethylenediaminetetraacetic acid
eggNOG	Evolutionary genealogy of genes: non-supervised orthologous groups
GH	Glycosyl hydrolase
GT	Glycosyl transferase
KEGG	Kyoto encyclopedia of genes and genomes
LPMO	Lytic polysaccharide monooxygenase
MAG	Metagenome-assembled genome
NCBI	National center for biotechnology information
ORFs	Open reading frames
PL	Polysaccharide lyase
SRA	Sequence read archive
SDS	Sodium dodecyl sulfate

✉ Faezah Mohd Salleh
 faezah@utm.my

✉ Chun Shiong Chong
 cschong@utm.my

Kheng Loong Chong
 klchong6@graduate.utm.my

Kok Jun Liew
 kokjunliew@gmail.com

¹ Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

Introduction

Mangroves are among the most carbon-rich ecosystems in the tropics, where fallen leaves, twigs, roots, and woody debris contribute to the organic matter pool (Priya et al. 2018). Plant-derived organic matter in mangroves is primarily composed of lignocellulose, a structural component of plant cell walls made up of cellulose, hemicellulose, and lignin (Zhang et al. 2021). The abundance of lignocellulosic material in mangrove ecosystem exerts selective pressure on indigenous microorganisms driving the evolution of genes and enzymes specialized for polymer degradation (Kathiresan 2019). The degradation of lignocellulose releases simple sugar molecules such as glucose, mannose, and xylose that serve as key energy source to support microbial growth and survival (Mai et al. 2021; Yousuf et al. 2020).

Lignocellulose degradation is mediated by carbohydrate-active enzymes (CAZymes), including cellulases, hemicellulases, lytic polysaccharide monooxygenases (LPMOs), and lignin-modifying enzymes (Zhai et al. 2022). CAZymes were categorized into six major domains: glycoside hydrolases (GHs), glycosyltransferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), auxiliary activities (AA), and carbohydrate-binding modules (CBMs) (Terapon et al. 2017). Among these, GH, CE, and AA families are particularly important for lignocellulose degradation (Bredon et al. 2018; Liew et al. 2022; Xiao et al. 2022). For instance, cellulases and hemicellulases are classified under GH and CE families, while LPMO and lignin-modifying enzymes fall within the AA family (Grgas et al. 2023; Gunjal et al. 2020).

Within the cellulase system, β -glucosidase (EC 3.2.1.21) performs the rate-limiting step of cellulose degradation by hydrolysing cellobiose to glucose (Sengupta et al. 2023). This reaction is crucial for efficient biomass saccharification, which converts lignocellulose into fermentable sugars for biofuel production (Sharma et al. 2022). However, β -glucosidases are often inhibited by accumulated glucose, reducing overall hydrolysis efficiency (Sengupta et al. 2023). Therefore, discovering new β -glucosidases with high glucose tolerance is essential to improve industrial biomass conversion (Muftaba et al. 2023).

Although numerous lignocellulolytic enzymes have been discovered through culture-dependent methods, a substantial proportion of microorganisms with lignocellulolytic potential remain unexplored. This is due to the limitations of culture-based techniques, as only approximately 1% of bacterial species are cultivated under laboratory conditions (Bodor et al. 2020). Metagenomics provides a culture-independent approach to analyse genetic material directly from environmental samples using bioinformatics tools and reference databases (CAZy, eggNOG, KEGG, and NCBI).

This approach allows access to previously overlooked bacterial communities and facilitates the identification of their functional traits and ecological roles (Sagar et al. 2024). Through metagenomics, novel lignocellulolytic gene pools have been uncovered from diverse environments, offering a valuable reservoir of enzymes with potential applications in various biotechnological processes (Liew et al. 2022; Rubio-Portillo et al. 2021; Zhao et al. 2019).

Despite Malaysia ranks among the top 10 countries with the largest mangrove coverage, metagenomics studies focusing specifically on lignocellulolytic genes in Malaysia mangrove soils remain limited (Priyashantha and Taufikurrahman 2020). In the present study, shotgun metagenomic sequencing was employed to investigate lignocellulose-degrading gene profiles from the Tanjung Piai mangrove soil microbiome. The aims of this study were: (1) to profile the functional genes of mangrove microbiome; (2) to characterize the contribution of bacterial communities to lignocellulose degradation; and (3) to clone, express, and biochemically characterize a GH1 β -glucosidase (BGL3_GH1) to determine its potential in biomass saccharification.

Materials and methods

Sample collection and soil physiochemical parameters

Soil samples were collected from the Tanjung Piai National Park, Malaysia (1°16'01.6"N, 103°30'27.6"E), following the procedure described by Priya et al. (2018). The sampling site was located within a coastal wetland area (Fig. 1). Nine individual soil cores were collected from a depth of 10 cm within a 1 m² plot and homogenized into one composite sample. The sampling was conducted in October 2022. The determination of physiochemical properties of soil samples was outsourced to Allied Chemists, Malaysia. Soil physiochemical parameters were measured for pH (7.4 ± 0.1), salinity ($1.5\% \pm 0.1$), and total carbon and nitrogen (C: 3.7%, N: 0.1%). The composite sample was transported to the laboratory on ice and stored at -80 °C until further process.

DNA extraction and shotgun metagenome sequencing

Total DNA was extracted from the composite sample using the DNeasy®PowerSoil® Pro Kit (QIAGEN, Germany), following the manufacturer's instruction. Approximately 250 mg of soil was homogenized using TissueLyser II (QIAGEN, Germany) at 25 Hz for three cycles. The quality of the extracted DNA was assessed by 1% (w/v) agarose gel electrophoresis, and DNA concentration was measured

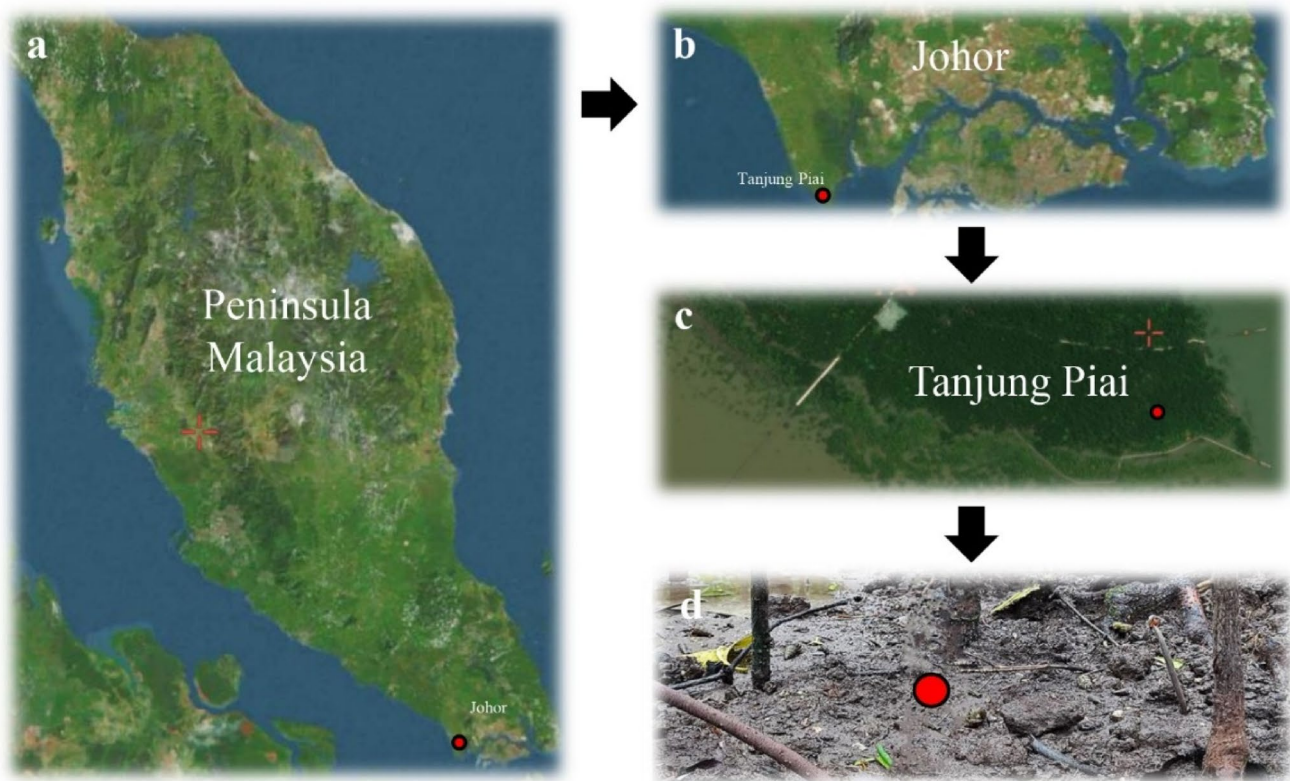


Fig. 1 Satellite map of Malaysia showing the location of sampling at Tanjung Piai National Park. The collection site is indicated by red circles. **a** Satellite image of Peninsular Malaysia, **b** satellite image of

Johor, **c** satellite image of Tanjung Piai, **d** Tanjung Piai mangrove forests sample collection site

using a NanoDrop 2000 spectrophotometer (ThermoScientific, Wilmington, DE, USA). Sequencing was performed by NovogeneAIT (Singapore) using the Illumina NovaSeq 6000 platform (paired end 150 bp). The sequencing data was deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA939948.

Metagenomics data analysis and selection of *BGL3_GH1*

The paired-end reads in FASTQ format were quality-filtered and adapter-trimmed using Trimmomatic V0.32, retaining reads with a minimum Phred score of 20 and length ≥ 50 bp (Bolger et al. 2014). High-quality reads were assembled de novo using MEGAHIT v1.2.9 with a minimum k-mer size of 21 (Li et al. 2015). To reduce artefacts associated with highly fragmented contigs, only assembled contigs ≥ 1000 bp were retained for downstream analysis. Open reading frames (ORFs) were predicted from these filtered contigs using Prodigal v2.6.3 in metagenome mode (Hyatt et al. 2010; Rho et al. 2010). Functional annotation of ORFs was performed using eggNOG-mapper v2.1.10 and Ghost-KOALA with default parameters, referencing the eggNOG and KEGG databases, respectively (Cantalapiedra et al.

2021; Kanehisa et al. 2016). Carbohydrate-active enzymes (CAZymes) were identified using run_dbcan v3.0.6 (Zhang et al. 2018). Only CAZymes that were supported by two out of three methods (HMMER v3.3.2, eCAMI, or DIAMOND v2.1.6) were retained for downstream analysis. Candidate sequences were further validated against the NCBI non-redundant, RCSB Protein Data Bank (PDB), and UniProt reviewed protein databases using Diamond v2.1.6 (Buchfink et al. 2015). Protein domains of selected CAZymes were analysed using InterProScan v5.61–93.0 (Jones et al. 2014). Taxonomy of predicted CAZymes was performed using DIAMOND BLASTP against the NCBI nr database with criteria of $\geq 50\%$ sequence identity and $\geq 50\%$ query coverage. Based on the annotation results, a β -glucosidase gene (*BGL3_GH1*) was selected from the metagenomic data for subsequent functional characterisation and sugarcane bagasse saccharification. Domain architecture was examined using Pfam database (Bateman et al. 2004). Phylogenetic analysis was conducted by aligning *BGL3_GH1* with representative GH 1 β -glucosidase retrieved from UniProt database and restricted to reviewed entries. Phylogenetic tree construction was performed using the Neighbor-Joining method with 1,000 bootstrap replicates via MEGA 11 (Tamura et al. 2021). The conserved catalytic

and substrate-binding residues of recombinant BGL3_GH1 and characterized GH1 β -glucosidase were identified by performing multiple sequence alignment in Jalview 2.11.5.1 using the MUSCLE algorithm with default parameters (Waterhouse et al. 2009). The nucleotide and protein sequence data of recombinant BGL3_GH1 were deposited in the NCBI database under accession number PX852031.

Production of the recombinant BGL3_GH1 protein

The *BGL3_GH1* gene was synthesized and codon-optimized for expression in *Escherichia coli* by Integrated DNA Technologies (IDT, Singapore). The synthesized gene and expression vector *pET-28a* (+) were digested with *EcoRI* and *XhoI* restriction enzyme, followed by ligation using T4 DNA ligase. The recombinant plasmid was transformed into chemically competent *E. coli* BL21 (DE3) cells and selected on Luria–Bertani (LB) agar plates containing 50 $\mu\text{g mL}^{-1}$ kanamycin. Positive transformant were cultured, harvested, and resuspended in 50 mM potassium phosphate buffer (pH 7.0). Cell lysis was carried out by sonication on ice, and the lysate was centrifuged at 4 °C, 10,000 $\times g$ for 20 min. The supernatant was filtered through a 0.45 μm filter and loaded onto a pre-equilibrated Protino Ni–NTA columns (Macherey–Nagel). The bound protein was eluted using a linear imidazole buffer gradient (10 mM to 250 mM). The eluted recombinant BGL3_GH1 protein was dialyzed against 50 mM potassium phosphate buffer (pH 7.0). Protein purity and integrity were assessed by 12% SDS-PAGE. The protein concentration was measured using Pierce™ BCA Protein Assay Kits (Thermo Scientific), according to the manufacturer's instructions (Liu et al. 2023).

Determination of β -glucosidase activity

β -glucosidase activity was assayed using *p*-nitrophenyl- β -D-glucopyranoside as substrate (*p*NPG; Sigma), following the procedure described by Jeilu et al. (2024), with slight modification. The reaction mixture consisted of 800 μL of 5 mM *p*NPG prepared in 50 mM potassium phosphate buffer (pH 7.0) and 200 μL purified enzyme. The mixture was incubated at 50 °C for 10 min, after which the reaction was terminated by addition of 100 μL of 1 M sodium carbonate. The release of *p*-nitrophenol (*p*NP) was quantified by measuring absorbance at 405 nm. One unit (U) of β -glucosidase activity was defined as the amount of enzyme required to release 1 μmol of *p*NP per minute under the assay condition.

Enzymatic characterisation of recombinant BGL3_GH1

The enzymatic characterization assays were performed using BGL3_GH1 at a final protein concentration of 0.083 mg mL^{-1} . The optimum pH of the enzyme was examined in different buffers: sodium acetate (pH 4.0–5.0); phosphate buffer (pH 6.0–7.0); Tris–HCl (pH 8.0–9.0); and Glycine–NaOH (pH 10.0–11.0). For pH stability assays, the enzyme was pre-incubated in the same buffer system without substrate for 240 min at room temperature, with aliquots collected every 30 min. Residual activity was quantified to a control sample without incubation.

The optimum temperature was evaluated by assaying the enzyme at temperature ranging from 5 °C to 80 °C in 5 °C increments. Thermostability assays were determined by pre-incubating the enzyme at temperature between 25 °C to 60 °C (5 °C interval) for 240 min. Aliquots were sampled every 30 min, and residual activity was measured to a control sample kept without heat treatment. To evaluate the protective effect of substrate, the enzyme was pre-incubated with *p*NPG at the same concentration as in the standard activity assay at 60 °C for 240 min. The residual activity was measured every 30 min.

The effect of glucose on enzyme activity was assessed by measuring in the presence of glucose concentrations from 0 to 3000 mM. Background control containing the substrate and glucose without enzyme were included to correct for any absorbance interference. Reaction rates were measured during the initial assay period at all glucose concentrations to confirm linearity. Relative activity was calculated by comparing to control sample without glucose. Salt tolerance was determined by assaying enzyme with NaCl concentration from 1 to 10% (w/v). The relative activity was measured by comparing to control without salt.

The influence of various metal ions on enzyme activity was evaluated by including KCl, NiSO_4 , MnCl_2 , CoCl_2 , AlCl_3 , FeCl_2 , ZnCl_2 , MgCl_2 , CaCl_2 , SrCl_2 , and CuCl_2 (final concentration of 5 mM) in the reaction mixture. The relative activity was calculated by comparing to a control without metal ions.

The effects of surfactants and inhibitors were studied by adding 5% (v/v) of chemical additives to the reaction mixture. The chemical additives include dimethyl sulfoxide (DMSO), glycerol, Tween-20, Tween-40, Tween-80, Triton X-100, Ethylenediaminetetraacetic acid (EDTA), Sodium dodecyl sulfate (SDS), and urea. The relative activity was determined by comparison with a control without chemical additives.

The kinetic parameters of the enzyme were determined by measuring enzyme activity at varying *p*NPG concentrations from 0.1 mM to 20 mM. Data were fitted to the

Michaelis–Menten equation using non-linear regression in GraphPad Prism with 95% confidence intervals. Since enzyme kinetics were determined on recombinant BGL3_GH1 that was purified by a single-step purification method, maximum reaction velocity (V_{max}) and turnover number (k_{cat}) were reported as apparent values to account for minor contaminants.

Two-step degradation of sugarcane bagasse

Sugarcane bagasse was collected from a local sugarcane juice vendor. The bagasse was washed thoroughly to remove impurities, and oven-dried at 60 °C to remove moisture. The dried sugarcane bagasse was ground and sieved to obtain a particle size of 1 mm. For enzymatic hydrolysis, 2.5% (w/v) of the processed biomass was suspended in 50 mM potassium phosphate buffer (pH 7.0). Enzymatic hydrolysis was conducted in commercial cellulase only, recombinant BGL3_GH1 only, and combined commercial cellulase and recombinant BGL3_GH1. All reactions were performed at 60 °C with shaking at 150 rpm for a total incubation time of 4 h. For the commercial cellulase only treatment, the biomass suspension was incubated with a commercial cellulase preparation, Cellic® CTec (Novozymes, Denmark) at an enzyme loading of 10 FPU g⁻¹ dry biomass for 4 h. The recombinant BGL3_GH1-only treatment was incubated with recombinant BGL3_GH1 at an enzyme loading of 10 U g⁻¹ dry biomass for 4 h. For the combined enzyme treatment, the biomass suspension was first hydrolyzed with commercial cellulase at 10 FPU g⁻¹ dry biomass for 2 h. The reaction mixture was then centrifuged at 10,000 × g for 5 min, and the supernatant was supplemented with recombinant BGL3_GH1 at 10 U g⁻¹ dry biomass. The reaction mixture was then incubated for an additional 2 h under the same condition. Control reaction without enzyme were included to account for non-enzymatic sugar release. All experiments were performed in triplicate. The concentration of reducing sugar in the final hydrolysate was determined using the 3,5-dinitrosalicylic acid (DNS) method as described by Gusakov et al. (2011). The overall workflow was illustrated in Fig. 2.

Statistical analysis

All experimental assays were performed using three independent biological replicates, with each biological replicates measured in technical triplicate. The data are presented as means ± standard deviation. Statistical significance among different experimental condition (pH, temperature, NaCl concentration, glucose concentration, and enzymatic treatments) was evaluated using one-way ANOVA with Tukey's

post hoc test. Differences were considered statistically significant at $p < 0.05$.

Results

Overview of metagenomics data

A total of 20.9 GB of raw shotgun metagenomic sequencing data in paired-end FASTQ format was obtained. The dataset comprised 67,646,094 pair-end reads. After adapter removal using Trimmomatic, 99.96% of the reads remained as clean data, with a GC content was 54.71%. Assembly of the clean reads using MEGAHIT produced 2,544,877 contigs with an N50 of 542 base pairs (Table 1). To improve reliability of downstream analysis, only contigs ≥ 1,000 bp were retained for ORF prediction and downstream analysis. The contigs length distribution was illustrated in Fig. S1. The filtering step resulted in 173,623 contigs with a total length of 272.9 Mb, and an average contig length of 1,571 base pairs. A total of 367,595 open reading frames (ORFs) were predicted from the filtered contigs. The ORF length distribution demonstrated that approximately 25% of the predicted genes exceed 300 amino acids (Fig. S2). This indicates the presence of long ORFs for functional annotation. Thus, the ORFs were used in subsequent analysis.

Functional analysis of metagenomics dataset

The functional gene profile of the bacterial community was determined by annotating predicted open-reading frames (ORFs) from contigs ≥ 1,000 bp against the KEGG and eggNOG databases. KEGG-based analysis revealed that most functional genes were assigned to the Metabolism category, accounting for 62.89% of the total annotated genes. This was followed by genetic information processing (16.04%) and environmental information processing (15.25%). The cellular processes category comprised 3.94% of the annotated genes, while human disease (1.23%) and organismal systems categories (0.64%) were the least represented (Fig. 3a). Further examination within the metabolism category revealed a diverse functional potential among the bacterial communities. Carbohydrate metabolism was the most abundant pathway (29.00%), followed by energy metabolism (15.93%) and amino acid metabolism (15.81%), respectively. The least represented metabolism pathway included xenobiotics biodegradation and metabolism (1.67%), metabolism of terpenoids and polyketides (1.56%), and biosynthesis of other secondary metabolites (0.38%) (Fig. 3b). Functional annotation using eggNOG exhibited a comparable distribution to the KEGG results (Fig. 3c). Functional categories such as energy production and conversion, amino acid transport and

Fig. 2 Overall workflow of mangrove lignocellulose degrading gene analysis and functional characterisation of GH 1 β -glucosidase (BGL3_GH1)

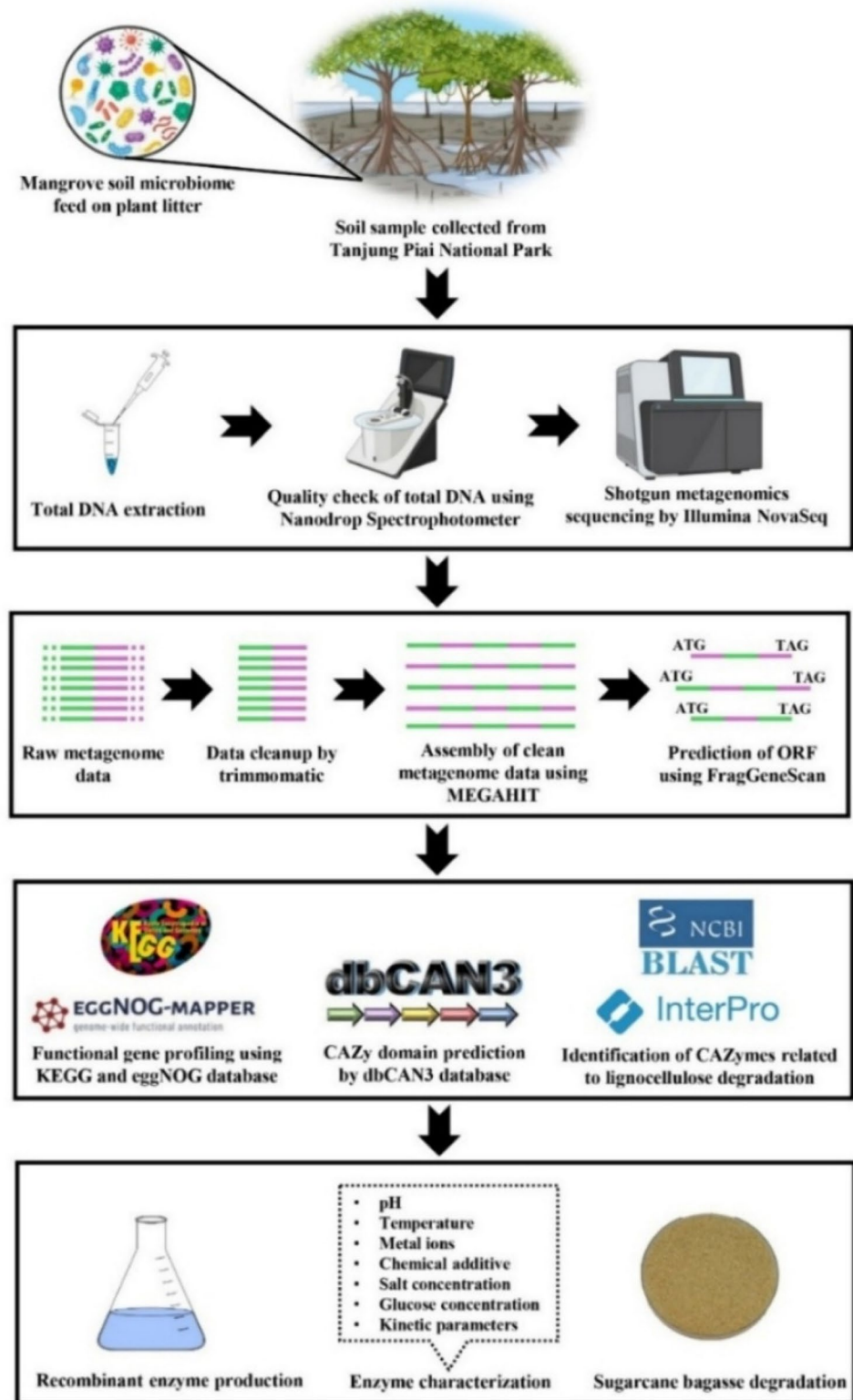


Table 1 Overview of metagenomics data

Parameters	TP mangrove soil sample
Total number of paired-end reads	67,646,094
Total number of clean paired-end reads	67,622,916
Q30 bases of clean reads	93.66%
Q20 bases of clean reads	97.81%
G+C content of clean reads	54.71%
Total number of contigs length ≥ 200 bp	2,544,877
Total assembly size (bp)	1,392,726,341
N50 (bp)	542
Total number of ORFs predicted from contigs $\geq 1,000$ bp	367,595
Total number of functional genes predicted from contigs $\geq 1,000$ bp	183,487
BioProject ID	PRJNA939948
BioSample ID	SAMN33549506
SRA accession number	SRR23652001

conversion, and carbohydrates transport and metabolism were among the most represented, comprising of 12.53%, 10.52%, and 5.15%, respectively.

Mining of carbohydrate degrading genes from metadata

A total of 3,035 CAZymes were identified from the filtered contigs which distributed among all six major classes. Among these CAZy families, GHs were highly represented, accounting for 47.31% of the annotated CAZymes, followed by the GTs at 38.10%. The CBM, CE, and AA families constituted 7.63%, 3.62%, and 2.97% respectively. PLs were the least abundant, comprising with only 0.37%. Further investigation identified 148 CAZymes putatively involved in lignocellulose degradation. These lignocellulolytic genes were distributed among GH (48 ORFs), CE (18 ORFs), and AA (82 ORFs) families (Fig. 4a). Based on their enzymatic functions, these genes were categorized into cellulases (20 genes), hemicellulases (46 genes), lignin-modifying enzymes (82 genes) (Fig. 4b). At the CAZy family level, cellulolytic genes were mainly represented by GH 1, 3, 5, and 6. The hemicellulases were determined with CE (CE 1, 4, 6, 7, and 15) and GH (GH 2, 10, 39, 51, 53, and 67) families. Lignin modification was primarily mediated by AA 2, 3, 4, 6, and 7 (Fig. 4c).

Classification of lignocellulose degrading genes

Analysis of lignocellulolytic genes revealed 20 CAZy families from the GH, CE, and AA classes associated with lignocellulose degradation. The cellulases group included of β -glucosidase (GH 1 and 3), endoglucanase (GH 5), and cellobiohydrolase (GH 6). These cellulases were encoded

by bacterial genera such as *Aurantiacibacter*, *Bacteroides*, *Paraglaciecola*, and *Pseudolabrys* (Table 2).

A total of 11 CAZy families were identified with of hemicellulolytic function, including acetyl xylan esterase (CE 1, 6, 7, and 15), polysaccharide deacetylase (CE 4), α -glucuronidase (GH 2, 67), α -arabinofuranosidase (GH 51), arabinogalactan endo-1,4- β -galactanase (GH 53), β -xylanase (GH 10), and β -xylosidase (GH 39). These genes were found in diverse bacteria genera such as *Aquisphaera*, *Bacteroides*, *Luteitalea*, *Mycobacterium*, *Mycolicibacterium*, *Paenibacillus*, *Pseudolabrys*, *Sphingomonas*, and *Trinickia* (Table 2).

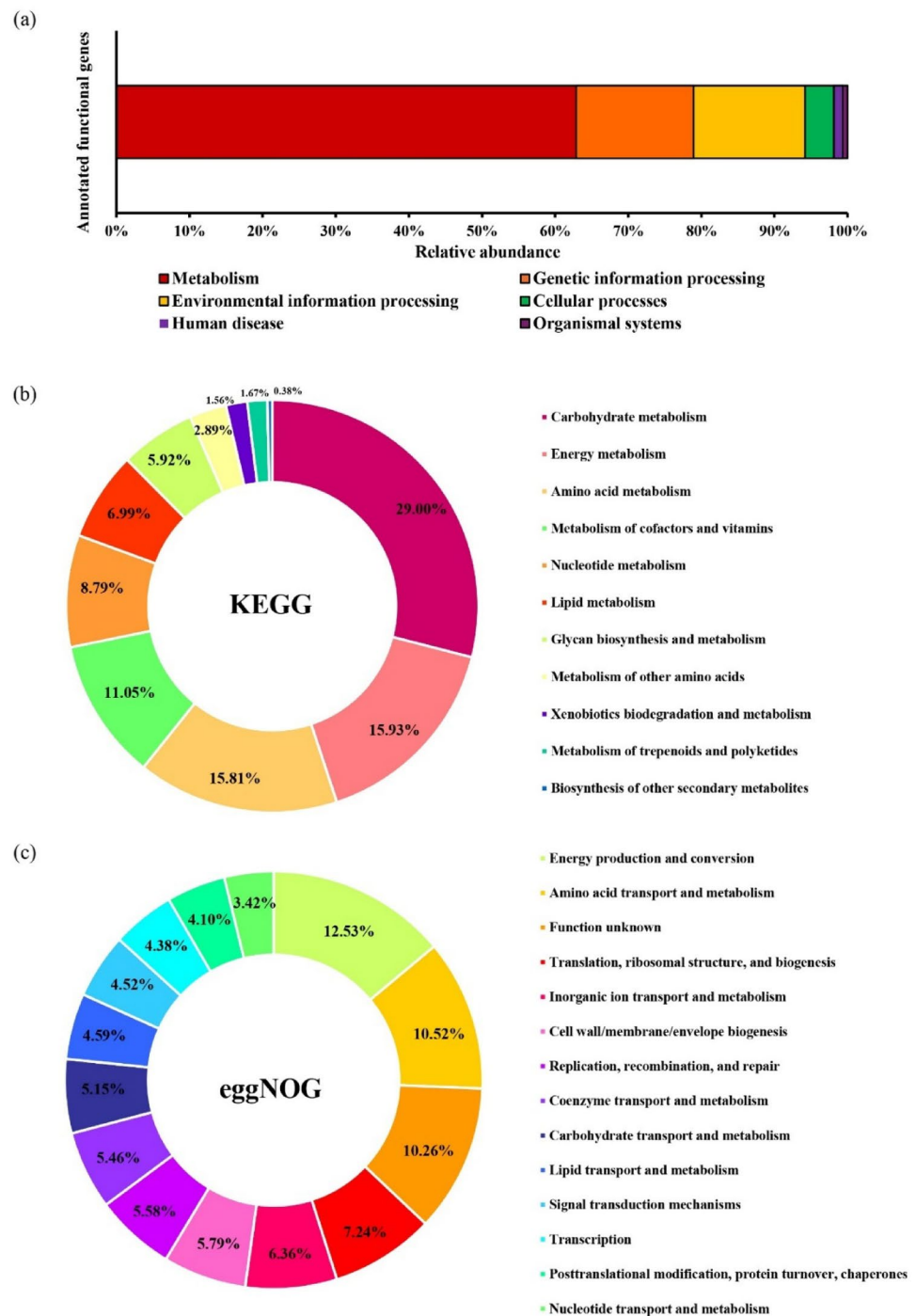
Lignin-modifying enzymes were represented by catalase-peroxidase (AA 2), GMC family oxidoreductase (AA 3), FAD-binding oxidoreductase (AA 4 and 7), and NADPH quinone oxidoreductase (AA 6). These enzymes were encoded by bacteria genera such as *Bauldia*, *Desulfobacca*, *Desulfofustis*, *Filomicrobium*, *Methyloceanibacter*, *Methylobacter*, *Mycobacterium*, *Nitrospira*, *Paraburkholderia*, *Pseudolabrys*, and *Thiohalobacter* (Table 2). These findings highlight the enzymatic profile of bacterial communities for lignocellulose degradation.

In this study, we defined the underexplored bacteria as (i) those that are unable to be assigned into genus level, (ii) genera consisting of limited species (< 5 species) listed in the LPSN database, or (iii) limited studies on protein functional characterization. Among the 148 CAZymes, 89.19% (132 genes) were categorized as underexplored bacteria, while the remaining 10.81% (16 genes) were identified from bacteria lineages that have been extensively studied (Fig. 5). Within the well-studied category, most genes were affiliated with the phylum Pseudomonadota (50.00%, 8 genes), followed by Actinomycetota (25.00%, 4 genes) and Bacteroidota (12.5%, 2 genes). In the underexplored category, Pseudomonadota was also the prominent phylum, accounting for 52.27% (69 genes), followed by Chloroflexota (22.00%, 29 genes), Acidobacteriota (7.58%, 10 genes), Actinomycetota (3.79%, 5 genes), and Bacteroidota (1.52%, 2 genes). The distribution of lignocellulolytic genes between well-studied and underexplored bacteria were illustrated at Fig. 5.

Synergy action of bacterial communities in lignocellulose degradation

Further analysis revealed that lignocellulolytic enzymes from diverse bacterial communities act synergistically to degrade lignocellulose. The enzymatic degradation of cellulose, hemicellulose, and lignin by mangrove bacterial communities were illustrated in Fig. 6. Five different bacteria phyla (Acidobacteriota, Actinomycetota, Bacteroidota, Chloroflexota, and Pseudomonadota) were found to participate in the degradation process. Pseudomonadota dominated

Fig. 3 Functional gene profile of TP mangrove soil microbiome based on ORFs predicted from contigs $\geq 1,000$ bp. **a** Relative abundance of annotated functional genes across major KEGG pathways. **b** Distribution of annotated genes among metabolic sub-pathways within the KEGG metabolism category. **c** Functional classification of metagenomic genes based on eggNOG-mapper annotation showing relative abundances of genes assigned to different functional categories



cellulose degradation, followed by Chloroflexota, Actinomycetota, and Acidobacteriota. Lignocellulolytic genes associated with Bacteroidota phyla were not detected in the cellulose degradation pathway. β -glucosidase (GH 3) were the dominant cellulase across five phyla. Notably, the cellobiohydrolase (GH 6) families were unique to Actinomycetota. The combined action of these cellulases can degrade cellulose into simple sugar, glucose (Fig. 6). Lignin degradation was primarily driven by Pseudomonadota

phylum (54 genes), followed by Chloroflexota (11 genes), Actinomycetota (2 genes), and Bacteroidota (1 gene), while Actinomycetota was not assigned with any AA families. Most lignin-degrading enzymes were discovered in AA 2 (catalase-peroxidase), followed by others auxiliary enzymes from AA 3, 4, 6, and 7. These lignin-modifying enzymes were able to oxidize lignin into monolignols such as sinapyl alcohol, coniferyl alcohol, and *p*-Coumaryl alcohol (Fig. 6).

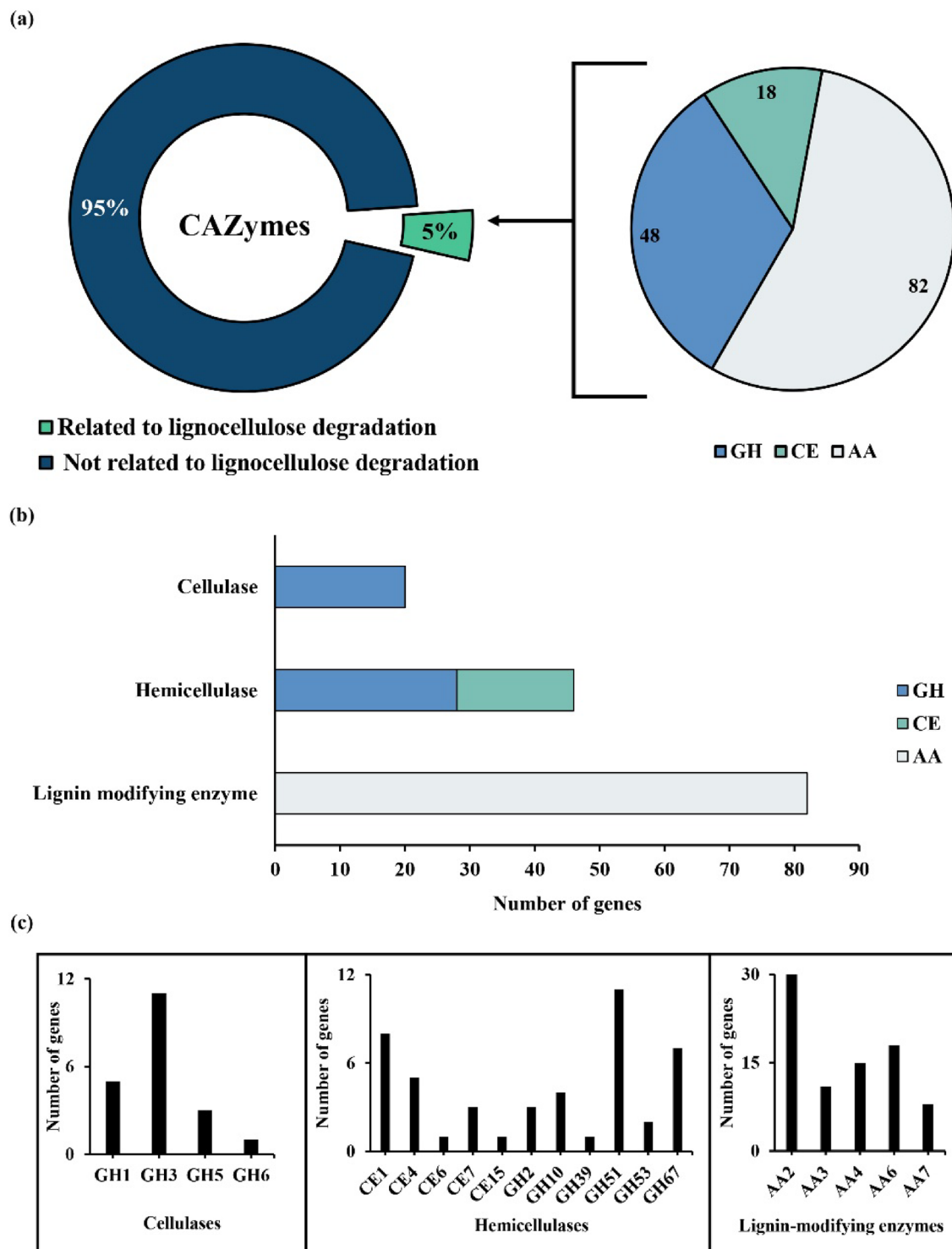


Fig. 4 CAZyme domains annotated in the metagenomic dataset with contigs $\geq 1,000$ bp. *GH* Glycosyl hydrolase, *AA* Auxiliary activity, *CE* Carbohydrate esterase. **a** Lignocellulose-related CAZyme domains identified based on dbCAN and InterProScan annotation; **b** Functional

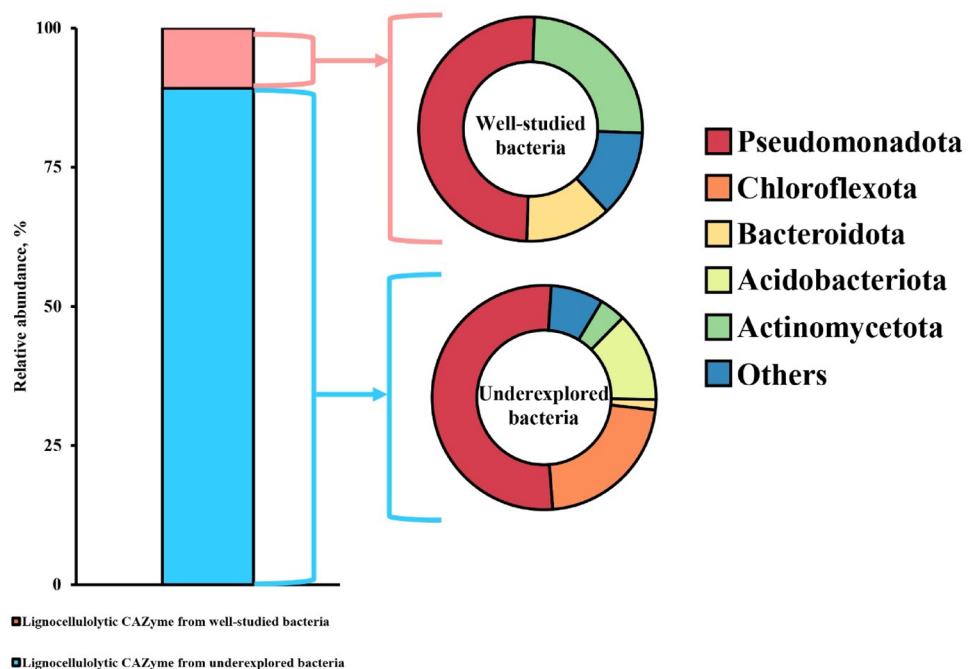
classification of lignocellulolytic CAZymes into cellulase, hemicellulase, and lignin-modifying enzyme; **c** CAZy family-level composition of cellulase, hemicellulase, and lignin-modifying enzyme in contribution to specific GH, CE, and AA families to lignocellulose degradation

Table 2 Taxonomic affiliation of lignocellulose degrading CAZymes identified from contigs $\geq 1,000$ bp based on DIAMOND BLASTP assignment ($\geq 50\%$ sequence identity and $\geq 50\%$ query coverage)

Enzymes	CAZy families	Detected bacteria communities
<i>Cellulases</i>		
β -glucosidase (EC 3.2.1.21)	GH 1, 3	<i>Anaerolineales</i> , <i>Acidimicrobiia</i> , <i>Aurantiacibacter</i> , <i>Bacteroides</i> , <i>Gemmatimonadales</i> , <i>Kiloniellaceae</i> , <i>Pseudolabrys</i> , <i>Streptosporangiales</i>
Endoglucanase (EC 3.2.1.4)	GH 5	<i>Anaerolineae</i> , <i>Deltaproteobacteria</i> , <i>Paraglaciecola</i>
Cellobiohydrolase (EC 3.2.1.91)	GH 6	<i>Streptosporangiales</i>
<i>Hemicellulases</i>		
Acetyl xylan esterase (EC 3.1.1.72)	CE 1, 6, 7, 15	<i>Acidobacteriota</i> , <i>Anaerolineae</i> , <i>Aquisphaera</i> , <i>Gemmatimonadales</i> , <i>Mycobacterium</i> , <i>Mycolicibacterium</i> , <i>Paenibacillus</i> , <i>Planctomycetota</i> , <i>Prolixibacteraceae</i> , <i>Verrucomicrobiales</i>
Polysaccharide deacetylase (EC 3.1.1.58)	CE 4	<i>Hyphomicrobiales</i> , <i>Hyphomicrobiaceae</i> , <i>Pseudolabrys</i>
α -glucuronidase (EC 3.2.1.39)	GH 2, 67	<i>Acidobacteriota</i> , <i>Anaerolineae</i> , <i>Sphingomonas</i> , <i>Thermoflexales</i>
α -arabinofuranosidase (EC 3.2.1.55)	GH 51	<i>Acidobacteriota</i> , <i>Ardenticatenaceae</i> , <i>Bacteroidales</i> , <i>Bacteroides</i> , <i>Chloroflexi</i> , <i>Deltaproteobacteria</i> , <i>Luteitalea</i>
Arabinogalactan endo-1,4- β -galactanase (EC 3.2.1.89)	GH 53	<i>Anaerolineaceae</i> , <i>Trinickia</i>
β -xylanase (EC 3.2.1.8)	GH 10	<i>Anaerolineae</i> , <i>Deltaproteobacteria</i> , <i>Pyrinomonadaceae</i>
β -xylosidase (EC 3.2.1.37)	GH 39	<i>Acidobacteriota</i>
<i>Lignin-modifying enzyme</i>		
Catalase-peroxidase (EC 1.11.1.21)	AA 2	<i>Anaerolineae</i> , <i>Acidithiobacillales</i> , <i>Acidimicrobiales</i> , <i>Bauldia</i> , <i>Deltaproteobacteria</i> , <i>Desulfobacteriales</i> , <i>Desulfofustis</i> , <i>Gammaproteobacteria</i> , <i>Methyloligella</i> , <i>Phycisphaerae</i> , <i>Thiohalobacter</i> , <i>Solirubrobacterales</i>
GMC family oxidoreductase (EC 1.1.99.B3)	AA 3	<i>Betaproteobacteria</i> , <i>Filomicrobium</i> , <i>Gammaproteobacteria</i> , <i>Hyphomicrobiales</i> , <i>Hyphomicrobiaceae</i> , <i>Mycobacterium</i> , <i>Rhodocyclaceae</i>
FAD-binding oxidoreductase (EC 1.1.5.9)	AA 4, 7	<i>Alphaproteobacteria</i> , <i>Anaerolineae</i> , <i>Chloroflexi</i> , <i>Deltaproteobacteria</i> , <i>Desulfobacteriales</i> , <i>Gammaproteobacteria</i> , <i>Gemmatimonadia</i> , <i>Hyphomicrobiales</i> , <i>Hyphomicrobiaceae</i> , <i>Methyloceanibacter</i> , <i>Mycobacterium</i> , <i>Paraburkholderia</i> , <i>Planctomycetota</i> , <i>Pseudolabrys</i>
NADPH quinone oxidoreductase (EC 1.6.5.5)	AA 6	<i>Burkholderiaceae</i> , <i>Deltaproteobacteria</i> , <i>Desulfobacteriales</i> , <i>Desulfobacca</i> , <i>Geminicoccaceae</i> , <i>Methyloceanibacter</i> , <i>Nitrospira</i> , <i>Nitrospirota</i>

GH Glycosyl hydrolase, AA Auxiliary activity, CE Carbohydrate esterase

Fig. 5 Taxonomic resolution of lignocellulolytic CAZymes identified from contigs $\geq 1,000$ bp. The bar chart shows the relative abundance of lignocellulolytic genes in both well-studied and underexplored bacteria lineages based on DIAMOND BLASTP thresholds ($\geq 50\%$ sequence identity and $\geq 50\%$ query coverage); The donut charts depict the taxonomy composition on phyla level of well-studied bacteria (top) and underexplored bacteria (bottom) taxa annotated from the lignocellulolytic CAZymes



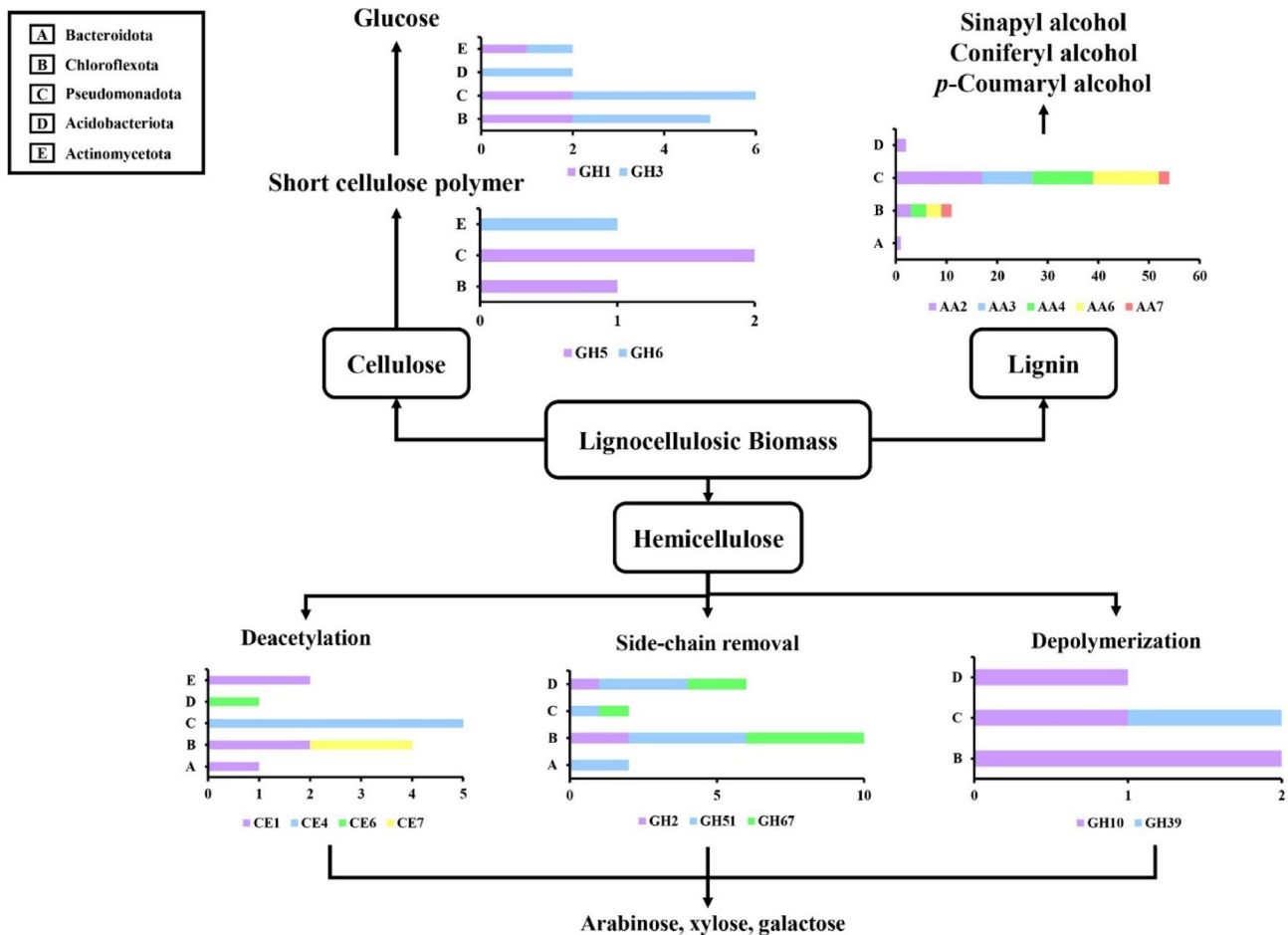


Fig. 6 Proposed synergistic model for lignocellulose degradation by mangrove bacterial phyla. *GH* Glycosyl hydrolase, *AA* Auxiliary activity, *CE* Carbohydrate esterase. The bar chart indicates the distribution

of CAZymes domains across different phyla involved in the degradation of specific lignocellulose polymers. y-axis: Bacterial phylum; x-axis: gene count

Furthermore, CAZymes linked to hemicellulose degradation were detected that involved in deacetylation, side-chain removal, and backbone depolymerization. Majority of the hemicellulase genes were related in removing side-chain and acetyl group from the hemicellulose backbone. Deacetylation was primarily mediated by CE 1, 4, 6, and 7, with CE 4 showing the highest contribution. The degradation of carbohydrate ester was predominant by CE 4 from the *Pseudomonadota* phylum (5 genes) and followed by *Chloroflexota* (4 genes). Notably, five bacterial phyla were detected in the deacetylation of hemicellulose. The side-chain removal was driven by GH 2, 51, and 67, which the predominant phyla was *Chloroflexota* (10 genes) and *Acidobacteriota* (6 genes). Hemicellulases in depolymerization of hemicellulose backbone were associated with *Chloroflexota*, *Pseudomonadota*, and *Acidobacteriota*. The primary enzymes in hemicellulose depolymerization were GH 10 and 39. The sequential degradation of hemicellulose by these hemicellulases will produce monosaccharides such as arabinose, xylose, and galactose. These results revealed the functional

specialization and cooperative roles of mangrove bacterial communities in lignocellulose degradation.

Production and characterisation of recombinant BGL3_GH1

Among the 148 lignocellulolytic CAZymes identified, a β -glucosidase gene belonging to the GH 1 family (designated *BGL3_GH1*) was selected for experimental characterization. The gene was selected due to (i) the gene was determined from the underexplored bacteria group (ii) industrial relevance as β -glucosidase plays an important role in cellulose degradation by catalyzing the terminal hydrolysis of celooligosaccharides into glucose. Thus, *BGL3_GH1* was chosen as the representative enzyme to validate the functional potential derived from the metagenomic analysis. The selected gene was predicted with the presence of a complete GH 1 catalytic domain based on InterProScan analysis. Recombinant BGL3_GH1 protein was successfully expressed and purified using a single-step affinity

chromatography approach. SDS-PAGE analysis on partially purified recombinant BGL3_GH1 protein revealed a dominant band at approximately 50 kDa which corresponds to the expected molecular mass of recombinant BGL3_GH1 (Fig. S3). This indicates that the recombinant BGL3_GH1 was suitable for subsequent biochemical characterization. In silico analysis further confirmed the functional annotation of recombinant BGL3_GH1, including the presence of a complete GH 1 catalytic domain (Fig. S4a), phylogenetic clustering with characterized GH 1 β -glucosidase (Fig.

S4b), and identification of key catalytic and substrate-binding residues of GH 1 enzymes (Fig. S4c).

Recombinant BGL3_GH1 demonstrated the highest activity at pH 7. The relative activity of the enzyme decreased at pH values below 6 and above 9, falling to less than 30% of the maximum activity (Fig. 7a). It remained most stable between pH 7 and 9, retaining around 70% of its residual activity. However, at pH 5 and 10, the enzyme activity decreases sharply to approximately 25% (Fig. 7b). The differences in activity across pH conditions were statistically significant ($p < 0.05$). Temperature profiling revealed

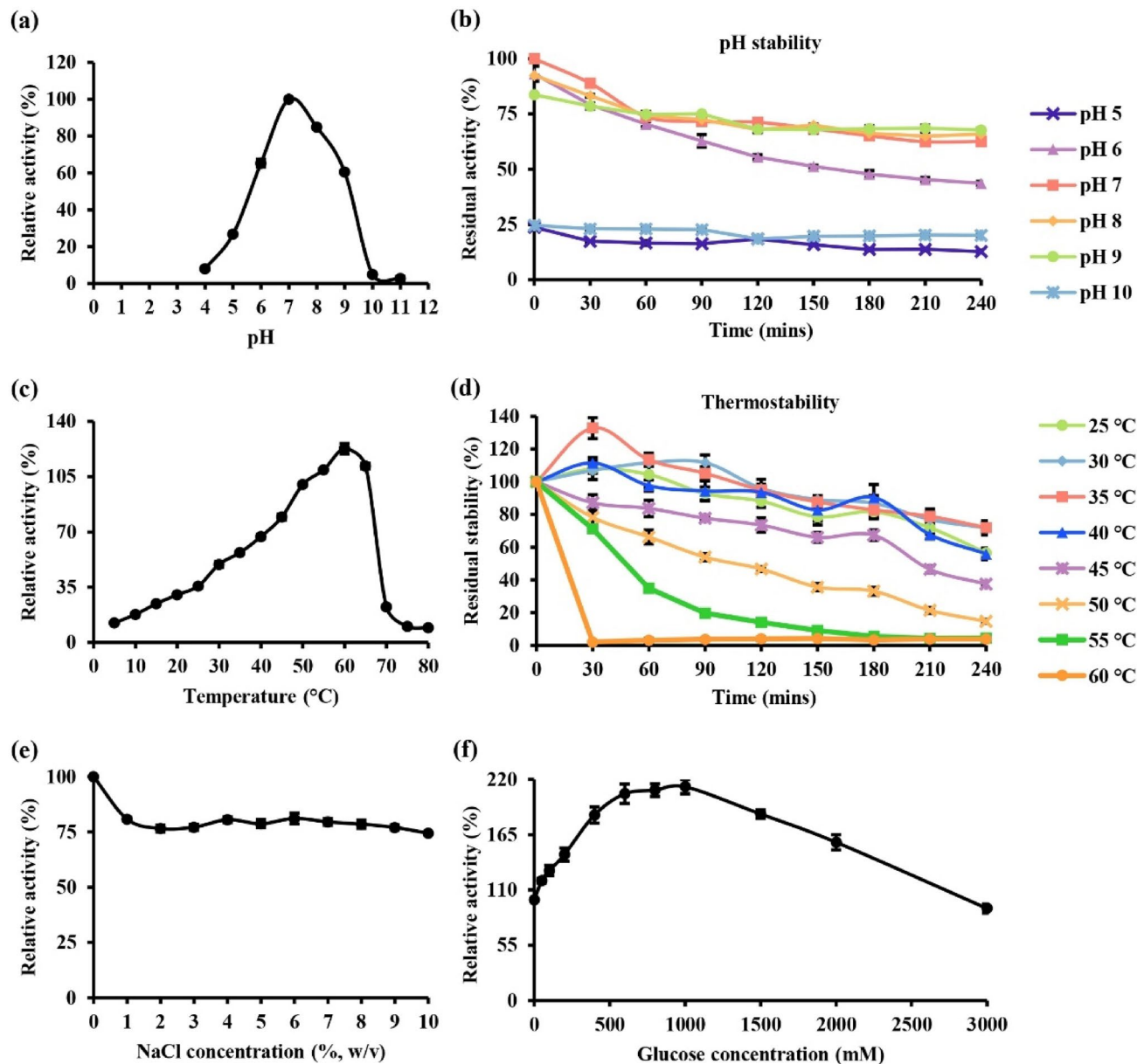


Fig. 7 Functional characterisation of recombinant BGL3_GH1 showing **a** optimal pH; **b** pH stability over a 240-min incubation period; **c** optimal temperature; **d** thermostability of over a 240-min incubation period; **e** effect of NaCl concentrations; **f** effect of glucose concen-

tration. Data were presented as mean $\pm$ standard deviation. Statistical significance among conditions was assessed where applicable using one-way ANOVA ($p < 0.05$)

that high enzymatic activity between 50 °C and 65 °C ($\geq 90\%$), with maximal activity at 60 °C. At temperature above 65 °C, the relative activity declined sharply to below 25% ($p < 0.05$) (Fig. 7c). Thermostability analysis showed that recombinant BGL3_GH1 was most stable at 30 °C and 35 °C, retaining more than 70% residual activity. However, the residual activity decreased when temperature increased, reaching approximately 4% at 60 °C (Fig. 7d). This suggests that elevated temperatures favor catalytic activity but compromise long-term enzyme stability.

Maximal relative activity was observed in the absence of salt. Notably, in the presence of NaCl concentration ranging from 1 to 10% (w/v), recombinant BGL3_GH1 retained over 75% of its maximal activity (Fig. 7e). Recombinant BGL3_GH1 exhibited increased enzymatic activity in the presence of glucose concentrations up to 2,000 mM, showing a 1.57-fold to 2.12-fold increase compared to glucose-free control ($p < 0.05$). The enzyme also tolerated glucose concentrations up to 3,000 mM, retaining approximately 90% of its maximal activity (Fig. 7f). No background absorbance was detected in enzyme-free controls, and reaction rates remained linear for all glucose concentrations tested.

The activity of recombinant BGL3_GH1 was evaluated in the presence of various metal ions and chemical additives (Table 3). Potassium (K^+) and magnesium (Mg^{2+}) have minimal effects on enzyme activity, with more than 90% relative activity retained and no significant different

compared to the control. In contrast, recombinant BGL3_GH1 activity was substantially inhibited by strontium (Sr^{2+}), calcium (Ca^{2+}), zinc (Zn^{2+}), copper (Cu^{2+}), iron (Fe^{2+}), nickel (Ni^{2+}), and aluminum (Al^{3+}), with relative activity reduced to below 30% ($p < 0.05$). Among the chemical additives tested, glycerol and the non-ionic surfactants such as Tween-20, Tween-40, Tween-80, and Triton X-100 enhanced recombinant BGL3_GH1 activity, resulting in approximately 1.6-fold to 2.2-fold increases compared to the control ($p < 0.05$). In contrast, the chelating agent, EDTA and anionic detergent, SDS significantly reduced the enzyme activity to below 20% ($p < 0.05$). The sensitivity of recombinant BGL3_GH1 to metal ions and EDTA suggests that metal-dependent interaction may influence enzyme structure or catalytic efficiency. As the enzyme preparation was obtained using a single-step affinity purification, the kinetic parameters were interpreted as apparent values. Apparent kinetic parameters were estimated by non-linear regression using *p*NPG at concentration ranging from 0.1 mM to 20 mM, with initial reaction rates pooled from three independent biological replicates ($n = 60$ total data points) (Fig. S5). Recombinant BGL3_GH1 exhibited an apparent maximum reaction velocity, V_{max} of $22.93 \mu\text{mol min}^{-1} \text{mg}^{-1}$ and an apparent Michaelis constant, K_m , of 2.16 mM. The apparent turnover number (k_{cat}) was calculated to be 276.2 s^{-1} , resulting in an apparent catalytic efficiency (k_{cat} / K_m) of $127.76 \text{ s}^{-1} \text{mM}^{-1}$.

Table 3 The effects of metal ions and chemical additives on the activity of recombinant BGL3_GH1

Chemicals	Relative activity (%)
Control	100.00 ± 0.01
K^+	99.74 ± 1.23
Mg^{2+}	90.55 ± 1.26
Co^{2+}	70.12 ± 0.62*
Mn^{2+}	51.82 ± 3.77*
Sr^{2+}	43.94 ± 3.75*
Ca^{2+}	42.69 ± 0.17*
Zn^{2+}	30.14 ± 0.92*
Cu^{2+}	25.20 ± 1.21*
Fe^{2+}	20.65 ± 4.10*
Ni^{2+}	19.55 ± 2.38*
Al^{3+}	12.02 ± 0.49*
Dimethyl sulfoxide	97.61 ± 8.45
Glycerol	222.86 ± 2.77*
Tween-20	177.92 ± 4.95*
Tween-40	170.10 ± 2.11*
Tween-80	142.17 ± 6.05*
Triton X-100	164.90 ± 0.67*
Ethylenediaminetetraacetic acid	18.90 ± 1.95*
Sodium dodecyl sulfate	6.13 ± 0.37*
Urea	80.07 ± 4.86*

Data are presented as mean ± standard deviation

Statistical significance relative to the control was determined using one-way ANOVA followed by Tukey's post hoc test, * $p < 0.05$

Sugarcane bagasse degradation by two-step degradation

The lignocellulolytic potential of recombinant BGL3_GH1 was evaluated using sugarcane bagasse as the substrate in a two-step hydrolysis process, consisting of an initial treatment by commercial cellulases followed by supplementation with BGL3_GH1. Recombinant BGL3_GH1 alone produced a saccharification yield of $0.0354 \pm 0.0041 \text{ g g}^{-1}$, while commercial cellulase alone resulted in a significantly higher saccharification yield of $0.0534 \pm 0.0011 \text{ g g}^{-1}$ ($p < 0.05$). The combined application of commercial cellulase and recombinant BGL3_GH1 produced the highest saccharification yield of $0.0674 \pm 0.0022 \text{ g g}^{-1}$, which was significantly greater than the commercial cellulase alone ($p < 0.05$) (Table 4). These results indicate that supplementation with BGL3_GH1 enhances cellulase-mediated saccharification efficiency, supporting its potential application in biomass hydrolysis.

Table 4 Saccharification yield from sugarcane bagasse degradation under different degradation condition by commercial cellulase and recombinant BGL3_GH1 in a two-step hydrolysis process

Degradation condition	Saccharification yield (g g ⁻¹ biomass)
Recombinant BGL3_GH1 only	0.0354 ± 0.0041
Commercial cellulase only	0.0534 ± 0.0011*
Commercial cellulase + recombinant BGL3_GH1	0.0674 ± 0.0022**

Data are presented as mean ± standard deviation

Asterisks * and ** indicate significant differences compared with recombinant BGL3_GH1 alone and commercial cellulase alone, respectively, using one-way ANOVA followed by Tukey's post hoc test, * $p < 0.05$

Discussion

Mangroves are intertidal forests that spread along the coastlines with periodic tidal flooding driving shifts in environmental conditions. This influences the mangrove bacterial communities and shaping both their taxonomic composition and functional capabilities (Imchen and Kumavath 2020). Previous culture-dependent methods have demonstrated that mangrove microbes possess diverse metabolic traits, including the ability to degrade lignocellulosic polymers through the production of cellulase, hemicellulase, and lignin modifying enzymes (Das et al. 2023; Lam et al. 2020; Pramono et al. 2021). However, majority of bacterial diversity and its functional potential remain inaccessible through conventional cultivation techniques (Bodor et al. 2020). In this study, a metagenomic approach was employed to investigate the functional gene repertoire and lignocellulolytic potential of mangrove soil microbiome from Tanjung Piai National Park, Johor.

Identifying the functional traits of bacterial communities in an environment is crucial for understanding ecosystem processes and predicting microbial responses to environmental changes. In this study, the functional gene profile of mangrove bacterial was determined using the KEGG, eggNOG, and CAZy databases. Functional annotation revealed that a significant number of ORFs were assigned to key metabolic pathways, particularly in carbohydrates metabolism. These functional genes primarily include housekeeping genes essential for maintaining bacterial life in the mangrove environment. This result aligns with other mangrove metagenomic studies from China and India, which also reported a high abundance of functional genes associated with carbohydrate metabolism (Ghosh et al. 2022; Priya et al. 2018; Zhao et al. 2019). The prominence of carbohydrate metabolism in the mangrove microbiome is likely due to the abundance of plant litter in the mangrove environment. Mangrove plant litters is rich in carbohydrates such as arabinose, galactose, glucose, and xylose, primarily in the form

of cellulose and hemicellulose (Mujtaba et al. 2023). As a result, mangrove ecosystems favor the growth of lignocellulose degrading bacteria as a natural means for processing the abundance organic matter, in order to maintain the health of the ecosystems (Kathiresan 2019).

The carbohydrate utilization potential of bacterial communities was further assessed by annotating the predicted ORFs using the CAZy database. This provides detailed insights into the carbohydrate metabolism of the bacterial communities. A total of 3,035 CAZy genes were identified and distributed across six different CAZy families. These annotated enzymes were associated with the degradation of various carbohydrate substrates, including alginate, carrageenan, chitin, pectin, starch, and lignocellulose. The identification of these CAZy families highlights the metabolic potential of the mangrove bacterial communities in processing a diverse range of complex carbohydrate (Zhang et al. 2023). As plant litters in mangrove are mainly made up of lignocellulose components comprises of cellulose, hemicellulose, and lignin (Krishna and Mohan 2017). In this study, 148 lignocellulolytic genes encoding cellulases, hemicellulases, and lignin-modifying enzymes from various GH, CE, and AA families were determined from the metagenomic dataset. The proportion of lignocellulolytic genes in our dataset (approximately 5%) and the dominance of GH and GT families are consistent with those reported in other mangrove soil metagenome studies, which also reported CAZy profiles enriched in GH and GT families (Wang et al. 2025; Yahaya et al. 2024). Despite the total CAZy gene count in this study was relatively lower than that reported in previous mangrove metagenomes, the comparable distribution of CAZy classes suggests that the mangrove bacterial communities retain a strong functional potential for plant litter degradation. Taxonomic assignment revealed that majority of the lignocellulolytic CAZy genes originated from phylum Pseudomonadota, Chloroflexota, Bacteroidota, Acidobacteriota, and Actinomycetota. Notably, out of 148 lignocellulolytic genes, 89.19% were derived from underexplored bacteria taxa. These included genera such as *Anaerolinea*, *Bauldia*, *Luteitalea*, and *Pseudolabrys*. The predominance of lignocellulose degrading genes from the underexplored bacteria indicates a valuable genetic resource for further biotechnological exploration.

The lignocellulolytic gene pools identified in this study reflects a coordinated enzymatic framework for lignocellulose biomass degradation. A set of cellulases, including β -glucosidase (GH 1 and 3), endoglucanase (GH 5), and cellobiohydrolase (GH 6) were detected in this study. These enzymes are essential for the efficient degradation of cellulose as they act on the β -(1,4)-glycosidic bond in cellulose and short cellulose polymer to produce glucose molecules as the final product (Soni et al. 2018). Similar findings

have been reported in other metagenomic studies, which β -glucosidase (GH 3) and endoglucanase (GH 5) families were typically detected as well (Ghose et al. 2024; Paixão et al. 2021; Parab et al. 2024; Priya et al. 2018; Zhang et al. 2023). However, cellobiohydrolase (GH 6) was poorly represented in those studies. Cellobiohydrolase is an important enzyme for degrading cellulose into cellobiose. A recent study on GH 6 cellobiohydrolase from salt marsh bacteria reported that the enzyme retained up to 87% activity in seawater, suggesting potential for application in seawater-based biorefinery (Leadbeater and Bruce 2024). Thus, the newly discovered cellobiohydrolase genes can serve as reference templates for future cloning and enzymatic characterisation.

Among the examined lignocellulolytic enzymes, hemicellulases were represented by a diverse array of GH and CE families. This reflects the structural complexity of hemicellulosic polymers. The GH families are responsible for cleaving α -(1,2)- and β -(1,3;1,4)-glycosidic bond linkages present in hemicellulose side chains and backbones (de Souza and Kawaguti 2021). Furthermore, the CE families hydrolyse ester linkage between acetyl side chain and lignin from hemicellulose backbone (Puchart and Biely 2023). The presence of side branches xylan in mangrove trees contribute to structural integrity and resistance to degradation by natural forces (Alves et al. 2020). This is significant because hemicellulose, particularly acetylated xylan, forms lignin-carbohydrates complexes that act as a protective barrier, preventing cellulose from enzymatic degradation (Curry et al. 2023). The removal of acetyl group and lignin side branches is crucial as these structural features can obstruct the activity of other GH enzymes during cellulose degradation (Nargotra et al. 2023). Notably, the Pseudomonadota phylum was identified as the major degrader for carbohydrate ester, with CE 4 as the key enzyme. In this study, CE 4 was identified affiliated with an underexplored bacteria genus, *Pseudolabrys* from the Pseudomonadota phylum. Currently, only one species was reported under the genus *Pseudolabrys* (Kämpfer et al. 2006). The lignocellulolytic abilities in this genus remain unknown. These findings suggest that *Pseudolabrys* may play an active role in carbohydrate ester degradation and require further exploration.

Among the three components of lignocellulose, lignin is the most recalcitrant polymers due to its complex and phenolic structure (Iram et al. 2021). The decomposition of lignin requires two main groups of lignin-modifying enzymes: ligninolytic enzymes and auxiliary enzymes that aid in the decomposition of lignin. Both enzyme groups were identified in this study and prominently represented by the Pseudomonadota phylum. Within the AA families, auxiliary enzymes from AA 3, 4, 6, and 7 are the main lignin-modifying enzymes, followed by the ligninolytic enzymes AA 2 family. These enzymes facilitate the oxidative reactions that

disrupt the lignin structure and generating reactive intermediate to support subsequent depolymerization process of lignocellulose (Ji et al. 2024). The high abundance of auxiliary enzymes indicate that mangrove bacteria actively participate in preventing repolymerization of phenoxyl radical and producing hydrogen peroxide for ligninolytic enzymes (Zhao et al. 2022). The Bacteroidota phylum that was recorded with the least amounts in lignin-modifying enzymes aligns with previous findings which reported that genes related to ligninolytic function were mostly found in Pseudomonadota, and less common in Bacteroidota (Grgas et al. 2023). Although previous metagenomic studies have highlighted carbohydrate-active enzymes involved in the degradation of hemicellulose, cellulose, and pectin, CAZY genes in the AA families were rarely discussed in their studies (Ghose et al. 2024; Paixão et al. 2021; Priya et al. 2018; Zhang et al. 2023; Zhao et al. 2019). Therefore, these findings provide an important insight into the discovery of new ligninolytic enzymes from mangrove bacteria, imposing further investigation to evaluate their potential for industrial applications.

The discovery of enzymes through functional metagenomics followed by biochemical characterisation plays a crucial role in linking genetic potential to practical industrial applications (Jeilu et al. 2024). While GH 3 β -glucosidase dominated the dataset, BGL3_GH1 from the GH 1 family was selected for functional characterisation because it is originated from the underexplored bacterial lineage. Moreover, GH 1 β -glucosidases are frequently reported to exhibit greater glucose tolerance and reduced end-product inhibition compared to the GH 3 enzymes, highlighting their greater relevance for industrial applications (Huang et al. 2023; Sun et al. 2024). Recombinant BGL3_GH1 exhibited optimal activity at pH 7.0 and 60 °C, with substantial activity retained across neutral to mild alkaline pH. Functional assays demonstrated that recombinant BGL3_GH1 exhibited considerable enzymatic activity and stability within a pH range of 7 to 9, indicating suitability for processes operating at neutral to mild alkaline condition. Notably, its activity was enhanced in the presence of non-ionic surfactants (Tween-20, Tween-40, Tween-80, and Triton X-100) further supports its industrial potential. These surfactants are used during the pre-treatment of lignocellulosic biomass to improve enzyme accessibility (Muñoz et al. 2022).

Recombinant BGL3_GH1 retained high catalytic activity across a wide temperature range, with optimal activity at 60 °C. However, thermostability assays revealed reduced residual activity in pre-incubation above 45 °C in the absence of substrate, highlighting the differences between optimal catalytic temperature and thermal stability. Previous studies have showed that ligand binding can enhance thermal stability of enzyme by increasing its conformational

rigidity and resistance to thermal denaturation (Gonzalez and Sreelatha 2024; Sampson et al. 2021). While the substrate-induced stabilization was not examined in the present study, the observed activity profile suggests that this mechanism may contribute to the optimal activity of recombinant BGL3_GH1 at 60 °C. However, further investigation will be required to assess the ligand-assisted thermostability. Acid and alkaline solution are commonly used as solvent in the pre-treatment of lignocellulosic biomass (Awogbemi and Von Kallon 2022). Upon neutralization, these solutions formed salts as byproducts and require an extra washing step to remove them. In this study, recombinant BGL3_GH1 exhibited tolerance to elevated NaCl concentration and retained high relative activity across the tested range. This characteristic is advantageous for biomass conversion processes that involved salt-generating pre-treatments, as it could reduce the requirement for extensive washing steps prior to enzymatic hydrolysis.

In addition, recombinant BGL3_GH1 also exhibited high tolerance to glucose, maintaining up to 90% activity at glucose concentration up to 3 M. In contrast, other reported glucose-tolerant GH 1 β -glucosidases retained only approximately 60% of their activity at high glucose concentration, such as 3 M for LQBG8 or 1 M for BgAr and ThBG1 (Huang et al. 2023; Sun et al. 2024; Yin et al. 2021). The glucose-enhanced activity observed for BGL3_GH1 is notable, as most GH 1 β -glucosidase are strongly inhibited by glucose. Although the underlying mechanism remains unclear, this behavior may be due to the presence of a secondary glucose-binding site on the enzyme. The binding of glucose on the secondary site may indirectly enhance catalytic activity rather than exerting an inhibitory effect (Yang et al. 2015). Notably, the high glucose tolerance observed in recombinant BGL3_GH1 underscores its potential for application in glucose-rich industrial processes. As the accumulation of glucose often leads to feedback inhibition of β -glucosidase which impaired cellulose hydrolysis (Sengupta et al. 2023). Notably, recombinant BGL3_GH1 exhibits apparent substrate affinity (2.16 mM) and catalytic efficiency ($127.76 \text{ s}^{-1} \text{ mM}^{-1}$), which are comparable to a reported thermophilic GH 1 β -glucosidase (BglAc) with substrate affinity (3.41 mM) and catalytic efficiency ($122.7 \text{ s}^{-1} \text{ mM}^{-1}$) (He et al. 2024). This suggests that recombinant BGL3_GH1 is functionally active at elevated temperatures and capable of hydrolyzing β -glucosidic substrates. The functional relevance of recombinant BGL3_GH1 was further demonstrated using sugarcane bagasse in a two-step saccharification assay. Supplementation with recombinant BGL3_GH1 improved the saccharification efficiency of sugarcane bagasse by 21% compared with commercial cellulase alone, indicating a clear synergistic effect. A similar trend was observed in another study that supplementation with

β -glucosidase resulted in approximately sixfold increase in saccharification yield (Qu et al. 2017). The higher improvement observed in that study is likely due to longer hydrolysis durations and higher enzyme loadings. Future work will focus on optimizing reaction conditions to further enhance saccharification efficiency. Collectively, these findings suggest that recombinant BGL3_GH1 is a promising candidate for various biotechnological applications, particularly in the biofuel industry.

Conclusion

In conclusion, this work provides the functional profile of mangrove soil bacteria, highlighting their synergistic roles in lignocellulose degradation as revealed through a shotgun metagenomic approach. The findings offer valuable insights into the bacterial consortia responsible for the degradation of lignocellulosic compounds in the mangrove ecosystems. A curated set of genes encoding enzymes involved in the degradation of cellulose, hemicellulose, and lignin was identified and bacteria from five different phyla were shown to collaborate in a synergistic network that degrades lignocellulose into sugar molecules. Notably, a large proportion of the lignocellulolytic CAZy genes were from affiliated with underexplored bacteria lineages. Such findings underscore the untapped lignocellulolytic potential of mangrove-associated bacteria. To experimentally validate the metagenomic prediction, a β -glucosidase gene, *BGL3_GH1* derived from an underexplored bacteria was cloned, expressed, and biochemically characterized. The recombinant enzyme exhibited maximum activity at 60 °C and pH 7.0, retained substantial activity at high salinity conditions, enhanced activity in the presence of non-ionic surfactants, and demonstrated tolerance to high glucose concentrations. Supplementation of recombinant BGL3_GH1 in the hydrolysis of sugarcane bagasse improved the saccharification yield, showing its complementary role in biomass saccharification. These favorable properties underscore its potential for industrial application in biomass degradation processes. In short, the identified 148 lignocellulolytic genes represent a valuable genetic resource for future biotechnological exploration.

Accession numbers

The data that supports the findings of this study are available in this article, or from the corresponding author on request. The sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession numbers PRJNA939948. The nucleotide and

protein sequence data of BGL3_GH1 were deposited in the NCBI database under accession number PX852031.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s13205-026-04788-x>.

Acknowledgements This work was supported by the Ministry of Higher Education Malaysia under Fundamental Research Grant Scheme (FRGS/1/2019/WAB13/UTM/02/1) as main sponsor. This work was also supported by Professional Development Research University Grant (Q.J130000.21A2.06E00). The authors are grateful for financial support from the UTM Nexus of Universiti Teknologi Malaysia for scholarship support. The authors would like to acknowledge Johor National Parks Corporation for sampling permit at Tanjung Piai National Park, Johor (CJB G No. 887005).

Author contributions Kheng Loong Chong: conceptualization, data curation, formal analysis, investigation, methodology, writing—original draft. Kok Jun Liew: conceptualization, data curation, software, validation, writing—review and editing, supervision. Chun Shiong Chong: conceptualization, resources, funding acquisition, writing—review and editing, supervision. Faezah Mohd Salleh: writing—review and editing, supervision.

Funding Open access funding provided by The Ministry of Higher Education Malaysia and Universiti Teknologi Malaysia

Data availability The data that supports the findings of this study are available in this article, or from the corresponding author on request. The sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession numbers PRJNA939948. The nucleotide and protein sequence data of BGL3_GH1 were deposited in the NCBI database under accession number PX852031.

Declarations

Conflict of interest The authors declare no conflicts of interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alves KJ, da Silva MCP, Cotta SR, Ottoni JR, van Elsas JD, de Oliveira VM, Andreote FD (2020) Mangrove soil as a source for novel xylanase and amylase as determined by cultivation-dependent and cultivation-independent methods. *Braz J Microbiol* 51:217–228. <https://doi.org/10.1007/s42770-019-00162-7>
- Awogbemi O, Von Kallon DV (2022) Pretreatment techniques for agricultural waste. *Case Stud Chem Environ Eng*. <https://doi.org/10.1016/j.csee.2022.100229>
- Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Sonnhammer EL (2004) The Pfam protein families database. *Nucleic Acids Res* 32(suppl_1):D138–D141. <https://doi.org/10.1093/nar/gkh121>
- Bodor A, Bounedjoum N, Vincze GE, Erdeiné Kis Á, Laczi K, Bende G, Rákhely G (2020) Challenges of unculturable bacteria: environmental perspectives. *Rev Environ Sci Biotechnol* 19:1–22. <https://doi.org/10.1007/s11157-020-09522-4>
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bredon M, Dittmer J, Noël C, Moumen B, Bouchon D (2018) Lignocellulose degradation at the holobiont level: teamwork in a key-stone soil invertebrate. *Microbiome* 6:1–19. <https://doi.org/10.1186/s40168-018-0536-y>
- Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12(1):59–60. <https://doi.org/10.1038/nmeth.3176>
- Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J (2021) eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* 38(12):5825–5829. <https://doi.org/10.1093/molbev/msab293>
- Curry TM, Peña MJ, Urbanowicz BR (2023) An update on xylan structure, biosynthesis, and potential commercial applications. *Cell Surf*. <https://doi.org/10.1016/j.tcsu.2023.100101>
- Das S, Sinha N, Sen M, Ghosh D (2023) Isolation and screening of dye degrading lignocellulolytic bacteria from Sundarban Mangrove ecosystem West Bengal India. *J Pure Appl Microbiol*. <https://doi.org/10.22207/JPAM.17.1.59>
- de Souza TS, Kawaguti HY (2021) Cellulases, hemicellulases, and pectinases: applications in the food and beverage industry. *Food Bioprocess Technol* 14(8):1446–1477. <https://doi.org/10.1007/s11947-021-02678-z>
- Ghose M, Parab AS, Manohar CS, Mohanan D, Toraskar A (2024) Unraveling the role of bacterial communities in mangrove habitats under the urban influence, using a next-generation sequencing approach. *J Sea Res* 198:102469. <https://doi.org/10.1016/j.seares.2024.102469>
- Ghosh A, Saha R, Bhadury P (2022) Metagenomic insights into surface water microbial communities of a South Asian mangrove ecosystem. *PeerJ* 10:e13169. <https://doi.org/10.7717/peerj.13169>
- Gonzalez A, Sreelatha A (2024) Utilizing thermal shift assay to probe substrate binding to selenoprotein O. *J vis Exp JoVE*. <https://doi.org/10.3791/67139>
- Grgas D, Rukavina M, Bešlo D, Štefanac T, Crnek V, Šikić T, Landeka Dragičević T (2023) The bacterial degradation of lignin—a review. *Water* 15(7):1272. <https://doi.org/10.3390/w15071272>
- Gunjal AB, Patil NN, Shinde SS (2020) Enzymes in degradation of the lignocellulosic wastes. Springer
- Gusakov AV, Kondratyeva EG, Sinitsyn AP (2011) Comparison of two methods for assaying reducing sugars in the determination of carbohydrate activities. *Int J Anal Chem* 2011(1):283658. <https://doi.org/10.1155/2011/283658>
- He J, Li Y, Sun X, Zuo D, Wang M, Zheng X, Shi P (2024) Characterization of a novel hyperthermophilic GH1 β -glucosidase from *Acidilobus* sp. and its application in the hydrolysis of soybean isoflavone glycosides. *Microorganisms* 12(3):533. <https://doi.org/10.3390/microorganisms12030533>
- Huang Y-Y, Lv Z-H, Zheng H-Z, Zhu Q, Liu M-T, Sang P, Yin Y-R (2023) Characterization of a thermophilic and glucose-tolerant GH1 β -glucosidase from hot springs and its prospective application

- in corn stover degradation. *Front Microbiol* 14:1286682. <https://doi.org/10.3389/fmicb.2023.1286682>
- Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:1–11. <https://doi.org/10.1186/1471-2105-11-119>
- Imchen M, Kumavath R (2020) Shotgun metagenomics reveals a heterogeneous prokaryotic community and a wide array of antibiotic resistance genes in mangrove sediment. *FEMS Microbiol Ecol* 96(10):fiaa173. <https://doi.org/10.1093/femsec/fiaa173>
- Iram A, Berenjian A, Demirci A (2021) A review on the utilization of lignin as a fermentation substrate to produce lignin-modifying enzymes and other value-added products. *Molecules* 26(10):2960. <https://doi.org/10.3390/molecules26102960>
- Jeilu O, Alexanderson E, Johansson E, Simachew A, Gessesse A (2024) A novel GH3- β -glucosidase from soda lake metagenomic libraries with desirable properties for biomass degradation. *Sci Rep* 14(1):10012. <https://doi.org/10.1038/s41598-024-60645-y>
- Ji T, Liaqat F, Khazi MI, Liaqat N, Nawaz MZ, Zhu D (2024) Lignin biotransformation: advances in enzymatic valorization and bio-production strategies. *Ind Crops Prod* 216:118759. <https://doi.org/10.1016/j.indcrop.2024.118759>
- Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, Nuka G (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9):1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>
- Kämpfer P, Young C-C, Arun A, Shen F-T, Jäckel U, Rosselló-Mora R, Rekha P (2006) *Pseudolabrys taiwanensis* gen. nov., sp. Nov., an alphaproteobacterium isolated from soil. *Int J Syst Evol Microbiol* 56(10):2469–2472. <https://doi.org/10.1099/ijs.0.64124-0>
- Kanehisa M, Sato Y, Morishima K (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 428(4):726–731. <https://doi.org/10.1016/j.jmb.2015.11.006>
- Kathiresan K (2019) Salt-tolerant microbes in mangroves: ecological role and bioprospecting potential. In: *Research developments in saline agriculture*, pp 237–255. https://doi.org/10.1007/978-981-13-5832-6_7
- Krishna M, Mohan M (2017) Litter decomposition in forest ecosystems: a review. *Energy Ecol Environ* 2:236–249. <https://doi.org/10.1007/s40974-017-0064-9>
- Lam MQ, Oates NC, Thevarajoo S, Tokiman L, Goh KM, McQueen-Mason SJ, Chong CS (2020) Genomic analysis of a lignocellulose degrading strain from the underexplored genus *Meridianimaribacter*. *Genomics* 112(1):952–960. <https://doi.org/10.1016/j.ygeno.2019.06.011>
- Leadbeater DR, Bruce NC (2024) Functional characterisation of a new halotolerant seawater active glycoside hydrolase family 6 cellobiohydrolase from a salt marsh. *Sci Rep* 14(1):3205. <https://doi.org/10.1038/s41598-024-53886-4>
- Li D, Liu C-M, Luo R, Sadakane K, Lam T-W (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31(10):1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>
- Liew KJ, Liang CH, Lau YT, Yaakop AS, Chan K-G, Shahar S, Goh KM (2022) Thermophiles and carbohydrate-active enzymes (CAZymes) in biofilm microbial consortia that decompose lignocellulosic plant litters at high temperatures. *Sci Rep* 12(1):2850. <https://doi.org/10.1038/s41598-022-06943-9>
- Liu Y-D, Yuan G, An Y-T, Zhu Z-R, Li G (2023) Molecular cloning and characterization of a novel bifunctional cellobiohydrolase/ β -xylosidase from a metagenomic library of mangrove soil. *Enzyme Microb Technol* 162:110141. <https://doi.org/10.1016/j.enzmictec.2022.110141>
- Mai Z, Ye M, Wang Y, Foong SY, Wang L, Sun F, Cheng H (2021) Characteristics of microbial community and function with the succession of mangroves. *Front Microbiol* 12:764974. <https://doi.org/10.3389/fmicb.2021.764974>
- Mujtaba M, Fraceto L, Fazeli M, Mukherjee S, Savassa SM, de Medeiros GA, Vilaplana F (2023) Lignocellulosic biomass from agricultural waste to the circular economy: a review with focus on biofuels, biocomposites and bioplastics. *J Clean Prod*. <https://doi.org/10.1016/j.jclepro.2023.136815>
- Muñoz SS, Balbino TR, Alba EM, Barbosa FG, de Pier FT, de Almeida ALM, Balagurusamy N (2022) Surfactants in biorefineries: role, challenges & perspectives. *Biores Technol* 345:126477. <https://doi.org/10.1016/j.biortech.2021.126477>
- Nargotra P, Sharma V, Lee Y-C, Tsai Y-H, Liu Y-C, Shieh C-J, Kuo C-H (2023) Microbial Lignocellulolytic enzymes for the effective valorization of lignocellulosic biomass: a review. *Catalysts* 13(1):83. <https://doi.org/10.3390/catal13010083>
- Paixão DAA, Tomazetto G, Sodre VR, Gonçalves TA, Uchima CA, Büchli F, Bragatto J (2021) Microbial enrichment and metagenomics analysis identify CAZymes from mangrove sediments with unique properties. *Enzyme Microb Technol* 148:109820. <https://doi.org/10.1016/j.enzmictec.2021.109820>
- Parab A, Ghose M, Jad V, Phakatkar S, Aiswarya K, Manohar C (2024) Comparative metagenomic analysis of bacterial communities and functional profiles in urban-influenced blue carbon ecosystems. *Authorea Preprints*. <https://doi.org/10.22541/au.171799699.91230799/v1>
- Pramono H, Mariana A, Ryandini D, Sudiana E (2021) Diversity of cellulolytic bacteria isolated from coastal mangrove sediment in Logending Beach, Kebumen, Indonesia. *Biodivers J Biol Divers* 22(4):10. <https://doi.org/10.13057/biodiv/d220433>
- Priya G, Lau N-S, Furusawa G, Dinesh B, Foong SY, Amirul A-AA (2018) Metagenomic insights into the phylogenetic and functional profiles of soil microbiome from a managed mangrove in Malaysia. *Agri Gene* 9:5–15. <https://doi.org/10.1016/j.aggene.2018.07.001>
- Priyashantha A, Taufikurrahman T (2020) Mangroves of Sri Lanka: distribution, status and conservation requirements. *Trop Plant Res* 7(3):654–668. <https://doi.org/10.22271/tp.2020.v7.i3.083>
- Puchart V, Biely P (2023) Microbial xylanolytic carbohydrate esterases. *Essays Biochem* 67(3):479–491. <https://doi.org/10.1042/EBC20220129>
- Qu X-S, Hu B-B, Zhu M-J (2017) Enhanced saccharification of cellulose and sugarcane bagasse by *Clostridium thermocellum* cultures with Triton X-100 and β -glucosidase/Cellic® CTec2 supplementation. *RSC Adv* 7(35):21360–21365. <https://doi.org/10.1039/C7RA02477K>
- Rho M, Tang H, Ye Y (2010) FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* 38(20):e191–e191. <http://s://doi.org/10.1093/nar/gkq747>
- Rubio-Portillo E, Martín-Cuadrado A-B, Ramos-Esplá AA, Antón J (2021) Metagenomics unveils *Posidonia oceanica* “Banquettes” as a potential source of novel bioactive compounds and carbohydrate active enzymes (CAZymes). *mSystems* 6(5):e00866–e1821. <https://doi.org/10.1128/mSystems.00866-21>
- Sagar K, Priti K, Gupta KK, Katara A, Yadav A, Chandra H (2024) The prospect of gene exploitation through soil metagenomics. *J Pure Appl Microbiol*. <https://doi.org/10.22207/JPAM.18.4.60>
- Sampson CD, Fàbregas Bellavista C, Stewart MJ, Mulligan C (2021) Thermostability-based binding assays reveal complex interplay of cation, substrate and lipid binding in the bacterial DASS transporter, VcINDY. *Biochem J* 478(21):3847–3867. <https://doi.org/10.1042/BCJ20210061>
- Sengupta S, Datta M, Datta S (2023) β -Glucosidase: Structure, function and industrial applications. In: *Glycoside hydrolases* (pp.

- 97–120). Academic Press. <https://doi.org/10.1016/B978-0-323-91805-3.00004-6>
- Sharma S, Tsai M-L, Sharma V, Sun P-P, Nargotra P, Bajaj BK, Dong C-D (2022) Environment friendly pretreatment approaches for the bioconversion of lignocellulosic biomass into biofuels and value-added products. *Environments* (Basel) 10(1):6. <https://doi.org/10.3390/environments10010006>
- Soni SK, Sharma A, Soni R (2018) Cellulases: role in lignocellulosic biomass utilization. *Cellulases Methods Protoc*. https://doi.org/10.1007/978-1-4939-7877-9_1
- Sun J, Wang W, Hao J (2024) A novel GH1 β -glucosidase from an Arctic bacterium: characterization and secretory expression in *Bacillus subtilis*. *Process Biochem* 140:108–116. <https://doi.org/10.1016/j.procbio.2024.02.016>
- Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol* 38(7):3022–3027. <https://doi.org/10.1093/molbev/msab120>
- Terrapon N, Lombard V, Drula E, Coutinho PM, Henrissat B (2017) The CAZy database/the carbohydrate-active enzyme (CAZy) database: principles and usage guidelines. A practical guide to using glycomics databases, pp 117–131. <https://doi.org/10.1093/nar/gkn663>
- Wang S, Zhao Z, Cheng R, Cui L, Wang J, Rubin-Blum M, Baltar F (2025) Phylogenetically and metabolically diverse active carbon-fixing microbes reside in mangrove sediments. *Microbiome* 13(1):194. <https://doi.org/10.1186/s40168-025-02177-9>
- Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9):1189–1191
- Xiao K, Liang X, Lu H, Li X, Zhang Z, Lu X, Luo W (2022) Adaptation of gut microbiome and host metabolic systems to lignocellulosic degradation in bamboo rats. *ISME J* 16(8):1980–1992. <https://doi.org/10.1038/s41396-022-01247-2>
- Yahaya N, Rehan MM, Hamdan NH, Nasaruddin SM (2024) Metagenomic data of microbiota in mangrove soil from Lukut River, Malaysia. *Data Brief* 53:110155. <https://doi.org/10.1016/j.dib.2024.110155>
- Yang Y, Zhang X, Yin Q, Fang W, Fang Z, Wang X, Xiao Y (2015) A mechanism of glucose tolerance and stimulation of GH1 β -glucosidases. *Sci Rep* 5(1):17296. <https://doi.org/10.1038/srep17296>
- Yin Y-R, Sang P, Xiao M, Xian W-D, Dong Z-Y, Liu L, Li W-J (2021) Expression and characterization of a cold-adapted, salt-and glucose-tolerant GH1 β -glucosidase obtained from *Thermobifida halotolerans* and its use in sugarcane bagasse hydrolysis. *Biomass Convers Biorefinery* 11:1245–1253. <https://doi.org/10.1007/s13399-019-00556-5>
- Yousuf A, Pirozzi D, Sannino F (2020) Fundamentals of lignocellulosic biomass. In *Lignocellulosic biomass to liquid biofuels* (pp. 1–15). Academic Press. <https://doi.org/10.1016/B978-0-12-815936-1.00001-0>
- Zhai R, Hu J, Jin M (2022) Towards efficient enzymatic saccharification of pretreated lignocellulose: Enzyme inhibition by lignin-derived phenolics and recent trends in mitigation strategies. *Biotechnol Adv* 61:108044. <https://doi.org/10.1016/j.biotechadv.2022.108044>
- Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Yin Y (2018) dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 46(W1):W95–W101. <http://doi.org/10.1093/nar/gky418>
- Zhang B, Gao Y, Zhang L, Zhou Y (2021) The plant cell wall: Biosynthesis, construction, and functions. *J Integr Plant Biol* 63(1):251–272. <https://doi.org/10.1111/jipb.13055>
- Zhang Z-F, Liu L-R, Pan Y-P, Pan J, Li M (2023) Long-read assembled metagenomic approaches improve our understanding on metabolic potentials of microbial community in mangrove sediments. *Microbiome* 11(1):188. <https://doi.org/10.1186/s40168-023-01630-x>
- Zhao H, Yan B, Mo S, Nie S, Li Q, Ou Q, Li N (2019) Carbohydrate metabolism genes dominant in a subtropical marine mangrove ecosystem revealed by metagenomics analysis. *J Microbiol* 57:575–586. <https://doi.org/10.1007/s12275-019-8679-5>
- Zhao L, Zhang J, Zhao D, Jia L, Qin B, Cao X, Liu F (2022) Biological degradation of lignin: a critical review on progress and perspectives. *Ind Crops Prod* 188:115715. <https://doi.org/10.1016/j.indcrop.2022.115715>