



Article

# The Role of *Petrimonas mucosa* ING2-E5A<sup>T</sup> in Mesophilic Biogas Reactor Systems as Deduced from Multiomics Analyses

Irena Maus <sup>1</sup>, Tom Tubbesing <sup>2</sup>, Daniel Wibberg <sup>1</sup>, Robert Heyer <sup>3,4</sup>, Julia Hassa <sup>1,5</sup>, Geizecler Tomazetto <sup>6</sup>, Liren Huang <sup>2</sup>, Boyke Bunk <sup>7</sup>, Cathrin Spröer <sup>7</sup>, Dirk Benndorf <sup>3,8,9</sup>, Vladimir Zverlov <sup>10,11</sup>, Alfred Pühler <sup>1</sup>, Michael Klocke <sup>12</sup>, Alexander Sczyrba <sup>2</sup> and Andreas Schlüter <sup>1,\*</sup>

<sup>1</sup> Center for Biotechnology (CeBiTec), Genome Research of Industrial Microorganisms, Bielefeld University, Universitätsstr. 27, 33615 Bielefeld, Germany; irena.maus@CeBiTec.Uni-Bielefeld.de (I.M.); dwibberg@cebitec.uni-bielefeld.de (D.W.); jhassa@cebitec.uni-bielefeld.de (J.H.); puehler@cebitec.uni-bielefeld.de (A.P.)

<sup>2</sup> Faculty of Technology, Bielefeld University, Universitätsstr. 25, 33615 Bielefeld, Germany; t.tubbesing@uni-bielefeld.de (T.T.); huanglr@cebitec.uni-bielefeld.de (L.H.); asczyrba@cebitec.uni-bielefeld.de (A.S.)

<sup>3</sup> Bioprocess Engineering, Otto von Guericke University Magdeburg, Universitätspl. 2, 39106 Magdeburg, Germany; heyer@mpi-magdeburg.mpg.de (R.H.); benndorf@mpi-magdeburg.mpg.de (D.B.)

<sup>4</sup> Database and Software Engineering Group, Department of Computer Science, Institute for Technical and Business Information Systems, Otto von Guericke University Magdeburg, Universitätspl. 2, 39106 Magdeburg, Germany

<sup>5</sup> Department of Bioengineering, Leibniz Institute for Agricultural Engineering and Bioeconomy, Max-Eyth-Allee 100, 14469 Potsdam, Germany

<sup>6</sup> Biological and Chemical Engineering Section (BCE), Department of Engineering, Aarhus University, 8000 Aarhus, Denmark; geizetomazetto@gmail.com

<sup>7</sup> Department Bioinformatics and Databases, Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Inhoffenstraße 7B, 38124 Braunschweig, Germany; Boyke.Bunk@dsmz.de (B.B.); ckc@dsmz.de (C.S.)

<sup>8</sup> Bioprocess Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany

<sup>9</sup> Microbiology, Anhalt University of Applied Sciences, Bernburger Straße 55, 06354 Köthen, Germany

<sup>10</sup> Chair of Microbiology, Technical University of Munich, Emil-Ramann-Str. 4, 85354 Freising, Germany; vladimir.zverlov@tum.de

<sup>11</sup> Institute of Molecular Genetics, National Research Centre «Kurchatov Institute», Kurchatov Sq. 2, 123128 Moscow, Russia

<sup>12</sup> Institute of Agricultural and Urban Ecological Projects Affiliated to Berlin Humboldt University (IASP), Philippstraße 13, 10115 Berlin, Germany; mklocke@posteo.de

\* Correspondence: aschluet@cebitec.uni-bielefeld.de; Tel.: +49-(0)-521-106-8757

Received: 12 November 2020; Accepted: 14 December 2020; Published: 17 December 2020



**Abstract:** Members of the genera *Proteiniphilum* and *Petrimonas* were speculated to represent indicators reflecting process instability within anaerobic digestion (AD) microbiomes. Therefore, *Petrimonas mucosa* ING2-E5A<sup>T</sup> was isolated from a biogas reactor sample and sequenced on the PacBio RSII and Illumina MiSeq sequencers. Phylogenetic classification positioned the strain ING2-E5A<sup>T</sup> in close proximity to *Fermentimonas* and *Proteiniphilum* species (family Dysgonomonadaceae). ING2-E5A<sup>T</sup> encodes a number of genes for glycosyl-hydrolyses (GH) which are organized in Polysaccharide Utilization Loci (PUL) comprising tandem *susCD*-like genes for a TonB-dependent outer-membrane transporter and a cell surface glycan-binding protein. Different GHs encoded in PUL are involved in pectin degradation, reflecting a pronounced specialization of the ING2-E5A<sup>T</sup> PUL systems regarding the decomposition of this polysaccharide. Genes encoding enzymes participating in amino acids

fermentation were also identified. Fragment recruitments with the ING2-E5A<sup>T</sup> genome as a template and publicly available metagenomes of AD microbiomes revealed that *Petrimonas* species are present in 146 out of 257 datasets supporting their importance in AD microbiomes. Metatranscriptome analyses of AD microbiomes uncovered active sugar and amino acid fermentation pathways for *Petrimonas* species. Likewise, screening of metaproteome datasets demonstrated expression of the *Petrimonas* PUL-specific component SusC providing further evidence that PUL play a central role for the lifestyle of *Petrimonas* species.

**Keywords:** carbohydrate-active enzymes; polysaccharide utilization loci; anaerobic digestion; biomethanation; metabolic pathway reconstruction; bioconversion

---

## 1. Introduction

Anaerobic digestion (AD) is commonly applied for treatment of organic wastes in order to achieve reduction of waste combined with parallel recovery of bioenergy [1,2]. The degradation of organic matter is a complex microbial process featuring interaction of several groups of microorganisms with different interacting or even interfering metabolic capabilities and growth requirements [3]. The microbial biogas network is complex and intertwined and therefore susceptible to environmental changes. Several factors causing process inhibition have been identified, such as temperature fluctuations, foaming, high concentrations in hydrogen sulfide or ammonia [4–7], imbalanced levels of light or heavy metals, the presence of interfering organic substances such as halogenated alkanes or aromatic compounds [8] as well as increased organic loading rates (OLR) [9,10]. Among these, the most frequent incident type described in literature resulting in process inhibition is the accumulation of ammonia [3]. Due to inhibition of syntrophic volatile fatty acid (VFA) oxidizers, high ammonia concentrations result in VFA accumulation which subsequently inhibits methanogenesis. Moreover, ammonia also directly inhibits methanogens [11].

In the past decades, remarkable progress has been made towards understanding of how disturbances affect the microbial community structures, biodiversity losses, and ecological processes [12–18]. Due to further improvements and developments in high-throughput sequencing technologies, deep insights into the microbial community became possible, allowing predictions on interrelations between the microbial community composition, interspecies interactions, and species as well as community dependencies concerning supplied AD substrates and deteriorating process parameters. In connection with the formation of VFA [19] and increased ammonium concentrations, it has been observed that members of the phylum Bacteroidetes, namely species of the families Dysgonomonadaceae (formerly Porphyromonadaceae) and Marinilabiaceae, are frequently abundant in mesophilic biogas reactor systems [20,21]. Often, corresponding biogas plants were fed with food waste or protein-rich and hardly digestible substrates such as straw showing unstable process flow with variations in biogas/methane yields and increased nitrogen/ammonia levels [13,22].

The family Dysgonomonadaceae comprises species that frequently are associated with a variety of human infections [23], but also species detected in different types of biogas reactor systems. The latter are closely related to *Proteiniphilum acetatigenes* [24], *Fermentimonas caenicola* [25], and *Petrimonas sulfuriphila* [26]. Currently, the family Dysgonomonadaceae comprises the genera *Dysgonomonas*, *Fermentimonas*, *Petrimonas*, and *Proteiniphilum* [27]. The genera *Fermentimonas* and *Petrimonas*, originally assigned to the family Porphyromonadaceae [25], were recently transferred to the family Dysgonomonadaceae and now represent a sister group of the genus *Proteiniphilum* [27]. Besides utilization of complex proteinaceous substrates such as yeast extract and peptone, members of the Dysgonomonadaceae family are able to ferment a wide range of mono- and disaccharides [23,25,28,29].

The impact of Dysgonomonadaceae members on biomass degradation in AD reactor systems operating at high OLR or increased nitrogen/ammonia levels caused by proteinaceous substrates has

been demonstrated recently [9,10]. Under mesophilic conditions, an increase in OLR had a positive effect on the abundance of Dysgonomonadaceae members of the genera *Proteiniphilum* and *Petrimonas* [9]. These observations led to the assumption that these genera may serve as microbial early warning indicators for an unbalanced process. They are therefore regarded as potential marker microorganisms featuring outstanding performance under stress conditions [9,10]. Moreover, a study published by He and coworkers [7] demonstrated the proliferation of members of the genus *Petrimonas* in biogas reactors that were operated under unfavorable or unstable conditions indicated by foaming. The relative abundance of *Petrimonas* was highest (29.34%) in foaming samples, and this value was almost seven times higher than those for non-foaming samples. Hence, it was speculated that the presence of Dysgonomonadaceae members, in particular of acid-producing *Petrimonas* bacteria, might help to optimize VFA compositions and to accelerate methane production by driving the formation of acetic acid and carbon dioxide [19].

To overcome the limited knowledge on microbial species affecting AD in a positive or negative way, a number of previously unknown bacterial species were isolated from biogas reactor samples and subsequently characterized regarding their physiological and genetic features [25,30–34]. Among these, *Petrimonas mucosa* ING2-E5A<sup>T</sup> was isolated from a mesophilic laboratory-scale biogas reactor [25]. Microbiological characterization revealed that this strain is able to utilize complex proteinaceous substrates, as well as to ferment mono- and disaccharides. Considering these phenotypic characteristics required for effective biomass conversion, the aim of this study was to establish and analyze the complete genome sequence of *P. mucosa* ING2-E5A<sup>T</sup> in order to understand the strain's functioning and competitiveness in biogas plant environments.

## 2. Materials and Methods

### 2.1. Strain Cultivation and DNA Isolation

*P. mucosa* ING2-E5A<sup>T</sup> was obtained from a sample taken from a mesophilic laboratory-scale biogas reactor continuously fed with a mixture of maize silage (10% of organic dry matter, ODM), pig (45% ODM) and cattle manure (45% ODM) as described previously [25]. The isolate ING2-E5A<sup>T</sup> was grown at 37 °C on Columbia Agar Base medium supplemented with 5% horse blood and 1.7 g/L sodium propionate. The extraction of genomic DNA was performed using the Genra Puregene Yeast/Bact. Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Unfortunately, required DNA purity criteria being A260/280 and A260/230 values greater than 1.8 were not achieved after DNA extraction with the Genra Puregene Yeast/Bact. Kit. Therefore, obtained DNA was additionally purified using the NucleoSpin<sup>®</sup> gDNA Clean-up kit (Macherey-Nagel, Düren, Germany).

### 2.2. PacBio Library Preparation, Sequencing, and Genome Assembly

Approximately 8 µg genomic DNA was used for the SMRTbell<sup>™</sup> template library construction, which was prepared according to the instructions from Pacific Biosciences (Menlo Park, CA, USA) following the 'Procedure and Checklist—Greater Than 10 kb Template Preparation'. SMRT sequencing was carried out on the PacBio RSII (Pacific Biosciences, MenloPark, CA, USA) taking one 240-min movie for two SMRT cells using the P6 Chemistry. Subsequently, SMRT Cell data was assembled using the 'RS\_HGAP\_Assembly.3' protocol included in the SMRT Portal version 2.3.0 using default parameters. The assembly revealed a circular chromosome. The 'RS\_Bridgmapper.1' protocol was applied to check the validity of the assembly. The chromosome was circularized and particularly artificial redundancies at the ends of the contigs were removed and adjusted to *dnaA* as the first gene. Finally, error-correction was performed by a mapping of paired-end MiSeq Illumina reads of 2 × 300 bp obtained previously for the strain ING2-E5AT [34] onto the finished genome using BWA [35] with subsequent variant and consensus calling using VarScan [36]. For further genome assembly details refer to [37].

The finished genome sequence was imported into the annotation platform GenDB [38] for automatic prediction of genes as described previously [29]. To identify phage-related genes and genomic islands (GI), the *P. mucosa* ING2-E5A<sup>T</sup> genome was uploaded in PHASTER [39] and IslandViewer 3.0 [40]. Finally, the program EDGAR 2.0 [41], a software tool for the comparative analysis of prokaryotic genomes, was applied in order to phylogenetically classify completely sequenced and annotated members of the order *Bacteroidales* as well as to analyze them in a comparative manner.

### 2.3. Reconstruction of Metabolic Pathways

To predict genes encoding carbohydrate-active enzymes, the carbohydrate-active enzyme database (CAZy) annotation web-server dbCAN [42] was applied. Metabolic pathways of interest were reconstructed based on EC numbers of enzymes and their assignment to KEGG pathway maps (<https://www.genome.jp/kegg/pathway.html>). Hydrogenases were identified using blastp and the database HydDB with a threshold of  $1 \times 10^{-100}$ . In addition, further classification of the predicted hydrogenases was also performed by means of HydDB [43].

### 2.4. Fragment Recruitment

To determine the distribution and abundance of *Petrimonas mucosa* in different biogas communities, fragment recruitments using the genome sequence of *Petrimonas mucosa* as template were performed by application of the bioinformatics tool SparkHit [44]. Corresponding computations were scaled-up and parallelized by using the de.NBI Cloud environment (<https://www.denbi.de/cloud>). As a fast and sensitive fragment recruitment tool, the so-called Sparkhit-recruiter was applied. This tool extends the FR-hit pipeline [45] and is implemented natively on top of the Apache Spark. In addition, it integrates a series of analytical tools and methods for various genomic applications. The fragment recruitment option implements the q-Gram algorithm to allow more mismatches than a regular read mapping during the alignment, so that extra information is provided for the metagenomic analysis. SparkHit was applied on metagenome FASTQ files from 257 biogas metagenome datasets that were downloaded from ENA ([www.ebi.ac.uk/](http://www.ebi.ac.uk/)). Randomly chosen two million reads of each FASTQ file were compared to the *P. mucosa* ING2-E5A<sup>T</sup> genome. The alignment identity threshold was set to > 50% similarity. The result of the fragment recruitments were visualized using Circos [46].

### 2.5. *Petrimonas* Transcriptional Profile as Deduced from Public Transcriptome Datasets

Metatranscriptome datasets pertaining to samples from different industrial- and laboratory-scale anaerobic digesters, detailed in Supporting Information Table S3, were reanalyzed to quantify transcripts of *Petrimonas mucosa*. Raw RNA-sequencing reads were trimmed with Trim Galore [47]. Subsequently, reads were filtered with SortMeRNA [48] to exclude ribosomal RNA from the downstream analysis. RnaSPAdes [49] was used to assemble transcripts from read datasets originating from the same biogas reactor or wastewater treatment plant. Reads from samples representing biological or technical replicates were pooled into sets. For each of these sets, the tool Kallisto [50] was used to assign read counts to the assembled metatranscriptome-contigs in case of paired-end reads. For single-end reads, Bowtie2 [51] was used to map reads to the metatranscriptome-contigs and uniquely mapping reads were counted for each contig. All contigs resulting from the transcriptome assembly were aligned to the genome of *P. mucosa* ING2-E5A<sup>T</sup> using nucleotide BLAST [52]. Metatranscriptome-contigs for which BLAST produced alignments covering at least 99% of the contig length with an identity of 97% or more were classified as transcripts originating from *P. mucosa*. The previously produced read counts for each of these contigs were assigned to features in the *P. mucosa* genome based on the annotated strain ING2-E5A<sup>T</sup> genome. Individual contigs may overlap multiple coding sequences of the genome. In such cases, the read counts that were computed for a contig were split up between all coding sequences which the contig is overlapping. A fraction  $X_i$  of total contig-read-counts  $X$  is assigned to

each coding sequence  $i$ .  $X_i$  is determined by dividing the length of the sequence  $l_i$  through the sum of the length of all sequences  $j$  which are overlapping the contig:

$$X_i = X * \frac{l_i}{\sum_j l_j}$$

In order to rank *P. mucosa* features based on their expression levels, transcripts per million (TPM) were calculated for each feature. Only transcripts assigned to *P. mucosa* were included in the TPM calculations.

### 2.6. Gene Expression Analysis of *Petrimonas* Species Based on Database Metaproteome Datasets

High-resolution metaproteomics datasets from ten industrial biogas plants and one laboratory-scale biogas fermenter (PRIDE ACCESSION: PXD009349) [53] were reanalyzed to quantify the abundance of *Petrimonas mucosa* proteins. For each of the biogas plants, the first replicate of the first sampling time point was searched with the database search engines X!Tandem [54] and OMMSA [55] implemented in the MetaProteomeAnalyser [56] (Version 3.0) against the metagenome database from the original study [57–60] extended by the protein sequences of *P. mucosa*. The remaining settings were used as described previously by Heyer and colleagues [56] with the minor change that redundant homologous proteins were grouped when they shared the same peptide.

### 2.7. Nucleotide Sequence Accession Number

The genome of *P. mucosa* ING2-E5A<sup>T</sup> was deposited in the EMBL-EBI database (European Bioinformatics Institute database; <https://www.ebi.ac.uk/ena/browser/view/LT608328>) under the accession number LT608328.

## 3. Results and Discussion

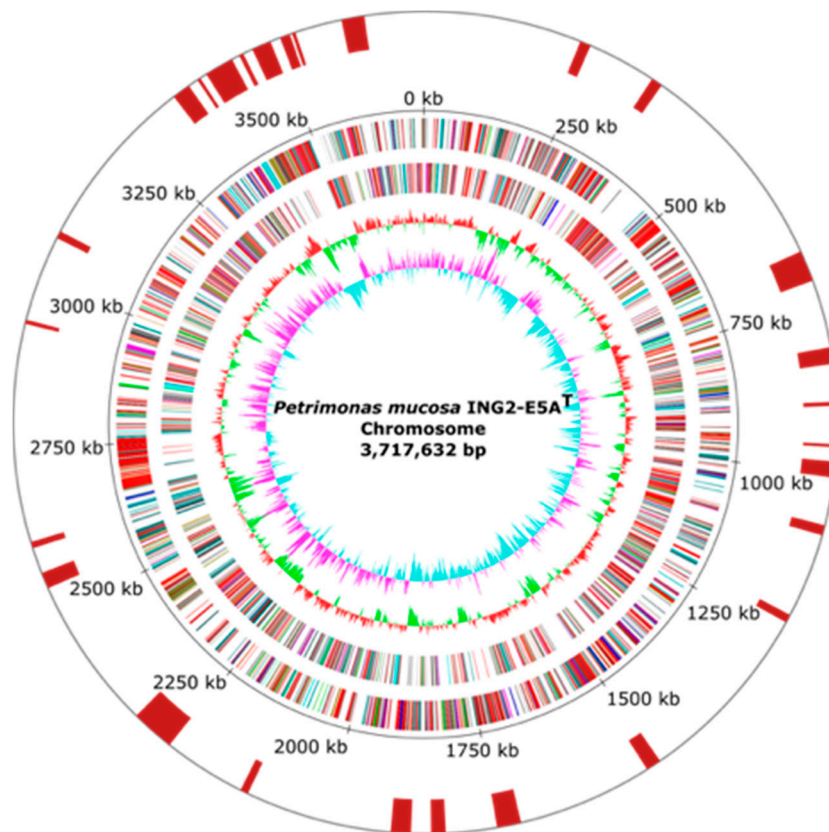
### 3.1. General Features of *Petrimonas mucosa* ING2-E5A<sup>T</sup> Genome

To deduce the importance of *Petrimonas* members in biogas-producing microbial communities, the genome of the *P. mucosa* strain ING2-E5A<sup>T</sup> (= DSM 28695<sup>T</sup> = CECT 8611<sup>T</sup>) was completely sequenced and analyzed in detail. General features of the *P. mucosa* ING2-E5A<sup>T</sup> genome are summarized in Table 1 and Figure 1. The sequencing approach yielded 115,545 reads with a mean read length of 8488 bp, accounting for 980,798,552 bp sequence information. Long read genome assembly followed by short read error correction resulted in one circular contig with 3,717,632 bp in size, featuring a GC-content of 47.97%. Gene prediction and annotation of the *P. mucosa* ING2-E5A<sup>T</sup> genome sequence resulted in the identification of 3000 coding sequences, 49 tRNA genes, and two *rrn* operons. Analysis of the *P. mucosa* ING2-E5A<sup>T</sup> genome for the presence of horizontally acquired DNA elements resulted in the identification of 27 genomic islands (Figure 1) and putative phage genes. However, no complete prophage cluster was predicted in the genome. ING2-E5A<sup>T</sup> possesses one CRISPR-*cas* system which may play a role in preventing invasion of phages and mobile genetic elements.

**Table 1.** General features of the *Petrimonas mucosa* ING2-E5A<sup>T</sup> genome.

Feature	Chromosome
Genome size (bp)	3,717,632
GC content (%)	47.97
Total genes	3055
Protein coding genes	3000
<i>rrn</i> operons	2
tRNA genes	49



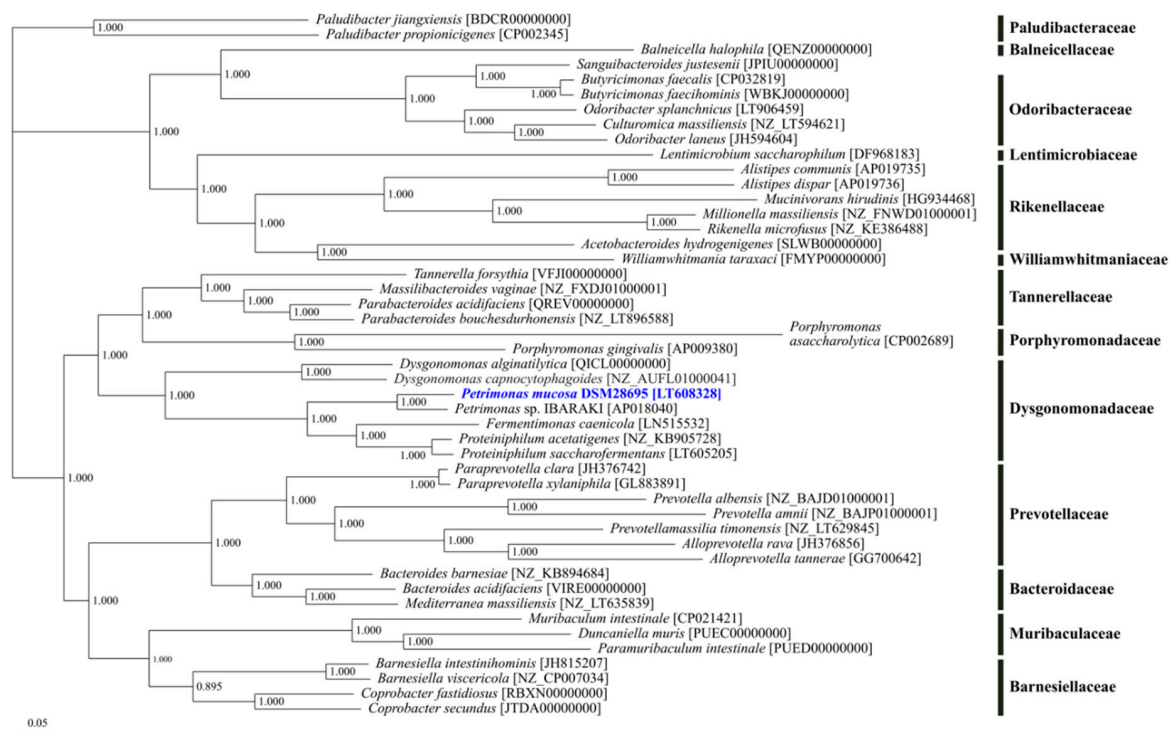


**Figure 1.** Circular representations of the *P. mucosa* ING2-E5A<sup>T</sup> chromosome. From the inner to the outer circle: Circle 1: GC skew; Circle 2: GC content; Circles 3 and 4: predicted protein-coding sequences (CDS) transcribed clockwise (outer part) or anticlockwise (inner part). The CDSs are colored according to the assigned Clusters of Orthologous Groups (COG) classes. Circle 5: genomic position in kb. Circle 6: predicted genomic islands (GI) in the ING2-E5A<sup>T</sup> genome. The replication initiation gene *dnaA* was chosen as the first gene of the chromosome.

### 3.2. Phylogenetic Classification of *Petrimonas mucosa* ING2-E5A<sup>T</sup>

Phylogenetic tree reconstruction was performed based on available genomes for 47 strains of the order Bacteroidales comprising 94 core genes in total (Figure 2). The analysis resulted in a distinct cluster of the genus *Petrimonas* together with the genera *Fermentimonas* and *Proteiniphilum*. Originally, these genera were assigned to the order Bacteroidales, and within this order, they were affiliated with to the family Porphyromonadaceae as deduced from 16S rRNA gene analyses [25,61]. However, in the obtained core gene-based phylogenetic tree, which is in accordance with the recent taxonomic revision within the phylum Bacteroidetes [27], *Porphyromonas* species are located in a separated cluster with notable distance to the *Petrimonas/Proteiniphilum* cluster which is much closer related to the genus *Dysgonomonas*. Therefore, members of the genera *Petrimonas*, *Proteiniphilum*, *Fermentimonas*, as well as *Dysgonomonas* constitute the current family Dysgonomonadaceae, which was earlier suggested by Ormerod et al. [62] and has only recently been validated [27].

The genus *Petrimonas* was firstly described by Grabowski et al. [26] based on cultivation, classification, and characterization of the type species *P. sulfuriphila*. So far, *P. mucosa* and *P. sulfuriphila* were the only validly described species for this genus. Moreover, within this genus, only for the type strain *P. mucosa* ING2-E5A<sup>T</sup> [25] the genome sequence is currently available. By 16S rRNA gene sequence analysis and DNA–DNA hybridization, *P. mucosa* ING2-E5A<sup>T</sup> appeared to be closely related to *P. sulfuriphila* BN3<sup>T</sup> with 97% 16S rRNA gene sequence identity and a DNA–DNA relatedness of 23.8–25.7% [25].

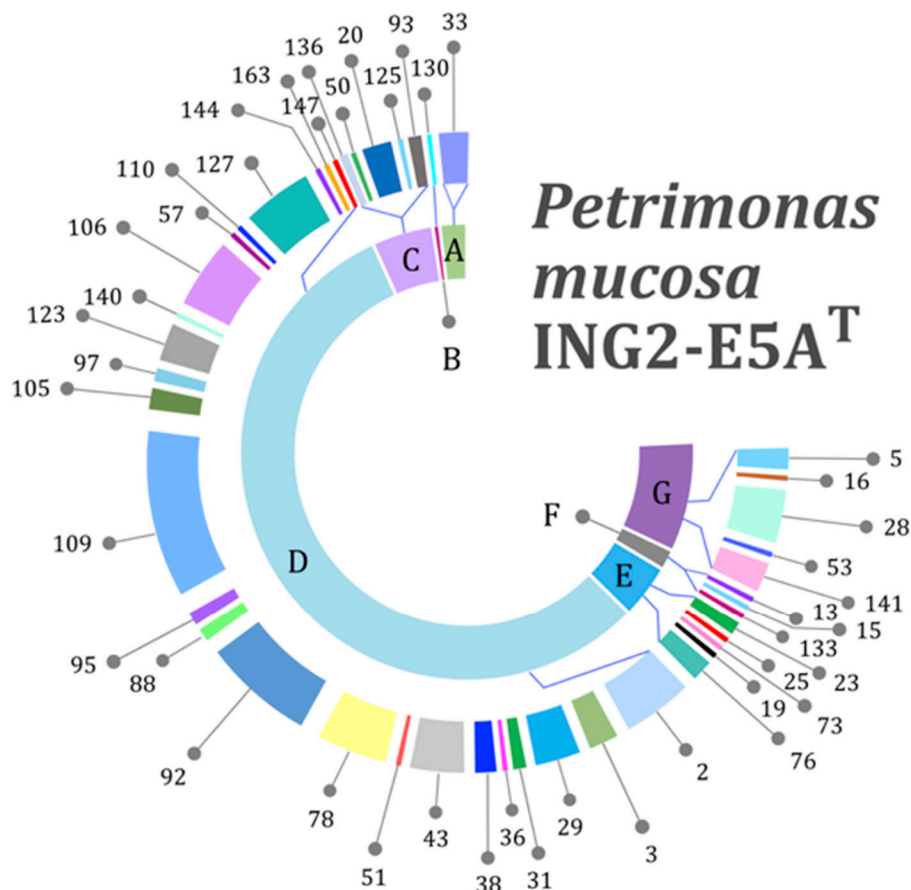


**Figure 2.** Phylogenetic tree of the order Bacteroidales. The phylogenetic tree is based on 94 core genes of the selected strains as determined by means of the comparative genomics tool EDGAR 2.0 [41]. Except for the metagenome assembled genome *Petrimonas* sp. IBARAKI [AP018040], genome datasets for type strains were used.

### 3.3. Genes Encoding Carbohydrate-Active Enzymes

To elucidate the *P. mucosa* ING2-E5A<sup>T</sup> genes involved in utilization of different carbohydrates and sugar components and to classify the strain's lifestyle, the carbohydrate-active enzyme database (CAZY) annotation web-server dbCAN [42] was applied. CAZymes are a class of enzymes which synthesize, modify, or break down different carbohydrates. Their classification comprises the following families: glycoside hydrolase families (GH), polysaccharide lyase families (PL), carbohydrate esterase families (CE), glycosyltransferase families (GT), auxiliary activity families (AA), and carbohydrate-binding module (CBM) families [63]. The proportion of CAZymes predicted in the ING2-E5A<sup>T</sup> genome was calculated to represent 7.6% of all gene products encoded in the genome, indicating the strain's ability to process various polysaccharides. Of a total of 233 CAZymes encoded in the ING2-E5A<sup>T</sup> genome, 154 GHs account for 44 different families (Figure 3). Previous studies on Gram-negative bacteria of the phylum Bacteroidetes also reported on a large variety of different GH families in these species (up to 62) [64]. Based on these observations, a broad biodegradation potential regarding carbohydrates is proposed for this phylum. The most dominant GH families in *P. mucosa* ING2-E5A<sup>T</sup> were the GH109, GH2, GH92, GH127, and GH106 families, listed by decreasing abundance. These families mostly encode glycosidases which hydrolyze single sugar residues from the non-reducing ends of di- and polysaccharides such as  $\beta$ -D-galactose (GH2), arabinosides (GH127), glycan (GH92),  $\alpha$ -L-rhamnose (GH106), and  $\alpha$ -(1 $\rightarrow$ 3)-N-acetylgalactosamine (GH109). Several other GH families identified in the ING2-E5A<sup>T</sup> genome are also associated with hydrolysis of single sugar residues indicating the importance of these substrates for this strain. Further results obtained for *P. mucosa* ING2-E5A<sup>T</sup> also showed that it encodes enzymes involved in cell wall degradation such as lysozyme (GH23, GH25, GH73) and chitinases (GH19, GH76). Hence, *P. mucosa* ING2-E5A<sup>T</sup> most probably can utilize peptidoglycan and chitin for growth. These polysaccharides are usually part of the bacterial cell wall or are found in fungi and insects, respectively, indicating the potential of *P. mucosa* ING2-E5A<sup>T</sup> for further biotechnological applications. As expected for a member of the phylum Bacteroidetes, no genes

encoding cohesin-containing putative scaffoldins and corresponding dockerin-containing glycoside hydrolases featuring a potential for cellulosome formation were identified. However, the genome of the strain ING2-E5A<sup>T</sup> encodes enzymes predicted to degrade cellulose (GH5) and hemicellulose (GH16, GH28, GH53, and GH141).



**Figure 3.** *P. mucosa* ING2-E5A<sup>T</sup> genes encoding for different glycoside hydrolases (GH) predicted to be involved in carbohydrate utilization as checked manually with the Carbohydrate Active Enzymes (CAZy) database (<http://www.cazy.org>). The outer ring represents the diversity of the GH families identified in the genome of the strain ING2-E5A<sup>T</sup>. The inner ring indicates the predicted function of the respective GH family as follows: A: Exo- $\alpha$ -sialidases, B: oligosaccharide phosphorylase, C: glycosidases (hydrolysis of single sugar residues), D: glycosidases (hydrolysis of single sugar residues from nonreducing ends), E: cell wall degradation, F: starch and glycogen hydrolases, G: heteropolymers and polysaccharides including hemicellulose.

### 3.4. Genes Encoding Polysaccharide Utilization Loci (PUL)

Degradation of the most abundant plant polysaccharide (cellulose) has usually been affiliated to the metabolic capabilities of members of the phylum Firmicutes [65,66]. According to the current literature, the Bacteroidetes are not associated with cellulose degradation. However, various members of this phylum encode a large diversity of polysaccharide utilization loci (PUL) which were suggested to be in combination with enzymes of the GH5 family acting on cellulose, as an alternative system for cellulose degradation in Bacteroidetes members, in addition to cellulosomes and free-enzyme systems [67,68]. PUL comprise clusters of genes encoding surface glycan binding proteins, transporters, and a variety of glycan-degrading enzymes. The genetic architecture is characterized by tandem *susCD*-like genes encoding a TonB-dependent outer-membrane transporter (TBDT) and an N-terminally lipidated cell surface glycan-binding protein (SGBP), respectively [68].



The *P. mucosa* ING2-E5A<sup>T</sup> genome contains 39 PUL (Figure 4), 25 of which comprise 62 GHs belonging to 24 distinct CAZyme families, representing 25.1% of all families predicted in the genome. The predominant families are GH28 and GH92 representing enzymes involved in pectin and glycan degradation (Figure 4). Besides GH28, other families related to pectin degradation are also associated with PUL, such as GH2, GH43, GH78, GH105, GH106, CE8, CE12, and PL1. Accordingly, 29.9% of all CAZymes located in PUL are involved in pectin degradation, reflecting a pronounced specialization of the enzymatic system for the decomposition of this polysaccharide. Other PUL were predicted to be involved in degradation of arabinan (e.g., PUL\_12), galactan (PUL\_39), and hemi-/cellulose (PUL\_11, PUL\_19).



**Figure 4.** Schematic overview of PUL (Polysaccharide Utilization Loci) predicted in the *P. mucosa* ING2-E5A<sup>T</sup> genome.

In addition to the genes for CAZymes, the *P. mucosa* ING2-E5A<sup>T</sup> PUL also are associated with peptidase genes. PUL\_05 and PUL\_06 (Figure 4) encode peptidases belonging to family M60 and S15, respectively. These enzymes were predicted to target glycoproteins [69] and serine peptidases [70]. Presence of peptidase genes in PUL suggests that strain ING2-E5A<sup>T</sup> may also be involved in protein degradation and subsequent amino acid fermentation [29,71]. However, it was recently reported that the strain ING2-E5A<sup>T</sup> was not able to grow on peptone or tryptone as sole carbon sources [25]. It may be speculated that growth on these substrates requires additional micronutrients that were not provided in the growth experiments conducted previously [25].

Approximately 46% of the *P. mucosa* PUL are associated with regulator system genes, for example genes encoding an extracytoplasmic function sigma factor (ECF- $\sigma$ ), an anti-sigma (*anti*- $\sigma$ ) factor, and a hybrid two-component system (HTCS). Involvement of corresponding regulators in controlling expression of PUL-associated genes remains to be determined [68]. Association of PUL with genes encoding ECF- $\sigma$ , anti- $\sigma$  pairs, HTCS, transcriptional regulators (e.g., LacI, AraC, GnTR), or sensor-regulator systems (SusR) is not unusual and has been reported for PUL of other Bacteroidetes species [68].

Similar characteristics were also observed in the CAZymes profile of *Petrimonas* sp. IBARAKI whose genome sequence is closely related to the one of strain ING2-E5A<sup>T</sup> (PULD database; [72]). Strain IBARAKI harbors 240 CAZymes, 33% of which are located in PUL. Approximately, 40% of its PUL encode enzymes assigned to CAZyme families involved in pectin degradation. Although there are only two genome sequences of *Petrimonas* species available, it can be assumed that this genus preferably degrades pectin.

In summary, genome sequencing of new Bacteroidetes species may elucidate how members of this phylum adapt PUL for decomposition of particular polysaccharides. Moreover, PUL-encoded enzymes provide exploitation possibilities for biotechnological purposes.

### 3.5. Genes Encoding Enzymes of the Central Fermentation Metabolism

Regarding its central metabolism, *P. mucosa* ING2-E5A encodes complete KEGG modules for glycolysis (Embden–Meyerhof pathway, M00001), pyruvate oxidation (M00307), the citrate cycle (TCA cycle, M00009), and the pentose phosphate pathway (M00006) allowing metabolism of glucose, pentoses, and other monosaccharides. As described above, several di-, oligo-, and polysaccharides potentially can be metabolized to the aforementioned common monosaccharides. *P. mucosa* was also predicted to perform degradation of proteins, oligo- and dipeptides since corresponding enzymatic activities are encoded in its genome. Amino acids resulting from peptide hydrolysis may feed fermentation pathways. For example, enzymes for conversion of L-asparagine via aspartate to fumarate are encoded in the *P. mucosa* genome (see Table 2). Citrate cycle enzymes catalyze the reactions from fumarate via malate to oxaloacetate which can then be decarboxylated to pyruvate by oxaloacetate decarboxylase (EC 4.1.1.112). Subsequently, ATP is generated via the phosphate-acetyltransferase/acetate kinase module (KEGG M00579) that is encoded in the *P. mucosa* genome.

**Table 2.** Catabolism of amino acids and formation of acetate and propionate in *P. mucosa* ING2-E5A<sup>T</sup>.

Locus Tag	EC No.	Enzyme	Reaction	Gene(s)
<b>Catabolism of Glutamine and Glutamate</b>				
ING2E5A_1298	3.5.1.2	Glutaminase	L-glutamine + H <sub>2</sub> O $\rightleftharpoons$ L-glutamate + NH <sub>3</sub>	<i>glsA</i>
ING2E5A_2384	1.4.1.13	Glutamate synthase (NADPH)	2 L-glutamate + NADP(+) $\rightleftharpoons$ L-glutamine + 2-oxoglutarate + NADPH	<i>gltD</i>
ING2E5A_0283, ING2E5A_2397	1.4.1.3, 1.4.1.2	Glutamate dehydrogenase (NAD(P) <sup>+</sup> ), Glutamate dehydrogenase	L-glutamate + H <sub>2</sub> O + NAD(P) <sup>+</sup> $\rightleftharpoons$ 2-oxoglutarate + NH <sub>3</sub> + NAD(P)H, L-glutamate + H <sub>2</sub> O + NAD <sup>+</sup> $\rightleftharpoons$ 2-oxoglutarate + NH <sub>3</sub> + NADH	<i>gdhA</i> <i>gdh</i>

Table 2. Cont.

Locus Tag	EC No.	Enzyme	Reaction	Gene(s)
ING2E5A_2385	1.4.7.1	Ferredoxin-dependent glutamate synthase 1	2 L-glutamate + 2 oxidized ferredoxin $\rightleftharpoons$ L-glutamine + 2-oxoglutarate + 2 reduced ferredoxin + 2 H <sup>+</sup>	<i>glfB</i>
ING2E5A_1056, ING2E5A_1057	1.2.7.3	2-oxoglutarate synthase	2-oxoglutarate + CoA + 2 oxidized ferredoxin $\rightleftharpoons$ succinyl-CoA + CO <sub>2</sub> + 2 reduced ferredoxin + 2 H <sup>+</sup>	<i>korB</i> , <i>korA</i>
ING2E5A_1055, ING2E5A_1054	1.2.4.2, 2.3.1.61	Oxoglutarate dehydrogenase (succinyl-transferring), Dihydrolipoyllysine-residue succinyltransferase	2-oxoglutarate $\rightarrow$ succinyl-CoA	<i>sucA</i> , <i>sucB</i>
ING2E5A_1515, ING2E5A_1517	5.4.99.2	Methylmalonyl-CoA mutase	(R)-methylmalonyl-CoA $\rightleftharpoons$ succinyl-CoA	<i>mutA</i> , <i>mutB</i>
ING2E5A_1046	5.1.99.1	Methylmalonyl-CoA epimerase	(R)-methylmalonyl-CoA $\rightleftharpoons$ (S)-2-Methylmalonyl-CoA	
ING2E5A_1045, ING2E5A_1864	6.4.1.3	Propionyl-CoA carboxylase	ATP + propanoyl-CoA + HCO <sub>3</sub> <sup>-</sup> $\rightleftharpoons$ ADP + phosphate + (S)-methylmalonyl-CoA	<i>pccB1</i> <i>pccB3</i>
ING2E5A_0730	2.3.1.8	Phosphate acetyltransferase	Acetyl-CoA + phosphate $\rightleftharpoons$ CoA + acetyl phosphate	<i>pta</i>
ING2E5A_0728	2.7.2.1	Acetate kinase	ATP + acetate $\rightleftharpoons$ ADP + acetyl phosphate	<i>ackA</i>
<b>Catabolism of Arginine and Aspartate</b>				
ING2E5A_2671	3.5.1.1	L-asparaginase 1	L-asparagine + H <sub>2</sub> O $\rightleftharpoons$ L-aspartate + NH <sub>3</sub>	<i>ansA</i>
ING2E5A_2287	4.3.1.1	Aspartate ammonia-lyase	L-aspartate $\rightleftharpoons$ fumarate + NH <sub>3</sub>	<i>aspA</i>
ING2E5A_3067	1.4.3.16	L-aspartate oxidase	L-aspartate + O <sub>2</sub> $\rightleftharpoons$ iminosuccinate + H <sub>2</sub> O <sub>2</sub>	<i>nadB</i>
ING2E5A_1308	2.6.1.1	Aspartate transaminase	L-aspartate + 2-oxoglutarate $\rightleftharpoons$ oxaloacetate + L-glutamate	<i>aspC</i>
ING2E5A_2589	3.5.3.6	Arginine deiminase	L-arginine + H <sub>2</sub> O $\rightleftharpoons$ L-citrulline + NH <sub>3</sub>	<i>arcA</i>
ING2E5A_0737	6.3.4.5	Argininosuccinate synthase	ATP + L-citrulline + L-aspartate $\rightleftharpoons$ AMP + diphosphate + N(omega)-(L-arginino)succinate	<i>argG</i>
ING2E5A_0731	4.3.2.1	Argininosuccinate lyase	2-(N(omega)-L-arginino)succinate $\rightleftharpoons$ fumarate + L-arginine	<i>argH</i>
ING2E5A_3031, ING2E5A_3032, ING2E5A_3033	1.3.5.1	Succinate dehydrogenase, Succinate dehydrogenase flavoprotein subunit	Succinate + a quinone $\rightleftharpoons$ fumarate + a quinol	<i>sdhA</i>
ING2E5A_0952, ING2E5A_0953	6.2.1.5	Succinate-CoA ligase (ADP-forming)	ATP + succinate + CoA $\rightleftharpoons$ ADP + phosphate + succinyl-CoA	<i>sucD</i> , <i>sucC</i>
ING2E5A_0662	4.2.1.2	Fumarate hydratase	(S)-malate $\rightleftharpoons$ fumarate + H <sub>2</sub> O	<i>fumA</i>
ING2E5A_2725	1.1.1.37	Malate dehydrogenase	(S)-malate + NAD <sup>+</sup> $\rightleftharpoons$ oxaloacetate + NADH	<i>mdh</i>
ING2E5A_1041, ING2E5A_1042	4.1.1.112	Oxaloacetate decarboxylase	Oxaloacetate $\rightleftharpoons$ pyruvate + CO <sub>2</sub>	<i>oadB2-1</i> , <i>oadB2-2</i>
ING2E5A_1825	4.1.1.112	Oxaloacetate decarboxylase	Oxaloacetate $\rightleftharpoons$ pyruvate + CO <sub>2</sub>	<i>oadB</i>
ING2E5A_0270, ING2E5A_1826	6.4.1.1	Pyruvate carboxylase	ATP + pyruvate + HCO <sub>3</sub> <sup>-</sup> $\rightleftharpoons$ ADP + phosphate + oxaloacetate	<i>pycB1</i> , <i>pycB3</i>
ING2E5A_1490, ING2E5A_1805, ING2E5A_1806,	1.2.4.1	Pyruvate dehydrogenase	Pyruvate + [dihydrolipoyllysine-residue acetyltransferase] lipoyllysine $\rightleftharpoons$ [dihydrolipoyllysine-residue acetyltransferase] S-acetyldihydrolipoyllysine + CO(2)	<i>pdhB1</i> , <i>pdhB3</i> , <i>pdhA</i>
ING2E5A_1158, ING2E5A_1264, ING2E5A_1491, ING2E5A_1804	2.3.1.12	Dihydrolipoyllysine-residue acetyltransferase	Acetyl-CoA + enzyme N(6)-(dihydrolipoyl)lysine $\rightleftharpoons$ CoA + enzyme N(6)-(S-acetyldihydrolipoyl)lysine	<i>pdhC1</i> , <i>dlcT</i> , <i>pdhC3</i> , <i>pdhC5</i>
ING2E5A_0730	2.3.1.8	Phosphate acetyltransferase	Acetyl-CoA + phosphate $\rightleftharpoons$ CoA + acetyl phosphate	<i>pta</i>
ING2E5A_0728	2.7.2.1	Acetate kinase	ATP + acetate $\rightleftharpoons$ ADP + acetyl phosphate	<i>ackA</i>

Table 2. Cont.

Locus Tag	EC No.	Enzyme	Reaction	Gene(s)
<b>Butanoate pathway enzymes</b>				
ING2E5A_0177	2.7.2.7	Butyrate kinase	ATP + butanoate $\rightleftharpoons$ ADP + butanoyl phosphate	<i>buk</i>
ING2E5A_0176	2.3.1.19	Phosphate butyryltransferase	Butanoyl-CoA + phosphate $\rightleftharpoons$ CoA + butanoylphosphate	<i>ptb</i>
ING2E5A_0729	1.1.1.157	3-hydroxybutyryl-CoA dehydrogenase	(S)-3-hydroxybutanoyl-CoA + NADP <sup>+</sup> $\rightleftharpoons$ 3-acetoacetyl-CoA + NADPH	<i>hbd</i>
<b>Malic Enzyme Module</b>				
ING2E5A_2598	1.1.1.40	Malate dehydrogenase (oxaloacetate-decarboxylating) (NADP <sup>+</sup> ).	(S)-malate + NADP <sup>+</sup> $\rightleftharpoons$ pyruvate + CO <sub>2</sub> + NADPH, Oxaloacetate $\rightleftharpoons$ pyruvate + CO <sub>2</sub>	<i>maeB</i>
ING2E5A_3022	2.7.9.1	Pyruvate, phosphate dikinase	ATP + pyruvate + phosphate $\rightleftharpoons$ AMP + phosphoenolpyruvate + diphosphate	<i>ppdK</i>

Likewise, glutamine and glutamate may be metabolized to yield succinyl-CoA via 2-oxoglutarate (refer to Table 2). *P. mucosa* encodes the key enzymes methylmalonyl-CoA mutase, methylmalonyl-CoA epimerase and propionyl-CoA carboxylase catalyzing the reactions of succinyl-CoA to (R)-methylmalonyl-CoA and subsequently to propionyl-CoA, respectively. The phosphate-acetyltransferase/acetate kinase module (KEGG M00579) is also able to convert propionyl-CoA to propionate via propionyl-phosphate yielding ATP through substrate-level phosphorylation. According to ExPASy (<https://enzyme.expasy.org>), acetate kinase (EC 2.7.2.1) can also use propanoate as a substrate. Acetate can be converted into the intermediate acetyl-CoA by coupling the reaction of acetate kinase with phosphate acetyltransferase (EC 2.3.1.8) activity. Both, this enzyme and acetate kinase also are important in the production of propanoate. Acetate, propionate, and carbon dioxide were previously identified as fermentation end-products in *P. mucosa* ING2-E5A<sup>T</sup> [25]. As well as the amino acids mentioned above, other amino acids such as proline, glycine, serine, and threonine potentially can be catabolized via glutamate or pyruvate since the corresponding enzymes are encoded in the *P. mucosa* genome.

Genes for the butanoate pathway enzymes butyrate kinase (EC 2.7.2.7) and phosphate butyryltransferase (EC 2.3.1.19) are present in the *P. mucosa* genome but the pathway is incomplete at least as predicted from the genome sequence information. Genes for the enzymes involved in conversion of 3-hydroxybutanoyl-CoA to crotonoyl-CoA (or *vice versa*) and crotonoyl-CoA to butanoyl-CoA could not be identified in the *P. mucosa* ING2-E5A<sup>T</sup> genome. However, it was shown that *P. mucosa* ING2-E5A<sup>T</sup> produced butyric acid in trace amounts when grown on glucose [25]. So far, not all enzymes necessary for butyric acid formation were identified in *P. mucosa*.

Very recently, importance of the phosphotransacetylase/acetate kinase pathway for ATP production was also described in *Porphyromonas gingivalis* which, similar to *P. mucosa*, belongs to the order Bacteroidales [73–76]. In the context of the central fermentation metabolism, it is interesting to note that *P. mucosa* encodes a malic enzyme (EC 1.1.1.40) and pyruvate, phosphate dikinase (EC 2.7.9.1) catalyzing the NAD(P)<sup>+</sup> dependent oxidative decarboxylation of malate to pyruvate [77] and the ATP-consuming conversion of pyruvate to phosphoenolpyruvate, respectively. Mostly, the reaction of malic enzyme contributes to NADPH-generation but under specific conditions, it may also function as anaplerotic node by maintaining suitable levels of TCA cycle intermediates [77]. For *P. mucosa*, the biological importance of the reaction catalyzed by malic enzyme remains to be determined.

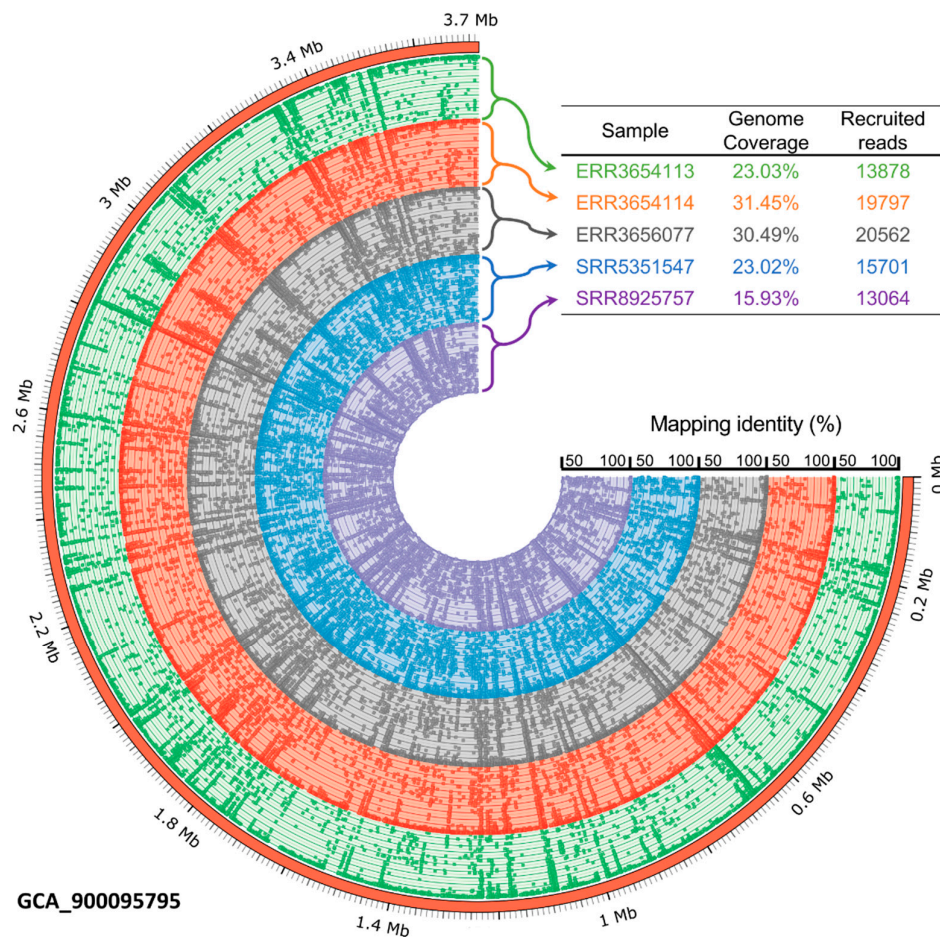
Furthermore, two hydrogenase gene regions were identified in the *P. mucosa* ING2-E5A<sup>T</sup> genome by means of BLASTp analyses and consulting the Hydrogenase Database HydDB [43]. The first hydrogenase, being classified as a HydA (ING2E5A\_0849), belongs to the [FeFe] Group C1; it is histidine kinase-linked and oxygen labile and its role remains unclear so far. The enzyme may serve as a hydrogen sensor in regulatory cascades. The second hydrogenase Hnd consists of four subunits (HndA, HndB, HndC, and HndD; ING2E5A\_0853-0856). The *hnd* genes are located in close proximity to *hydA*. A previous study on the Hnd hydrogenase isolated from *Desulfovibrio fructosovorans* described



the electron-bifurcating function of this enzyme coupling an exergonic redox reaction to an endergonic one allowing energy conservation in anaerobic microorganisms [78].

### 3.6. The Occurrence of *Petrimonas* Relatives in Biogas-Producing Microbial Communities as Deduced from Publicly Available Metagenome Data

To investigate the prevalence and occurrence of *P. mucosa* in biogas-producing microbial communities, metagenome sequences obtained from 257 publicly available datasets originating from the RefSeq collection were mapped onto the ING2-E5A<sup>T</sup> genome applying the program Sparkhit [44]. The mapping approach was aimed at the reconstruction of *Petrimonas* genomes from the previously published metagenome data. A further objective of this approach was to determine the degree of relatedness of species within the analyzed biogas plant microbiomes to the genome of strain ING2-E5A<sup>T</sup>. Fragment recruitments were done by mapping of at maximum ten million sequences from the corresponding metagenome representing the biogas microbial community. Only five samples (accession numbers at the NCBI: ERR3656077, ERR3654114, SRR8925757, ERR3654113, SRR5351547) showing metagenome fragment mapping results of greater than 1% were visualized in the fragment recruitment plot (Figure 5). The complete report is provided as Supporting Information Table S1.



**Figure 5.** Fragment recruitment of publicly available metagenome sequences derived from five different mesophilic biogas-producing microbial communities and mapped onto the *P. mucosa* ING2-E5A<sup>T</sup> genome sequence. Selected samples showing metagenome fragment mapping results of greater than 1% were visualized. The outer ring represents the *P. mucosa* ING2-E5A<sup>T</sup> genome in Mega base pairs (Mb), while remaining colored rings show the coverage of the mapped metagenome reads on the ING2-E5A<sup>T</sup> genome sequence. Abbreviations: GCA, genome collections accession; ERR and SRR stand for run accessions.

The genus *Petrimonas* was detected in 146 out of 257 metagenomes analyzed ( $\geq 0.1\%$  of all metagenome sequences) with different abundance values. Particularly high abundance values ranging between 1.03% and 3.16% of all metagenome sequences were detected in 14 out of 257 different biogas reactor metagenomes (Figure 5). In these metagenomes, the genome of *P. mucosa* ING2-E5A<sup>T</sup> was not only abundant, but also fully covered. The highest number of mapped sequences onto the genome of *P. mucosa* (3.16% of all sequences) was found for the microbiome originating from a production-scale biogas plant located in Germany (Ger-06), operated under mesophilic (40 °C) temperature regime and fed with manure (48%), maize (30%), dung (17%), and grain corn (7%) [79]. Brandt and colleagues [79] also described the widespread occurrence of the genus *Petrimonas* in microbiomes of biogas reactors analyzed within their study. They detected *Petrimonas* spp. in 13 out of 20 German and Swedish biogas plants (BGPs) and presented abundance values of up to 10% (Ger-06) of the entire community.

*Petrimonas* members were also detected in microbiomes originating from a pilot-scale dry anaerobic digester fed with residual biomass from the bovine husbandry (64.2% manure, 18.2% oxidation lagoon water, 3% fresh rumen fluid, 9.6% wood chips, 4.0% corn stover, and 0.9% dust mill) and operated under mesophilic conditions [80]. Interestingly, the highest number of sequences assigned to the *P. mucosa* genome in this reactor (2.87% of all metagenome sequences) was detected a short time after a process disturbance that occurred due to a pump failure during leachate recirculation. Most probably, *P. mucosa* was able to cope with corresponding stress conditions and displaced other competing bacteria that are less stress resistant. This is in line with previous studies that link the genus *Petrimonas* with process disorders [7,9]. Our mapping results also showed an abundance decrease of the genus *Petrimonas* in reactor samples taken after the process disorder (from 1.41% to 1.03%), indicating suppression by other competitive bacteria in the stabilization phase of the biogas process.

Information obtained from the fragment recruitment approach leads to the assumption that *Petrimonas* and its relatives are widespread in biogas plants. *Petrimonas* members probably proliferate when reactor conditions become unstable and the community undergoes a structural change. Under these conditions, they may play an important role in conversion of lignocellulosic biomass and also proteins for subsequent methane production.

### 3.7. The Transcriptional Profile of *Petrimonas* Species

Transcriptome sequence data reflect the transcriptional activity of the species of interest in their respective habitats. To study the *Petrimonas* transcriptional activity in microbiomes in which this species was found with a relative abundance of above 1% (for more details see above), corresponding metatranscriptome data were searched in the publicly available sequence databases. Unfortunately, no metatranscriptome data exist for the metagenomes presented in Figure 5 referring to the fragment recruitment analyses (see above). To gain insights into the transcriptional profile of *Petrimonas* members and their relatives and to understand their role in AD in general, public databases were searched for metatranscriptome datasets representing AD environments. Not all of these datasets also comprise metagenome data. For metagenome data for which also metatranscriptome data was available, only fractions of the *P. mucosa* genome could be assembled from the DNA reads. For this reason, a solely RNA-based analysis was carried out. Transcriptome assemblies were conducted for communities from 13 different biogas and wastewater treatment plants. Sequence reads of the corresponding studies were grouped into 29 datasets, each representing a distinct condition or sampling time point in one of the analyzed reactors. For each of these 29 datasets, reads were mapped back to the assembled transcripts. The ING2-E5A<sup>T</sup> genome was used to assign transcripts that originate from *P. mucosa*. Based on the number of reads that could be mapped to each *P. mucosa* transcript, counts and TPM values were determined for each dataset. Detailed information on process parameters and used substrates recorded for the analyzed biogas reactors were summarized in Supporting Information Table S2 as far as corresponding metadata were documented in the literature.

Transcripts assigned to *Petrimonas* and its relatives were detected in 20 out of 29 metatranscriptome datasets analyzed. The highest number of transcripts (0.23–0.53% of all sequences) that could be

mapped onto the ING2-E5A<sup>T</sup> genome was observed for a mesophilic laboratory-scale biogas plant fed with maize silage and separated digestate from other mesophilic full-scale research biogas plants [81] (Supporting Information Table S3). Grohmann and coworkers [81] reported that the analyzed biogas plant suffered from a process disturbance leading to a strong accumulation of VFA. Acetic and propionic acid concentrations exceeded critical values and were on an alarming level. During the experiment, the reactor was continuously sampled. *Petrimonas* transcripts were consistently detected in all metatranscriptome datasets belonging to the latter study (on average 0.31% of all metatranscriptome sequences taking into account all metagenomes from this study). However, the highest number of transcripts (0.53%) were observed for the sample that was taken when the process had nearly stabilized.

Within the list of highly transcribed genes (Table 3 and Supporting Information Table S4), those encoding proteins associated with hydrolysis of 6-phosphogluconolactone to 6-phosphogluconic acid via the pentose phosphate pathway, such as 6-phosphogluconolactonase Pgl (EC: 3.1.1.31), were observed. Genes involved in oligopeptide utilization such as the ATP-dependent protease *clp* (EC: 3.4.21.-) and the peptidase IV *sppA* gene (EC: 3.4.21.-) were also transcribed. Furthermore, *P. mucosa* transcribed the *mscL* gene encoding the large-conductance mechanosensitive channel. This pore-forming membrane protein is an important component of the cells' management during osmotic stress [82]. Moreover, these nanopores can be utilized for the passage of membrane-impermeable compounds or peptides [83]. The latter molecules were predicted to represent fermentation substrates of *Petrimonas* species. Furthermore, *P. mucosa* transcribed the *susC* gene encoding the TonB-dependent outer-membrane transporter, which is part of PUL clusters presented in Figure 4. Further copies of this gene were also highly transcribed. Transcripts encoding proteins involved in mandatory functions (basic house-keeping functions) such as transcription and translation were not further considered.

**Table 3.** Most actively transcribed genes for enzymes involved in carbohydrate or amino acid utilization in *P. mucosa* which were ranked by TPM values from highest to lowest as deduced from publicly available metatranscriptome dataset (accession numbers: ERR3010911, ERR3010910, ERR3010909, ERR3010916).

Ranking	GenID	Gene Length	Putative Gene Product	Gene	EC Number	Transcriptome Reads Mapped	Normalized Number of Transcripts (TPM)
7	ING2E5A_1308	1191	Aspartate aminotransferase	<i>ast</i>	2.6.1.1	739	21,591
8	ING2E5A_1484	1179	6-phosphogluconolactonase	<i>pgls</i>	3.1.1.31	178	5255
21	ING2E5A_1407	1500	Propionyl-CoA:succinate CoA transferase	<i>scpC</i>	2.8.3.-	186	4328
26	ING2E5A_0536	720	Peptidase E	<i>pepE</i>	3.4.13.21	76	3688
32	ING2E5A_3029	1191	Aspartate aminotransferase	<i>ast</i>	2.6.1.1	113	3315
37	ING2E5A_0388	657	Uracil phosphoribosyltransferase	<i>uprt</i>	2.4.2.9	57	3049
43	ING2E5A_2638	1521	Periplasmic serine endoprotease	<i>degP</i>	3.4.21.107	121	2774
50	ING2E5A_0405	1452	6-phosphogluconate dehydrogenase, decarboxylating	<i>gndA</i>	1.1.1.44	97	2344
85	ING2E5A_0914	3792	TonB-dependent outer-membrane transporter	<i>susC</i>	n.d.	205	1887
94	ING2E5A_0692	1014	ATP-dependent 6-phosphofructokinase	<i>pfk</i>	2.7.1.11	50	1725

n.d. = not determined.

Based on genome interpretation and transcriptional profiling, a lifestyle based on carbohydrate decomposition and oligopeptide fermentation is proposed for *P. mucosa* ING2-E5A<sup>T</sup>. Furthermore, results obtained from metatranscriptome analyses are in line with previous findings showing increased abundance of *Petrimonas* members in reactors operating under stressed conditions followed by accumulation of VFA [10,19]. Results obtained within this study also indicated that *Petrimonas* members are competent in fermenting sugars and amino acids, with acetate and propionic acid as the major fermentation products.

### 3.8. Metaproteome Analysis Revealed the Presence of *Petrimonas* spp. Proteins in Various Biogas-Producing Communities

Metaproteomics became a powerful technique to study the biological functions of microbial communities, to analyze and correlate the taxonomic and functional profiles of microbiomes [84] and also to evaluate the responses of microbial taxa to environmental changes [85,86]. In order to evaluate which proteins/enzymes were expressed by *P. mucosa*, metaproteome data originating from mesophilic and thermophilic biogas reactors (previously published by [53]) were screened for proteins/enzymes that were assigned to this taxon. Further details on operational and process parameters of the investigated BGPs as well as fed substrates are summarized in Figure 6. In total, 11 biogas microbiomes were investigated using SDS-PAGE for prefractionation of proteins and subsequent liquid chromatography (LC) coupled to a high-resolution Orbitrap Elite tandem mass spectrometer (MS/MS). Spectra assigned to *P. mucosa* were identified using the MetaProteomeAnalyzer software [87]. The available datasets were searched with X!tandem and OMSSA against the same protein database complemented by the protein sequences deduced from the *Petrimonas mucosa* genome sequence.

Based on abundances of spectral counts, peptides originating from *P. mucosa* proteins after tryptic digestion were detected in all biogas samples analyzed, showing abundance values in the range of 0.2–1.4% of all spectra identified. The highest number of spectra assigned to *P. mucosa* was detected in BGP\_03\_T1 (1335) representing a mesophilic plug-flow biogas reactor operating at stable process conditions as deduced from the process parameters recorded for this reactor. Among the highest expressed proteins are many of those annotated as house-keeping proteins e.g., uncharacterized outer membrane proteins, and ribosomal proteins of the 30S and 50S ribosomal subunits. Furthermore, several copies of the TonB-dependent outer-membrane transporters (SusC) were found among the most abundant spectra (all are part of different PUL clusters presented in Figure 4: PUL\_02, PUL\_10, PUL\_11, PUL\_24, PUL\_29, PUL\_31) and represent an essential component of *P. mucosa*'s PUL required for the saccharification of complex carbohydrates. This finding indicates the importance of PUL-encoded enzymes for *P. mucosa*'s lifestyle and provides evidence that corresponding enzymes play a central role for this species.



Biogas plant	Reactor type	Temperature (°C)	Substrate								Process parameters					Biogas production								Acid content								Spectra			
			Maize silage (t)	Poultry manure (t)	Rye silage (t)	Liquid manure (m <sup>3</sup> )	Cow dung (t)	Grass silage (t)	Dry chicken faeces (t)	Pig manure (m <sup>3</sup> )	total (t)	OLR (kg <sub>VS</sub> m <sub>r</sub> <sup>-3</sup> d <sup>-1</sup> )	HRT (d)	CH <sub>4</sub> (%)	CO <sub>2</sub> (%)	H <sub>2</sub> S (ppm)	Biogas production (m <sup>3</sup> d <sup>-1</sup> )	Biogas productivity (m <sub>b</sub> <sup>3</sup> m <sub>r</sub> <sup>-3</sup> d <sup>-1</sup> )	pH value (-)	TVFA/TA (-)	TAN (g L <sup>-1</sup> )	Acid content (g L <sup>-1</sup> )	Acetate (%)	Propionate (%)	n-Butyrate (%)	iso-Butyrate (%)	n-Valeriate (%)	iso-Valeriate (%)	Measured spectra (#)	Identified spectra (#)	Spectra assigned to <i>Petrimonas mucosa</i> (#)	Abundance in the microbial community (%)			
BGP_02_T1	CSTR	43.4	18.5		10.0					28.5	3.9	39.6	56.1	39.2	20	3209	2.8	7.9	0.3	2.2	0.8	84	13.7								226164	59822	413	0.7%	
BGP_03_T1	Plug-flow	42.0	6.0	2.2	7.0	1.4		0.5	17.0	5.6	15.8	51.4	47.1	122	2342	8.7	7.8	0.3	1.9	1.9	71	22.5							283551	106828	1335	1.2%			
BGP_04_T1	CSTR	41.7	20.5	3.6	8.2	5.6			37.9	2.7	39.1	51.4	44.8	10	2970	1.0	8.0	0.1	2.5	1.5	73	7.3	4.7	2.7	9.3				257251	94605	681	0.7%			
BGP_05a_T1	Plug-flow	42.2	11.5	2.6					14.1	5.3	23.1	51.0	45.8	78	6605	20.3	8.5	0.6	5.1	10.7	56	25.7	4.2	3.9	1.3	8.6				347811	128715	436	0.3%		
BGP_05b_T1	Plug-flow	42.3	13.1	2.6					15.7	6.1	20.7	51.6	45.6	27	7844	24.1	8.4	0.6	4.7	7.3	50	27.5	3.6	6.1	1.6	11.0				364242	130047	518	0.4%		
BGP_07_T1	CSTR	40.3	16.4	1.8	12.3	3.6			34.0	1.4	51.4	51.7	48.3		4934	2.8	7.8	0.2	2.3	0.5	97	2.9							305821	64472	904	1.4%			
BGP_09_T1	CSTR	40.3	15.3		10.0		7.7		33.0	1.3	49.2	51.2		46	6338	2.8	8.3	0.3	5.9	0.6	84	10.9							333163	86938	305	0.4%			
BGP_10_T1	CSTR	39.7	9.7		7.7	1.4			18.8	3.4	44.3	52.5	44.9	3	2620	1.9	7.8	0.3	3.0	0.4	100								258411	59490	168	0.3%			
BGP_X1_T1	CSTR	40.0	32.0	16.0		32.0			80.0	3.3	37.5	56.0			22800	7.6	8.9		6.0	0.2									271211	77200	191	0.2%			
BGP_X2_T1	CSTR	40.0	26.0					11.0	37.0	4.0	60.0				4315		7.6		0.2										317854	82454	210	0.3%			
BGP_X3_T1	CSTR	43.0	0.000006						6 tons	2.0	84.0						7.4	0.2	0.1 mg/kg	0.1									296114	17679	211	1.2%			

**Figure 6.** Abundance of *P. mucosa* ING2-E5A<sup>T</sup> in biogas-producing communities as deduced from metaproteome analysis. Metaproteome datasets originating from eight different biogas plants represented by 11 reactors published previously [53] were investigated. The measurements were searched with X!tandem and OMSSA against the protein database derived from original data extended with the sequences of *Petrimonas mucosa* proteins. Subsequently, an FDR of 1% was applied and redundant peptide identifications were grouped when they shared the same peptide set. The main process parameters for the investigated biogas plants were adopted from the original publication [53]. Blue and red bars mark the mesophilic or thermophilic process temperature, respectively. Green bars visualize the values of the different parameters. CSTR: continuous stirred tank reactor; OLR: organic loading rate; HRT: hydraulic retention time; VS: volatile solids; TS: total solids; TVFA/TA: total volatile fatty acids to total alkalinity; TAN: total ammonia nitrogen; m<sub>b</sub><sup>3</sup>: cubic meter biogas; m<sub>r</sub><sup>3</sup>: cubic meter reactor volume; kg<sub>VS</sub>: kilogram VS.

#### 4. Conclusions

Integrated omics analyses are useful for a detailed characterization of both, entire microbiomes and individual microbial species enabling new ways to access the microbiological basis for AD. In this study, the genome sequence of *P. mucosa* ING2-E5A<sup>T</sup> isolated from a mesophilic laboratory-scale reactor was analyzed in detail including genome-based metabolic reconstruction. In addition, a polyphasic approach exploiting publicly available metagenome, metatranscriptome, and metaproteome datasets complemented the genome interpretation of the strain. Genome-based phylogenetic classification of *P. mucosa* ING2-E5A<sup>T</sup> positioned this strain in close proximity to *Fermentimonas* and *Proteiniphilum* species belonging to the family Dysgonomonadaceae which is in accordance with the recent taxonomic reclassification within the phylum Bacteroidetes. *Dysgonomonas* species are located in a sister group of the *Petrimonas*/*Fermentimonas*/*Proteiniphilum* cluster whereas *Porphyromonas* species are more distantly related to the Dysgonomonadaceae members.

Insights into the strain's genetic repertoire and metabolic potential revealed that *P. mucosa* ING2-E5A<sup>T</sup> encodes a diverse set of glycosyl-hydrolyses (GH) involved in carbohydrate metabolism as well as pectin degradation. Most of the corresponding genes are organized in Polysaccharide Utilization Loci (PUL) also comprising tandem *susCD*-like genes for a TonB-dependent outer-membrane transporter and a cell surface glycan-binding protein. Furthermore, ING2-E5A<sup>T</sup> possesses genetic determinants needed to utilize amino acids as carbon sources. Its function in the AD process most probably is associated with acidogenesis, which is supported by reconstruction of corresponding metabolic pathways leading to production of acetate, propionate, and carbon dioxide as fermentation end-products.

Finally, publicly available metagenome, metatranscriptome, and metaproteome data were used to follow up the question on spreading of *Petrimonas* and its relatives in different biogas plants and subsequently to study its biological functions in the corresponding environment. The genus *Petrimonas* was prevalent in approximately half of the datasets investigated, supporting the importance of this genus for the AD process. Moreover, fragment recruitment results linked the genus *Petrimonas* with process disorders. Most probably, *P. mucosa* was able to cope with corresponding stress conditions and displaced other competing bacteria that are less stress resistant. Functional profiles of *Petrimonas* members shed new light on the importance of PUL-encoded enzymes for *P. mucosa*'s lifestyle, leading to the evidence that corresponding proteins are indispensable for this species.

In the context of biogas production, an integrated -omics approach is believed to provide deeper insights into functional roles of AD community members, thus facilitating the implementation of new optimization strategies for the biogas process. The knowledge gained from multi-omics analyses will provide the basis for application of the strain in demanding biotechnological processes. *P. mucosa* may also be used as indicator species for unbalanced AD processes in the future. However, further information on the performance of the strain in different AD microbiomes exposed to unfavorable process conditions is needed to precisely determine its functional behavior.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2076-2607/8/12/2024/s1>, Supporting Information Table S1: List of publicly available metagenome datasets originating from the RefSeq used for fragment recruitment analysis; Supporting Information Table S2: Characteristics and fed substrates of the 13 different biogas plants analyzed in the scope of the metatranscriptome analysis; Supporting Information Table S3: Statistics of the obtained and processed metatranscriptome sequences analyzed in order to investigate the transcriptional profile of *Petrimonas* species; Supporting Information Table S4: *P. mucosa* ING2-E5A<sup>T</sup> most actively transcribed genes for enzymes involved in carbohydrate or protein utilization in 29 analyzed datasets originating from different biogas plants.

**Author Contributions:** Conceptualization, A.S. (Andreas Schlüter), A.S. (Alexander Sczyrba), D.W., J.H., I.M., M.K.; Methodology, G.T., I.M., L.H., R.H., T.T., V.Z.; Software, D.W., L.H., R.H., T.T.; Validation, A.S. (Andreas Schlüter), A.S. (Alexander Sczyrba), I.M., Formal Analysis, D.W., G.T., I.M., J.H., R.H., T.T., V.Z., B.B.; Investigation, A.S. (Andreas Schlüter), I.M.; Resources, A.P., A.S. (Alexander Sczyrba), A.S. (Andreas Schlüter), C.S., D.B.; Data Curation, D.W., I.M., L.H., T.T., R.H.; Writing—Original Draft Preparation, A.S. (Andreas Schlüter), G.T., I.M., M.K.; Writing—Review and Editing, all participants; Visualization, D.W., G.T., I.M., L.H., R.H., M.K.; Supervision, A.S. (Andreas Schlüter), A.S. (Alexander Sczyrba), M.K., A.P.; Project Administration, A.P., A.S. (Andreas Schlüter), M.K.; Funding Acquisition, A.P., A.S. (Andreas Schlüter), M.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded in the context of the joint project BIOGAS-CORE granted by the German Federal Ministry of Food and Agriculture (grant no. 22017111) and supported by the German Fachagentur Nachwachsende Rohstoffe. A.S. and J.H. gratefully acknowledge funding by the Federal Ministry of Food and Agriculture (BMP-3, GeneMining). A.S. (Alexander Sczyrba) and A.S. (Andreas Schlüter) acknowledge funding by the European Union's Horizon 2020 Research and Innovation Programme under the Grant agreement No. 818431 (SIMBA, Sustainable Innovation of Microbiome Applications in Food Systems). The bioinformatics support of the BMBF-funded project 'Bielefeld-Gießen Center for Microbial Bioinformatics-BiGi (grant nos. 031A533 and 031L0103)' within the German Network for Bioinformatics Infrastructure (de.NBI) is also gratefully acknowledged. This work was also supported by the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) (031A532B, 031A533A, 031A533B, 031A534A, 031A535A, 031A537A, 031A537B, 031A537C, 031A537D, 031A538A).

**Acknowledgments:** The authors thank Simone Severitt and Nicole Heyer for excellent technical assistance regarding PacBio genome sequencing.

**Conflicts of Interest:** The authors declare that there is no conflict of interest.

## References

1. Venkiteshwaran, K.; Bocher, B.; Maki, J.; Zitomer, D. Relating Anaerobic Digestion Microbial Community and Process Function. *Microbiol. Insights* **2015**, *8*, 37–44. [[PubMed](#)]
2. Schnürer, A. Biogas Production: Microbiology and Technology. *Adv. Biochem. Eng. Biotechnol.* **2016**, *156*, 195–234. [[PubMed](#)]
3. Westerholm, M.; Dolfing, J.; Schnürer, A. Growth Characteristics and Thermodynamics of Syntrophic Acetate Oxidizers. *Environ. Sci. Technol.* **2019**, *53*, 5512–5520. [[CrossRef](#)] [[PubMed](#)]
4. Rajagopal, R.; Saady, N.M.C.; Torrijos, M.; Thanikal, J.V.; Hung, Y.-T. Sustainable Agro-Food Industrial Wastewater Treatment Using High Rate Anaerobic Process. *Water* **2013**, *5*, 292. [[CrossRef](#)]
5. Niu, Q.; Hojo, T.; Qiao, W.; Qiang, H.; Li, Y.-Y. Characterization of Methanogenesis, Acidogenesis and Hydrolysis in Thermophilic Methane Fermentation of Chicken Manure. *Chem. Eng. J.* **2014**, *244*, 587–596. [[CrossRef](#)]
6. Lv, Z.; Wu, X.; Zhou, B.; Wang, Y.; Sun, Y.; Wang, Y.; Chen, Z.; Zhang, J. Effect of One Step Temperature Increment from Mesophilic to Thermophilic Anaerobic Digestion on the Linked Pattern between Bacterial and Methanogenic Communities. *Bioresour. Technol.* **2019**, *292*, 121968. [[CrossRef](#)]
7. He, Q.; Li, L.; Zhao, X.; Qu, L.; Wu, D.; Peng, X. Investigation of Foaming Causes in Three Mesophilic Food Waste Digesters: Reactor Performance and Microbial Analysis. *Sci. Rep.* **2017**, *7*, 13701. [[CrossRef](#)]
8. Yenigün, O.; Demirel, B. Ammonia Inhibition in Anaerobic Digestion: A Review. *Process. Biochem.* **2013**, *48*, 901–911. [[CrossRef](#)]
9. Maus, I.; Klocke, M.; Derenkó, J.; Stolze, Y.; Beckstette, M.; Jost, C.; Wibberg, D.; Blom, J.; Henke, C.; Willenbücher, K.; et al. Impact of Process Temperature and Organic Loading Rate on Cellulolytic/Hydrolytic Biofilm Microbiomes during Biomethanation of Ryegrass Silage Revealed by Genome-Centered Metagenomics and Metatranscriptomics. *Environ. Microbiome* **2020**, *15*, 7. [[CrossRef](#)]
10. He, Q.; Li, L.; Peng, X. Early Warning Indicators and Microbial Mechanisms for Process Failure Due to Organic Overloading in Food Waste Digesters. *J. Environ. Eng.* **2017**, *143*, 04017077. [[CrossRef](#)]
11. Rajagopal, R.; Massé, D.I.; Singh, G. A Critical Review on Inhibition of Anaerobic Digestion Process by Excess Ammonia. *Bioresour. Technol.* **2013**, *143*, 632–641. [[CrossRef](#)] [[PubMed](#)]
12. Campanaro, S.; Treu, L.; Kougias, P.G.; De Francisci, D.; Valle, G.; Angelidaki, I. Metagenomic Analysis and Functional Characterization of the Biogas Microbiome Using High Throughput Shotgun Sequencing and a Novel Binning Strategy. *Biotechnol. Biofuels* **2016**, *9*, 26. [[CrossRef](#)] [[PubMed](#)]
13. Müller, B.; Sun, L.; Westerholm, M.; Schnürer, A. Bacterial Community Composition and Fhs Profiles of Low- and High-Ammonia Biogas Digesters Reveal Novel Syntrophic Acetate-Oxidising Bacteria. *Biotechnol. Biofuels* **2016**, *9*, 48. [[CrossRef](#)] [[PubMed](#)]
14. Treu, L.; Kougias, P.G.; Campanaro, S.; Bassani, I.; Angelidaki, I. Deeper Insight into the Structure of the Anaerobic Digestion Microbial Community; the Biogas Microbiome Database Is Expanded with 157 New Genomes. *Bioresour. Technol.* **2016**, *216*, 260–266. [[CrossRef](#)]
15. Stolze, Y.; Bremges, A.; Maus, I.; Pühler, A.; Sczyrba, A.; Schlüter, A. Targeted in Situ Metatranscriptomics for Selected Taxa from Mesophilic and Thermophilic Biogas Plants. *Microb. Biotechnol.* **2018**, *11*, 667–679. [[CrossRef](#)]

16. Maus, I.; Koeck, D.E.; Cibis, K.G.; Hahnke, S.; Kim, Y.S.; Langer, T.; Kreubel, J.; Erhard, M.; Bremges, A.; Off, S.; et al. Unraveling the Microbiome of a Thermophilic Biogas Plant by Metagenome and Metatranscriptome Analysis Complemented by Characterization of Bacterial and Archaeal Isolates. *Biotechnol. Biofuels* **2016**, *9*, 171. [[CrossRef](#)]
17. Fischer, M.A.; Güllert, S.; Refai, S.; Künzel, S.; Deppenmeier, U.; Streit, W.R.; Schmitz, R.A. Long-term Investigation of Microbial Community Composition and Transcription Patterns in a Biogas Plant Undergoing Ammonia Crisis. *Microb. Biotechnol.* **2018**, *12*, 305–323. [[CrossRef](#)]
18. Campanaro, S.; Treu, L.; Rodriguez-R, L.M.; Kovalovszki, A.; Ziels, R.M.; Maus, I.; Zhu, X.; Kougias, P.G.; Basile, A.; Luo, G.; et al. The Anaerobic Digestion Microbiome: New Insights from the Biogas Microbiome by Comprehensive Genome-Resolved Metagenomics of Nearly 1600 Species Originating from Multiple Anaerobic Digesters. *Biotechnol. Biofuels* **2020**, *13*, 25. [[CrossRef](#)]
19. Zhang, J.; Loh, K.-C.; Lee, J.; Wang, C.-H.; Dai, Y.; Wah Tong, Y. Three-Stage Anaerobic Co-Digestion of Food Waste and Horse Manure. *Sci. Rep.* **2017**, *7*, 1269. [[CrossRef](#)]
20. Sun, L.; Pope, P.B.; Eijsink, V.G.H.; Schnürer, A. Characterization of Microbial Community Structure during Continuous Anaerobic Digestion of Straw and Cow Manure. *Microb. Biotechnol.* **2015**, *8*, 815–827. [[CrossRef](#)]
21. Moset, V.; Poulsen, M.; Wahid, R.; Højberg, O.; Møller, H.B. Mesophilic versus Thermophilic Anaerobic Digestion of Cattle Manure: Methane Productivity and Microbial Ecology. *Microb. Biotechnol.* **2015**, *8*, 787–800. [[CrossRef](#)] [[PubMed](#)]
22. Goux, X.; Calusinska, M.; Lemaigre, S.; Marynowska, M.; Klocke, M.; Udelhoven, T.; Benizri, E.; Delfosse, P. Microbial Community Dynamics in Replicate Anaerobic Digesters Exposed Sequentially to Increasing Organic Loading Rate, Acidosis, and Process Recovery. *Biotechnol. Biofuels* **2015**, *8*, 122. [[CrossRef](#)] [[PubMed](#)]
23. Hofstad, T.; Olsen, I.; Eribe, E.R.; Falsen, E.; Collins, M.D.; Lawson, P.A. *Dysgonomonas* Gen. Nov. to Accommodate *Dysgonomonas Gadei* Sp. Nov., an Organism Isolated from a Human Gall Bladder, and *Dysgonomonas Capnocytophagoides* (Formerly CDC Group DF-3). *Int. J. Syst. Evol. Microbiol.* **2000**, *50*, 2189–2195. [[CrossRef](#)] [[PubMed](#)]
24. Chen, S.; Dong, X. *Proteiniphilum Acetatigenes* Gen. Nov., Sp. Nov., from a UASB Reactor Treating Brewery Wastewater. *Int. J. Syst. Evol. Microbiol.* **2005**, *55*, 2257–2261. [[CrossRef](#)] [[PubMed](#)]
25. Hahnke, S.; Langer, T.; Koeck, D.E.; Klocke, M. Description of *Proteiniphilum Saccharofermentans* Sp. Nov., *Petrimonas Mucosa* Sp. Nov. and *Fermentimonas Caenicola* Gen. Nov., Sp. Nov., Isolated from Mesophilic Laboratory-Scale Biogas Reactors, and Emended Description of the Genus *Proteiniphilum*. *Int. J. Syst. Evol. Microbiol.* **2016**, *66*, 1466–1475. [[CrossRef](#)] [[PubMed](#)]
26. Grabowski, A.; Tindall, B.J.; Bardin, V.; Blanchet, D.; Jeanthon, C. *Petrimonas Sulfuriphila* Gen. Nov., Sp. Nov., a Mesophilic Fermentative Bacterium Isolated from a Biodegraded Oil Reservoir. *Int. J. Syst. Evol. Microbiol.* **2005**, *55*, 1113–1121. [[CrossRef](#)]
27. García-López, M.; Meier-Kolthoff, J.P.; Tindall, B.J.; Gronow, S.; Woyke, T.; Kyrpides, N.C.; Hahnke, R.L.; Göker, M. Analysis of 1000 Type-Strain Genomes Improves Taxonomic Classification of Bacteroidetes. *Front. Microbiol.* **2019**, *10*, 2083. [[CrossRef](#)]
28. Sakamoto, M. The Family Porphyromonadaceae. In *The Prokaryotes: Other Major Lineages of Bacteria and the Archaea*; Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 811–824.
29. Tomazetto, G.; Hahnke, S.; Wibberg, D.; Pühler, A.; Klocke, M.; Schlüter, A. *Proteiniphilum Saccharofermentans* Str. M3/6T Isolated from a Laboratory Biogas Reactor Is Versatile in Polysaccharide and Oligopeptide Utilization as Deduced from Genome-Based Metabolic Reconstructions. *Biotechnol. Rep.* **2018**, *18*, e00254. [[CrossRef](#)]
30. Koeck, D.E.; Zverlov, V.V.; Liebl, W.; Schwarz, W.H. Comparative Genotyping of *Clostridium Thermocellum* Strains Isolated from Biogas Plants: Genetic Markers and Characterization of Cellulolytic Potential. *Syst. Appl. Microbiol.* **2014**, *37*, 311–319. [[CrossRef](#)]
31. Cibis, K.G.; Gneipel, A.; König, H. Isolation of Acetic, Propionic and Butyric Acid-Forming Bacteria from Biogas Plants. *J. Biotechnol.* **2016**, *220*, 51–63. [[CrossRef](#)]
32. Tejerizo, G.T.; Kim, Y.S.; Maus, I.; Wibberg, D.; Winkler, A.; Off, S.; Pühler, A.; Scherer, P.; Schlüter, A. Genome Sequence of *Methanobacterium Congolense* Strain Buetzberg, a Hydrogenotrophic, Methanogenic Archaeon, Isolated from a Mesophilic Industrial-Scale Biogas Plant Utilizing Bio-Waste. *J. Biotechnol.* **2017**, *247*, 1–5. [[CrossRef](#)] [[PubMed](#)]



33. Pechtl, A.; Rückert, C.; Maus, I.; Koeck, D.E.; Trushina, N.; Kornberger, P.; Schwarz, W.H.; Schlüter, A.; Liebl, W.; Zverlov, V.V. Complete Genome Sequence of the Novel Cellulolytic, Anaerobic, Thermophilic Bacterium *Herbivorax Saccincola* Type Strain GGR1, Isolated from a Lab Scale Biogas Reactor as Established by Illumina and Nanopore MinION Sequencing. *Genome Announc.* **2018**, *6*, e01493-17. [[CrossRef](#)] [[PubMed](#)]
34. Maus, I.; Bremges, A.; Stolze, Y.; Hahnke, S.; Cibis, K.G.; Koeck, D.E.; Kim, Y.S.; Kreubel, J.; Hassa, J.; Wibberg, D.; et al. Genomics and Prevalence of Bacterial and Archaeal Isolates from Biogas-Producing Microbiomes. *Biotechnol. Biofuels* **2017**, *10*, 264. [[CrossRef](#)] [[PubMed](#)]
35. Li, H.; Durbin, R. Fast and Accurate Long-Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **2010**, *26*, 589–595. [[CrossRef](#)] [[PubMed](#)]
36. Koboldt, D.C.; Zhang, Q.; Larson, D.E.; Shen, D.; McLellan, M.D.; Lin, L.; Miller, C.A.; Mardis, E.R.; Ding, L.; Wilson, R.K. VarScan 2: Somatic Mutation and Copy Number Alteration Discovery in Cancer by Exome Sequencing. *Genome Res.* **2012**, *22*, 568–576. [[CrossRef](#)] [[PubMed](#)]
37. Nelkner, J.; Henke, C.; Lin, T.W.; Pätzold, W.; Hassa, J.; Jaenicke, S.; Grosch, R.; Pühler, A.; Sczyrba, A.; Schlüter, A. Effect of Long-Term Farming Practices on Agricultural Soil Microbiome Members Represented by Metagenomically Assembled Genomes (MAGs) and Their Predicted Plant-Beneficial Genes. *Genes* **2019**, *10*, 424. [[CrossRef](#)]
38. Meyer, F.; Goesmann, A.; McHardy, A.C.; Bartels, D.; Bekel, T.; Clausen, J.; Kalinowski, J.; Linke, B.; Rupp, O.; Giegerich, R.; et al. GenDB—An Open Source Genome Annotation System for Prokaryote Genomes. *Nucleic Acids Res.* **2003**, *31*, 2187–2195. [[CrossRef](#)]
39. Arndt, D.; Grant, J.R.; Marcu, A.; Sajed, T.; Pon, A.; Liang, Y.; Wishart, D.S. PHASTER: A Better, Faster Version of the PHAST Phage Search Tool. *Nucleic Acids Res.* **2016**, *44*, W16–W21. [[CrossRef](#)]
40. Dhillon, B.K.; Laird, M.R.; Shay, J.A.; Winsor, G.L.; Lo, R.; Nizam, F.; Pereira, S.K.; Waglechner, N.; McArthur, A.G.; Langille, M.G.I.; et al. IslandViewer 3: More Flexible, Interactive Genomic Island Discovery, Visualization and Analysis. *Nucleic Acids Res.* **2015**, *43*, W104–W108. [[CrossRef](#)]
41. Blom, J.; Kreis, J.; Spänig, S.; Juhre, T.; Bertelli, C.; Ernst, C.; Goesmann, A. EDGAR 2.0: An Enhanced Software Platform for Comparative Gene Content Analyses. *Nucleic Acids Res.* **2016**, *44*, W22–W28. [[CrossRef](#)]
42. Yin, Y.; Mao, X.; Yang, J.; Chen, X.; Mao, F.; Xu, Y. DbCAN: A Web Resource for Automated Carbohydrate-Active Enzyme Annotation. *Nucleic Acids Res.* **2012**, *40*, W445–W451. [[CrossRef](#)] [[PubMed](#)]
43. Søndergaard, D.; Pedersen, C.N.S.; Greening, C. HydDB: A Web Tool for Hydrogenase Classification and Analysis. *Sci. Rep.* **2016**, *6*, 34212. [[CrossRef](#)] [[PubMed](#)]
44. Huang, L.; Krüger, J.; Sczyrba, A. Analyzing Large Scale Genomic Data on the Cloud with Sparkhit. *Bioinformatics* **2018**, *34*, 1457–1465. [[CrossRef](#)] [[PubMed](#)]
45. Niu, B.; Zhu, Z.; Fu, L.; Wu, S.; Li, W. FR-HIT, a Very Fast Program to Recruit Metagenomic Reads to Homologous Reference Genomes. *Bioinformatics* **2011**, *27*, 1704–1705. [[CrossRef](#)] [[PubMed](#)]
46. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An Information Aesthetic for Comparative Genomics. *Genome Res.* **2009**, *19*, 1639–1645. [[CrossRef](#)] [[PubMed](#)]
47. Trim Galore. Available online: <https://www.psc.edu/user-resources/software/trim-galore> (accessed on 20 October 2020).
48. Kopylova, E.; Noé, L.; Touzet, H. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **2012**, *28*, 3211–3217. [[CrossRef](#)]
49. Bushmanova, E.; Antipov, D.; Lapidus, A.; Prjibelski, A.D. RnaSPAdes: A de Novo Transcriptome Assembler and Its Application to RNA-Seq Data. *Gigascience* **2019**, *8*, giz100. [[CrossRef](#)]
50. Bray, N.L.; Pimentel, H.; Melsted, P.; Pachter, L. Near-Optimal Probabilistic RNA-Seq Quantification. *Nat. Biotechnol.* **2016**, *34*, 525–527. [[CrossRef](#)]
51. Langmead, B.; Salzberg, S.L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
52. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and Applications. *BMC Bioinform.* **2009**, *10*, 421. [[CrossRef](#)]
53. Heyer, R.; Schallert, K.; Siewert, C.; Kohrs, F.; Greve, J.; Maus, I.; Klang, J.; Klocke, M.; Heiermann, M.; Hoffmann, M.; et al. Metaproteome Analysis Reveals That Syntrophy, Competition, and Phage-Host Interaction Shape Microbial Communities in Biogas Plants. *Microbiome* **2019**, *7*, 69. [[CrossRef](#)] [[PubMed](#)]
54. Craig, R.; Beavis, R.C. TANDEM: Matching Proteins with Tandem Mass Spectra. *Bioinformatics* **2004**, *20*, 1466–1467. [[CrossRef](#)] [[PubMed](#)]

55. Geer, L.Y.; Markey, S.P.; Kowalak, J.A.; Wagner, L.; Xu, M.; Maynard, D.M.; Yang, X.; Shi, W.; Bryant, S.H. Open Mass Spectrometry Search Algorithm. *J. Proteome Res.* **2004**, *3*, 958–964. [[CrossRef](#)] [[PubMed](#)]
56. Heyer, R.; Schallert, K.; Büdel, A.; Zoun, R.; Dorl, S.; Behne, A.; Kohrs, F.; Püttker, S.; Siewert, C.; Muth, T.; et al. A Robust and Universal Metaproteomics Workflow for Research Studies and Routine Diagnostics Within 24 h Using Phenol Extraction, FASP Digest, and the MetaProteomeAnalyzer. *Front. Microbiol.* **2019**, *10*, 1883. [[CrossRef](#)] [[PubMed](#)]
57. Rademacher, A.; Zakrzewski, M.; Schlüter, A.; Schönberg, M.; Szczepanowski, R.; Goesmann, A.; Pühler, A.; Klocke, M. Characterization of Microbial Biofilms in a Thermophilic Biogas System by High-Throughput Metagenome Sequencing. *FEMS Microbiol. Ecol.* **2012**, *79*, 785–799. [[CrossRef](#)] [[PubMed](#)]
58. Hanreich, A.; Schimpf, U.; Zakrzewski, M.; Schlüter, A.; Benndorf, D.; Heyer, R.; Rapp, E.; Pühler, A.; Reichl, U.; Klocke, M. Metagenome and Metaproteome Analyses of Microbial Communities in Mesophilic Biogas-Producing Anaerobic Batch Fermentations Indicate Concerted Plant Carbohydrate Degradation. *Syst. Appl. Microbiol.* **2013**, *36*, 330–338. [[CrossRef](#)]
59. Stolze, Y.; Bremges, A.; Ruming, M.; Henke, C.; Maus, I.; Pühler, A.; Sczyrba, A.; Schlüter, A. Identification and Genome Reconstruction of Abundant Distinct Taxa in Microbiomes from One Thermophilic and Three Mesophilic Production-Scale Biogas Plants. *Biotechnol. Biofuels* **2016**, *9*, 156. [[CrossRef](#)]
60. Schlüter, A.; Bekel, T.; Diaz, N.N.; Dondrup, M.; Eichenlaub, R.; Gartemann, K.-H.; Krahn, I.; Krause, L.; Krömeke, H.; Kruse, O.; et al. The Metagenome of a Biogas-Producing Microbial Community of a Production-Scale Biogas Plant Fermenter Analysed by the 454-Pyrosequencing Technology. *J. Biotechnol.* **2008**, *136*, 77–90. [[CrossRef](#)]
61. Pikuta, E.V.; Lyu, Z.; Hoover, R.B.; Liu, Y.; Patel, N.B.; Busse, H.J.; Lawson, P.A. *Williamwhitmania Taraxaci* Gen. Nov., Sp. Nov., a Proteolytic Anaerobe with a Novel Type of Cytology from Lake Untersee in Antarctica, Description of *Williamwhitmaniaceae* Fam. Nov., and Emendation of the Order Bacteroidales Krieg 2012. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 4132–4145. [[CrossRef](#)]
62. Ormerod, K.L.; Wood, D.L.A.; Lachner, N.; Gellatly, S.L.; Daly, J.N.; Parsons, J.D.; Dal’Molin, C.G.O.; Palfreyman, R.W.; Nielsen, L.K.; Cooper, M.A.; et al. Genomic Characterization of the Uncultured Bacteroidales Family S24-7 Inhabiting the Guts of Homeothermic Animals. *Microbiome* **2016**, *4*, 36. [[CrossRef](#)]
63. CAZyedia Consortium. Ten Years of CAZyedia: A Living Encyclopedia of Carbohydrate-Active Enzymes. *Glycobiology* **2018**, *28*, 3–8. [[CrossRef](#)] [[PubMed](#)]
64. Chernysheva, N.; Bystritskaya, E.; Stenkova, A.; Golovkin, I.; Nedashkovskaya, O.; Isaeva, M. Comparative Genomics and CAZyme Genome Repertoires of Marine *Zobellia Amurskyensis* KMM 3526T and *Zobellia Laminariae* KMM 3676T. *Mar. Drugs* **2019**, *17*, 661. [[CrossRef](#)] [[PubMed](#)]
65. Leis, B.; Held, C.; Bergkemper, F.; Dennemarck, K.; Steinbauer, R.; Reiter, A.; Mechelke, M.; Moerch, M.; Graubner, S.; Liebl, W.; et al. Comparative Characterization of All Cellulosomal Cellulases from *Clostridium Thermocellum* Reveals High Diversity in Endoglucanase Product Formation Essential for Complex Activity. *Biotechnol. Biofuels* **2017**, *10*, 240. [[CrossRef](#)] [[PubMed](#)]
66. Artzi, L.; Bayer, E.A.; Morais, S. Cellulosomes: Bacterial Nanomachines for Dismantling Plant Polysaccharides. *Nat. Rev. Microbiol.* **2017**, *15*, 83–95. [[CrossRef](#)] [[PubMed](#)]
67. Naas, A.E.; Mackenzie, A.K.; Mravec, J.; Schückel, J.; Willats, W.G.T.; Eijsink, V.G.H.; Pope, P.B. Do Rumen Bacteroidetes Utilize an Alternative Mechanism for Cellulose Degradation? *mBio* **2014**, *5*, e01401-14. [[CrossRef](#)] [[PubMed](#)]
68. Grondin, J.M.; Tamura, K.; Déjean, G.; Abbott, D.W.; Brumer, H. Polysaccharide Utilization Loci: Fueling Microbial Communities. *J. Bacteriol.* **2017**, *199*, e00860-16. [[CrossRef](#)] [[PubMed](#)]
69. Nakjang, S.; Ndeh, D.A.; Wipat, A.; Bolam, D.N.; Hirt, R.P. A Novel Extracellular Metallopeptidase Domain Shared by Animal Host-Associated Mutualistic and Pathogenic Microbes. *PLoS ONE* **2012**, *7*, e30287. [[CrossRef](#)]
70. Rawlings, N.D.; Barrett, A.J. Families of Serine Peptidases. *Methods Enzym.* **1994**, *244*, 19–61.
71. Accetto, T.; Avguštin, G. Polysaccharide Utilization Locus and CAZYme Genome Repertoires Reveal Diverse Ecological Adaptation of *Prevotella* Species. *Syst. Appl. Microbiol.* **2015**, *38*, 453–461. [[CrossRef](#)]
72. Terrapon, N.; Lombard, V.; Drula, É.; Lapébie, P.; Al-Masaudi, S.; Gilbert, H.J.; Henrissat, B. PULDB: The Expanded Database of Polysaccharide Utilization Loci. *Nucleic Acids Res.* **2018**, *46*, D677–D683. [[CrossRef](#)]
73. Yoshida, Y.; Sato, M.; Nonaka, T.; Hasegawa, Y.; Kezuka, Y. Characterization of the Phosphotransacetylase-Acetate Kinase Pathway for ATP Production in *Porphyromonas Gingivalis*. *J. Oral Microbiol.* **2019**, *11*, 1588086. [[CrossRef](#)] [[PubMed](#)]

74. Hendrickson, E.L.; Xia, Q.; Wang, T.; Lamont, R.J.; Hackett, M. Pathway Analysis for Intracellular *Porphyromonas Gingivalis* Using a Strain ATCC 33277 Specific Database. *BMC Microbiol.* **2009**, *9*, 185. [[CrossRef](#)] [[PubMed](#)]
75. Takahashi, N.; Sato, T.; Yamada, T. Metabolic Pathways for Cytotoxic End Product Formation from Glutamate- and Aspartate-Containing Peptides by *Porphyromonas Gingivalis*. *J. Bacteriol.* **2000**, *182*, 4704–4710. [[CrossRef](#)] [[PubMed](#)]
76. Takahashi, N.; Sato, T. Preferential Utilization of Dipeptides by *Porphyromonas Gingivalis*. *J. Dent. Res.* **2001**, *80*, 1425–1429. [[CrossRef](#)]
77. Spaans, S.K.; Weusthuis, R.A.; van der Oost, J.; Kengen, S.W.M. NADPH-Generating Systems in Bacteria and Archaea. *Front. Microbiol.* **2015**, *6*, 742. [[CrossRef](#)]
78. Kpebe, A.; Benvenuti, M.; Guendon, C.; Rebai, A.; Fernandez, V.; Le Laz, S.; Etienne, E.; Guigliarelli, B.; García-Molina, G.; de Lacey, A.L.; et al. A New Mechanistic Model for an O<sub>2</sub>-Protected Electron-Bifurcating Hydrogenase, Hnd from *Desulfovibrio Fructosovorans*. *Biochim. Biophys. Acta Bioenerg.* **2018**, *1859*, 1302–1312. [[CrossRef](#)]
79. Brandt, C.; Bongcam-Rudloff, E.; Müller, B. Abundance Tracking by Long-Read Nanopore Sequencing of Complex Microbial Communities in Samples from 20 Different Biogas/Wastewater Plants. *Appl. Sci.* **2020**, *10*, 7518. [[CrossRef](#)]
80. Senés-Guerrero, C.; Colón-Contreras, F.A.; Reynoso-Lobo, J.F.; Tinoco-Pérez, B.; Siller-Cepeda, J.H.; Pacheco, A. Biogas-Producing Microbial Composition of an Anaerobic Digester and Associated Bovine Residues. *Microbiol. Open* **2019**, *8*, e00854. [[CrossRef](#)]
81. Grohmann, A.; Fehrmann, S.; Vainshtein, Y.; Haag, N.L.; Wiese, F.; Stevens, P.; Naegele, H.-J.; Oechsner, H.; Hartsch, T.; Sohn, K.; et al. Microbiome Dynamics and Adaptation of Expression Signatures during Methane Production Failure and Process Recovery. *Bioresour. Technol.* **2018**, *247*, 347–356. [[CrossRef](#)]
82. Bremer, E.; Krämer, R. Responses of Microorganisms to Osmotic Stress. *Annu. Rev. Microbiol.* **2019**, *73*, 313–334. [[CrossRef](#)]
83. Booth, I.R.; Blount, P. The MscS and MscL Families of Mechanosensitive Channels Act as Microbial Emergency Release Valves. *J. Bacteriol.* **2012**, *194*, 4802–4809. [[CrossRef](#)] [[PubMed](#)]
84. Heyer, R.; Kohrs, F.; Reichl, U.; Benndorf, D. Metaproteomics of Complex Microbial Communities in Biogas Plants. *Microb. Biotechnol.* **2015**, *8*, 749–763. [[CrossRef](#)] [[PubMed](#)]
85. Kleiner, M. Metaproteomics: Much More than Measuring Gene Expression in Microbial Communities. *mSystems* **2019**, *4*, e00115-19. [[CrossRef](#)] [[PubMed](#)]
86. Chiapello, M.; Zampieri, E.; Mello, A. A Small Effort for Researchers, a Big Gain for Soil Metaproteomics. *Front. Microbiol.* **2020**, *11*, 88. [[CrossRef](#)]
87. Muth, T.; Behne, A.; Heyer, R.; Kohrs, F.; Benndorf, D.; Hoffmann, M.; Lehtevä, M.; Reichl, U.; Martens, L.; Rapp, E. The MetaProteomeAnalyzer: A Powerful Open-Source Software Suite for Metaproteomics Data Analysis and Interpretation. *J. Proteome Res.* **2015**, *14*, 1557–1565. [[CrossRef](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).