Journal of Biotechnology 339 (2021) 53–64

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Genome-centric investigation of anaerobic digestion using sustainable second and third generation substrates

Roland Wirth [a,b], Bernadett Pap [a], Dénes Dudits [a], Balázs Kakuk [b,1], Zoltán Bagi [b], Prateek Shetty [a], Kornél L. Kovács [b,c], Gergely Maróti [a,d,*]

[a] Institute of Plant Biology, Biological Research Centre, Szeged, Hungary
[b] Department of Biotechnology, University of Szeged, Szeged, Hungary
[c] Department of Oral Biology and Experimental Dental Research, Faculty of Dentistry, University of Szeged, Szeged, Hungary
[d] Faculty of Water Sciences, University of Public Service, Baja, Hungary

ARTICLE INFO

Keywords:
Willow shrub
Co-Digestion
Microalgal biomass
Metagenomics
Biogas
Waste biomasses

ABSTRACT

Biogas production through co-digestion of second and third generation substrates is an environmentally sustainable approach. Green willow biomass, chicken manure waste and microalgae biomass substrates were combined in the anaerobic digestion experiments. Biochemical methane potential test showed that biogas yields of co-digestions were significantly higher compared to the yield when energy willow was the sole substrate. To scale up the experiment continuous stirred-tank reactors (CSRTs) are employed, digestion parameters are monitored. Furthermore, genome-centric metagenomics approach was employed to gain functional insight into the complex anaerobic decomposing process. This revealed the importance of Firmicutes, Actinobacteria, Proteobacteria and Bacteroidetes phyla as major bacterial participants, while Methanomicrobia and Methanobacteria represented the archaeal constituents of the communities. The bacterial phyla were shown to perform the carbohydrate hydrolysis. Among the representatives of long-chain carbohydrate hydrolysing microbes Bin_61: Clostridia is newly identified metagenome assembled genome (MAG) and Bin_13: *DTU010 sp900018335* is common and abundant in all CSTRs. Methanogenesis was linked to the slow-growing members of the community, where hydrogenotrophic methanogen species *Methanoculleus* (Bin_10) and *Methanobacterium* (Bin_4) predominate. A sensitive balance between H_2 producers and consumers was shown to be critical for stable biomethane production and efficient waste biodegradation.

1. Background

The global demand for clean energy has led to an increased attention on the sustainability of energy supplies. Biogas is a promising candidate through bio-waste utilization, nutrient recycling and decrease of greenhouse gas (GHG) emissions, therefore biogas production has been proposed as an important component of circular economy strategies (Kougias and Angelidaki, 2018). Biogas can be directly used for power generation, while the digestate is a valuable fertilizer or soil conditioner in agriculture. Compared to other renewable energy sources (i.e. wind, solar energy), biogas production is independent of seasonal fluctuations and can be securely produced since biomass is locally available and abundant. Thus, biomass utilization through technically flexible anaerobic digestion (AD) is considered a competitive renewable energy generation approach (Kougias and Angelidaki, 2018).

Based on feedstock and method of production, biomasses used as a source of biogas are classified in different groups named as first, second and third generation (Alalwan et al., 2019; Zabed et al., 2019). First generation biogas sources are edible biomasses rich in starch and sugar, which increases the cost of production and causes inefficient utilization of resources and energy spent in cultivating crops. Specifically, using edible biomass competes with food crops, requires significant amount of fertilizer, pesticides and water, large areas of cropland (Kakuk et al., 2021; Rulli et al., 2016). Because of these disadvantages policymakers and plant operators are pursued to replace this kind of biomasses (Xue et al., 2020). Second generation biogas sources are more renewable alternatives by utilizing inedible lignocellulosic materials such as crop wastes, sawdust, low-priced woods (Neshat et al., 2017). This

* Corresponding author at:
E-mail address: maroti.gergely@brc.hu (G. Maróti).
[1] Present address: Department of Medical Biology, University of Szeged, Szeged, Hungary

https://doi.org/10.1016/j.jbiotec.2021.08.002
Received 6 April 2021; Received in revised form 23 July 2021; Accepted 3 August 2021
Available online 8 August 2021

generation overcomes the drawbacks of the first generation, like the net emitted/consumed carbon is neutral or even negative and can be produced in non-cultivable lands. Microalgae represent third-generation biomasses (Montingelli et al., 2015; Wirth et al., 2018b). Microalgae have excellent potential to produce special nutritional products like vitamins and lipids via CO_2 consumption due to their photosynthetic activity (Dębowski et al., 2013). Moreover, microalgae have a higher biomass productivity than that of terrestrial crops and can be cultivated all year around (Klassen et al., 2016).

To ensure sustainability and efficiency of biogas production using second and third generation biomasses two criteria must be taken into account: productivity and degradability (Kafle and Kim, 2013). Hybrid energy willow as a second generation biomass is resistant to diseases and has a fast growing ability (Cseri et al., 2020; Dudits et al., 2016). However, AD of woody biomass is not considered technically feasible as monosubstrate, due to many factors that influence anaerobic digestibility, such as particle size and proportion of compact carbohydrate structures (i.e. hemicellulose and lignin). In order to improve digestibility different biological (microbial, enzymatic), chemical (acid, alkaline) and thermal (steam explosion, extrusion) pre-treatment methods have been tested. These processes require additional energy inputs beside being complicated and costly (Abraham et al., 2020). Along with its structural features, the high C/N ratio (stem: 40–90, leaf: 10–20 depends on its age and genotype) of willow biomass also limits its efficient use in biogas producing processes (Paul and Dutta, 2018). An optimal C/N ratio (20−30:1) of biomass promotes efficient methane production. In a previous experiment linking to the present study, two key factors affected the degradability of willow biomass: the low lignin and high "soluble" (oligosaccharide and protein) contents (Kakuk et al., 2021). Co-digestion with nitrogen-rich biomass can provide optimal C/N ratio and additional nutrient content, thus contributes to increasing the stability of biogas production through synergistic effects (Neshat et al., 2017). Third-generation microalgal biomass (C/N: 5–20; depending on the microalga strain and cultivation parameters) is a promising candidate for co-digestion to achieve optimal C/N ratio and to ensure efficient methane production (Solé-Bundó et al., 2019). The synergistic effect is mainly attributed to the wide selection and availability of nutrients and increased buffering capacity in co-digestion (Wirth et al., 2018b). The critical elements of energetic use of microalgal biomass are the high cultivation cost and efficiency (Lam and Lee, 2012). The use of liquid wastes as alternative nutrient sources is an environment-friendly solution because these culture media can contribute to the sustainable microalgal biomass production as well as wastewater treatment (Wirth et al., 2020). *Chlorella vulgaris* is a common eukaryotic microalgae species found in various natural freshwaters, the algae have a relatively small cell size, thin cell wall, fast growing rate and short reproduction time (Coronado-Reyes et al., 2020), these features make this microalga suitable for cultivation in wastewater and for combined biomass generation (Wirth et al., 2020). The rapid growth of the poultry industry produces huge volumes of manure, an environmentally hazardous bio-waste. Our previous results showed that simple water extraction of solid chicken manure yields a supernatant that is a suitable medium for efficient photo-heterotrophic cultivation of microalgae (Böjti et al., 2017). The treated chicken manure (TCM) can be recycled in anaerobic digestion process, whereas the purified water can be reused for further chicken manure pre-treatment (Wirth et al., 2020). The exploitability of TCM in AD co-digestion process has been demonstrated. Improved methane production was achieved (percentage of improvement: ~17 %) in co-digestion with the crop waste corn stover compared to TCM as mono-substrate (Böjti et al., 2017).

A great number of various microorganisms are involved in the AD process, where different species have distinct roles, the interacting bacteria and archaea compose a highly diverse and specialized microbiome (Kougias and Angelidaki, 2018). The kinetics of this process is highly dependent on the substrate. In case of lignocellulosic biomass, the hydrolysis step is generally considered rate limiting (Kumar et al., 2008). Therefore, a clear understanding of the organization and behaviour of this multifarious community is crucial for optimization of their performance and attainment of the stable operation of the process. In order to optimally exploit the biomass sources it is required to understand not only which microorganisms are present, but also their metabolic potentials. The past decade has brought important technical breakthroughs to reveal the compositions of diverse microbial communities. The "next generation sequencing" technologies allow the study of genetic material recovered directly from environmental samples. These methods employ various chemical reactions for the rapid and accurate determination of DNA sequences (Jünemann et al., 2017). The ongoing rapid development of high-throughput shotgun sequencing molecular tools and bioinformatics allow the reconstruction of genomes of individual species (or MAGs: metagenome assembled genomes) from mixed cultures named as genome-centric metagenomics (Turaev and Rattei, 2016). This genome-centric approach provides a powerful method to understand the phylogenetic and metabolic diversity of AD communities without relying on culture-dependent techniques, thus offering the opportunity to discover novel taxonomic groups. Increasing number of studies reported MAGs recovered from laboratory-scale reactors and state-of-the-art biogas plants (Campanaro et al., 2020).

In the present study, we examined and compared different co-digestion mixtures of green willow biomass (GWB), microalgal-bacterial biomass grown on wastewater (chicken manure effluent) (MABA) and pre-treated chicken manure (TCM). Both MABA and TCM were co-digested with GWB. Synergistic effects were monitored along different process parameter measurements during AD in biochemical methane potential (BMP) tests and experimental continuous stirred-tank reactors (CSTR).

2. Materials and methods

2.1. Preparation of substrates

2.1.1. Green willow biomass (GWB)

Green willow biomass plants (*Salix viminalis* 'Energo') were harvested in August 2019 in a 1-year-old short rotation plantation in Szeged, Hungary (Kakuk et al., 2021). The harvest was carried out manually with pruning shears and the chopped shoots were collected in plastic bags (normal cutting length 10 cm). Leaves were collected separately (in plastic bags). The chopped woody willow shoots were ground with Retsch SM100 (Retsch, Haan, Germany) cutting mill (particle size 1 cm). Leaves were manually chopped with scissors into 1 cm pieces. Grinded woody willow and chopped leaves were mixed based on the original leaf to stalk ratio (leaf: 36 % VS, stalk: 64 % VS; VS = volatile solid). The GWB was stored at −20 °C before use in the experiment.

2.1.2. Treated chicken manure (TCM)

Chicken manure (CM) was collected from a commercial broiler poultry farm (Hungerit Corp., Csengele, Hungary). The free-range poultry houses use wheat straw bedding. Water extraction comprised of soaking 5 *g* CM in 100 mL tap water (5 v/v %) for 4 h at room temperature followed by separation of the liquid and solid phases by centrifugation (10,000 rpm for 8 min). The solid fraction was air dried and stored at −20 °C. This treated chicken manure (TCM) was used in AD experiments (Böjti et al., 2017).

2.1.3. Microalgal-bacterial biomass (MABA)

The *Chlorella vulgaris* MACC-360 microalga was obtained from the Mosonmagyaróvár Algal Culture Collection (MACC) of Hungary. *C. vulgaris* was maintained and cultivated on TAP (TRIS-Acetate-Phosphate) plates, then TAP liquid medium (500 mL) was used for the pre-growth of microalgal colonies. The TAP plates and liquid media were incubated at 50 μmol m^{-2} s^{-1} light intensity at 25 °C for 4 days (OD750: 4 ± 0.2). The microalgal stock solution was equally distributed in 17−17 mL portions into 50 mL Falcon tubes with a final optical density

(OD750) of 0.7 ± 0.1. Microalgal biomass was separated by centrifugation (10,000 rpm for 5 min) from the medium and used for inoculation (microalgal dry mass content: ~100 mg/L). The liquid phase of water extraction of chicken manure (CMS: chicken manure supernatant) was used as alternative cultivation media for microalgae (Wirth et al., 2020). Cultivation was performed in 2 L bottles (Simax, ISO bottle blue cap, 12, 112,420) with liquid volume of 2 L and stirred on a magnetic stirrer tray. Cultivation time was 4 days. Bottles were sealed with paper plugs. Media were incubated at 50 μmol m^{-2} s^{-1} light intensity at 25 °C. The final produced biomass was separated by centrifugation (10,000 rpm for 10 min) and dried (40 °C). Information about the used biomasses are summarized in Table 1.

2.2. Anaerobic digestion experiments (AD)

Biochemical methane potential (BMP) measurements were employed to determine methane potential and biodegradability of biomass. Experiments were carried out in 160 mL reactor vessels (Wheaton glass serum bottle, Z114014 Aldrich) containing 60 mL liquid phase at mesophilic temperature (37 ± 0.5 °C). All fermentations were done in triplicates according to the VDI 4630 protocol (Vereins Deutscher Ingenieure 4630, 2006). The inoculum sludge originated from an operating biogas plant (Zöldforrás Ltd., Hungary) fed with maize silage (68 % VS) and pig manure slurry (15 % VS) and maintained in semi-CSTR digesters. The inoculum was filtered (>1 mm particles filtered out) and used as negative control in BMP test and the "base" inoculum. Maize silage, which is a common substrate in biogas plants (as first-generation substrate) served as positive control. The batch reactors were supplied with GWB, TCM and MABA in mono- and co-digestions (Supplementary picture 1). The wet mass C and N contents of each substrate were measured (Table 1). In mono-digestion the VDI protocol was followed. In co-digestion, alsothis standard has been followed, and the C/N was set to ~22. The formula below was used to calculate the co-substrates' biomass ratio, which determines the amount of specific biomasses to be measured in BMP tests.

$$\left(\frac{\frac{X_{VS}*X_{TS}}{OLR}}{1000}\right) * Y_S = Biomass\ ratio$$

$$\frac{OLR * Biomass\ ratio}{X_{VS} * X_{TS}} * 10000 = Y_S;$$

$$\left(\frac{X_C}{100}\right) * Y_S = Z_C\ ;\ \left(\frac{X_N}{100}\right) * Y_S = Z_N\ ;\ \frac{(Z_C)+(Z_C)+(Z_C)}{(Z_N)+(Z_N)+(Z_N)} = Co - substrate\frac{C}{N}ratio$$

Where X means the substrate (GWB, TCM or MABA), the OLR is the organic loading rate, Y_S means the amount of fed substrate, C = carbon, N = nitrogen. The amount of substrate feeding depended on the organic loading rate, biomass ratio, as well as the substrate volatile solid (*Xvs*) and total solid (X_{TS}) contents. Each batch fermentation experiment lasted for 30 days in triplicates. The results of BMP measurements are summarized in Supplementary Table 1.

To scale up the experiment, anaerobic digestions were carried out in three 5 L semi-CSTRs in fed-batch operational mode. The experimental design and time course followed the scheme of the Supplementary Fig. 1. Following the 1-month long "incubation" phase (to allow for the disappearance of the residual biogas potential), the 3 reactors were fed as follows: one reactor received GWB and TCM biomasses (Fermentor 1: F1), one reactor was supplied with GWB and MABA (Fermentor 2: F2) and the third one with GWB + TCM + MABA (Fermentor 3: F3). Biomasses were fed at a loading rate of 1 g VS L^{-1} (VS = volatile solid) at C/N ratio of ~22 into each fermenter (for calculations see above and Table 1). The reactors were fed daily with the specific co-substrate. Temperature was maintained at 37 ± 1.0 °C by an electronically heated jacket which surrounded the cylindrical apparatus. The pH was between 7.5 and 8.2, and the redox potential was less than −500 mV. Heating, stirring, pH, and redox potential values were monitored and collected on-line (Kovács et al., 2013). The accumulated gas volume was recorded in every 4 h with thermal mass flow controllers (DMFC, Brooks) in the experimental phase. Data are summarized in Supplementary Table 1 under "CSTR methane" tab.

2.3. Determination of anaerobic digestion parameters

The total solids (TS) content was quantified by drying the biomass at 105 °C overnight and weighing the residue. Further heating of this residue at 550 °C for 1 h provided the volatile solids (VS) content.

To determine C/N, an Elementar Analyzer Vario MAX CN (Elementar Group, Hanau, Germany) was employed. The approach is based on the principle of catalytic tube combustion under O_2 supply at high temperatures (combustion temperature: 900 °C, post-combustion temperature: 900 °C, reduction temperature: 830 °C, column temperature: 250 °C). The desired components were separated from each other using specific adsorption columns (containing Sicapent (Merck, Billerica, USA), in C/N mode) and were determined in succession with a thermal conductivity detector. Helium served as flushing and carrier gas.

The fibre composition, including the neutral detergent fibre (NDF), acid detergent fibre (ADF) and acid detergent lignin (ADL) of the biomass samples (GWB, MABA and TCM) were determined with a VELP Scientific FIWE 3 Fibre Analyzer (VELP Scientifica Srl, Usmate, Italy) according to the manufacturer's guidelines using the Van Soest method. Hemicellulose was estimated as NDF-ADF, while cellulose as ADF-ADL.

For the determination of total ammonia nitrogen (TAN = ammonium ions and dissolved ammonia) content, the Merck Spectroquant Ammonium test (1.00683.0001) (Merck, Billerica, USA) was used.

The volatile organic acid and total inorganic carbon (VOAs, TIC) measurements process were carried out using a Pronova FOS/TAC 2000 Version 812−09.2008 automatic titrator (Pronova, Berlin, Germany). 5 g of sample was taken for the analysis and diluted to 20 g with distilled water. 0.1 N H_2SO_4 was used for the titration.

The samples were centrifuged (13,000 rpm for 10 min) and the supernatant was filtered through polyethersulfone (PES) centrifugal filter (PES 516−0228, VWR) at 16,000 *g* for 20 min. The concentrations of volatile organic acids were measured with HPLC (Hitachi LaChrome Elite) equipped with refractive index detector L2490. The separation

Table 1
Characteristics of the biomass types used in the experiment. The "Biomass ratio" indicates the co-substrates' ratios. TS = total solid, VS = volatile solid. The table contains the TAN concentrations (measured from the liquid phase of digestion sample) and the methane yields of specific co-substrates (CH_4 mL_N gVS^{-1}).

Parameters	GWB	MABA	TCM
Lignin (% TS)	18.78	0	7.12
Cellulose+Hemicell. (% TS)	46.88	15.95	26.52
Solubles (% TS)	34.4	84.05	76.36
Wet mass C (mg g^{-1})	21.18	46.86	44.04
Wet mass N (mg g^{-1})	0.58	9.28	3.74
C/N	36.83	5.05	11.8
TS (%)	39.74	93.88	95.63
VS (%)	95.13	83.84	84.61
	GWB + MABA	GWB + TCM	GWB + MABA + TCM
Co-digest. biomass ratio	0.9 + 0.1	0.7 + 0.3	0.8 + 0.05 + 0.15
Co-digest. C/N	22.13	22.68	22.4
TAN (g L^{-1})	1.75 ± 0.1	1.7 ± 0.1	1.8 ± 0.1
CH_4 (mL_N gVS^{-1})	189 ± 6	186 ± 4	187 ± 5

was performed on an ICSep ICE−COREGEL—64H column. The temperature of the column and detector was 50 and 41 °C, respectively. 0.01 M H_2SO_4 (0.8 mL min^{-1}) was used as an eluent. Acetate, propionate, and butyrate were determined in a detection range of 0.01−10 $g\ L^{-1}$. Propionate and butyrate were present in traces relative to acetate and therefore these are not reported in the results section.

The CH_4 content was determined with an Agilent 6890 N GC (Agilent Technologies) equipped with an HP Molsieve 5 Å (30 m ×0.53 mm ×25 µm) column and a TCD detector. The temperature of the injector was 150 °C and split mode 0.2:1 was applied. The column temperature was maintained at 60 °C. The carrier gas was Linde HQ argon 5.0 with the flow rate set at 16.8 mL/min. The temperature of TCD detector was set to 150 °C.

2.4. Total DNA isolation for metagenome analysis

The composition of the microbial community was investigated five times during the experimental period: at the beginning of feeding with the selected co-substrates (start), one week later (day 7), when the system was working at full capacity (day 35 and day 63), and at the end of the process (day 84). For total community DNA isolation 2 mL of samples were used from each cultivation media type. DNA extractions were carried out using the Zymo Research Fecal/Soil DNA kit (D6010, Zymo Research, Irvine, USA). After lysis (bead beating was performed by Vortex Genie 2, bead size: 0.1 mm; beating time: 15 min, beating speed: max), the Zymo Research kit protocol was followed. The quantity of DNA was estimated using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, USA) and a Qubit 2.0 Fluorimeter (Life Technologies, Carlsbad, USA). DNA purity was tested by agarose gel electrophoresis and on an Agilent 2200 TapeStation instrument (Agilent Technologies, Santa Clara, USA).

2.5. Total metagenome sequencing

The recommendations of the Illumina sequencing platform were closely followed (Illumina Inc., USA). DNA samples were used to sequence preparation applied by the NEBNext Ultra II Library Prep Kit. The metagenomics sequencing was performed by Illumina with MiSeq chemistry (MiSeq Reagent kit v2). The characteristic fragment parameters were summarized in Supplementary table 3. Raw sequences are available on NCBI Sequence Read Archive (SRA) under the accession number: PRJNA695133.

2.6. Raw sequence filtering

Galaxy Europe server was employed to pre-process the raw sequences (i.e., sequence filtering, mapping, quality checking) (Afgan et al., 2016). Low-quality reads were filtered by Prinseq (min. length: 100; min. score: 15; quality score threshold to trim positions: 20; sliding window used to calculated quality score:1). Filtered sequences were checked with FastQC (Supplementary table 2).

2.7. Metagenome co-assembly, gene calling, binning and phylogenomic tree reconstruction

The filtered sequences produced by Prinseq were co-assembled with Megahit (Li et al., 2015) (min. contig length: 2000; min k-mer size: 21; max k-mer size: 141). After simplifying the header of contig FASTA file using the Anvi'o script "reformat-fasta" Bowtie2 was equipped to map back the original reads to the contigs (Langmead and Salzberg, 2012). Then we used Anvi'o (v6: Esther) following the "metagenomics" workflow (Eren et al., 2015). Briefly, during the first step contig database was generated, where open reading frames (ORFs) were identified by Prodigal and each contig k-mer frequencies were computed (Hyatt et al., 2010). Then Hidden Markov Modell (HMM) of single-copy genes (SCGs) were aligned by HMMER using GTDB database SCG collection. We used InterProScan v5.31–70 on Pfam for the functional annotation of contigs (Finn et al., 2017). The outputs were imported into the contig database by using the "anvi-import" command. The functional profiles are supplemented with NCBI COGs (Clusters of Orthologous Groups) categories. BAM files made by Bowtie2 were used to profile contig database, in this way we obtained sample-specific information about the contigs (i.e. mean coverage) (Langmead and Salzberg, 2012). This sample-specific information was merged together using the "anvi-merge" command. Three automated binning programs, namely CONCOCT, METABAT2 and MAXBIN2 were employed to reconstruct microbial genomes from the contigs (Alneberg et al., 2014; Kang et al., 2015; Wu et al., 2015). The combination of Metawrap and Anvi'o human-guided refine option were used to improve the quality of metagenome assembled genomes (MAGs). For taxonomic assignment of bins three different genomic databases were utilized: GTDB, Progenomes2 and Miga (Mende et al., 2020; Parks et al., 2020; Rodriguez-R et al., 2018). Consensus or highest taxonomic rank match results were applied to name the specific MAGs. The reconstructed bins were also compared to the Biogasmicrobiome database in Miga database (Campanaro et al., 2020). Binning statistics was summarized in Supplementary table 2. Phylogenomic tree was build up by GenomeTreeTK program using Single-copy Core Genes (SCGs) of bacterial and archaeal protein genes (inferred with GTDB) (https://github.com/dparks1134/GenomeTreeTk). The interactive Tree of Life (iTOL) tool was employed to visualize the phylogenomic tree and the binning results (https://itol.embl.de/).

2.8. Statistical analyses

Statistical Analysis of Metagenomics Profiles (STAMP) was used to investigate microbial communities and calculate principal component analysis (PCA) (using Tukey-Kramer as post-hoc test and Storey false discovery rate correction). To compare significantly different MAGs in reactors (average percent recruitment) fed with different substrates two-sided *t*-test was calculated (Parks and Beiko, 2010).

2.9. Identification of CAZymes

The ORFs were identified by Prodigal and enzymes involved in carbohydrates utilization were collected using the combination of Pfam functional profiles and the carbohydrate-active enzyme database (CAZy) (Lombard et al., 2014). MAGs belonging to the domain Bacteria were involved in the analysis (Supplementary table 3). The results are visualized by Circos software (https://github.com/vigsterkr/circos).

2.10. Estimation of functional pathway completions

Prokka was employed to translate and map protein sequences (create protein FASTA file of the translated protein coding sequences) (Seemann, 2014). For the calculation of module completion ratio (MCR) MAPLE 2.3.1 (Metabolic And Physiological potentiaL Evaluator) was used (https://maple.jamstec.go.jp/maple/maple-2.3.1/). This automatic system is mapping genes on an individual genome and calculating the MCR in each functional module defined by Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Guto, 2000) (Supplementary table 4). MAGs have medium or above quality were included in the calculation. Results were visualized by iTOL web tool (https://itol.embl.de/).

2.11. Calculation of replication indexes

In replication index calculations the genome size and the total number of reads mapped on each MAG were considered. Bowtie2 program was used to determine the coverage of each MAG. The MAGs having completeness higher than 75 %, contamination lower than 5%, a number of scaffolds per Mbp lower than 175 and a coverage value higher than five, were selected in order to determine their index of replication

(iRep) applying the iRep software (Brown et al., 2016) (Supplementary table 5). The number of replication origins in archaeal genomes was inspected using Ori-Finder 2 software and those having none or more than one were excluded from further analyses (http://tubic.tju.edu.cn/Ori-Finder2/). The results of the calculations have been visualized by R ggplot2 package.

3. Results and discussion

3.1. Stable and efficient AD using co-substrates

In our experiments standardized batch BMP tests and semi-continuously operated CSTRs were employed to investigate the efficiency and stability of co-digestion of various biomasses. Previous work concluded that although the methane yield was lower, the extractable energy potential per ha of GWB reached and surpassed the corresponding performance values of maize silage (Supplementary Table 1 "Batch methane" tab) (Kakuk et al., 2021). The key factor was the high "soluble" content. Both TCM and MABA had a low C/N ratio compared to GWB, but contains high amount of additional nutrients (Table 1). Lignocellulosic biomasses are difficult to degrade in AD (Paul and Dutta, 2018). Low C/N biomasses, could result in excessive ammonia release, often along with high levels of volatile fatty acids (VFAs) accumulation, which could lead to process failure (Shi et al., 2016). However, by co-digesting these biomasses balanced and efficient anaerobic digestion can be achieved (Wirth et al., 2018b). In previous experiments anaerobic batch tests were conducted to evaluate the effects of adding corn straw as co-substrate to the digestion of Taihu blue algae (Zhong et al., 2012). The co-digestion of this mixed biomass at a C/N ratio of 20:1 significantly increased the methane yield (algae digestion alone: 201 mL g^{-1} VS^{-1}; co-digestion: 325 mL g^{-1} VS^{-1}) (Zhong et al., 2012). In another study algal-bacterial biomass grown on wastewater, was used to adjust the C/N ratio of the cellulose substrate in batch tests. It was found that the optimal C/N ratio for co-digestion of these biomass sources were in the range of 20−30:1 (Bohutskyi et al., 2018). In our work, the BMP of mono and co-digestion of TCM, MABA and GWB biomasses were compared first. Based on previous literature data, the C/N ratio of co-digestions were set to 22 (Table 1). The BMP tests indicated that the co-digestions gave similar CH_4 yields, which were an average 21 % higher relative to the CH_4 potential of GWB biomass alone in mono-digestion (mean of co-digestions: 194 ± 4 mL_N g^{-1} VS^{-1}; GWB in monodigestion: 152 ± 9 mL_N g^{-1} VS^{-1}) (Supplementary Table 1). Therefore, we used this C/N ratio to scale up the experiments in CSTRs.

Temperature (37 ± 1.0 °C), mixing speed (10 rpm min^{-1}) and pH (7.5–8.2) of the AD process were continuously monitored in the CSTR scale-up experiments. Gas production data were collected for 12 weeks (84 days). Cumulative CH_4 production data are plotted in Supplementary Table 1 showing similar dynamics in various co-digestions. The CH_4 content in the evolved gas was between 42–50 %, and the yields were 186−189 mL_N g^{-1} VS^{-1}, respectively (Table 1). The volatile organic acids and total inorganic carbon (VOAs and TIC) is a reliable indicator of the AD process. In our experiments, low VOAs (average: 0.2 g L^{-1}) and high TIC (average: 10 g $CaCO_3$ L^{-1}) values were observed (Fig. 1A). The amount of total ammonia nitrogen (TAN), which is a combination of dissolved ammonia and ammonium ions is also a critical indicator of the AD process stability (Yenigün and Demirel, 2013). TAN is a key macronutrient and therefore a certain concentration is essential for proper microbial growth, moreover it also contributes to the elevation of buffer capacity maintaining the pH stable. The TAN inhibition level in AD has been reported a huge disparity in the inhibitory limits. However, it was observed that pH, temperature, and TAN concentration together are the main factors affecting inhibition of AD (Capson-Tojo et al., 2020). The hydrogenotrophic methanogens are the most resistant to high TAN concentration, although depends on the used biomass, inoculum, microbial community and acclimation (Capson-Tojo et al., 2020; Nielsen and Angelidaki, 2008). In our reactors the TAN concentration was between 1.5 and 2 g L^{-1} (Table 1), which was comparable to previous experiments where TCM and maize silage were co-digested in 1:1 ratio in CSTR digesters (Böjti et al., 2017). The MABA biomass did not cause ammonia accumulation, because of the balanced C/N ratio (Wirth et al., 2019). The acetate concentrations exhibited a tendency similar to that of VOAs (Fig. 1A VOAs and 1B). Low acetate concentrations were observed during the experiment (Fig. 1B). It can be noted that, because of the low VOAs and acetate concentration there was no organic overloading. This implies that most of the accumulated solids were recalcitrant in the digester as reflected by the high buffering capacity (Fig. 1A TIC). Similar AD process parameters have been reported previously (Estevez et al., 2014). Low VOAs (0.32 g L^{-1}) and NH_4^+-N (1 g L^{-1}) contents have been observed during mesophilic anaerobic co-digestion of steam-exploded willow and cow manure (willow: 40 % TS + manure: 60 % TS diluted with water) in laboratory-scale CSTR. The CH_4 yield was 185 mL g^{-1} VS^{-1}. This CH_4 yield is comparable with our data (186−189 ml g^{-1} VS^{-1}). However, our approach avoided steam-explosion pre-treatment of willow biomass. Recirculation of the digestate is possible to improve the lignocellulosic biomass digestion (Estevez et al., 2014), but may lead to VOAs accumulation and process inhibition (Nordberg et al., 2007).

3.2. Metagenomes of AD microbial community

Rational management of AD requires a comprehensive understanding of biogas producing microbiome. Monitoring participant microbes and their specific activities during the AD process will help maintain and design sustainable operation process (Bozan et al., 2017). Shotgun sequencing and genome-centric metagenomics approaches were applied to reveal the microbial composition and the ongoing metabolic activities

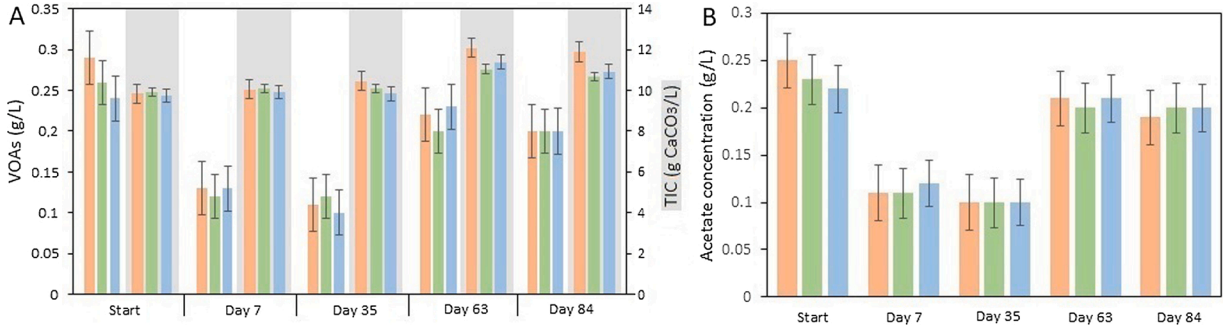

Fig. 1. Anaerobic digestion process parameters. The different colours represent the digesters fed with the distinct substrates: orange columns: energy willow (GWB) + treated chicken manure (TCM); green columns: energy willow (GWB) + microalgal-bacterial biomass (MABA); blue columns: GWB + TCM + MABA. (A) VOAs and TIC values. Grey background indicates the TIC data. (B) Measured acetate concentrations. The samples for AD parameter analysis were taken at the starting point (Start) and at days 7, 35, 63 and 84, respectively.

of the co-digestion systems operating with GWB, MABA and TCM substrates.

3.2.1. Genome-centric metagenome data

84.4 million sequence reads passed the filtering step (average: 5.6 million reads/sample) (Supplementary table 2). The filtered reads of each of the 15 samples were assembled by Megahit creating a total of 108,103 contigs (minimum contig length: 1500 bp). The subsequent binning approach with the combination of automated and human-guided genome-centric metagenomics resulted 178 MAGs (metagenome assembled genomes, or genome parts) from which 75 were high quality MAGs (Completion (C): >90 %, Redundancy (R) : <5%), 64 medium (C: ≥50 %, R: <10 %) and 39 low (C: <50 %, R: <10 %), based on the MIMAG initiative (Bowers et al., 2017). For taxonomic assignment of MAGs three different genomic datasets were employed (Fig. 3). The reconstructed genomes were also compared to the Bio-Gas Microbiome database in Miga, containing MAGs identified from 134 samples collected from biogas plants and lab-scale reactors in seven countries (Campanaro et al., 2020). The results showed that 64.4 % of the MAGs were assigned to known species, 1.1 % to genera, 10.1 % to families, 2.8 % to orders, 16.9 % to classes and 2.8 % of the total MAGs could be linked only at phylum level (Supplementary table 2). The taxonomic distribution of binning results showed that 165 MAGs belonged in the domain Bacteria and 13 MAGs in the domain Archaea (Supplementary table 2). Most of the reconstructed genomes in the domain Bacteria were associated to the phylum Firmicutes (119 MAGs) followed by Bacteroidetes (19 MAGs), Proteobacteria (7 MAGs) and Actinobacteria (6 MAGs), which is similar to the previously identified microbial community in BMP tests using GWB (Kakuk et al., 2021). Van der Lelie et al. (2012) in their work also observed similar microbial composition in the AD of poplar wood chips (van der Lelie et al., 2012). Their binning results revealed that species belonging in the phyla Firmicutes, Bacteroidetes and Proteobacteria predominated the lignocellulosic biomass-degrading communities. Also a related bacterial community was revealed by genome-centric metagenomics in AD of co-digesting pig manure together with ensiled meadow grass. The phyla Firmicutes and

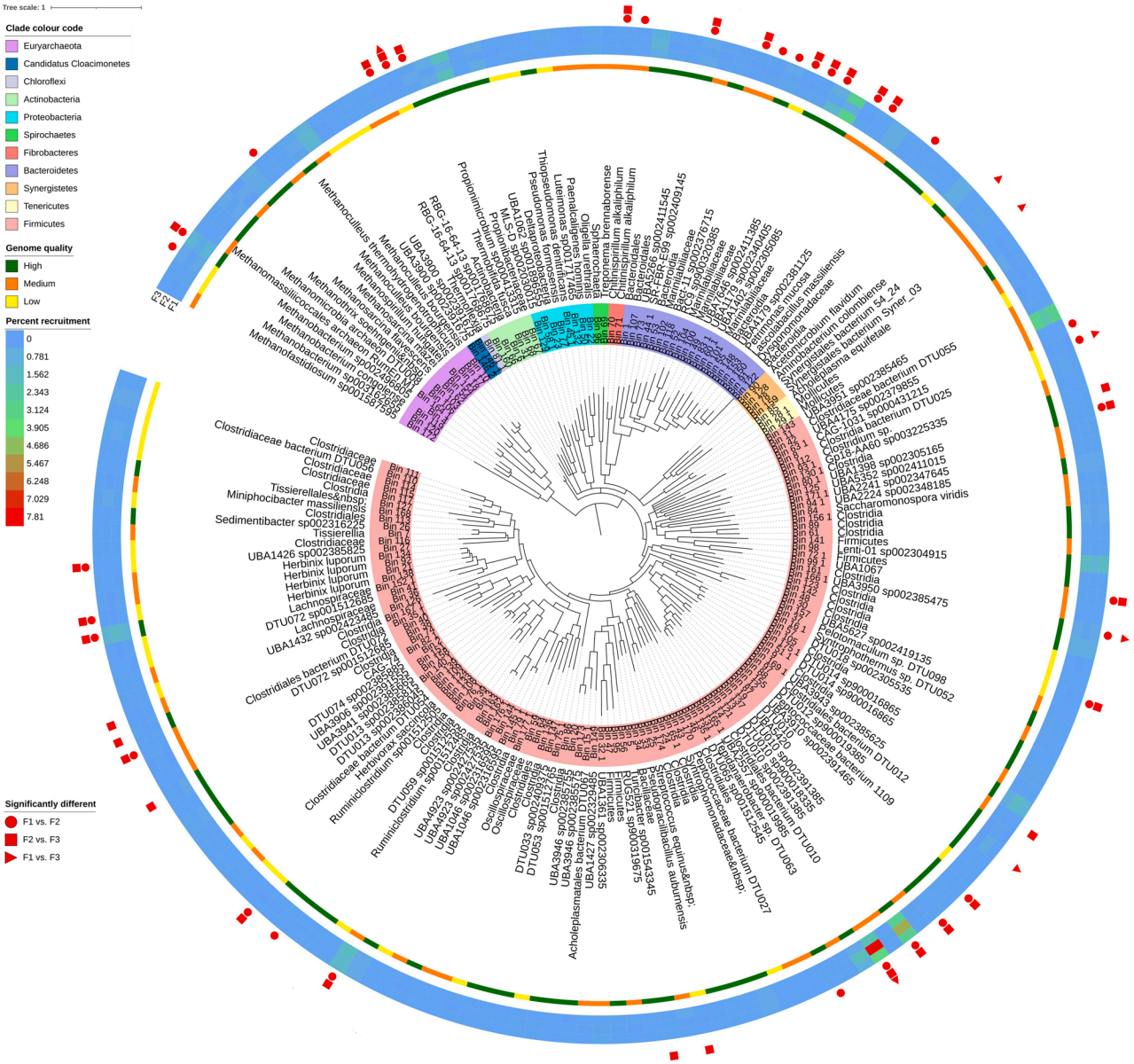

Fig. 3. The phylogenomic tree of recruited MAGs based on bacterial and archaeal SCGs. The first ring shows the bin names and the background colours indicate the phyla where they belong to. The next section contains the names of specific bins based on consensus of different genomics databases. The quality and size of reconstructed MAGs are indicated. The outer rings show the average recruitments/coverage of MAGs in F1, F2 and F3 digesters.

Bacteroidetes predominated the bacterial community (Kougias et al., 2018). Interestingly, we recruited three high-quality MAGs (Bin_61: Clostridia, Bin_67: *Propionibacteriaceae* and Bin_107: Bacteroidales; C: >95 %, R: <5%) creating stable community in all AD digesters and representing 1.7 % of our binning data could not be identified below the taxonomic level of family. These MAGs potentially represent newly discovered microbes based on MiGA Bio-Gas Microbiome database (average amino-acid identity: <60 %) (Supplementary table 2). This indicates that the genome-centric investigations allow us to fill gaps in our knowledge on biogas producing microbial community (Kougias et al., 2018). Microbial species belonging to the Candidate Phyla Radiation (CPR) have also been reconstructed from our digesters (Bin_164_1 and 164_2 were assigned to the *Cloacimonadaceae* family). Besides, in this work the principal component analysis (PCA) showed clear alteration in the AD communities (Fig. 2A). The microorganisms of the starter inoculum successfully adapted to the fed biomasses. Although the community structure covers 82 % similarity, slight differences can be observed around from week 5 (day 35) (Fig. 2A). The developed microbiomes of digesters fed with GWB + TCM (F1) and GWB + TCM + MABA (F3) were more similar to each other, than to the GWB + MABA (F2) biomass degrading community (Fig. 2A and B). The calculation of significant differences resulted that the phyla Bacteroidetes and Fibrobacteres are higher in F1 and F3, while Firmicutes in F2 (p-value <0.05) (Fig. 2C). MAGs which were different between digesters are marked with red symbols in Fig. 3. There are 41 MAGs which are altered between F2 and the other two digesters and only 7 found significantly different between F1 and F3 (Fig. 3). For example, the above mentioned tendency can be observed in two MAGs (Bin_13 and Bin_14 from the phylum Firmicutes), which were highly covered in all the three co-digesting reactors and detected significantly higher in F2 (p-value <0.05). The methanogenic AD community was primarily represented by the phylum *Euryarchaeota* (13 MAGs) (Fig. 3). The orders Methanomicrobiales (8 MAGs), Methanobacteriales (3 MAGs) predominated. In the order Methanomicrobiales the genus *Methanoculleus* (Bin_10), while in Methanobacteriales the genus *Methanobacterium* (Bin_4) were found in considerable abundances (Fig. 3 and Supplementary table 2 "Percent recruitment" tab). In addition, one MAG was aligned to the class Thermoplasmata (Bin_164_1) and one to the genus *Candidatus Methanofastidiosum* (Bin_172). Overall the phylum Euryarchaeota did not differ between the AD digesters (Fig. 2C), although 3 MAGs (two from the order Methanobacteriales: Bin_119 and Bin_4, and one from Methanomicrobiales: Bin_154) were observed in significantly lower quantity in GWB + MABA fed digesters (Fig. 3). Therefore, TCM and MABA made clear influence on the microbial community, which shaped the pattern of clustering results (Fig. 2A).

3.3. Functional analysis of MAGs

The mixed biomass used as substrate in these experiments was rich in complex carbohydrate molecules. Therefore, it was essential to analyse the presence of carbohydrate-active enzymes (CAZymes) in the identified MAGs, to get deeper insight into the rate-limiting hydrolysis step of the biogas producing food chain (Kumar et al., 2008). Genome-centric data were analysed to better understand the metabolic processes. The metabolic pathway reconstruction of medium and high quality MAGs were used for the identification of functional units of KEGG modules.

3.3.1. The identified CAZymes and their distribution

Overall, 85 distinct CAZymes were detected in the investigated microbial communities (Supplementary table 3). These enzymes were grouped in 4 CAZyme classes. The glycoside hydrolases (GHs) represent the most widespread group, followed by glycosyltransferases (GTs), carbohydrate binding domains (CBMs) and polysaccharide lyases (PLs) (Fig. 4A). GHs and GTs have been found in all observed bacterial phyla. Nevertheless, 69 % of all identified CAZymes were linked to the phyla Firmicutes, while 16 % of all identified CAZymes could be rendered to Bacteroidetes. Moreover, 86 % of CBMs were detected in Firmicutes and Bacteroidetes, while PLs, which are responsible for the breakdown of C-

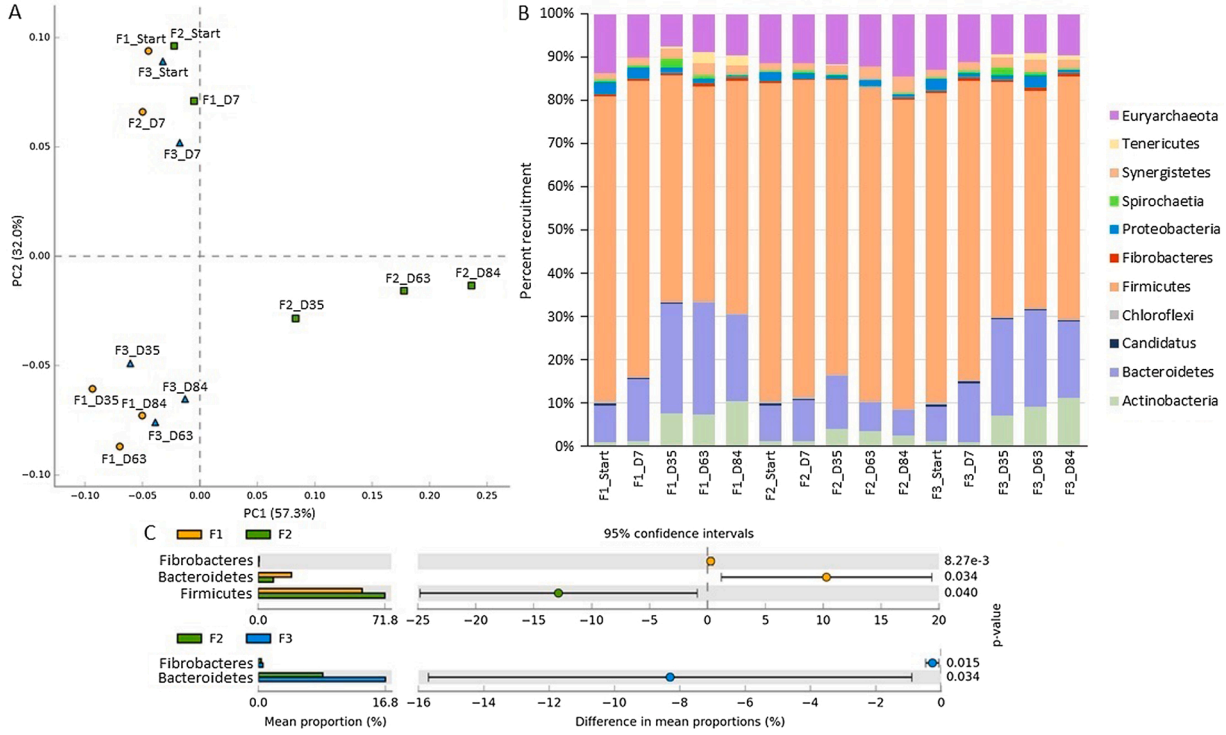

Fig. 2. Genome-centric results. (A) Principal component analysis of microbial community of all samples. Each symbol is related to a specific sample/digester (GWB + TCM = F1, GWB + MABA = F2 and GWB + TCM + MABA = F3; time-series: start of the experiment, D7, 35, 63, and 84 are days, when the samples were collected). (B) Taxonomic distribution of the microbiome, i.e., the taxa present in all 15 metagenomes (digester and time-series) at phylum level. (C) Significantly different phyla between the average microbiome of digesters ($p \leq 0.05$, see in section 2.8).

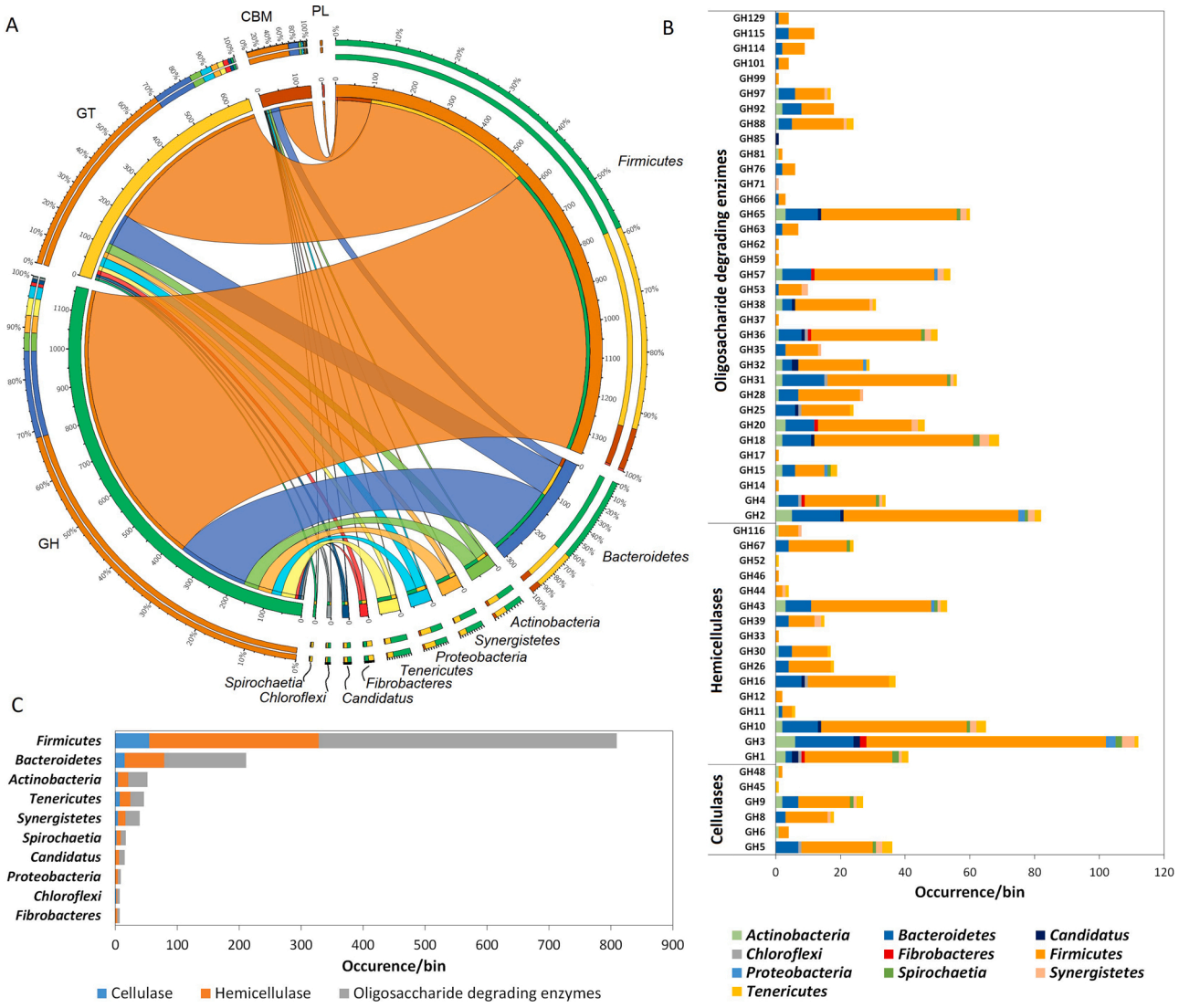

Fig. 4. Identified CAZymes and associated phyla. (A) The Circos plot illustrates the found CAZyme classes and their distribution across the bacterial phyla. (B) The identified glucoside hydrolase (GH) enzyme families and their occurrence in specific phyla. (C) Summary of the occurrence of GH families in specific phyla.

6 carboxylated polysaccharides could be uniquely found in the representatives of Firmicutes (namely: Bin_13, _61, _80_1, _9 and _98), from which Bin_61 seems to be a newly detected MAG in AD community. Furthermore, one of them is a common and abundant MAG in all AD digester (Bin 13) which belongs to the class Clostridia (Fig. 2A and Fig. 3). This indicates that the degradation of lignocellulosic biomass is a complex microbial activity, where the hydrolysing enzymes rely on each other. The phyla Firmicutes and Bacteroidetes are frequently provide the majority of polysaccharide degraders in biogas producing communities (Güllert et al., 2016). However, our data demonstrated that members of other phyla like Actinobacteria (4% of the identified CAZymes), Tenericutes (3% of all identified CAZymes) and Synergistetes (3% of all identified CAZymes) also had complex carbohydrate degrading capacity, thus played notable role in the lignocellulosic biomass degradation (Fig. 4B). As the dominance of glycoside hydrolases (GHs) suggests, numerous cellulose-. hemicellulose- and oligosaccharide-degrading enzymes are present (Fig. 4C). GH5, GH9 and GH8 represent the cellulose degrading enzymes. The phyla Firmicutes, Bacteroidetes and Tenericutes predominate among microbes expressing these GH families. It is noteworthy that these families have been recognised in numerous complex carbohydrate degrading ecosystems, suggesting an important and general role in cellulose degradation (Güllert et al., 2016). Among hemicellulases GH3, GH10, GH43, GH1 and GH16 are the most widespread families possessing xylanase and xylosidase activities. The GH families responsible for these tasks are dominated by the phyla Firmicutes, Bacteroidetes, Actinobacteria and Tenericutes. Apparently, the oligosaccharide-degrading enzymes had the most diverse range of catalytic activities for carbohydrate breakdown. Inside this group the GH3, GH10, GH43, GH1 and GH16 are the largest families. In general, oligosaccharide-degrading activities predominates in the phyla Firmicutes, Bacteroidetes, Actinobacteria and Synergistetes (Fig. 4C). GH9, GH5, GH3, GH43, GH10 and GH16 cellulases and hemicellulases were shown to play important role in the degradation of lignocellulosic materials. These CAZymes have also been observed to be abundant and active in other biogas reactors fed with similar substrates and even in cow rumen (Güllert et al., 2016; Jia et al., 2018; Wirth et al., 2018a). No alterations were found in MAGs (Bin_157 and _15_1 from the phylum Firmicutes) harbouring these cellulases and hemicellulases between the GWB, MABA and TCM fed fermenters. Although, other representatives of phyla Firmicutes (Bin_13, 63_1, 80_1, _138), and certain Bacteroidetes (Bin_104), Actinobacteria (Bin_31) and Synergistetes (Bin_128) MAGs harbouring wide variety of CAZymes (≥20) have shown substantial differences in their relative abundances (Fig. 3). Therefore, the fed co-substrates contributed to the changes in CAZyme production which manifested in synergistic effect (Wirth et al., 2018b).

3.3.2. Metabolic reconstruction of MAGs

In order to further characterize the biogas producing food chain, module completion ratio (MCR) was calculated using MAPLE software on KEGG database. Overall 139 MAGs (78 % of all) met the criteria to investigate their metabolic pathways (medium or high quality MAGs). 66 % of the overall KEGG modules were represented partly or wholly in quality filtered MAGs, suggesting a specialized functional repertoire of the mixed substrate digesting community (Supplementary table 4). From the identified KEGG modules 81 were identified as "complete" at least in one MAG and 37 modules of these were detected as "frequent modules" (average: ≥50 % MCR/MAGs), which were common in all of the observed MAGs (Fig. 5). Our "frequent modules" included the "core modules" observed previously. The core modules were the following ones: "PRPP biosynthesis" (average: 86 % MCR/MAGs), "glycolysis, core module involving three-carbon compounds" (average: 69 % MAGs/MCR), and "C1-unit interconversion" (average: 69 % MCR/MAGs) (Campanaro et al., 2020). Among the central carbohydrate metabolism modules, the "pyruvate-oxidation" (average: 86 % MCR/MAGs), "Embden-Meyerhof pathway" (average: 65 % MCR/MAGs) and "pentose-phosphate pathway" (average: 55 % MCR/MAGs) were also highly represented in our study, creating additional core functions. This suggested a preferred decomposition pathway where the sugars originated from the hydrolysis step were converted to pyruvate through the Embden-Meyerhof and pentose phosphate pathways generating CO_2 and electrons. These initial steps (acidogenesis) are equivalent with the observed activity of glucose and avicel degrading microbes (Zhu et al., 2019). Additionally, acetyl-CoA to acetate (average: 59 % MCR/MAGs) conversion was also detected among the "frequent modules" related to carbon fixation as previously observed in manure supplemented reactors (Campanaro et al., 2020). Functions associated to "fatty acid

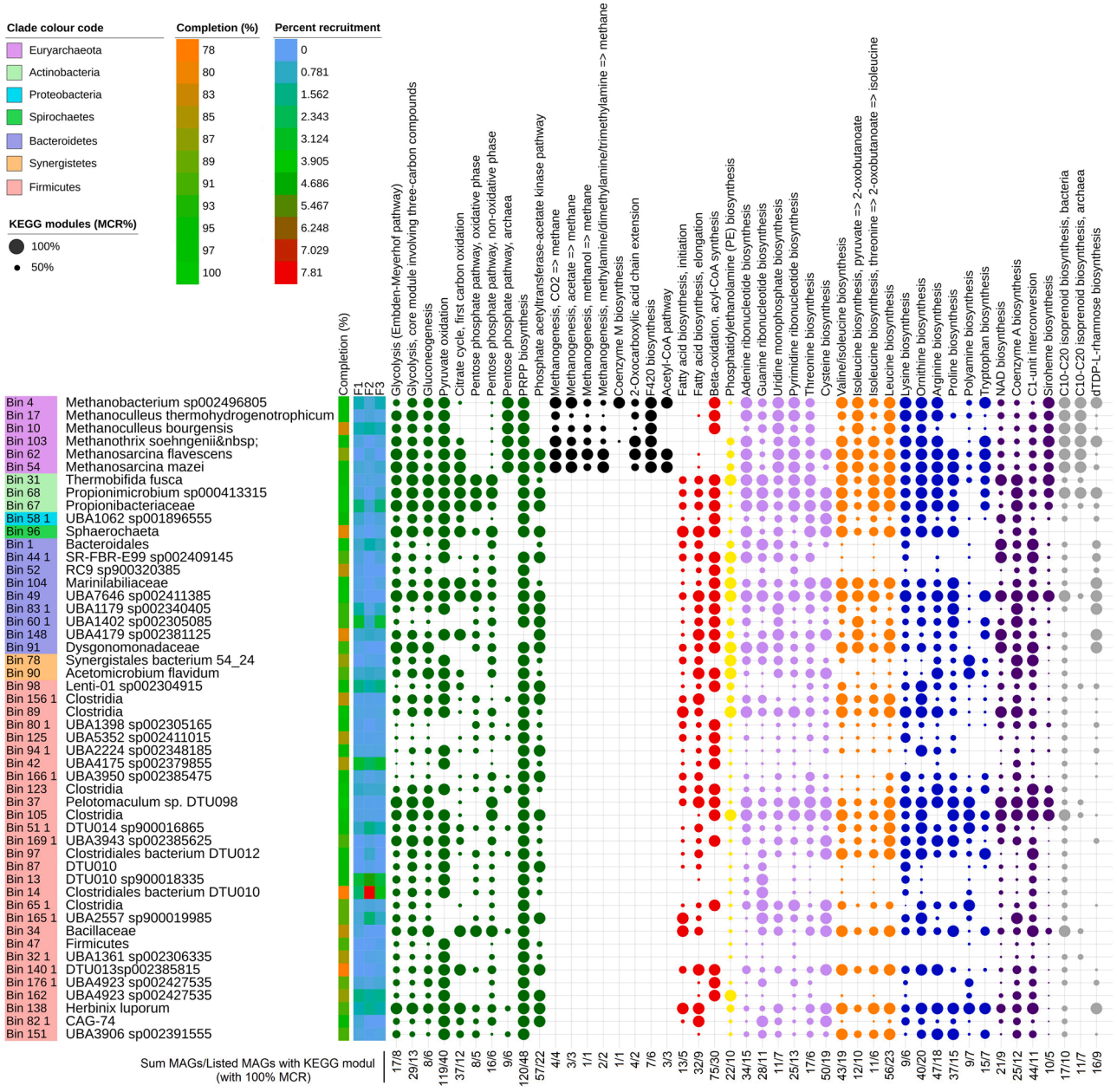

Fig. 5. The KEGG modules identified in high and medium quality MAGs. The top 50 most abundant MAGs were selected. The left part of the figure shows the MAGs name, its completion and average percent recruitment in the given digester (F1, F2 and F3). The bubble chart represents the results of MCR calculations; the 37 "frequent modules" supplemented with the methanogenesis module (or related). At the bottom of the figure the total number of complete KEGG modules of specific pathways (100 % MCR) are shown, presented as all filtered MAGs/the listed MAGs.

biosynthesis" (average: 58 % MCR/MAGs) and "beta oxidation" (average: 54 % MCR/MAGs), supplemented with the low acetate concentration, were observed indicating a slow but sufficient AD of co-substrates. This was supported by the numerous modules associated with energy, amino acid and cofactor production, which are all essential for a well-functioning biogas generation ecosystem (Fig. 5 and Supplementary table 4). However, substantial differences were observed in modules that can be linked to sugar and fatty acid metabolism. MAGs that are detected significantly different between fermenters showed that the Embden-Meyerhof pathway is more characteristic in F2 (average: 74 % MCR/sign. dif. MAGs), pyruvate oxidation in F2 and F3 (100 % MCR/sign. dif. MAGs), while beta-oxidation in F3 (average: 75 % MCR/sign. dif. MAGs). Methanogenesis was among the "rare modules" (average: 4.4 % MCR/MAGs) (Fig. 5). These modules are associated with the conversion of CO_2/H_2, acetate, methylamines and methanol into CH_4. Hydrogenotroph methanogenesis was identified in 5 MAGs (Bin_4, _10, _17, _18 and _154), methylotroph and acetotroph in one of each two MAGs (Bin 106 and Bin 103) while mixed hydrogenotroph, acetotroph, and methylotroph methanogenesis was detected in 2 MAGs (Bin_62 and Bin_54). Besides, a H_2-dependent methylotrophic CH_4 production pathway was detected in two MAGs (Bin_164 and Bin_172) according to the MCR calculations (Fig. 5) (Vanwonterghem et al., 2016). The dominance of CO_2/H_2 dependent methanogenesis and F420 biosynthesis modules suggested that the hydrogenotrophic pathway was the main route of CH_4 formation in digesters decomposing complex polysaccharide. Moreover, two MAGs – Bin_4 and _154 – were observed in elevated quantity in F2 and F3, therefore the balanced nutrients in co-substrates may contribute to the development of hydrogenotrophic methanogen archaeal community.

3.4. The replication indices of MAGs

In an unsynchronized population of various microbes, the rate of genome replication shows high alterations between the community members. To determine the replication index of specific MAGs, involved in particular steps in AD mechanisms across multiple samples, iRep program was employed. The program uses an algorithm to calculate the index of replication (iRep) based on the sequence coverage. The coverage values are correlating with the rate of bi-directional genome replication initiated from a single origin of replication (Brown et al., 2016). Therefore, those Euryarchaeota MAGs, which have none or multiple replication origins have been excluded from the analysis (see section 2.11.) (Fig. 6). In total, 39 MAGs met the conditions of iRep measurements (Supplementary table 5). Values close to median of iRep values were found for 38 MAGs (median iRep: 1.6), only 1 MAG could be considered as "fast growing" microbe (iRep ≥2), this MAG related to *Mollicutes* (phylum Tenericutes: Bin 29_1: iRep 2.4). Additionally, 11 MAGs were detected with above median iRep values (iRep ≥1.6: Bin_49, _104, _60_1, _83_1, _123, _87, _98, _47, _65_1, _32_1 and _90), these belong to the phyla Bacteroidetes, Firmicutes and Synergistetes (Fig. 6), from which Bin_49, _83_1 and _104 (from phylum Bacteroidetes) were found significantly higher in F1 and F3. The comparison of iRep values to the independently calculated MCR values demonstrated that the MAGs involved in polymer degradation had low replication indices (iRep 1.4–1.5) representing slow growing cellulolytic microflora. Similar observations were described in a previous study (Campanaro et al., 2020). These results emphasize that the complex polysaccharide substrate degradation can be linked to a slowly reproducing microbial community, which also may contribute to the development of microbial dynamics estimation in mathematical models (Weinrich et al., 2019). As expected, the methanogen microbes associated with the phylum Euryarchaeota represented the "slow growing" community in the GWB, MABA and TCM digesters (iRep ≤1.5) (Campanaro et al., 2020). The replication indices of archaeal species were under the median value (iRep 1.4). Nevertheless, the iRep of a *Methanobacterium* MAG (Bin_4: iRep 1.5, a hydrogenotroph methanogen) was found to be higher than

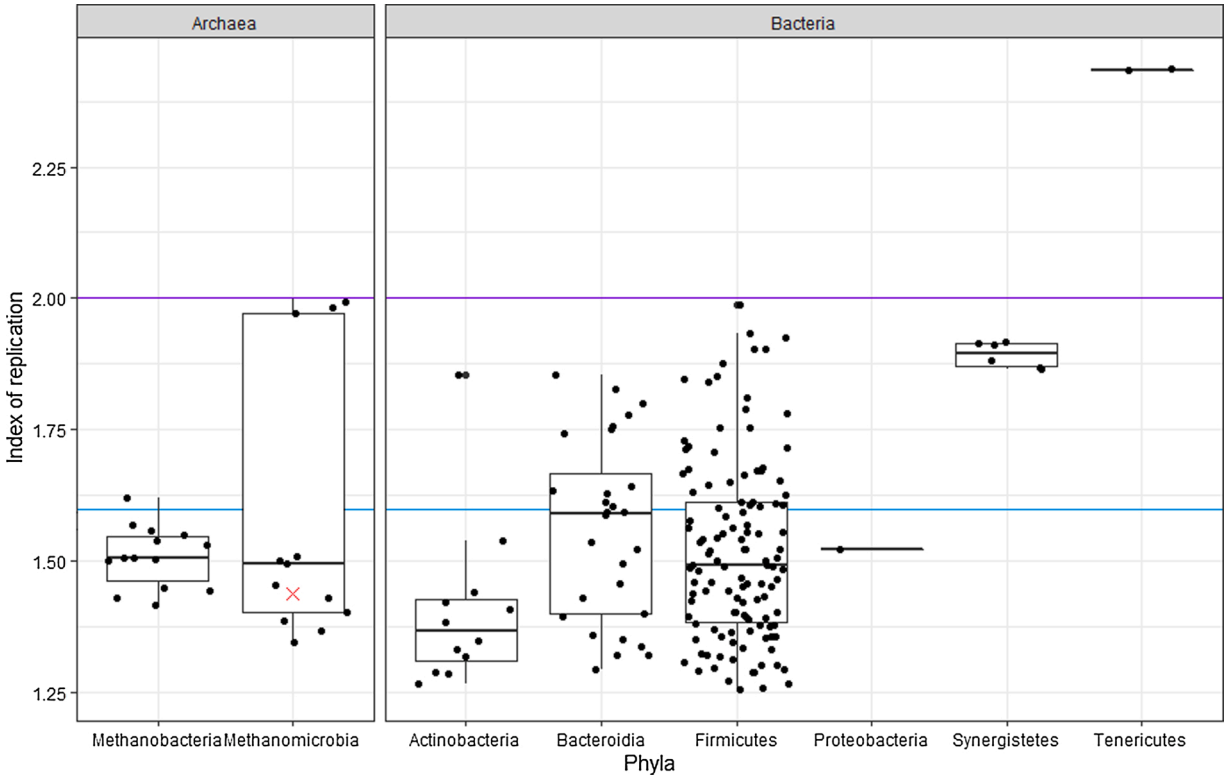

Fig. 6. Box plot displays the index of replication calculations at the phylum level. The specific iREP values (iREP values of MAGS/time/digester) are plotted in dots. The red X symbol indicates the mean replication index of *Methanotrix soehngenii*. The purple line marks high iREP values (2), and the blue line indicates median iREP values (1.6) measured in this experiment.

the average replication index of archaeal species. As previously described, this MAG also has been found in elevated quantities in F2 and F3, further strengthening the positive contribution of co-substrates in CH_4 formation. In our study, *Methanothrix soehngenii* (Bin_103: iRep 1.4) MAG (acetotroph methanogen) showed slower replication than that observed in a previous comprehensive AD microbiome study (iRep >2) (Fig. 6) (Campanaro et al., 2020).

4. Conclusions

Second generation green energy willow biomass was co-digested with low C/N ratio substrates, containing low lignin and high soluble content, (chicken manure and microalgae biomass). The initial lignocellulose degrading microbiome successfully adapted to the decomposition of mixed substrates. Several CAZymes were found in the rate limiting hydrolysing step confirming the evolution of complex polysaccharide digesting microbial communities. Hydrogenotrophic methanogenesis was clearly shown to be the predominant CH_4 producing pathway. It is concluded that a balance between H_2 producers and consumers and the appropriate selection and combination of sustainable co-substrates are critical for the efficient development and operation of the biogas microbial community relying on waste materials and lignocellulosic substrates.

CRediT authorship contribution statement

Roland Wirth designed and performed the bioinformatics analyses and composed the manuscript. **Bernadett Pap**, **Dénes Dudits**, **Balázs Kakuk**, and **Zoltán Bagi** performed the experiments and analytical measurements. **Prateek Shetty** contributed to the metagenome analyses. **Kornél L. Kovács** and **Gergely Maróti** designed the study, composed the manuscript and thoroughly discussed the relevant literature. All authors read and approved the final manuscript.

Funding

This study has been supported in part by the Hungarian National Research, Development and Innovation Fund projects GINOP- 2.2.1-15-2017-00081, EFOP- 3.6.2-16-2017-00010 and 2020-1.1.2-PIACI-KFI-2020-00117. RW, BZ and GM received support from the Hungarian NKFIH fund projects PD132145, FK123902 and FK123899. This work was also supported by the Lendület-Programme (GM) of the Hungarian Academy of Sciences (LP2020-5/2020).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jbiotec.2021.08.002.

References

Abraham, A., Mathew, A.K., Park, H., Choi, O., Sindhu, R., Parameswaran, B., Pandey, A., Park, J.H., Sang, B.I., 2020. Pretreatment strategies for enhanced biogas production from lignocellulosic biomass. Bioresour. Technol. 301, 122725 https://doi.org/10.1016/j.biortech.2019.122725.

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Grüning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., Goecks, J., 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res. 44, W3–W10. https://doi.org/10.1093/nar/gkw343.

Alalwan, H.A., Alminshid, A.H., Aljaafari, H.A.S., 2019. Promising evolution of biofuel generations. Subject review. Renew. Energy Focus 28, 127–139. https://doi.org/10.1016/j.ref.2018.12.006.

Alneberg, J., Bjarnason, B.S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson, A.F., Quince, C., 2014. Binning metagenomic contigs by coverage and composition. Nat. Methods 11, 1144–1146. https://doi.org/10.1038/nmeth.3103.

Bohutskyi, P., Phan, D., Kopachevsky, A.M., Chow, S., Bouwer, E.J., Betenbaugh, M.J., 2018. Synergistic co-digestion of wastewater grown algae-bacteria polyculture biomass and cellulose to optimize carbon-to-nitrogen ratio and application of kinetic models to predict anaerobic digestion energy balance. Bioresour. Technol. 269, 210–220. https://doi.org/10.1016/j.biortech.2018.08.085.

Böjti, T., Kovács, K.L., Kakuk, B., Wirth, R., Rákhely, G., Bagi, Z., 2017. Pretreatment of poultry manure for efficient biogas production as monosubstrate or co-fermentation with maize silage and corn stover. Anaerobe. https://doi.org/10.1016/j.anaerobe.2017.03.017.

Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B.K., Schulz, F., Jarett, J., Rivers, A.R., Eloe-Fadrosh, E.A., Tringe, S.G., Ivanova, N. N., Copeland, A., Clum, A., Becraft, E.D., Malmstrom, R.R., Birren, B., Podar, M., Bork, P., Weinstock, G.M., Garrity, G.M., Dodsworth, J.A., Yooseph, S., Sutton, G., Glöckner, F.O., Gilbert, J.A., Nelson, W.C., Hallam, S.J., Jungbluth, S.P., Ettema, T.J. G., Tighe, S., Konstantinidis, K.T., Liu, W.T., Baker, B.J., Rattei, T., Eisen, J.A., Hedlund, B., McMahon, K.D., Fierer, N., Knight, R., Finn, R., Cochrane, G., Karsch-Mizrachi, I., Tyson, G.W., Rinke, C., Lapidus, A., Meyer, F., Yilmaz, P., Parks, D.H., Eren, A.M., Schriml, L., Banfield, J.F., Hugenholtz, P., Woyke, T., 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. 35, 725–731. https://doi.org/10.1038/nbt.3893.

Bozan, M., Akyol, Ç., Ince, O., Aydin, S., Ince, B., 2017. Application of next-generation sequencing methods for microbial monitoring of anaerobic digestion of lignocellulosic biomass. Appl. Microbiol. Biotechnol. 1–16. https://doi.org/10.1007/s00253-017-8438-7.

Brown, C.T., Olm, M.R., Thomas, B.C., Banfield, J.F., 2016. Measurement of bacterial replication rates in microbial communities. Nat. Biotechnol. 34, 1256–1263. https://doi.org/10.1038/nbt.3704.

Campanaro, S., Treu, L., Rodriguez-R, L.M., Kovalovszki, A., Ziels, R.M., Maus, I., Zhu, X., Kougias, P.G., Basile, A., Luo, G., Schlüter, A., Konstantinidis, K.T., Angelidaki, I., 2020. New insights from the biogas microbiome by comprehensive genome-resolved metagenomics of nearly 1600 species originating from multiple anaerobic digesters. Biotechnol. Biofuels 13, 1–18. https://doi.org/10.1186/s13068-020-01679-y.

Capson-Tojo, G., Moscoviz, R., Astals, S., Robles, Steyer, J.P., 2020. Unraveling the literature chaos around free ammonia inhibition in anaerobic digestion. Renewable Sustainable Energy Rev. 117, 109487 https://doi.org/10.1016/j.rser.2019.109487.

Coronado-Reyes, J.A., Salazar-Torres, J.A., Juárez-Campos, B., González-Hernández, J. C., 2020. Chlorella vulgaris, a microalgae important to be used in Biotechnology: a review. Food Sci. Technol. 2061, 1–11. https://doi.org/10.1590/fst.37320.

Cseri, A., Borbély, P., Poór, P., Fehér, A., Sass, L., Jancsó, M., Penczi, A., Rádi, F., Gyuricza, C., Digruber, T., Dudits, D., 2020. Increased adaptation of an energy willow cultivar to soil salinity by duplication of its genome size. Biomass Bioenergy 140. https://doi.org/10.1016/j.biombioe.2020.105655.

Dębowski, M., Zieliński, M., Grala, A., Dudek, M., 2013. Algae biomass as an alternative substrate in biogas production technologies - Review. Renewable Sustainable Energy Rev. 27, 596–604. https://doi.org/10.1016/j.rser.2013.07.029.

Dudits, D., Török, K., Cseri, A., Paul, K., Nagy, A.V., Nagy, B., Sass, L., Ferenc, G., Vankova, R., Dobrev, P., Vass, I., Ayaydin, F., 2016. Response of organ structure and physiology to autotetraploidization in early development of energy willow Salix viminalis. Plant Physiol. 170, 1504–1523. https://doi.org/10.1104/pp.15.01679.

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., Delmont, T. O., 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3, e1319. https://doi.org/10.7717/peerj.1319.

Estevez, M.M., Sapci, Z., Linjordet, R., Schnürer, A., Morken, J., 2014. Semi-continuous anaerobic co-digestion of cow manure and steam-exploded Salix with recirculation of liquid digestate. J. Environ. Manage. 136, 9–15. https://doi.org/10.1016/j.jenvman.2014.01.028.

Finn, R.D., Attwood, T.K., Babbitt, P.C., Bateman, A., Bork, P., Bridge, A.J., Chang, H.Y., Dosztanyi, Z., El-Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G.L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., Mistry, J., Natale, D.A., Necci, M., Nuka, G., Orengo, C.A., Park, Y., Pesseat, S., Piovesan, D., Potter, S.C., Rawlings, N.D., Redaschi, N., Richardson, L., Rivoire, C., Sangrador-Vegas, A., Sigrist, C., Sillitoe, I., Smithers, B., Squizzato, S., Sutton, G., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Wu, C.H., Xenarios, I., Yeh, L.S., Young, S.Y., Mitchell, A.L., 2017. InterPro in 2017-beyond protein family and domain annotations. Nucleic Acids Res. 45, D190–D199. https://doi.org/10.1093/nar/gkw1107.

Güllert, S., Fischer, M.A., Turaev, D., Noebauer, B., Ilmberger, N., Wemheuer, B., Alawi, M., Rattei, T., Daniel, R., Schmitz, R.A., Grundhoff, A., Streit, W.R., 2016. Deep metagenome and metatranscriptome analyses of microbial communities affiliated with an industrial biogas fermenter, a cow rumen, and elephant feces reveal major differences in carbohydrate hydrolysis strategies. Biotechnol. Biofuels 9, 121. https://doi.org/10.1186/s13068-016-0534-x.

Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11. https://doi.org/10.1186/1471-2105-11-119.

Jia, Y., Ng, S.K., Lu, H., Cai, M., Lee, P.K.H., 2018. Genome - centric metatranscriptomes and ecological roles of the active microbial populations during cellulosic biomass

anaerobic digestion. Biotechnol. Biofuels 11, 1–15. https://doi.org/10.1186/s13068-018-1121-0.
Jünemann, S., Kleinbölting, N., Jaenicke, S., Henke, C., Hassa, J., Nelkner, J., Stolze, Y., Albaum, S.P., Schlüter, A., Goesmann, A., Sczyrba, A., Stoye, J., 2017. Bioinformatics for NGS-based metagenomics and the application to biogas research. J. Biotechnol. 261, 10–23. https://doi.org/10.1016/j.jbiotec.2017.08.012.
Kafle, G.K., Kim, S.H., 2013. Effects of chemical compositions and ensiling on the biogas productivity and degradation rates of agricultural and food processing by-products. Bioresour. Technol. 142, 553–561. https://doi.org/10.1016/j.biortech.2013.05.018.
Kakuk, B., Bagi, Z., Rákhely, G., Maróti, G., Dudits, D., Kovács, K.L., 2021. Methane production from green and woody biomass using short rotation willow genotypes for bioenergy generation. Bioresour. Technol. 333, 125223 https://doi.org/10.1016/j.biortech.2021.125223.
Kanehisa, M., Guto, S., 2000. Kegg: kyoto encyclopedia of genes and genomes. Infect. Genet. Evol. 44, 313–317. https://doi.org/10.1016/j.meegid.2016.07.022.
Kang, D.D., Froula, J., Egan, R., Wang, Z., 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3, e1165. https://doi.org/10.7717/peerj.1165.
Klassen, V., Blifernez-Klassen, O., Wobbe, L., Schlüter, A., Kruse, O., Mussgnug, J.H., 2016. Efficiency and biotechnological aspects of biogas production from microalgal substrates. J. Biotechnol. 234, 7–26. https://doi.org/10.1016/j.jbiotec.2016.07.015.
Kougias, P.G., Angelidaki, I., 2018. Biogas and its opportunities — a review. Front. Environ. Sci. 12, 1–22.
Kougias, P.G., Campanaro, S., Treu, L., Tsapekos, P., Armani, A., Angelidaki, I., 2018. Spatial distribution and diverse metabolic functions of lignocellulose-degrading uncultured bacteria as revealed by genomecentric metagenomics. Appl. Environ. Microbiol. 84, 1–14. https://doi.org/10.1128/AEM.01244-18.
Kovács, K.L., Ács, N., Kovács, E., Wirth, R., Rákhely, G., Strang, O., Herbel, Z., Bagi, Z., 2013. Improvement of biogas production by bioaugmentation. Biomed Res. Int. 2013 https://doi.org/10.1155/2013/482653.
Kumar, R., Singh, S., Singh, O.V., 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. J. Ind. Microbiol. Biotechnol. 35, 377–391. https://doi.org/10.1007/s10295-008-0327-8.
Lam, M.K., Lee, K.T., 2012. Microalgae biofuels: a critical review of issues, problems and the way forward. Biotechnol. Adv. 30, 673–690. https://doi.org/10.1016/j.biotechadv.2011.11.008.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923.
Li, D., Liu, C.-M., Luo, R., Sadakane, K., Lam, T.-W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31 (10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033.
Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, 490–495. https://doi.org/10.1093/nar/gkt1178.
Mende, D.R., Letunic, I., Maistrenko, O.M., Schmidt, T.S.B., Milanese, A., Paoli, L., Hernández-Plaza, A., Orakov, A.N., Forslund, S.K., Sunagawa, S., Zeller, G., Huerta-Cepas, J., Coelho, L.P., Bork, P., 2020. ProGenomes2: an improved database for accurate and consistent habitat, taxonomic and functional annotations of prokaryotic genomes. Nucleic Acids Res. 48, D621–D625. https://doi.org/10.1093/nar/gkz1002.
Montingelli, M.E., Tedesco, S., Olabi, A.G., 2015. Biogas production from algal biomass: a review. Renewable Sustainable Energy Rev. 43, 961–972. https://doi.org/10.1016/j.rser.2014.11.052.
Neshat, S.A., Mohammadi, M., Najafpour, G.D., Lahijani, P., 2017. Anaerobic co-digestion of animal manures and lignocellulosic residues as a potent approach for sustainable biogas production. Renewable Sustainable Energy Rev. 79, 308–322. https://doi.org/10.1016/j.rser.2017.05.137.
Nielsen, H.B., Angelidaki, I., 2008. Strategies for optimizing recovery of the biogas process following ammonia inhibition. Bioresour. Technol. 99, 7995–8001. https://doi.org/10.1016/j.biortech.2008.03.049.
Nordberg, Å., Jarvis, Å., Stenberg, B., Mathisen, B., Svensson, B.H., 2007. Anaerobic digestion of alfalfa silage with recirculation of process liquid. Bioresour. Technol. 98, 104–111. https://doi.org/10.1016/j.biortech.2005.11.027.
Parks, D.H., Beiko, R.G., 2010. Identifying biologically relevant differences between metagenomic communities. Bioinformatics 26, 715–721. https://doi.org/10.1093/bioinformatics/btq041.
Parks, D.H., Chuvochina, M., Chaumeil, P.A., Rinke, C., Mussig, A.J., Hugenholtz, P., 2020. A complete domain-to-species taxonomy for Bacteria and Archaea. Nat. Biotechnol. 38, 1079–1086. https://doi.org/10.1038/s41587-020-0501-8.
Paul, S., Dutta, A., 2018. Challenges and opportunities of lignocellulosic biomass for anaerobic digestion. Resour. Conserv. Recycl. 130, 164–174. https://doi.org/10.1016/j.resconrec.2017.12.005.
Rodriguez-R, L.M., Gunturu, S., Harvey, W.T., Rosselló-Mora, R., Tiedje, J.M., Cole, J.R., Konstantinidis, K.T., 2018. The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. Nucleic Acids Res. 46, W282–W288. https://doi.org/10.1093/nar/gky467.
Rulli, M.C., Bellomi, D., Cazzoli, A., De Carolis, G., D'Odorico, P., 2016. The water-land-food nexus of first-generation biofuels. Sci. Rep. 6, 1–10. https://doi.org/10.1038/srep22521.
Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
Shi, X., Lin, J., Zuo, J., Li, P., Li, X., Guo, X., 2016. Effects of free ammonia on volatile fatty acid accumulation and process performance in the anaerobic digestion of two typical bio-wastes. J. Environ. Sci. China (China) 55, 1–9. https://doi.org/10.1016/j.jes.2016.07.006.
Solé-Bundó, M., Passos, F., Romero-Güiza, M.S., Ferrer, I., Astals, S., 2019. Co-digestion strategies to enhance microalgae anaerobic digestion: a review. Renewable Sustainable Energy Rev. 112, 471–482. https://doi.org/10.1016/j.rser.2019.05.036.
Turaev, D., Rattei, T., 2016. High definition for systems biology of microbial communities: metagenomics gets genome-centric and strain-resolved. Curr. Opin. Biotechnol. 39, 174–181. https://doi.org/10.1016/j.copbio.2016.04.011.
van der Lelie, D., Taghavi, S., McCorkle, S.M., Li, L.-L., Malfatti, S.A., Monteleone, D., Donohoe, B.S., Ding, S.-Y., Adney, W.S., Himmel, M.E., Tringe, S.G., 2012. The metagenome of an anaerobic microbial community decomposing poplar wood chips. PloS One 7 (5). https://doi.org/10.1371/journal.pone.0036740.
Vanwonterghem, I., Evans, P.N., Parks, D.H., Jensen, P.D., Woodcroft, B.J., Hugenholtz, P., Tyson, G.W., 2016. Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. Nat. Microbiol. 1, 1–9. https://doi.org/10.1038/nmicrobiol.2016.170.
Weinrich, S., Koch, S., Bonk, F., Popp, D., Benndorf, D., Klamt, S., Centler, F., 2019. Augmenting biogas process modeling by resolving intracellular metabolic activity. Front. Microbiol. 10, 1–14. https://doi.org/10.3389/fmicb.2019.01095.
Wirth, R., Kádár, G., Kakuk, B., Maróti, G., Bagi, Z., Szilágyi, Á., Rákhely, G., Horváth, J., Kovács, K.L., 2018a. The planktonic core microbiome and core functions in the cattle rumen by next generation sequencing. Front. Microbiol. 9, 1–19. https://doi.org/10.3389/fmicb.2018.02285.
Wirth, R., Lakatos, G., Böjti, T., Maróti, G., Bagi, Z., Rákhely, G., Kovács, K.L., 2018b. Anaerobic gaseous biofuel production using microalgal biomass – a review. Anaerobe 52, 1–8. https://doi.org/10.1016/j.anaerobe.2018.05.008.
Wirth, R., Böjti, T., Lakatos, G., Maróti, G., Bagi, Z., Rákhely, G., Kovács, K.L., 2019. Characterization of core microbiomes and functional profiles of mesophilic anaerobic digesters fed with chlorella vulgaris green microalgae and maize silage. Front. Energy Res. 7 https://doi.org/10.3389/fenrg.2019.00111.
Wirth, R., Pap, B., Böjti, T., Shetty, P., Lakatos, G., 2020. Chlorella vulgaris and its phycosphere in wastewater : microalgae-bacteria interactions during nutrient removal. Front. Bioeng. Biotechnol. 8, 1–15. https://doi.org/10.3389/fbioe.2020.557572.
Wu, Y.W., Simmons, B.A., Singer, S.W., 2015. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32, 605–607. https://doi.org/10.1093/bioinformatics/btv638.
Xue, S., Song, J., Wang, X., Shang, Z., Sheng, C., Li, C., Zhu, Y., Liu, J., 2020. A systematic comparison of biogas development and related policies between China and Europe and corresponding insights. Renewable Sustainable Energy Rev. 117, 109474 https://doi.org/10.1016/j.rser.2019.109474.
Yenigün, O., Demirel, B., 2013. Ammonia inhibition in anaerobic digestion: a review. Process Biochem. 48, 901–911. https://doi.org/10.1016/j.procbio.2013.04.012.
Zabed, H.M., Akter, S., Yun, J., Zhang, G., Awad, F.N., Qi, X., Sahu, J.N., 2019. Recent advances in biological pretreatment of microalgae and lignocellulosic biomass for biofuel production. Renewable Sustainable Energy Rev. 105, 105–128. https://doi.org/10.1016/j.rser.2019.01.048.
Zhong, W., Zhang, Z., Luo, Y., Qiao, W., Xiao, M., Zhang, M., 2012. Biogas productivity by co-digesting Taihu blue algae with corn straw as an external carbon source. Bioresour. Technol. 114, 281–286. https://doi.org/10.1016/j.biortech.2012.02.111.
Zhu, X., Campanaro, S., Treu, L., Kougias, P.G., Angelidaki, I., 2019. Novel ecological insights and functional roles during anaerobic digestion of saccharides unveiled by genome-centric metagenomics. Water Res. 151, 271–279. https://doi.org/10.1016/j.watres.2018.12.041.