



# Genomics and cellulolytic, hemicellulolytic, and amylolytic potential of *Iocasia fonsfrigidae* strain SP3-1 for polysaccharide degradation

Sobrone Heng<sup>1</sup>, Sawanee Sutheworapong<sup>2</sup>, Verawat Champreda<sup>3</sup>, Ayaka Uke<sup>4</sup>, Akihiko Kosugi<sup>4</sup>, Patthra Pason<sup>1,5</sup>, Rattiya Waeonukul<sup>1,5</sup>, Ruben Michael Ceballos<sup>6,7</sup>, Khanok Ratanakhanokchai<sup>1,5</sup> and Chakrit Tachaapaikoon<sup>1,5</sup>

<sup>1</sup> School of Bioresources and Technology, King Mongkut's Institute of Technology Thonburi, Bangkok, Thailand

<sup>2</sup> Pilot Plant Development and Training Institute, King Mongkut's Institute of Technology Thonburi, Bangkok, Thailand

<sup>3</sup> National Center for Genetic Engineering and Biotechnology, Thailand Science Park, Klong Luang, Pathumthani, Thailand

<sup>4</sup> Biological Resources and Post-harvest Division, Japan International Research Center for Agricultural Sciences, Ibaraki, Japan

<sup>5</sup> Excellent Center of Enzyme Technology and Microbial Utilization, Pilot Plant Development and Training Institute, King Mongkut's Institute of Technology Thonburi, Bangkok, Thailand

<sup>6</sup> Department of Biological Sciences, University of Arkansas, Fayetteville, AR, United States of America

<sup>7</sup> Arkansas Center for Space & Planetary Sciences, University of Arkansas, Fayetteville, AR, United States of America

## ABSTRACT

**Background.** Cellulolytic, hemicellulolytic, and amylolytic (CHA) enzyme-producing halophiles are understudied. The recently defined taxon *Iocasia fonsfrigidae* consists of one well-described anaerobic bacterial strain: NS-1<sup>T</sup>. Prior to characterization of strain NS-1<sup>T</sup>, an isolate designated *Halocella* sp. SP3-1 was isolated and its genome was published. Based on physiological and genetic comparisons, it was suggested that *Halocella* sp. SP3-1 may be another isolate of *I. fonsfrigidae*. Despite being geographic variants of the same species, data indicate that strain SP3-1 exhibits genetic, genomic, and physiological characteristics that distinguish it from strain NS-1<sup>T</sup>. In this study, we examine the halophilic and alkaliphilic nature of strain SP3-1 and the genetic substrates underlying phenotypic differences between strains SP3-1 and NS-1<sup>T</sup> with focus on sugar metabolism and CHA enzyme expression.

**Methods.** Standard methods in anaerobic cell culture were used to grow strains SP3-1 as well as other comparator species. Morphological characterization was done via electron microscopy and Schaeffer-Fulton staining. Data for sequence comparisons (e.g., 16S rRNA) were retrieved via BLAST and EzBioCloud. Alignments and phylogenetic trees were generated via CLUTAL\_X and neighbor joining functions in MEGA (version 11). Genomes were assembled/annotated via the Prokka annotation pipeline. Clusters of Orthologous Groups (COGs) were defined by eegNOG 4.5. DNA-DNA hybridization calculations were performed by the ANI Calculator web service.

**Results.** Cells of strain SP3-1 are rods. SP3-1 cells grow at NaCl concentrations of 5-30% (w/v). Optimal growth occurs at 37 °C, pH 8.0, and 20% NaCl (w/v). Although phylogenetic analysis based on 16S rRNA gene indicates that strain SP3-1 belongs to the

Submitted 13 April 2022

Accepted 19 September 2022

Published 19 October 2022

Corresponding author  
Chakrit Tachaapaikoon,  
chakrit.tac@kmutt.ac.th

Academic editor  
Pedro Silva

Additional Information and  
Declarations can be found on  
page 21

DOI 10.7717/peerj.14211

© Copyright  
2022 Heng et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

genus *Iocasia* with 99.58% average nucleotide sequence identity to *Iocasia fonsfrigidae* NS-1<sup>T</sup>, strain SP3-1 is uniquely an extreme haloalkaliphile. Moreover, strain SP3-1 ferments D-glucose to acetate, butyrate, carbon dioxide, hydrogen, ethanol, and butanol and will grow on L-arabinose, D-fructose, D-galactose, D-glucose, D-mannose, D-raffinose, D-xylose, cellobiose, lactose, maltose, sucrose, starch, xylan and phosphoric acid swollen cellulose (PASC). D-rhamnose, alginate, and lignin do not serve as suitable culture substrates for strain SP3-1. Thus, the carbon utilization profile of strain SP3-1 differs from that of *I. fonsfrigidae* strain NS-1<sup>T</sup>. Differences between these two strains are also noted in their lipid composition. Genomic data reveal key differences between the genetic profiles of strain SP3-1 and NS-1<sup>T</sup> that likely account for differences in morphology, sugar metabolism, and CHA-enzyme potential. Important to this study, *I. fonsfrigidae* SP3-1 produces and extracellularly secretes CHA enzymes at different levels and composition than type strain NS-1<sup>T</sup>. The high salt tolerance and pH range of SP3-1 makes it an ideal candidate for salt and pH tolerant enzyme discovery.

**Subjects** Biodiversity, Biotechnology, Microbiology, Molecular Biology, Taxonomy

**Keywords** Halophilic alkaliphilic anaerobic bacterium, Cellulolytic enzyme, Hemicellulolytic enzyme, Amylolytic enzyme, Carbohydrate-binding module, *Iocasia fonsfrigidae*

## INTRODUCTION

Starch-based biomass, such as brewery-spent grains, cassava pulp, rice bran, sago pith residues, and wheat bran, is a by-product of agro-industrial and agricultural operations (Hoang & Nghiem, 2021). Starch-based biomass is produced in significant amounts around the world. For example, more than 174.1 MT/year of sugarcane, cassava, rice, and palm are produced in Thailand alone (Jusakulvijit, Bezama & Thrän, 2021). These feedstocks are composed of polysaccharides, including: cellulose, hemicellulose, and starch, which serve as low-price raw materials for bioproducts. Such starch-based biomass is hydrolyzed by cellulolytic, hemicellulolytic, and amylolytic (CHA) enzymes to yield monosaccharides and oligosaccharides (Cheawchanlertfa et al., 2021). Monosaccharides can be converted to bioethanol, organic acids, or other value-added products while oligosaccharides can be used as prebiotics (Phakeenuya et al., 2020).

Anaerobic bacteria are a proven natural source for the identification and isolation of novel CHA enzymes (Cheawchanlertfa et al., 2021). The bioprospecting of extremophiles, including halophilic anaerobic bacteria, has also yielded novel enzymes with unique properties for commercial applications (Kivistö & Karp, 2011). In this study, we examine CHA enzymes derived from the halophilic anaerobic bacterium designated as *Halocella* sp. SP3-1 (Heng et al., 2019), which is renamed as described below.

*Iocasia* was recently proposed as a new genus with *I. fonsfrigidae* NS-1<sup>T</sup> as the archetype (Zhang et al., 2021). This strain was isolated from cold seep sediment of the South China Sea. *I. fonsfrigidae* NS-1<sup>T</sup> was shown to metabolize several carbohydrates, including: starch, xylan, alginate, carboxymethyl cellulose, and a polymer of the aromatic compound lignin (Zhang et al., 2021). *I. fonsfrigidae* NS-1<sup>T</sup> is a moderate halophile, which will readily grow in 1.25–15.0% NaCl.

Prior to reporting the discovery of *I. fonsfrigidiae* NS-1<sup>T</sup>, our laboratory published the complete genome of an isolate originally designated as *Halocella* sp. SP3-1 (Heng et al., 2019), which we identify and rename in the present study as *Iocasia fonsfrigidiae* strain SP3-1. The strain SP3-1 was isolated from the soil of a salt evaporation pond (13°28′37.55″N; 100°7′8.27″E) in the Samut Sakhon province of Thailand (Heng et al., 2019). SP3-1 readily grows on cellulose, hemicellulose, or starch under higher salt content (5–30% NaCl) than strain NS-1<sup>T</sup>. The complete genome sequences for *I. fonsfrigidiae* strains SP3-1 and NS-1<sup>T</sup> were analyzed. The average nucleotide identity (ANI) between the two strains is 97.64% (Zhang et al., 2021), supporting the “same species” determination based on a suggested ANI threshold of 95–96% (Richter & Rosselló-Móra, 2009). Although, we conclude that the SP3-1 isolate is a strain of *I. fonsfrigidiae*, genetic, physiological, and biochemical properties between SP3-1 and NS-1<sup>T</sup> vary. For example, NS-1<sup>T</sup> encodes several genes not found in SP3-1, including genes encoding for proteins related to carbohydrate metabolism, ABC transporters, PTS sugar transporters, type II secretion systems, type I-B and type III-B CRISPR associated proteins, and clusters of proteins related to ethanolamine and propanediol metabolism (Zhang et al., 2021). Conversely, strain SP3-1 exhibits higher salt tolerance and encodes genes not found in strain NS-1<sup>T</sup>. These include genes that code for endo-  $\beta$ -1,4-galactanase, xylan-  $\alpha$ -1,2-glucuronosidase,  $\beta$ -xylosidase,  $\alpha$ -L-arabinofuranosidase, and  $\beta$ -L-arabinofuranosidase (*i.e.*, hemicellulolytic enzymes) and oligo-  $\alpha$ -1,6-glucosidase (amylolytic enzymes). The unique CHA enzymes constituency of the strain SP3-1 proteome motivate this study. Specifically, we distinguish strains SP3-1 and NS-1<sup>T</sup> based on phylogenetics, physiology, and biochemistry with particular interest in differentiating the respective proteomes based on enzymes content. Our analysis demonstrates that *I. fonsfrigidiae* strain SP3-1 expresses a suite of CHA enzymes with potential for optimal functionality under high salinity conditions.

## MATERIAL AND METHODS

### Bacterial strains and media

Strain SP3-1 was isolated from the soil sample (Heng et al., 2019), whereas *Halocella cellulositytica* DSM 7362<sup>T</sup> was purchased from the Leibniz Institute, DSMZ-German Collection of Microorganisms and Cell Cultures GmbH. Strain SP3-1 was deposited at Thailand Institute of Scientific and Technological Research Culture Collection (TISTR) and Korean Collection for Type Cultures (KCTC) under accession numbers TISTR 2992 and KCTC 25333, respectively. Strain SP3-1 was cultured in the basal medium, pH 8.0 (BM) composed of (per liter): 200 g NaCl, 1.5 g KH<sub>2</sub>PO<sub>4</sub>, 2.9 g K<sub>2</sub>HPO<sub>4</sub>, 2.1 g urea, 4.5 g yeast extract, 0.5 g cysteine-HCl, 0.001 g resazurin, and 200  $\mu$ L mineral solution (25.0 g/L MgCl<sub>2</sub>·6H<sub>2</sub>O, 37.5 g/L CaCl<sub>2</sub>·2H<sub>2</sub>O and 0.3 g/L FeSO<sub>4</sub>·6H<sub>2</sub>O). For shorter term storage, strain SP3-1 was preserved at –20 °C in liquid media with 25% of glycerol. Two methods were used for longer term storage: storages at –80 °C in liquid media with 25% of glycerol and *via* lyophilization.

Strain *H. cellulositytica* DSM7362<sup>T</sup> was cultured in DSMZ medium 702 (pH 7.0) ([https://www.dsmz.de/microorganisms/medium/pdf/DSMZ\\_Medium702.pdf](https://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium702.pdf)). Both media

were anaerobically prepared in bottles sealed with butyl rubber stoppers under an atmosphere of high-purity N<sub>2</sub> and sterilized by autoclaving at 121 °C for 15 min. Phosphoric acid swollen cellulose (PASC), prepared from Avicel PH-101, as previously described by [Zhang et al. \(2006\)](#), xylan, and starch were used as the sole carbon source to observe the ability of bacteria to degrade cellulose, hemicellulose (*i.e.*, xylan), and starch, respectively.

### 16S rRNA gene analysis and phylogenetic tree

Genomic DNA was extracted as previously described by [Heng et al. \(2019\)](#). The 16S rRNA gene was amplified *via* PCR with the universal primers 8F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3') ([Chimtung et al., 2014](#)). The PCR amplification was performed with 1 µL of a DNA template, followed by 5 µL of 10X Ex-Taq buffer, 1 µL of 10 mM dNTPs, 1 µL of each 10 µM primer, and 0.25 µL of Ex-Taq DNA polymerase (Promega Corp., Madison, WI, USA). PCR conditions consisted of an initial denaturation step at 95 °C for 30 s, followed by 30 cycles at 95 °C for 30 s, annealing at 68 °C for 30 s, and extension at 68 °C for 1 min. The final extension step was 10 min at 68 °C. PCR product was purified by a QIAquick PCR purification kit (QIAGEN, Hilden, Germany). The nearly complete 16S rRNA gene sequence was compiled with the BioEdit software ([Hall, Biosciences & Carlsbad, 2011](#)). The 16S rRNA gene sequence of strain SP3-1 was compared to taxa which were retrieved through BLAST ([Altschul et al., 1997](#)) and EzBioCloud databases ([Yoon et al., 2017](#)). The 16S rRNA sequence of strain SP3-1 and correlated taxa were aligned by *via* CLUTAL\_X in MEGA software version 11 and phylogenetic trees were constructed using the neighbor-joining (NJ) method ([Tamura, Stecher & Kumar, 2021](#)). Confidence values for phylogenetic tree branches were determined *via* bootstrap analyses on 1,000 replicates ([Rzhetsky & Nei, 1992](#)).

### Physiological and biochemical analysis

Cells of strain SP3-1 were harvested during late exponential growth phase (3 days). Cell cultures were filtered through a nucleopore membrane filter (0.6 µm pore size), dehydrated by a series of graded ethanol solutions, and dried to the critical point with liquid CO<sub>2</sub> ([Pason, Kyu & Ratanakhanokchai, 2006](#)). Samples were then examined under a scanning electron microscope (SEM, SU800 Hitachi, Japan) to identify morphological features, such as flagella ([Belas & Colwell, 1982](#)). Bacterial motility was assessed by colony growth characteristics on semi-solid media (0.4% agar plates) *via* microscopy. Gram staining was performed as previously described ([Manafi & Kneifel, 1990](#)) and spore formation was examined using the Schaeffer-Fulton staining method ([Oktari et al., 2017](#)). Physiological characterization of strain SP3-1 was performed using standard protocols commonly used in bacterial systems ([Zhang et al., 2021](#); [Doetsch, 1981](#)). For all experiments, cells were incubated at various temperatures under static conditions in triplicate. The growth range and optima of strain SP3-1 were monitored *via* absorbance spectroscopy (*e.g.*, optical density readings) at 600 nm in anaerobic Hungate tubes under various conditions, including a range of: temperatures 25–70 °C (25, 37, 40, 45, 50, 55, 60, 65, 70 °C); pH of 5.0–10.0 (pH 5.0, 6.0, 7.0, 8.0, 9.0, 10.0); and, NaCl concentrations of 0–30% (w/v) (0, 2.5, 5, 7.5, 10, 15, 20 and 30%).

System pH was controlled by varying buffer systems. For pH 5.0–6.5, an acetate buffer was used. A MOPS buffer was used for pH 6.5–8.0. A TABS buffer was used for pH 8.0–9.0. A CHES buffer was used for pH 9.0–10.0. Optical density (OD<sub>600nm</sub>) readings were taken every 24 hrs for three days until achieving the late exponential phase. Fermentation products were identified *via* a gas chromatograph equipped with a flame ionization detector and a Carbowax 20M column (GC-14A; Shimadzu, Japan). The column, injector, and detector temperatures were 170 °C, 230 °C, and 230 °C, respectively. Carbon source use was tested by growing strain SP3-1 in BM containing 0.5% (w/v) of one the following: L-arabinose, D-fructose, D-galactose, D-glucose, D-mannose, D-raffinose, D-rhamnose, D-xylose, cellobiose, lactose, maltose, sucrose, PASC, xylan, starch, alginate, or lignin.

### Chemotaxonomic analysis

The cell wall peptidoglycan was determined as described by [Komagata & Suzuki \(1987\)](#). Cellular fatty acids were extracted, methylated, and analyzed using the standard microbial identification system protocol (Sherlock Microbial Identification System, version 6.1), whereas fatty acids were identified using the TSBA6 database of the microbial identification system ([Sasser, 1990](#)). Polar lipids were analyzed from freeze-dried cells by two-dimensional thin-layer chromatography (TLC), as described by [Minnikin et al. \(1984\)](#). Appropriate detection reagents were used for visualizing TLC bands: phosphomolybdic acid reagent 5% (w/v) solution in ethanol (Sigma-Aldrich, Saint Louis, MO, USA) was used to detect total polar lipids; ninhydrin reagent (0.2% solution; Sigma-Aldrich Saint Louis, MO, USA) was used to detect amino lipids; the Dittmer and Lester reagent (molybdenum blue, 1.3%; Sigma-Aldrich Saint Louis, MO, USA) was used to detect phospholipids; and, Dragendorff's reagent (Sigma-Aldrich Saint Louis, MO, USA) was used to detect phosphatidylcholine.

### Cultivation and enzyme production

Strain SP3-1 was cultivated in 1 L of BM containing 0.5% (w/v) PASC for 3 days at 37 °C, pH 8.0 under static conditions in an anaerobic chamber (Bactron II, USA). The culture supernatant was collected by centrifugation at 10,000×g for 15 min at 4 °C. Culture supernatant was subsequently concentrated using a hollow fiber cartridge with a 10 kDa cutoff membrane (GE Healthcare, USA). The retentate (approximately 40-times concentrated) was then used as the crude enzyme.

### Enzyme assays and protein determination of crude enzyme of *I. fonsfrigidae* strain SP3-1

Enzyme activity was determined using 50 µL of enzyme (containing 250 µg protein) mixed with 50 µL of substrate in a 50 mM sodium phosphate buffer (pH 7.0) incubated at 50 °C for 15 min. Enzymatic activity on 1% (w/v) PASC, birchwood xylan, or soluble starch were assayed by determining the amount of reducing sugar by the DNS method ([Hu et al., 2008](#)). One unit (U) of enzyme activity is defined as the amount of enzyme releasing 1 µmol of reducing sugar in 1 min. The Lowry method was used for measurement of the protein concentration and using bovine serum albumin as a standard ([Lowry et al., 1951](#)).

### Library preparation and genome sequencing of *I. fonsfrigidae* strain SP3-1

Genome sequencing and library preparation were performed as described in [Heng et al. \(2019\)](#). Briefly, strain SP3-1 was cultured at 37 °C in BM containing 1% (w/v) cellobiose and 20% (w/v) NaCl until late exponential growth phase under anaerobic conditions. Cell culture was collected and used for genomic DNA extraction *via* the DNeasy blood and tissue kit (Qiagen, Hilden, Germany). Subsequently, the SMRTbell template prep kit 1.0 (Pacific Biosciences, Menlo Park, CA, USA) was used to construct sequencing libraries. Polymerase reads were trimmed using high-quality regions, with a minimum subread length (500 bp), a minimum polymerase read quality (0.80), and a minimum polymerase read length (100 bp).

### Genome assembly and annotation of *I. fonsfrigidae* strain SP3-1

Genome annotation was conducted with a Prokka annotation pipeline ([Seemann, 2014](#)). The rRNA and tRNA genes were identified with the RNAmmer ([Lagesen et al., 2007](#)) and the Aragorn software ([Laslett & Canback, 2004](#)). Functional classification of protein-coding genes in strain SP3-1 was done by assigning Cluster of Orthologous Groups (COGs) codes to each gene, using eggNOG-Mapper ([Huerta-Cepas et al., 2017](#)) and eggNOG version 4.5 ([Huerta-Cepas et al., 2016](#)). The *in silico* DNA-DNA hybridization (DDH) calculation between *I. fonsfrigidae* strain SP3-1 and *I. fonsfrigidae* NS-1<sup>T</sup> was done using the Genome-to-Genome Distance Calculator (GGDC 3.0) ([Rigden & Fernández, 2022](#)). Online tools predict genome-to-genome distances between pairs of completely or partially sequenced genomes. Average Nucleotide Identity (ANI) was performed by the ANI Calculator web service (<https://www.ezbiocloud.net/tools/ani>) ([Yoon et al., 2017](#)).

### Comparative genome analysis with comparison of CHA enzyme and carbohydrate-binding module (CBM) genes between *I. fonsfrigidae* strain SP3-1 and *I. fonsfrigidae* NS-1<sup>T</sup>

The *I. fonsfrigidae* SP3-1 genome (NCBI GenBank accession number [CP032760](#)) was compared with closely related strain *I. fonsfrigidae* NS-1<sup>T</sup> (NCBI GenBank accession number [CP046640](#)). The polysaccharide-degrading enzyme-related genes of strain SP3-1 were grouped into CHA enzyme-encoding and CBM-encoding families using HMMER `hmmsearch` with Pfam\_Is HMMs (full-length models) to identify complete matches to the family, which were named per the CAZy nomenclature scheme ([Cantarel et al., 2009](#)). All hits with E-values below  $10^{-4}$  were calculated and their sequences were further analyzed. For CHA enzymes and CBM families, which currently do not have the Pfam HMM, representative sequences were selected from the CAZy website per [Warnecke et al. \(2007\)](#). In this case, BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) was used to identify these CHA enzymes, CBM families, and percent nucleotide sequence identity of all genes.

## RESULTS

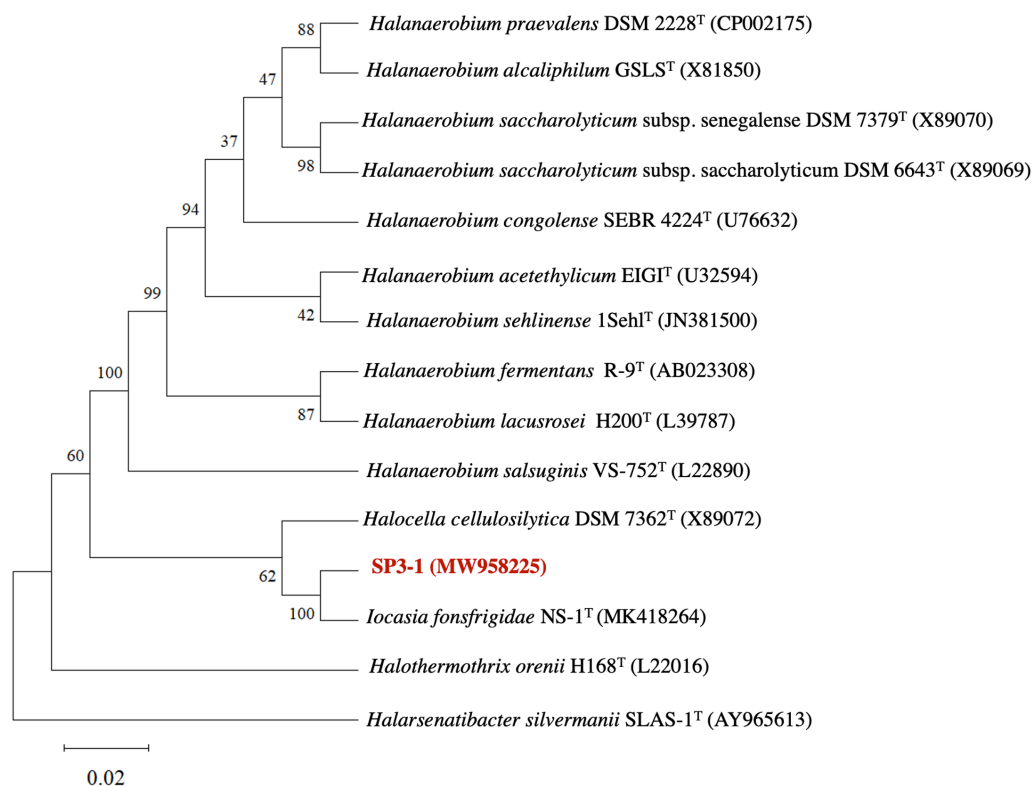
### 16S rRNA analysis of *Halocella* sp. SP3-1 suggests relatedness to *I. fonsfrigidae* NS-1<sup>T</sup>

The genetic profile of *I. fonsfrigidae* isolate SP3-1 was examined by 16S rRNA sequence analysis. Phylogenetic analysis of 16S rRNA indicates SP3-1 falls within genus *Iocasia* of the family Halanaerobiaceae, order Halanaerobiales. *I. fonsfrigidae* NS-1<sup>T</sup> (Zhang et al., 2021), *H. cellulositytica* DSM 7362<sup>T</sup> (Simankova et al., 1993), and *Halothermothrix orenii* H168<sup>T</sup> (Cayol et al., 1994) are all closely related to strain SP3-1 as indicated by 99.58%, 92.77%, and 90.61% nucleotide sequence identity, respectively (Fig. 1).

Analysis of 16S rRNA genes from isolates SP3-1, NS-1<sup>T</sup>, and other species in the family Halanaerobiaceae indicate that isolate SP3-1 may be the same species as strain NS-1<sup>T</sup> based on a 98.65% threshold for species determination (Kim et al., 2014). This agrees with a recent report noting isolates SP3-1 and NS-1<sup>T</sup> as geographic variants of the same species (Zhang et al., 2021). Although 16S rRNA is used to determine relationships between genera, 16S analysis alone is not sufficient for making a same species call in microbiology (Lee et al., 1998).

### Comparative genome analysis of *I. fonsfrigidae* NS-1<sup>T</sup> and isolate SP3-1 suggests they are geographic variants of the same species

Whole genome sequences were used to examine relatedness between isolates SP3-1 and NS-1<sup>T</sup>. Whole genomic sequence analysis shows 97.64% of ANI, which corresponds to a previously reported value (Zhang et al., 2021). DNA-DNA hybridization (DDH) was also conducted to further explore the relatedness between isolate SP3-1 and NS-1<sup>T</sup>. The *in-silico* DDH values between NS-1<sup>T</sup> and isolate SP3-1 (79.9%) exceeded the accepted threshold of 70% for a distinct species call. Based on sequence similarity (Kim et al., 2014) and DDH (Qin et al., 2014), our results also indicate that NS-1<sup>T</sup> and SP3-1 are the same species. Further data support these results. Specifically, *I. fonsfrigidae* NS-1<sup>T</sup> was reported to have a 3,926,493 base pair (bp) genome and a total G+C content of 35.72 mol% (Zhang et al., 2021). A complete, gapless, circular genome assembly was generated for isolate SP3-1 yielding a 4,035,760 bp genome (Fig. S1, Table S1) with a total G+C content of 35.1 mol% (Heng et al., 2019). Predictions from annotated genome assemblies suggest that strain NS-1<sup>T</sup> encodes 3,671 proteins and SP3-1 encodes 3,729 proteins. Results from an extensive genome comparison between isolates SP3-1 and NS-1<sup>T</sup> reveal nucleotide sequence identities between homologous genes at: 20–30% for 200 genes; 40–50% for 122 genes; 60–70% for 76 genes; 80–90% for 2,477 genes; and 100% for 470 genes (see Table 1 and Table S2). No plasmids were detected and the origin of duplication in isolate SP3-1 was determined based on GC skew analysis. Despite the determination that isolate SP3-1 and strain NS-1<sup>T</sup> both fall under the taxon *I. fonsfrigidae*, significant differences exist beyond their genomic profiles.



**Figure 1** Phylogenetic tree of *I. fonsfrigidae* strain SP3-1, with other members in the genera *Halanaerobium*, *Halocella*, *Halothermothrix*, *Halarsenatibacter*, and *Iocasia*. The phylogenetic tree is based on comparing a nearly complete 16S rRNA g.

Full-size DOI: 10.7717/peerj.14211/fig-1

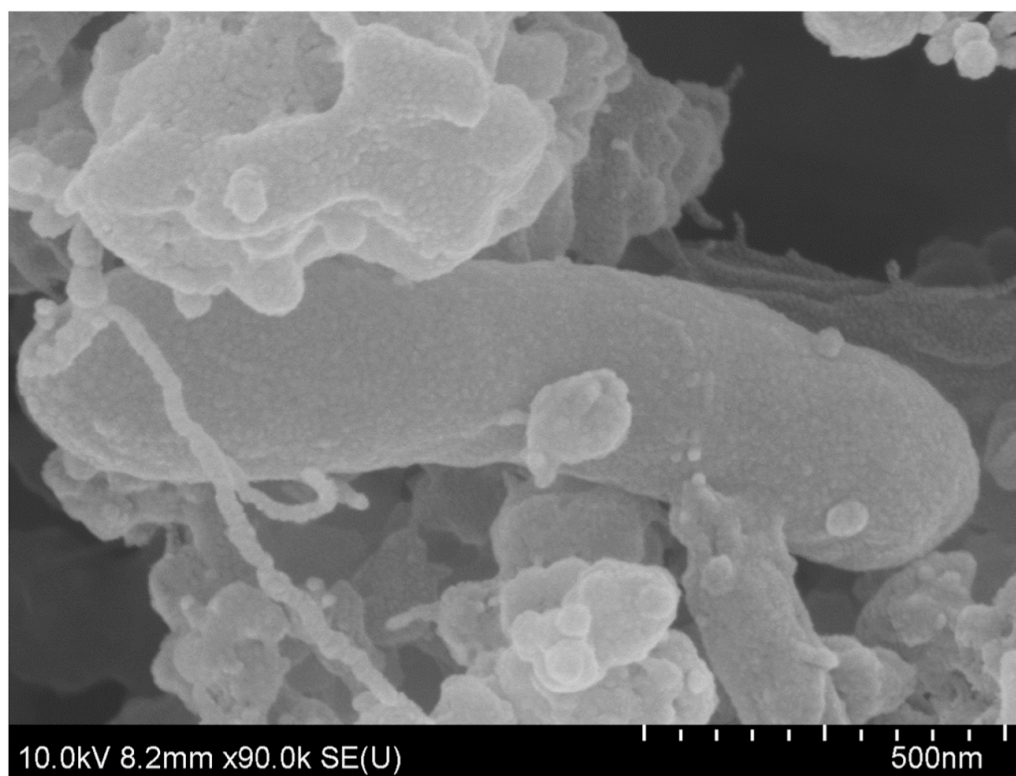
### Cell morphology, physiology, and biochemistry reveal differences between *I. fonsfrigidae* NS-1<sup>T</sup> and isolate SP3-1 suggesting that they are distinct strains of the same species

On solid media, isolate SP3-1 forms small, white colonies with a well-bounded smooth surface. SEM of SP3-1 reveals rod-shaped cells approximately 0.4  $\mu\text{m}$  dia.  $\times$  1.3  $\mu\text{m}$  in length (Fig. 2). In contrast electron microscopy of strain NS-1<sup>T</sup> reveals longer rod-shape cells approximately 0.2–0.3  $\mu\text{m}$  dia.  $\times$  6.0–10.0  $\mu\text{m}$  in length. Strain NS-1<sup>T</sup> also exhibits multiple, long (5–10  $\mu\text{m}$ ) flagella extending from one end (*i.e.*, unipolar) of the major axis of the cell (Zhang *et al.*, 2021). Both *I. fonsfrigidae* NS-1<sup>T</sup> and SP3-1 are Gram-negative and do not form non-endospores. When cultured in BM with 1% (w/v) cellobiose as a carbon source but with different conditions of pH (5.0–10.0) and temperature (25–70  $^{\circ}\text{C}$ ), isolate SP3-1 exhibits a wider range of viable growth compared to strain NS-1<sup>T</sup> (see Table 2). Although optimal growth of isolate SP3-1 occurs at pH 8.0, 37  $^{\circ}\text{C}$ , it readily grows at pH 10.0, confirming that it is a mesophilic alkaliphile. Beyond growth optima, the viable temperature range for isolate SP3-1 is 25–55  $^{\circ}\text{C}$  and the viable pH range is 5.5–10.0 (Table 2). Comparatively, *I. fonsfrigidae* NS-1<sup>T</sup> grows between 20–45  $^{\circ}\text{C}$  at pH values between 6.5–8.0 with optima at 37  $^{\circ}\text{C}$  and pH 7.0 (Zhang *et al.*, 2021). Moreover, SP3-1 grows on all of the same carbon sources as strain NS-1<sup>T</sup>, except D-rhamnose and alginate, and



**Table 1** All genes indicated high similarity and coverage between strain SP3-1 and NS-1<sup>T</sup>. The percentage identity of genes >70% indicated high similarity and high coverage.

Strain		SP3-1	NS-1 <sup>T</sup>
<b>High similarity and high coverage genes</b>		<b>2933</b>	<b>2899</b>
Identity of genes (%)	90–100	2814	2792
	80–90	80	75
	70–80	39	32
<b>The number of different genes</b>		<b>796</b>	<b>772</b>
<b>Total genes</b>		<b>3729</b>	<b>3671</b>



500 nm

**Figure 2** The SEM of *I. fonssfrigidiae* strain SP3-1 grown on a basal medium containing 0.5% (w/v) PASC as the sole carbon source with 20% (w/v) NaCl, scale bar = 500 nm.

[Full-size !\[\]\(e3f8612927870f2e0f9f5989e6dd3064\_img.jpg\) DOI: 10.7717/peerj.14211/fig-2](https://doi.org/10.7717/peerj.14211/fig-2)

lignin. Suitable carbon sources include: L-arabinose, D-fructose, D-galactose, D-glucose, D-mannose, D-raffinose, D-xylose, cellobiose, lactose, maltose, sucrose, xylan, and starch (Table 2). We note that strain SP3-1 also readily grows on PASC.

The main metabolic products of isolate SP3-1 are acetic acid, butyric acid, carbon dioxide, hydrogen, ethanol, and butanol (Table 2). *I. fonssfrigidiae* NS-1<sup>T</sup> produces lactate when grown on glucose (Zhang *et al.*, 2021); however, isolate SP3-1 does not produce

**Table 2** Differentiation of the *I. fonsfrigidae* SP3-1 from related members in the family *Halanaerobiaceae* Strains: 1, *Iocasia fonsfrigidae* SP3-1 (this study); 2, *Iocasia fonsfrigidae* NS-1<sup>T</sup>; 3, *Halocella cellulolytica* DSM7362<sup>T</sup>; 4, *Halothermothrix orenii* H168<sup>T</sup>.

Characteristic	1	2	3	4
Isolation source	Salt evaporation pond	Sediment, cold seep, South China Sea	Hypersaline lagoons of Lake Sivash	Sediment of a Tunisian salted lake
Morphology	Rods	Long rods	Rods	Rods
Cell size (μm)	0.4 × 1.3	0.2–0.3 × 6.0–10.0	0.4–0.6 × 3.8–12.0	0.4–0.6 × 10.0–20.0
Flagella	None	Flagella	Flagella	Flagella
NaCl concentration range (%)	5–30	1.25–15.0	5–20	4–20
Optimum NaCl (%)	20	2.5–7.5	15	5–10
Temperature growth range (°C)	25–55 (opt. 37)	20–45 (opt. 37)	20–50 (opt. 39)	45–68 (opt. 60)
pH growth range	5.5–10.0 (opt. 8.0)	6.5–8.0 (opt. 7.0)	5.5–8.5 (opt. 7.0)	5.5–8.2 (opt. 6.5–7.0)
Carbon sources (0.5%)				
L-Arabinose	+	+	–	+
D-Fructose	+	+	+	+
D-Galactose	+	+	–	+
D-Glucose	+	+	+	+
D-Mannose	+	+	NA	–
D-Raffinose	+	NA	+	+
D-Rhamnose	–	+	NA	NA
D-Xylose	+	+	–	+
Cellobiose	+	NA	+	+
Lactose	+	NA	+	+
Maltose	+	+	NA	–
Sucrose	+	+	+	–
PASC	+	NA	NA	NA
Xylan	+	+	–	NA
Starch	+	+	+	+
Alginate	–	+	NA	NA
Lignin	–	+	NA	NA
Antibiotic sensitivity	Ampicillin (100 μg/mL), ceftriaxone (60 μg/mL), ciprofloxacin (4 μg/mL), clindamycin (10 μg/mL), rifampicin (50 μg/mL), and vancomycin (30 μg/mL),	Ampicillin (100 μg/mL), erythromycin (20 μg/mL) and rifampicin (50 μg/mL)	** Ampicillin (100 μg/mL), amikacin (8 μg/mL), amoxycillin (5 μg/mL), ceftriaxone (60 μg/mL), ciprofloxacin (4 μg/mL), clindamycin (10 μg/mL), kanamycin (100 μg/mL), rifampicin (50 μg/mL), trimethoprim (20 μg/mL), and vancomycin (30 μg/mL)	NA

(continued on next page)

Table 2 (continued)

Characteristic	1	2	3	4
Antibiotic resistance	Amikacin (8 µg/mL), amoxicillin (5 µg/mL), colistin (8 µg/mL), kanamycin (100 µg/mL), and trimethoprim (20 µg/mL)	Kanamycin (100 µg/mL), gentamicin (20 µg/mL), chloramphenicol (20 µg/mL), streptomycin (30 µg/mL), and vancomycin (30 µg/mL)	** Colistin (8 µg/mL)	NA
DNA G+C content (mol%)	35.1	35.72	29.0	39.6
Fermentation products*	Acetate, butyrate, CO <sub>2</sub> , H <sub>2</sub> , ethanol, butanol	Acetate, butyrate, ethanol, lactate, propionate H <sub>2</sub> , CO <sub>2</sub>	Acetate, ethanol, lactate, CO <sub>2</sub> , H <sub>2</sub> ,	Acetate, ethanol, CO <sub>2</sub> , H <sub>2</sub>
Reference	This study	Zhang et al. (2021)	Simankova et al. (1993)	Cayol et al. (1994)

**Notes.**

NA, indicates not available; “+,” growth; “-”, no growth; “\*” used glucose as carbon source, “a\*\*” purchase strain DSM 7362<sup>T</sup> to culture for check the antibiotic sensitivity and resistance; Opt. is optimum.

lactate during growth on glucose. One of the most distinguishing features of isolate SP3-1 is that it readily grows in high salt. Specifically, isolate SP3-1 grows in 5–30% (w/v) NaCl with an optimum at 20% (w/v) NaCl. *I. fonsfrigidae* NS-1<sup>T</sup> grows in a salt concentration range of 1.25–15.0% (Zhang et al., 2021). The ability of isolate SP3-1 to grow in a medium containing up to 30% (w/v) NaCl distinguishes it as more halophilic than other related strains, including *I. fonsfrigidae* NS-1<sup>T</sup>.

In addition to morphological and physiological differences between *I. fonsfrigidae* NS-1<sup>T</sup> and isolate SP3-1 in terms of: size and shape; substrate utilization; and, temperature, pH, and salinity ranges - there are notable differences in antibiotic resistance between the two isolates. SP3-1 is sensitive to ampicillin (100 µg/mL), ceftriaxone (60 µg/mL), ciprofloxacin (4 µg/mL), clindamycin (10 µg/mL), rifampicin (50 µg/mL), and vancomycin (30 µg/mL); however, this isolate is resistant to amikacin (8 µg/mL), amoxicillin (5 µg/mL), colistin (8 µg/mL), kanamycin (100 µg/mL), and trimethoprim (20 µg/mL). This differs from *I. fonsfrigidae* NS-1<sup>T</sup>, which in addition to ampicillin (100 µg/mL) and rifampicin (50 µg/mL) sensitivity, is also sensitive to erythromycin (20 µg/mL). Like isolate SP3-1, *I. fonsfrigidae* NS-1<sup>T</sup> is resistant to kanamycin (100 µg/mL). Unlike isolate SP3-1, *I. fonsfrigidae* NS-1<sup>T</sup> is also resistant to gentamicin (20 µg/mL), chloramphenicol (20 µg/mL), streptomycin (30 µg/mL) and vancomycin (30 µg/mL) (Table 2).

Despite genetic similarity indicating that *I. fonsfrigidae* NS-1<sup>T</sup> and isolate SP3-1 are the same species, the numerous macromorphological, physiological, and biochemical differences suggest that isolate SP3-1 is not merely a geographic variant but indeed a distinct strain of *I. fonsfrigidae*.

### Chemotaxonomic analysis reveals major differences in polar lipid composition between *I. fonsfrigidae* NS-1<sup>T</sup> and isolate SP3-1

To further study differences in strain SP3-1 versus NS-1<sup>T</sup>, chemical profiles were investigated. In general, fatty acid profiles are suitable for differentiating between closely related genera. Indeed, it is suggested that quantitative cellular fatty acid composition differentiate species (Tindall et al., 2010). The cell wall peptidoglycan of strain SP3-1 contains meso-diaminopimelic acid (meso-DAP) as the diagnostic diamino acid. The

cellular fatty acid profile of strain SP3-1 is shown in [Table S3](#). The cellular fatty acid composition of strain SP3-1 reveals: aiC<sub>15:0</sub> (21.5%), nC<sub>14:0</sub> (13.8%), iC<sub>14:0</sub> (12.6%), iC<sub>15:0</sub> (11.7%), and nC<sub>16:0</sub> (11.4%). This differs from the fatty acid composition of many related type strains in the family Halanaerobiaceae. For example, aiC<sub>15:0</sub> is absent in *Halothermothrix orenii* H168<sup>T</sup> ([Cayol et al., 1994](#)) and *Haloanaerobium praevalens* GSL<sup>T</sup> ([Zeikus et al., 1983](#)). However, aiC<sub>15:0</sub> is present in *Halocella cellulositytica* (DSM 7362<sup>T</sup>) but only at 7.6% ([Simankova et al., 1993](#)). Interestingly, strain NS-1<sup>T</sup> features an aiC<sub>15:0</sub> composition of 23.7% ([Zhang et al., 2021](#)), which more closely resembles that of strain SP3-1. Significant variability in predominant cellular fatty acid content among members of the Halanaerobiaceae has been previously reported ([Oren, 2015](#)). In the case of polar lipid profiles, strain SP3-1 contains diphosphatidylglycerol (DPG), phosphatidylethanolamine (PE), phosphatidylglycerol (PG), phosphatidyl-*N*-methyl ethanolamine (PME), phosphatidylcholine (PC), an unidentified amino lipid (*i.e.*, AL1), an unidentified polar lipid (*i.e.*, L1), and unidentified glycolipids (GL2, GL3, and GL4) ([Fig. S2](#)). Unlike strain SP3-1, the composition of polar lipids was found in strain NS-1<sup>T</sup> consists only of diphosphatidylglycerol (DPG), phosphatidylglycerol (PG), an unidentified phosphoglycolipid (*i.e.*, PGL), and two unidentified glycolipids (GL) ([Zhang et al., 2021](#)).

These cellular fatty acids and polar lipid profiles support the suggestion that strain SP3-1 is the same species as NS-1<sup>T</sup> while highlighting major biochemical differences between these two different strains. Significantly different polar lipid profiles between members of related (or the same) taxa often indicated adaptation to different niches and thus different functionality ([Dlugosch et al., 2022](#)).

### Functional genomics analysis of *I. fonsfrigidae* strain SP3-1 indicates a complex carbohydrate metabolism

To elucidate differences in metabolic potential between *I. fonsfrigidae* NS-1<sup>T</sup> and strain SP3-1 fully annotated genomes were analyzed to detect clusters of orthologous genes that drive distinct metabolic functions. From the ~4.0 Mbp genome, 4,044 genes are predicted with 3,875 protein-coding sequences (CDS), 12 rRNA sequences, and 59 tRNA sequences ([Table S1](#)). In comparison, for the ~3.9 Mbp genome of *I. fonsfrigidae* NS-1<sup>T</sup> 3,774 genes were predicted with 3,671 protein-coding sequences (CDS), 12 rRNA, and 58 tRNA sequences ([Zhang et al., 2021](#)). Notably, the NCBI Prokaryotic Genome Annotation Pipeline (version 4.6) was also employed and predicted 3,885 total genes for strain SP3-1 with 3,729 CDS (see [Table S1](#)). These differences are due to the distinct algorithms used in the in-house *versus* the NCBI pipeline. For this study, we used values from the in-house pipeline (*i.e.*, 3,875 CDS) for further analyses.

For strain SP3-1 ~95% (*i.e.*, 3,666 of 3,875) protein-coding sequences were identified as members of COG functional categories ([Tables S1 and S4](#)). Conversely, ~78% (*i.e.*, 2,889 out of 3,671) of the protein-coding sequences of *I. fonsfrigidae* NS-1<sup>T</sup> fall into COG functional categories. Major COG functional categories identified include: amino acid transport and metabolism (strain SP3-1 and strain NS-1<sup>T</sup>; E:238 and E:174, respectively); translation, ribosomal structure, and biogenesis (strain SP3-1 and strain NS-1<sup>T</sup>; J:144 and J:160); replication, recombination, and repair (strain SP3-1 and NS-1<sup>T</sup>; L:195 and

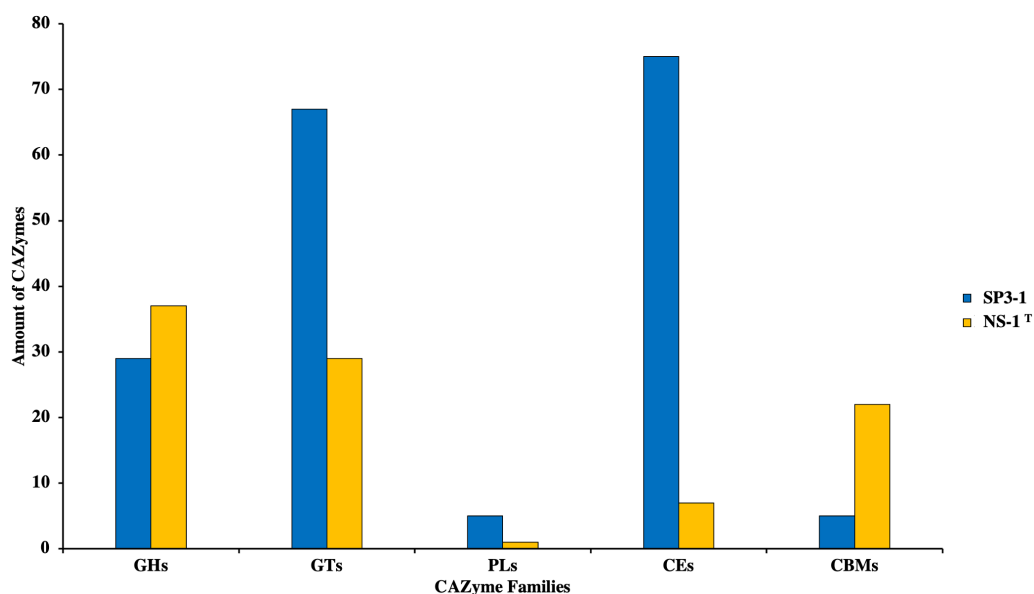
**Table 3** The comparison of genes encoding for cellulolytic, hemicellulolytic, and amyolytic enzymes in the genome of *I. fonsfrigidae* strain SP3-1 and *I. fonsfrigidae* NS-1<sup>T</sup>.

Enzymes	EC number	Strains	
		SP3-1	NS-1 <sup>T</sup>
<b>Cellulolytic enzyme</b>			
$\beta$ -Glucosidase	3.2.1.21	4	2
Endoglucanase	3.2.1.4	2	1
	<b>Total</b>	<b>6</b>	<b>3</b>
<b>Hemicellulolytic enzyme</b>			
Endo- $\beta$ -1,4-galactanase	3.2.1.89	1	0
Xylan 1,4- $\beta$ -xylosidase	3.2.1.37	1	1
$\beta$ -Galactosidase	3.2.1.23	1	1
Xylan- $\alpha$ -1,2-glucuronosidase	3.2.1.131	1	0
$\beta$ -Xylosidase	3.2.1.37	1	0
$\alpha$ -Xylosidase	3.2.1.177	1	1
$\alpha$ -L-Arabinofuranosidase	3.2.1.55	1	0
$\beta$ -L-Arabinofuranosidase	3.2.1.185	1	0
	<b>Total</b>	<b>8</b>	<b>3</b>
<b>Amyolytic enzyme</b>			
$\alpha$ -Amylase	3.2.1.1	1	2
Oligo- $\alpha$ -1,6-glucosidase	3.2.1.10	1	0
Pullulanase	3.2.1.41	1	1
	<b>Total</b>	<b>3</b>	<b>3</b>

L:176, respectively); transcription (strain SP3-1 and strain NS-1<sup>T</sup>; K:268 and K:284); cell wall/ membrane/ envelope biogenesis (strain SP3-1 and strain NS-1<sup>T</sup>; M:201 and M:189, respectively); and, of interest for this study, carbohydrate transport and metabolism (strain SP3-1 and NS-1<sup>T</sup>; G:439 and G: 365, respectively). In addition to differences in carbohydrate metabolism, a remarkable difference in COG calls with unknown function is present (strain SP3-1 and NS-1<sup>T</sup>; S:1,019 and S:538, respectively). These predicted COGs may represent additional, and perhaps novel, proteins related to carbohydrate metabolism since this category dominates the COG profile (Olsen, Choffnes & Mack, 2012). Annotation of the strain SP3-1 genome reveals 17 CHA enzymes (compared to 9 CHA enzymes for strain NS-1<sup>T</sup>), which justifies further investigation of the carbohydrate degradation potential of strain SP3-1 (Table 3). Indeed, many of the unknown functional genes may be involved in polysaccharide deconstruction processes and the transport of sugar products (Olsen, Choffnes & Mack, 2012; Kaznadzey, Shelyakin & Gelfand, 2017).

### ***I. fonsfrigidae* strain SP3-1 encodes for and produces CHA enzymes with a notable range of cellulolytic, hemicellulolytic, and amyolytic activities**

Both the range of carbon sources promoting viable growth of strain SP3-1 as well as the detection of known (and expected) carbohydrate metabolism-related COGs identified in the strain SP3-1 genome justify further examination of the CHA enzymes potential



**Figure 3** Comparative analysis of CAZymes in *I. fonsfrigidae* strain SP3-1 and *I. fonsfrigidae* NS-1<sup>T</sup>. Number of CAZymes in both strains and their distributions among different families. GHs, glycoside hydrolases; GTs, glycosyl transferases; PLs.

Full-size DOI: [10.7717/peerj.14211/fig-3](https://doi.org/10.7717/peerj.14211/fig-3)

of strain SP3-1. To gauge the extent to which strain SP3-1 serves as a suitable prospect for unique carbohydrate deconstruction enzyme discovery, more detailed analyses of genes (and proteins) involved in carbohydrate metabolism was conducted. Because strain SP3-1 grows rapidly on PASC (as well as xylan and starch), a quantitative comparison was performed between strain SP3-1 and an available relative, *H. cellulosilytica* (DSM 7362<sup>T</sup>). *H. cellulosilytica* DSM 7362<sup>T</sup> grows slowly on PASC and starch. It does not grow on xylan. Residual PASC (dry weight) values after 3 days of growth for SP3-1 and *H. cellulosilytica* DSM 7362<sup>T</sup> were 46.53% and 75.73%, respectively (Fig. S3).

Growth on PASC is important since many industrial-use feedstocks are acid pretreated either prior to enzymatic hydrolysis steps of biomass deconstruction. Carbohydrate Active Enzymes (CAZymes) are involved in the breakdown, biosynthesis, and/or modification of glycoconjugates, oligosaccharides, and polysaccharides (Yang et al., 2019). On swollen cellulosic substrates, endoglucanase (and  $\beta$ -glucosidases) are key to effective cellulose substrate deconstruction. Although *I. fonsfrigidae* NS-1<sup>T</sup> was not available for side-by-side quantitative comparison of PASC utilization, it is noted that strain SP3-1 has at least double the number of endoglucanases and  $\beta$ -glucosidases (see Table 3). Indeed, results from advanced COG analysis reveal 181 genes in SP3-1 that are likely involved in carbohydrate breakdown. This includes genes encoding: GHs (glycoside hydrolases, 29), GTs (glycosyl transferases, 67), PLs (polysaccharide lyase, 5), CEs (carbohydrate esterases, 75), and CBMs (carbohydrate-binding modules, 5) (see Fig. 3). CAZymes in *I. fonsfrigidae* NS-1<sup>T</sup> are about half as many. Specifically, *I. fonsfrigidae* NS-1<sup>T</sup> has only 96 genes encoding for CAZymes including: GHs, GTs, PLs, CEs, and CBMs (Fig. 3).

**Table 4** Enzymatic activities of the crude enzyme from *I. fonsfrigidae* strain SP3-1 grown on the BM containing 0.5% (w/v) PASC as a carbon source.

Enzyme	Specific activity (U/g protein)
Cellulase	5.86 ± 0.03
Xylanase	4.43 ± 0.04
Amylase	4.71 ± 0.02

Several CHA enzymes are found in the genome of *I. fonsfrigidae* strain SP3-1 (Table S5). Cellulolytic enzymes were found in the *I. fonsfrigidae* strain SP3-1 loci: AZO94579.1 and AZO94980.1 (endoglucanases); and, AZO93138.1, AZO94802.1, AZO96211.1, and AZO96438.1 ( $\beta$ -glucosidases). Hemicellulolytic enzymes were detected in the genome of strain SP3-1 at loci: AZO93139.1 ( $\beta$ -xylosidase), AZO94803.1 ( $\beta$ -galactosidase), AZO95768.1 ( $\alpha$ -xylosidase), AZO96253.1 (xylan-  $\alpha$ -1,2-glucuronosidase), AZO96255.1 (xylan 1,4-  $\beta$ -xylosidase), AZO96369.1 (endo-  $\beta$ -1,4-galactanase), AZO96448.1 ( $\alpha$ -L-arabinofuranosidase), and AZO96465.1 ( $\beta$ -L-arabinofuranosidase) (see Table S5). Amylolytic enzymes are also encoded in the *I. fonsfrigidae* strain SP3-1 genome, including genes at loci: AZO93170.1 ( $\alpha$ -amylase), AZO93840.1 (oligo-  $\alpha$ -1,6-glucosidase), and AZO96695.1 (pullulanase), (Table S5). These are endo-acting, exo-acting, and debranching amylases, respectively. When compared to *I. fonsfrigidae* NS-1<sup>T</sup>, strain SP3-1 has a higher number of CHA genes suggesting higher potential for starch-based biomass deconstruction.

To determine the fecundity of *I. fonsfrigidae* strain SP3-1 was cultivated in the BM containing PASC (0.5% w/v) as a carbon source at pH 8.0 and 37 °C. Culture supernatant was harvested after 3 days (at late exponential growth phase) and concentrated. The crude supernatant concentrates exhibit cellulase activity of 5.86 U/g protein on PASC. This CHA gene expression appears to be inducible. In addition, xylanase (4.43 U/g protein) and amylase (4.71 U/g protein) activities were also detected (Table 4). These are likely constitutive enzymes involved in the breakdown of xylan and starch, respectively. Thus, supernatant concentrate derived from strain SP3-1 cultures contains CHA enzymes which readily degrade PASC, xylan, and starch, which are present in the starch-based biomass.

## DISCUSSION

Starch-based biomass is a by-product of agro-industrial operations worldwide with millions of tons produced annually (Ceballos, 2017). This biomass consists predominantly of polysaccharides such as cellulose, hemicellulose, and starch—all of which can be used as low-price raw materials to produce secondary products. These substrates can be degraded by enzymes. Enzymatic deconstruction of these polysaccharides can be an “environment-friendly” approach to process feedstock by reducing the use of hazardous chemicals (e.g., strong acids and bases) for biomass conversion (Ceballos et al., 2015; Ceballos, 2017). Specifically, cellulose, hemicellulose, and starch are hydrolyzed by CHA enzymes to yield simple sugars (i.e., monosaccharides and oligosaccharides). Thus, this type of biomass may be used as a renewable resource to produce high-value-added products (Cheawchanlertfa et al., 2021) as long as effective, low-cost enzymes and enzyme technologies are available

(Ceballos et al., 2014). Bioprospecting for enzymes capable of efficiently degrading starch-based biomass is an ongoing scientific endeavor and extremophiles are one focus for novel enzyme discovery. Recently, a bacterium strain NS-1<sup>T</sup> was isolated from deep-sea cold seeps in the South China Sea. The strain NS-1<sup>T</sup> was reported to have a high similarity to *H. cellulositytica* DSM 7362<sup>T</sup>. Ultimately, this halophilic isolate NS-1<sup>T</sup> showed phylogenetic, genomic, and physiological traits unique enough to establish a novel genus within the family Halanaerobiaceae. Thus, isolate NS-1<sup>T</sup> became the archetype of *I. fonsfrigidae*. *I. fonsfrigidae* NS-1<sup>T</sup> exhibited the ability to metabolize a diverse array of carbohydrates (Zhang et al., 2021). In this study, the archetype *I. fonsfrigidae* NS-1<sup>T</sup> (Zhang et al., 2021) and related species were contrasted with the more halophilic *I. fonsfrigidae* strain SP3-1, which was initially designated *Halocella* sp. SP3-1 (Heng et al., 2019).

### **Morphological differences between *I. fonsfrigidae* NS-1<sup>T</sup> and strain SP3-1 underscore divergent adaptations in locomotive strategies**

Despite the fact that our data support a recent report identifying strain SP3-1 as the same species as *I. fonsfrigidae* NS-1<sup>T</sup> based on accepted thresholds for genetic similarity (Zhang et al., 2021), we note remarkable differences between strain NS-1<sup>T</sup> and strain SP3-1 in terms of morphology, physiology, and biochemistry. In terms of morphological differences, the most notable feature of strain SP3-1 (when compared to strain NS-1<sup>T</sup>) is the absence of developed and functional flagella. Although the genome of strain SP3-1 contains many genes related to formation of the basal body of flagella (*i.e.*, *fliE*, *fliF*, *fliG*, *fliH*, *fliI*, *fliJ*, *fliK*, *fliL*, *fliM*, *fliN*, *fliO*, *flp*, *fliQ*, *fliR*, and *flhB*) and hook proteins (*i.e.*, *flgA*, *flgB*, *flgC*, *flgD*, *flgF*, *flgG*, *flgH*, and *flgI*), it is missing *flhA* and *flgJ*, which are key genes for basal body and hook formation, respectively (see Table 5 and Table S6). Importantly, the absence or interruption of the *flhA* gene in Gram-negative bacteria leads to nonmotile cells, which lack flagella and are incapable of exporting flagella-related proteins (Bange et al., 2010). The absence of the *flgJ* gene prevents proper assembly of the hook-filament junction in flagella. The *flgJ* gene product is also critical for stabilizing protein-protein interactions between basal structures (*e.g.*, L-ring formation) in flagella (Cohen & Hughes, 2014).

### **Physiological enhancements demonstrate the ability of strain SP3-1 to survive and thrive in a broader range of environments than strain NS-1<sup>T</sup> and other related species**

Beyond the ability of strain SP3-1 to robustly grow at higher salt concentration and higher pH (when compared to NS-1<sup>T</sup> and other members of the family Halanaerobiaceae), strain SP3-1 grows on PASC, which makes it a viable candidate for bioprospecting carbohydrate-degrading enzymes. The ability of strain SP3-1 to grow in alkaline environments while utilizing an acid-treated substrate as a carbon source underscores the strain's tolerance to pH as well as its halophilic nature. This physiological profile is supported by a set of salt stress and high pH tolerance genes identified within the strain SP3-1 genome. For example, the strain SP3-1 genome includes a molybdate ABC transporter substrate-binding gene (*i.e.*, *modA*; gene loci: AZO94203.1, AZO94204.1) (Table 5). The marine bacterium *Staphylococcus* sp. strain P-TSB-70, which readily grows in saline media with up to 20% NaCl (Das et al., 2020), is also endowed with a similar molybdate ABC transporter. Strain



**Table 5** Genes present only one strain, indicating the differences in phenotypes between strain SP3-1 and NS-1<sup>T</sup>.

SP3-1	Length	Function	NS-1 <sup>T</sup>	Length	Function	% Identity
<b>Genes coding for the flagellum</b>						
–	–	–	QTL99899.1	679	Flagellar biosynthesis protein FlhA	–
–	–	–	QTL96989.1	327	Peptidoglycan hydrolase FlgJ	–
<b>Genes coding for lactate utilization</b>						
AZO93337.1	233		QTL99594.1	231	PFAM Bacterial regulatory proteins, gntR family	98.68
AZO93433.1	229		QTL99676.1	229	Transcriptional regulator (GntR)	96.07
AZO93498.1	244		–	–	–	–
AZO93635.1	244		–	–	–	–
AZO93658.1	239		–	–	–	–
AZO93770.1	228		QTL96707.1	228	Transcriptional regulator (GntR)	97.36
AZO94208.1	231		QTL97135.1	231	SMART regulatory protein GntR HTH	99.56
AZO94455.1	244	GntR family transcriptional regulator: HTH-type transcriptional regulator genes ( <i>lutR</i> )	QTL97393.1	220	Transcriptional regulator, GntR family	99.09
AZO94576.1	125		–	–	–	–
AZO94675.1	379		QTL97604.1	358	PFAM regulatory protein GntR HTH	98.60
AZO95757.1	234		–	–	–	–
AZO95845.1	226		–	–	–	–
AZO95907.1	237		–	–	–	–
AZO96195.1	232		QTL98974.1	232	SMART regulatory protein GntR HTH	100
AZO96359.1	243		–	–	–	–
<b>Genes coding for the Raffinose degradation</b>						
AZO94804.1	742	Alpha-galactosidase	–	–	–	–
<b>Genes coding for D-Rhamnose degradation</b>						
AZO93797.1	332	Rhamnose ABC transporter substrate-binding protein ( <i>RhaS</i> )	QTL96735.1	332	Rhamnose ABC transporter substrate-binding protein ( <i>RhaS</i> )	98.44
AZO93795.1	335	ABC transporter permease ( <i>RhaP</i> )	QTL96733.1	335	ABC transporter permease ( <i>RhaP</i> )	98.81
AZO93796.1	328	ABC transporter permease ( <i>RhaQ</i> )	QTL96734.1	328	ABC transporter permease ( <i>RhaQ</i> )	98.48
–	–	–	QTL96732.1	505	ATP-binding cassette domain-containing protein rhamnose transport system ( <i>RhaT</i> )	–

(continued on next page)

Table 5 (continued)

SP3-1	Length	Function	NS-1 <sup>T</sup>	Length	Function	% Identity
<b>Genes coding for salt stress/tolerance</b>						
<a href="#">AZO94203.1</a>	259	Molybdate ABC transporter	–	–	–	–
<a href="#">AZO94204.1</a>	268	substrate-binding gene ( <i>modA</i> )	–	–	–	–
<a href="#">AZO94733.1</a>	155	Glycine/sarcosine/betaine reductase complex	–	–	–	–
<a href="#">AZO94734.1</a>	441		–	–	–	–
<a href="#">AZO94735.1</a>	339	selenoprotein A	–	–	–	–
<b>Genes coding for alkaliphilic</b>						
<a href="#">AZO93678.1</a>	430		QTL96621.1	430		99.30
<a href="#">AZO93681.1</a>	435	Na <sup>+</sup> /H <sup>+</sup> antiporter genes	QTL96624.1	345	Na <sup>+</sup> /H <sup>+</sup> antiporter genes	99.77
<a href="#">AZO94987.1</a>	401		QTL97942.1	401		99.5
<a href="#">AZO95163.1</a>	427		QTL98098.1	427		98.12
<a href="#">AZO96039.1</a>	238		–	–		–
<a href="#">AZO94275.1</a>	517	Na <sup>+</sup> /H <sup>+</sup> antiporter (NhaC)	–	–	–	–
<a href="#">AZO96817.1</a>	476		–	–	–	–
–	–	–	QTL98140.1	463	Na <sup>+</sup> /H <sup>+</sup> antiporter (NhaC)	–

**Notes.**

“–” gene absent in the other strains.

SP3-1 also contains three genes encoding a glycine/sarcosine/betaine reductase complex (gene loci: [AZO94733.1](#), [AZO94734.1](#), and [AZO94735.1](#)) that are absent in strain NS-1<sup>T</sup> (Table 5). The glycine/sarcosine/betaine reductase complex includes selenoprotein A. Selenoprotein A is involved in betaine utilization as reported by [Manzoor et al. \(2015\)](#), which demonstrates the growth of selenoprotein A-expressing *Syntrophaceticus schinkii* strain Sp3 growth on betaine. This is notable since accumulation and utilization of betaine from culture media is essential in high salinity conditions as reported by [Nyyssölä \(2001\)](#) in the study of the extreme halophile *Actinopolyspora halophila*.

Na<sup>+</sup>/H<sup>+</sup> antiporters are also reported to play an essential role in allowing halophilic bacteria to thrive in high salinity environments ([Das et al., 2020](#); [Su et al., 2021](#)). Strains SP3-1 and NS-1<sup>T</sup> both harbor Na<sup>+</sup>/H<sup>+</sup> antiporter genes; however, strain SP3-1 possesses a larger suite of such genes (gene loci: [AZO93678.1](#), [AZO93681.1](#), [AZO94987.1](#), [AZO95163.1](#), [AZO96039.1](#), [AZO94275.1](#), and [AZO96817.1](#)) than NS-1<sup>T</sup> (see Table 5).

In addition to facilitating salt tolerance, Na<sup>+</sup>/H<sup>+</sup> antiporters (as well as ABC transporters) are known to be expressed in alkaliphilic halophiles. This has been demonstrated in several species including *B. halotolerans* KKD1 ([Cheng et al., 2016](#)), *Bacillus firmus* OF4 ([Ito et al., 1997](#)), and *Bacillus* sp. G1 ([Liew et al., 2007](#)). The fact that strain SP3-1 is endowed with a more diverse repertoire of genes related to halophilic and alkaliphilic growth suggests a link between genetics and phenotype that permit strain SP3-1 to tolerate higher salt concentrations and higher pH when compared to strain NS-1<sup>T</sup>.

## Genomic analyses highlight genetic substrates that underlie differences carbohydrate metabolism between *I. fonsfrigidae* strains SP3-1 and NS-1<sup>T</sup>

Results from carbon utilization studies provided an initial indication that carbohydrate metabolism, biochemistry, and perhaps genetics, are different between *I. fonsfrigidae* strain SP3-1 and NS-1<sup>T</sup>. For example, strain NS-1<sup>T</sup> grown on glucose yields detectable lactate in the culture supernatant. However, strain SP3-1 does not indicating a difference in sugar metabolism between the strains. Lactate dehydrogenase (LDH) is a key enzyme in the last step of glycolysis that plays a key role in pyruvate-to-lactate reactions ([Hamadneh et al., 2021](#)). Although both strains SP3-1 and NS-1<sup>T</sup> contain LDH genes (gene loci at [AZO94114.1](#) and [QTL97032.1](#), respectively), strain SP3-1 has 15 genes encoding GntR family transcriptional regulators, including an HTH-type transcriptional regulator gene *lutR*. In contrast, strain NS-1<sup>T</sup> only contains seven GntR genes (see [Table 5](#)). The more extensive repertoire of GntR genes may underlie the ability of strain SP3-1 to metabolize lactate. For example, it is known that operons associated with lactate metabolism are controlled by the GntR family ([Augustiniene & Malys, 2022](#)). Specifically, HTH-type transcriptional regulators (*lutR* genes) have been shown to regulate genes involved in lactate utilization ([Wang et al., 2019](#)). Thus, strain NS-1<sup>T</sup> with fewer GntR genes may exhibit reduced efficiency in metabolizing lactate, which SP3-1 readily metabolizes this by-product of glucose utilization.

It is notable that SP3-1 can also uptake raffinose as a sole carbon source. This is not surprising since the SP3-1 genome encodes an  $\alpha$ -galactosidase gene (locus: [AZO94804.1](#)), which is absent in the genome of strain NS-1<sup>T</sup> (see [Table 5](#)). It is known that  $\alpha$ -galactosidase degrades raffinose. This was demonstrated in both *Pseudobalsamia microspore* ([Yang et al., 2015](#)) and *Saccharomyces cerevisiae* ([Álvarez Cao et al., 2019](#)).

Conversely, strain NS-1<sup>T</sup> encodes an ATP-binding cassette domain-containing protein rhamnose transport system gene (*i.e.*, *rhaT*), which is lacking in SP3-1. This may underlie the inability of strain SP3-1 to use rhamnose as a carbon source (see [Richardson, Hynes & Oresnik, 2004](#)).

## Biochemical (and metabolic) properties highlight the CHA enzymes potential of *I. fonsfrigidae* strain SP3-1

Differences in carbohydrate metabolism extend well beyond simple absence/presence of genes for utilization of simple sugars as carbon sources. *I. fonsfrigidae* strain SP3-1 also stands out for its suite of genes encoding CHA enzymes. Not only does the strain SP3-1 genome encode more CHA enzymes than NS-1<sup>T</sup>, but the range of exo-acting, endo-acting, and side chain-acting enzymes are more expansive. The complement of CHA enzymes in strain SP3-1 suggests that its carbohydrate metabolism is more advanced than that of strain NS-1<sup>T</sup>.

Since these classes of enzymes work in concert (and often synergistically) to degrade polysaccharides ([Linares-Pasten, Andersson & Karlsson, 2014](#)), the extensive repertoire of CHA enzymes found in the strain SP3-1 genome justifies exploration of this halophilic alkaliphile as a potential source for novel enzyme discovery. Our data on strain SP3-1

show a suite of endoglucanases, which promote cleavage at internal sites within cellulose molecular structures, as well as  $\beta$ -glucosidases, which act on short-chain oligosaccharides and cellobiose to produce glucose (Baramée *et al.*, 2017). Genomic data for strain SP3-1 also show a gene encoding for CBM. CBM is reported to assist in the binding of enzymes to insoluble substrates to promote the efficient degradation of cellulosic substrates (Limsakul *et al.*, 2021).

Strain SP3-1 also contains genes encoding for hemicellulolytic enzymes. This suite of enzymes includes: xylan 1,4-  $\beta$ -xylosidase,  $\beta$ -xylosidase,  $\alpha$ -xylosidase, xylan-  $\alpha$ -1,2-glucuronosidase,  $\alpha$ -L-arabinofuranosidase,  $\beta$ -L-arabinofuranosidase, endo-  $\beta$ -1,4-galactanase, and  $\beta$ -galactosidase (Table S5). Such enzymes are key in the conversion of hemicellulose fractions to simpler sugars (*i.e.*, monomeric carbohydrates). Hemicellulolytic enzymes are useful in biofuels production (Ceballos, 2017) and the synthesis of prebiotics (Phakeenuya *et al.*, 2020).

In addition to cellulolytic and hemicellulolytic enzymes, strain SP3-1 harbors a suite of amyolytic genes that include endo-acting, exo-acting and debranching amylases. Amyolytic enzymes catalyze the cleavage of  $\alpha$ -D-1,4- and  $\alpha$ -D-1,6-glycosidic linkages of starch and related oligosaccharides producing short-chain oligosaccharides and glucose (Sidar *et al.*, 2020). Oligosaccharides from starch are used as prebiotics to promote the growth of healthy gut microflora (Belorkar & Gupta, 2016). Furthermore, amyolytic enzymes can be used in starch liquefaction as well as in paper, food, pharmaceutical, and sugar production operations (Pervez *et al.*, 2014).

Since CHA enzymes are the primary enzymes for the breakdown of polysaccharides in starch-based biomass, strain SP3-1 is an attractive and promising microorganism for the novel discovery of CHA enzymes and conversion of starch-based biomass into value-added products.

## CONCLUSIONS

*Halocella* sp. SP3-1 was isolated from a high salt evaporation pond in Samut Sakhon, Thailand as described by Heng *et al.* (2019). The whole-genome sequence of strain SP3-1 was deposited at NCBI GenBank under the accession number CP032760. It was later found to be a species of a new taxon, *Iocasia fonsfrigidiae*, which includes a characterized type strain: NS-1<sup>T</sup> (Zhang *et al.*, 2021). Strain SP3-1, which was isolated from a salt evaporation pond, readily grows in higher salt content of 30% NaCl and higher pH than the type strain NS-1<sup>T</sup>. The halophilic and alkaliphilic nature of strain SP3-1 prompted a studied focused on its potential to metabolize simple and complex carbohydrates as well as the genetic/genomic substrates that underlie its phenotypic nature.

In this study, we have demonstrated that despite the same species determination for strain SP3-1 (Heng *et al.*, 2019) and type strain NS-1<sup>T</sup> (Zhang *et al.*, 2021), *I. fonsfrigidiae* strain SP3-1 is more halophilic and alkaliphilic than strain NS-1<sup>T</sup> and that there are genetic differences that can account for phenotypic (*e.g.*, morphological and physiological) differences between these two strains. Our analyses demonstrate that strain SP3-1 expresses and secretes a suite of CHA-enzymes which is distinct from that of strain NS-1<sup>T</sup>. Given

the adaptation of strain SP3-1, to higher salinity and higher pH environments, this strain serves as a suitable candidate for novel enzyme discovery. Although both strains of *Iocasia fonsfrigidiae* are likely limited in their ability to degrade lignocellulosic substrates due to the absence of ligninolytic enzymes such as: lignin peroxidase, manganese peroxidase, versatile peroxidase, laccase, phenoloxidases, and auxiliary enzymes, which play a key role in the degradation of lignin (Biko, Bloom & Van Zyl, 2020), the prospect of discovering or engineering high salt- and high pH- tolerant enzymes from the strain SP3-1 proteome is promising and the subject of ongoing research.

## ACKNOWLEDGEMENTS

We acknowledge members of Enzyme Technology Laboratory, King Mongkut's University of Technology Thonburi, Thailand for the technical assistance.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

The authors received financial support from the King Mongkut's University of Technology Thonburi through the Excellent Center of Enzyme Technology and Microbial Utilization (Grant number 7601.24/4054). Sobrone Heng received support from the Development of High-Quality Research Graduates in Science and Technology Petchra Pra Jom Klao Ph.D. Research Scholarship (KMUTT-NSTDA) (Grant number 4/2562). The authors received support for their student and faculty exchange activities (including Sobrone Heng) between KMUTT and the University of Arkansas from the the U.S. National Science Foundation (NSF) through an NSF RCN UBE (award no. 1624171; PI-Ceballos) and an NSF INFEWS/T3 grant (award no. 1856091; PI-Ceballos). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

The King Mongkut's University of Technology Thonburi through the Excellent Center of Enzyme Technology and Microbial Utilization: 7601.24/4054.

The Development of High-Quality Research Graduates in Science and Technology Petchra Pra Jom Klao Ph.D. Research Scholarship (KMUTT-NSTDA): 4/2562.

The U.S. National Science Foundation (NSF): 1624171.

NSF INFEWS/T3 grant: 1856091.

### Competing Interests

The authors declare there are no competing interests.

### Author Contributions

- Sobrone Heng conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

- Sawanee Sutheeworapong conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Verawat Champreda conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Ayaka Uke conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Akihiko Kosugi conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Patthra Pason conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Rattiya Waeonukul conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Ruben Michael Ceballos conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Khanok Ratanakhanokchai conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Chakrit Tachaapaikoon conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

### **DNA Deposition**

The following information was supplied regarding the deposition of DNA sequences:

The 16S rRNA gene sequence of *Iocasia fonsofrigidiae* strain SP3-1 is available at GenBank: [MW958225](#). The whole-genome sequence of *Iocasia fonsofrigidiae* strain SP3-1 is also available at GenBank: [CP032760](#).

### **Data Availability**

The following information was supplied regarding data availability:

The raw data are available in the [Supplemental Files](#).

### **New Species Registration**

The following information was supplied regarding the registration of a newly described species:

The strain SP3-1 is deposited in Thailand and South Korea under deposit codes TISTR 2992 (Thailand Institute of Scientific and Technological Research) and KCTC 25333 (Korean Collection for Type Cultures), respectively.

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.14211#supplemental-information>.

## REFERENCES

- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**:3389–3402 DOI [10.1093/nar/25.17.3389](https://doi.org/10.1093/nar/25.17.3389).
- Álvarez Cao ME, Cerdán ME, González-Siso MI, Becerra M. 2019. Optimization of *Saccharomyces cerevisiae*  $\alpha$ -galactosidase production and application in the degradation of raffinose family oligosaccharides. *Microbial Cell Factories* **18**:1–17 DOI [10.1186/s12934-019-1222-x](https://doi.org/10.1186/s12934-019-1222-x).
- Augustiniene E, Malys N. 2022. Identification and characterization of L- and D-lactate-inducible systems from *Escherichia coli* MG1655, *Cupriavidus necator* H16 and *Pseudomonas* species. *Scientific Reports* **12**:1–11 DOI [10.1038/s41598-022-06028-7](https://doi.org/10.1038/s41598-022-06028-7).
- Bange G, Kümmerer N, Engel C, Bozkurt G, Wild K, Sinning I. 2010. FlhA provides the adaptor for coordinated delivery of late flagella building blocks to the type III secretion system. *Proceedings of the National Academy of Sciences of the United States of America* **107**:11295–11300 DOI [10.1073/pnas.100138310](https://doi.org/10.1073/pnas.100138310).
- Baramee S, Teeravivattanakit T, Phitsuwan P, Waeonukul R, Pason P, Tachapaikoon C, Kosugi A, Sakka K, Ratanakhanokchai K. 2017. A novel GH6 cellobiohydrolase from *Paenibacillus curdolanolyticus* B-6 and its synergistic action on cellulose degradation. *Applied Microbiology and Biotechnology* **101**:1175–1188 DOI [10.1007/s00253-016-7895-8](https://doi.org/10.1007/s00253-016-7895-8).
- Belas MR, Colwell RR. 1982. Scanning electron microscope observation of the swarming phenomenon of *Vibrio parahaemolyticus*. *Journal of Bacteriology* **150**:956–959 DOI [10.1128/jb.150.2.956-959.1982](https://doi.org/10.1128/jb.150.2.956-959.1982).
- Belorkar SA, Gupta AK. 2016. Oligosaccharides: a boon from nature's desk. *AMB Express* **6**:1–11 DOI [10.1186/s13568-016-0253-5](https://doi.org/10.1186/s13568-016-0253-5).
- Biko ODV, Bloom MV, Van Zyl WH. 2020. Microbial lignin peroxidases: applications, production challenges and future perspectives. *Enzyme and Microbial Technology* **141**:109669 DOI [10.1016/j.enzmictec.2020.109669](https://doi.org/10.1016/j.enzmictec.2020.109669).
- Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Research* **37**:D233–D238 DOI [10.1093/nar/gkn663](https://doi.org/10.1093/nar/gkn663).
- Cayol JL, Ollivier B, Patel BK, Prensier G, Guezennec J, Garcia JL. 1994. Isolation and characterization of *Halothermothrix orenii* gen. nov. sp. nov. a halophilic, thermophilic, fermentative, strictly anaerobic bacterium. *International Journal of Systematic and Evolutionary Microbiology* **44**:534–540 DOI [10.1099/00207713-44-3-534](https://doi.org/10.1099/00207713-44-3-534).
- Ceballos RM. 2017. *Bioethanol and natural resources: substrates, chemistry and engineered systems*. Boca Raton: CRC Press DOI [10.1201/9781315154299](https://doi.org/10.1201/9781315154299).

- Ceballos RM, Batchenkova NA, Chan MK, Duffing-Romero AX, Nelson AE, Man S. 2015. Bioethanol: feedstock alternatives, pretreatments, lignin chemistry, and the potential for green value-added lignin co-products. *Journal of Environmental Analytical Chemistry* 2:1–24 DOI 10.4172/2380-2391.1000164.
- Ceballos RM, Ceballos Jr RM, Rani A, Morales CT, Batchenkova NA. 2014. Improved hydrolysis of pretreated lignocellulosic biomass using mobile enzyme sequestration platforms. *Recent Advances in Energy, Environment, and Materials* 47–54.
- Cheawchanlertfa P, Tongasuk P, Sutheeworapong S, Waeonukul R, Pason P, Poomputsa K, Ratanakhanokchai K, Kosugi A, Tachaapaikoon C. 2021. A novel amylolytic/xylanolytic/cellulolytic multienzyme complex from *Clostridium manihotivorum* that hydrolyzes polysaccharides in cassava pulp. *Applied Microbiology and Biotechnology* 105:6719–6733 DOI 10.1007/s00253-021-11521-6.
- Cheng B, Meng Y, Cui Y, Li C, Tao F, Yin H, Yang C, Xu P. 2016. Alkaline response of a halotolerant alkaliphilic *Halomonas* strain and functional diversity of its Na<sup>+</sup> (K<sup>+</sup>)/H<sup>+</sup> antiporters. *Journal of Biological Chemistry* 291:26056–26065 DOI 10.1074/jbc.M116.751016.
- Chimtung S, Tachaapaikoon C, Sornyotha S, Pason P, Waeonukul R, Kosugi A, Ratanakhanokchai K. 2014. Symbiotic behavior during co-culturing of *Clostridium thermocellum* NKP-2 and *Thermoanaerobacterium thermosaccharolyticum* NOI-1 on corn hull. *BioResources* 9:2471–2483 DOI 10.15376/biores.9.2.2471-2483.
- Cohen EJ, Hughes KT. 2014. Rod-to-hook transition for extracellular flagellum assembly is catalyzed by the L-ring-dependent rod scaffold removal. *Journal of Bacteriology* 196:2387–2395 DOI 10.1128/JB.01580-14.
- Das P, Behera BK, Chatterjee S, Das BK, Mohapatra T. 2020. De novo transcriptome analysis of halotolerant bacterium *Staphylococcus* sp. strain P-TSB-70 isolated from East coast of India: in search of salt stress tolerant genes. *PLOS ONE* 2:e0228199 DOI 10.1371/journal.pone.0228199.
- Dlugosch L, Poehlein A, Wemheuer B, Pfeiffer B, Badewien TH, Daniel R, Simon M. 2022. Significance of gene variants for the functional biogeography of the near-surface Atlantic Ocean microbiome. *Nature Communications* 13:456 DOI 10.1038/s41467-022-28128-8.
- Doetsch RN. 1981. Determinative methods of light microscopy. In: *Manual of methods for general bacteriology*. Washington, D.C.: American Society of Microbiology, 21–33.
- Hall T, Biosciences I, Carlsbad C. 2011. BioEdit: an important software for molecular biology. *GERF Bulletin of Biosciences* 2:60–61.
- Hamadneh L, Al-Lakkis L, Alhusban AA, Tarawneh S, Abu-Irmaileh B, Albustanji S, Al-Bawab AQ. 2021. Changes in lactate production, lactate dehydrogenase genes expression and DNA methylation in response to tamoxifen resistance development in MCF-7 cell line. *Gene* 12:777 DOI 10.3390/genes12050777.
- Heng S, Sutheeworapong S, Prommeenate P, Cheevadhanarak S, Kosugi A, Pason P, Waeonukul R, Ratanakhanokchai K, Tachaapaikoon C. 2019. Complete genome



- sequence of *Halocella* sp. strain SP3-1, an extremely halophilic, glycoside hydrolase- and bacteriocin-producing bacterium isolated from a salt evaporation pond. *Microbiology Resource Announcements* 8:e01696–18 DOI 10.1128/MRA.01696-18.
- Hoang TD, Nghiem N. 2021.** Recent developments and current status of commercial production of fuel ethanol. *Fermentation* 7:314 DOI 10.3390/fermentation7040314.
- Hu R, Lin L, Liu T, Ouyang P, He B, Liu S. 2008.** Reducing sugar content in hemi-cellulose hydrolysate by DNS method: a revisit. *Journal of Biobased Materials and Bioenergy* 2:156–161 DOI 10.1166/jbmb.2008.306.
- Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Von Mering C, Bork P. 2017.** Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular Biology and Evolution* 34:2115–2122 DOI 10.1093/molbev/msx148.
- Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ. 2016.** eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Research* 44:D286–D293 DOI 10.1093/nar/gkv1248.
- Ito M, Guffanti AA, Zemsky J, Ivey DM, Krulwich TA. 1997.** Role of the nhaC-encoded Na<sup>+</sup>/H<sup>+</sup> antiporter of alkaliphilic *Bacillus firmus* OF4. *Journal of Bacteriology* 179:3851–3857 DOI 10.1128/jb.179.12.3851-3857.1997.
- Jusakulvijit P, Bezama A, Thrän D. 2021.** The availability and assessment of potential agricultural residues for the regional development of second-generation bioethanol in Thailand. *Waste and Biomass Valorization* 11:6091–6118 DOI 10.1007/s12649-021-01424-y.
- Kaznadzey A, Shelyakin P, Gelfand MS. 2017.** Sugar Lego: gene composition of bacterial carbohydrate metabolism genomic loci. *Biology Direct* 12:1–4 DOI 10.1186/s13062-017-0200-7.
- Kim M, Oh HS, Park SC, Chun J. 2014.** Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology* 64:346–351 DOI 10.1099/ijs.0.059774-0.
- Kivistö AT, Karp MT. 2011.** Halophilic anaerobic fermentative bacteria. *Journal of Biotechnology* 152:114–124 DOI 10.1016/j.jbiotec.2010.08.014.
- Komagata K, Suzuki K. 1987.** Lipid and cell-wall analysis in bacterial systematics. *Methods in Microbiology* 19:161–203.
- Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW. 2007.** RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research* 35:3100–3108 DOI 10.1093/nar/gkm160.
- Laslett D, Canback B. 2004.** ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research* 32:11–16 DOI 10.1093/nar/gkh152.
- Lee IM, Gundersen-Rindal DE, Davis RE, Bartoszyk IM. 1998.** Revised classification scheme of phytoplasmata based on RFLP analyses of 16S rRNA and ribosomal protein

- gene sequences. *International Journal of Systematic and Evolutionary Microbiology* **48**:1153–1169 DOI [10.1099/00207713-48-4-1153](https://doi.org/10.1099/00207713-48-4-1153).
- Liew CW, Illias RM, Mahadi NM, Najimudin N. 2007.** Expression of the Na<sup>+</sup>/H<sup>+</sup> antiporter gene (g1-nhaC) of alkaliphilic *Bacillus* sp. G1 in *Escherichia coli*. *FEMS Microbiology Letters* **276**:114–122 DOI [10.1111/j.1574-6968.2007.00925.x](https://doi.org/10.1111/j.1574-6968.2007.00925.x).
- Limsakul P, Phitsuwan P, Waeonukul R, Pason P, Tachaapaikoon C, Poomputsa K, Kosugi A, Ratanakhanokchai K. 2021.** A novel multifunctional arabinofuranosidase/endo-xylanase/ $\beta$ -xylosidase GH43 from *Paenibacillus curdlanolyticus* B-6 and its synergistic action to produce arabinose and xylose from cereal arabinoxylan. *Applied and Environmental Microbiology* **87**:e01730–21 DOI [10.1128/AEM.01730-21](https://doi.org/10.1128/AEM.01730-21).
- Linares-Pasten JA, Andersson M, Karlsson EN. 2014.** Thermostable glycoside hydrolases in biorefinery technologies. *Current Biotechnology* **3**:26–44 DOI [10.2174/22115501113026660041](https://doi.org/10.2174/22115501113026660041).
- Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. 1951.** Protein measurement with the Folin phenol reagent. *Journal of Biological Chemistry* **193**:265–275 DOI [10.1016/S0021-9258\(19\)52451-6](https://doi.org/10.1016/S0021-9258(19)52451-6).
- Manafi M, Kneifel W. 1990.** Rapid methods for differentiating Gram-positive from Gram-negative aerobic and facultative anaerobic bacteria. *Journal of Applied Bacteriology* **69**:822–827 DOI [10.1111/j.1365-2672.1990.tb01579.x](https://doi.org/10.1111/j.1365-2672.1990.tb01579.x).
- Manzoor S, Müller B, Niazi A, Schnürer A, Bongcam-Rudloff E. 2015.** Working draft genome sequence of the mesophilic acetate oxidizing bacterium *Syntrophaceticus schinkii* strain Sp3. *Standards in Genomic Sciences* **10**:1–8 DOI [10.1186/s40793-015-0092-z](https://doi.org/10.1186/s40793-015-0092-z).
- Minnikin DE, O'donnell AG, Goodfellow M, Alderson G, Athalye M, Schaal A, Parlett JH. 1984.** An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *Journal of Microbiological Methods* **2**:233–241 DOI [10.1016/0167-7012\(84\)90018-6](https://doi.org/10.1016/0167-7012(84)90018-6).
- Nyyssölä A. 2001.** Pathways of glycine betaine synthesis in two extremely halophilic bacteria, *Actinopolyspora halophila* and *Ectothiorhodospira halochloris*. D. Phil. Thesis, Helsinki University of Technology.
- Oktari A, Supriatin Y, Kamal M, Syafrullah H. 2017.** The bacterial endospore stain on schaeffer fulton using variation of methylene blue solution. *Journal of Physics: Conference Series* **812**:012066 DOI [10.1088/1742-6596/812/1/012066](https://doi.org/10.1088/1742-6596/812/1/012066).
- Olsen L, Choffnes ER, Mack A. 2012.** *The social biology of microbial communities: Workshop summary*. Washington, D.C.: Institute of Medicine, The National Academies Press DOI [10.17226/13500](https://doi.org/10.17226/13500).
- Oren A. 2015.** Halanaerobiaceae. *Bergey's Manual of Systematics of Archaea and Bacteria* **17**:1–5 DOI [10.1002/9781118960608.fbm00148](https://doi.org/10.1002/9781118960608.fbm00148).
- Pason P, Kyu KL, Ratanakhanokchai K. 2006.** *Paenibacillus curdlanolyticus* strain B-6 xylanolytic-cellulolytic enzyme system that degrades insoluble polysaccharides. *Applied and Environmental Microbiology* **72**:2483–2490 DOI [10.1128/AEM.72.4.2483-2490.2006](https://doi.org/10.1128/AEM.72.4.2483-2490.2006).

- Pervez S, Aman A, Iqbal S, Siddiqui NN, Qader SA. 2014.** Saccharification and liquefaction of cassava starch: an alternative source for the production of bioethanol using amylolytic enzymes by double fermentation process. *BMC Biotechnology* **14**:1–10 DOI [10.1186/1472-6750-14-49](https://doi.org/10.1186/1472-6750-14-49).
- Phakeenuya V, Ratanakhanokchai K, Kosugi A, Tachaapaikoon C. 2020.** A novel multifunctional GH9 enzyme from *Paenibacillus curdlanolyticus* B-6 exhibiting endo/exo functions of cellulase, mannanase and xylanase activities. *Applied Microbiology and Biotechnology* **104**:2079–2096 DOI [10.1007/s00253-020-10388-3](https://doi.org/10.1007/s00253-020-10388-3).
- Qin QL, Xie BB, Zhang XY, Chen XL, Zhou BC, Zhou J, Oren A, Zhang YZ. 2014.** A proposed genus boundary for the prokaryotes based on genomic insights. *Journal of Bacteriology* **196**:2210–2215 DOI [10.1128/JB.01688-14](https://doi.org/10.1128/JB.01688-14).
- Richardson JS, Hynes MF, Oresnik IJ. 2004.** A genetic locus necessary for rhamnose uptake and catabolism in *Rhizobium leguminosarum* bv. trifolii. *Journal of Bacteriology* **186**:8433–8442 DOI [10.1128/JB.186.24.8433–8442.2004](https://doi.org/10.1128/JB.186.24.8433-8442.2004).
- Richter M, Rosselló-Móra R. 2009.** Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences of the United States of America* **106**:19126–19131 DOI [10.1073/PNAS.0906412106](https://doi.org/10.1073/PNAS.0906412106).
- Rigden DJ, Fernández XM. 2022.** The 2022 Nucleic Acids Research database issue and the online molecular biology database collection. *Nucleic Acids Research* **7**:D1–D10 DOI [10.1093/nar/gkab1195](https://doi.org/10.1093/nar/gkab1195).
- Rzhetsky A, Nei M. 1992.** A simple method for estimating and testing minimum-evolution trees. *Molecular Biology and Evolution* **9**:945–967 DOI [10.1093/oxfordjournals.molbev.a040771](https://doi.org/10.1093/oxfordjournals.molbev.a040771).
- Sasser M. 1990.** Identification of bacteria by gas chromatography of cellular fatty acids. Available at [http://midi-inc.com/pdf/MIS\\_Technote\\_101.pdf](http://midi-inc.com/pdf/MIS_Technote_101.pdf) (accessed on 11 March 2022).
- Seemann T. 2014.** Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**:2068–2069 DOI [10.1093/bioinformatics/btu153](https://doi.org/10.1093/bioinformatics/btu153).
- Sidar A, Albuquerque ED, Voshol GP, Ram AF, Vijgenboom E, Punt PJ. 2020.** Carbohydrate binding modules: diversity of domain architecture in amylases and cellulases from filamentous microorganisms. *Frontiers in Bioengineering and Biotechnology* **8**:871 DOI [10.3389/fbioe.2020.00871](https://doi.org/10.3389/fbioe.2020.00871).
- Simankova MV, Chernych NA, Osipov GA, Zavarzin GA. 1993.** *Halocella cellulolytica* gen. nov. sp. nov. a new obligately anaerobic, halophilic, cellulolytic bacterium. *Systematic and Applied Microbiology* **16**:385–389 DOI [10.1016/S0723-2020\(11\)80270-5](https://doi.org/10.1016/S0723-2020(11)80270-5).
- Su Z, Wang S, Yang S, Yin Y, Cao Y, Li G, Ma T. 2021.** Genetic and comparative genome analysis of *Exiguobacterium aurantiacum* SW-20, a petroleum-degrading bacteria with salt tolerance and heavy metal-tolerance isolated from produced water of Changqing oilfield, China. *Microorganisms* **10**:66 DOI [10.3390/microorganisms10010066](https://doi.org/10.3390/microorganisms10010066).
- Tamura K, Stecher G, Kumar S. 2021.** MEGA 11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38**:3022–3027 DOI [10.1093/molbev/msab120](https://doi.org/10.1093/molbev/msab120).

- Tindall BJ, Rosselló-Móra R, Busse HJ, Ludwig W, Kämpfer P. 2010.** Notes on the characterization of prokaryote strains for taxonomic purposes. *International Journal of Systematic and Evolutionary Microbiology* **60**:249–266 DOI [10.1099/ijms.0.016949-0](https://doi.org/10.1099/ijms.0.016949-0).
- Wang Y, Zhang C, Liu G, Ju J, Yu B, Wang L. 2019.** Elucidating the role and regulation of a lactate permease as lactate transporter in *Bacillus coagulans* DSM1. *Applied and Environmental Microbiology* **85**:e00672–19 DOI [10.1128/AEM.00672-19](https://doi.org/10.1128/AEM.00672-19).
- Warnecke F, Luginbühl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, Cayouette M, McHardy AC, Djordjevic G, Aboushadi N, Sorek R. 2007.** Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450**:560–565 DOI [10.1038/nature06269](https://doi.org/10.1038/nature06269).
- Yang D, Tian G, Du F, Zhao Y, Zhao L, Wang H, Ng TB. 2015.** A fungal alpha-galactosidase from *Pseudobalsamia microspora* capable of degrading raffinose family oligosaccharides. *Applied Biochemistry and Biotechnology* **176**:2157–2169 DOI [10.1007/s12010-015-1705-0](https://doi.org/10.1007/s12010-015-1705-0).
- Yang Y, Liu X, Cai J, Chen Y, Li B, Guo Z, Huang G. 2019.** Genomic characteristics and comparative genomics analysis of the endophytic fungus *Sarocladium brachiariae*. *BMC Genomics* **20**:1–20 DOI [10.1186/s12864-019-5911-y](https://doi.org/10.1186/s12864-019-5911-y).
- Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017.** Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology* **67**:1613–1611 DOI [10.1099/ijsem.0.001755](https://doi.org/10.1099/ijsem.0.001755).
- Zeikus JG, Hegge PW, Thompson TE, Phelps TJ, Langworthy TA. 1983.** Isolation and description of *Haloanaerobium praevalens* gen. nov. and sp. nov. an obligately anaerobic halophile common to Great Salt Lake sediments. *Current Microbiology* **9**:225–233 DOI [10.1007/BF01567586](https://doi.org/10.1007/BF01567586).
- Zhang J, Zhang Y, Liu R, Cai R, Liu F, Sun C. 2021.** *Iocasia fonsfrigidiae* NS-1 gen. nov. sp. nov. a novel deep-sea bacterium possessing diverse carbohydrate metabolic pathways. *Frontiers in Microbiology* **12**:725159 DOI [10.3389/fmicb.2021.725159](https://doi.org/10.3389/fmicb.2021.725159).
- Zhang YH, Cui J, Lynd LR, Kuang LR. 2006.** A transition from cellulose swelling to cellulose dissolution by *o*-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromolecules* **7**:644–648 DOI [10.1021/bm050799c](https://doi.org/10.1021/bm050799c).