

Genome-centric metagenomic insights into the impact of alkaline/acid and thermal sludge pre-treatment on digestion sludge microbiome

Shanquan Wang (✉ wangshanquan@mail.sysu.edu.cn)

Sun Yat-Sen University <https://orcid.org/0000-0002-7153-6684>

Zhiwei Liang

Sun Yat-Sen University

Jiangjian Shi

Sun Yat-Sen University

Chen Wang

Sun Yat-Sen University

Junhui Li

Northern Arizona University

Zhili He

Sun Yat-Sen University

Research

Keywords: sludge pre-treatment, digestion sludge microbiome, metagenome, APAD

Posted Date: January 9th, 2020

DOI: <https://doi.org/10.21203/rs.2.20429/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Wastewater treatment generates large amounts of waste activated sludge (WAS), which mainly consist of recalcitrant microbial cells and particulate organic matter. WAS pre-treatment is an effective way to destabilize sludge floc structure and release cellular macromolecules and other organic matter for improvement of digestion efficiency. Nonetheless, impacts of WAS pre-treatment on the complex digestion sludge microbiome, as well as mechanistic insight into how sludge pre-treatment improve digestion performance, remain to be elucidated.

Results: In this study, genome-centric metagenomic approach was employed to investigate the digestion sludge microbiome in four methanogenic sludge digesters with different feeding sludge: APAD, WAS pre-treated with 0.25 mol/L alkaline/acid; HS-APAD, WAS pre-treated with 0.8 mol/L alkaline/acid; Thermal-AD, thermal pre-treated WAS; Control-AD, fresh WAS. We retrieved 254 metagenomic-assembled genomes (MAGs) to identify the key functional populations involved in methanogenic digestion process. These MAGs span 28 phyla with 69 of them as yet-to-be-cultivated lineages, and 30 novel lineages were characterized with metabolic potential associated with hydrolysis and fermentation. Interestingly, functional populations involving carbohydrate digestion were overrepresented in APAD and HS-APAD, while lineages related to protein and lipid fermentation were overrepresented in Thermal-AD, reflecting different digestion substrates released from alkaline/acid and thermal pre-treatments. Of the three major functional populations, i.e., fermentative bacteria, acetogenic syntrophs and methanogens, significant correlations between genome sizes of the fermentative bacteria and their abundance were observed, particularly in the APAD and HS-APAD with improved digestion performance.

Conclusion: These genome-centric metagenomic insights advance our understanding of sludge pre-treatment on digestion sludge microbiomes, shedding light on future optimization of methanogenic sludge digestion and resource recovery.

Introduction

Waste activated sludge (WAS) from wastewater treatment plants contains high levels of organic matter in forms of cells, extracellular polymeric substances (EPS) and macromolecules generated from cell lysis, as well as pathogens and other biohazards [1, 2]. Anaerobic digestion as a sustainable sludge treatment technology can convert these organic substances into biogas via a multiple-step process consisting of hydrolysis, fermentation, acetogenesis and methanogenesis [3]. The organic matter (e.g., microbial cells and EPS), together with metals and other ions, in WAS form stable and complicated sludge flocs. These sludge flocs are recalcitrant to anaerobic digestion and consequently require sludge pre-treatment to destabilize their structure and release organic matter for improving digestion efficiency [4]. Many pre-treatment techniques (e.g., thermal, ultrasonic, microwave and alkaline pre-treatments) have been developed to effectively release sludge organic matter [5, 6, 7]. For example, thermal hydrolysis (> 100 °C), ultrasonication, microwave and alkaline pre-treatments could enhance sludge fermentation by 4–20%, 1–18%, 1–15% and 3–35%, respectively, compared to their non-pretreated controls [7]. In our previous

studies, a new alkaline/acid pre-treatment and anaerobic digestion (APAD) process was developed, in which organic carbon removal achieved $52.8 \pm 1.7\%$ [2]. By contrast, the organic carbon removals were significantly lower in other digesters under the same operational conditions but with different influent sludge, i.e., $42.4 \pm 1.6\%$ in Thermal-AD with thermal sludge pre-treatment and $30.9 \pm 2.2\%$ in Control-AD with fresh WAS [2]. Further increasing alkaline/acid concentrations from 0.25 mol/L (APAD) to 0.8 mol/L (HS-APAD) did not show notable changes in digestion performance and community taxonomic composition in the sludge digesters [8]. Consequently, dissolved organic compounds (DOC) derived from sludge pre-treatments, rather than salinity, could be a predominant selective pressure driving the performance and microbiome changes in both APAD and HS-APAD, compared to the Control-AD [8]. Nonetheless, the detailed mechanistic insights into the digestion improvement and impact of WAS pre-treatment on digestion sludge microbiome remain unknown.

The conversion of WAS organic matter into biogas highly relies on the complex and tightly coupled synergistic interactions of microbial communities in digestion sludge [9]. Previous studies based on 16S rRNA gene amplicon sequencing showed that relative abundance of bacteria and archaea in digestion sludge microbiomes generally accounted for $> 95\%$ and $< 5\%$, respectively [10, 11], being further confirmed by recent metagenomics analyses [9, 12]. In view of the bacterial community, Bacteroidetes, Proteobacteria, Spirochaetes and Firmicutes were generally the dominant phyla in sludge digesters, and most of them were fermentative bacteria with high compositional and functional redundancy [9, 13–16]. By contrast, the slow-growing acetogenic syntrophs and methanogenic archaea in digestion sludge were limited to several lineages [9, 11]. Accumulating experimental evidences suggested that a variety of factors including feeding substrates and operational parameters (e.g., pH, temperature and ammonia) might change community composition and function of the digestion sludge microbiome [11, 12, 16, 17]. The information generated from the 16S rRNA gene-based analyses was largely limited to community composition and succession [18]. Metagenome sequencing could theoretically obtain all microbial genome information within a sample, which might provide direct access to the metabolic potential and networks in the highly complex digestion sludge microbiome. Early metagenomic studies mainly relied on gene-centric analyses, which were biased towards existing databases [9, 19, 20]. Current advances in both high-throughput sequencing technologies and population genome binning algorithms allowed the development of genome-centric approaches for the assessment of complex microbiomes [12, 18, 21]. For example, the genome-centric metagenomics analysis was employed to recover 101 population genomes and revealed their metabolic potential and interactions in a cellulose-degrading digester [16]. Very recently, a collection of 1,635 metagenome-assembled genomes were recovered from publicly available datasets derived from different methanogenic digesters, showing high species diversity related to methane generation [12]. Also, metagenomics changed the pace of virus discovery by enabling the accurate identification of viral genome sequences without requiring isolation of viruses [22, 23]. In contrast to increasing metagenomic data on anaerobic digestion, no information was available on impacts of WAS pre-treatment on digestion sludge microbiome. It would be rational to assume that the genome-centric and strain-resolved metagenomic approaches could provide a systematic understanding

of digestion sludge microbiomes, particularly the yet-to-be elucidated impact of WAS pre-treatment on the digestion sludge microbiome.

In this study, we employed metagenomic approach to explore prokaryotic and DNA viral community composition and function of digestion sludge microbiomes in four sludge digesters (i.e., APAD, HS-APAD, Thermal-AD and Control-AD). Co-assembly of the four metagenomes followed by genomic binning resulted in the recovery of 254 population genomes that constituted the majority of the digestion sludge microbiome. Their metabolic potential and networks were further reconstructed to reveal impacts of WAS pre-treatment on the digestion sludge microbiome. The results provided the first genome-centric insight into how WAS pre-treatment change community composition and function, as well as metabolic networks of key players and their genomic traits in the digestion sludge microbiome.

Methods

Sample collection, DNA extraction and metagenome sequencing

Four mesophilic anaerobic sludge digesters were setup to treat fresh WAS (Control-AD), thermal pre-treated sludge (Thermal-AD), and sludge pre-treated with 0.25 mol/L alkaline/acid (APAD) and 0.8 mol/L alkaline/acid (HS-APAD), as described [2, 8]. Sludge samples for metagenomic sequencing were collected from steady-state digesters (sampling time points were set on day 192 for APAD, Thermal-AD Control-AD, and on day 142 for HS-APAD), and their gDNA was extracted using FastDNA Spin Kit for Soil (MP Biomedicals, Carlsbad, CA, USA) [24]. The quality and quantity of the DNA extract were evaluated with gel electrophoresis and Quantus Fluorometer (Promega, Madison, WI, USA). The DNA sequencing libraries and subsequent Illumina HiSeq sequencing services were provided by BGI (Shenzhen, China).

Metagenome assembly and genome binning

All metagenomic sequencing raw data were filtered to remove low quality bases/reads using Sickle [25], with the parameters set to “-q = 20 and -l = 100”. Then the four metagenomes were combined and de novo co-assembled using SPAdes (version 3.12.0) [26] with the following parameters, -k 33,55,77 -meta. To generate high-quality metagenome-assembled genomes (MAGs), contigs with length < 1000 bp in the assembly were removed. Three binning methods, i.e., MetaBAT [27], Maxbin [28] and Concoct [29], were employed and compared for the binning of the contigs into population genomes as described [30]. To improve binning quality, RefineM [31] was used to filter scaffolds with divergent genomic properties, incongruent taxonomic classification and 16S rRNA genes. Manual curation of these population genomes was executed using single copy genes, k-mer frequency distribution, contigs coverage and GC content. To get the optimal genome quality, clean reads for each bin were recruited using BBMap [32] with following parameters: k = 15 minid = 0.9 build = 1. Then, genome bins were reassembled by SPAdes (version 3.12.0) [26] with the following parameters: -careful -k 21,33,55,77. The completeness,

contamination and strain heterogeneity of each bin were evaluated using CheckM [33]. Finally, 254 bins with > 70% completeness and < 5% contamination were obtained for subsequent analyses.

Genome annotation and metabolic reconstruction

Protein coding sequences (CDS) were determined using prokka [34] with the “–quiet” option for all genome bins. Functional annotations were conducted based on comparisons with the KEGG [35], and eggnog [36] databases using DIAMOND [37] with an E-value threshold of < 1e-5. Carbohydrate-active enzymes were identified using the carbohydrate-active enZYmes (CAZy) database [38]. A subset of CAZy genes involving in extracellular polysaccharide degradation pathways [39] was selected for further analysis. Peptidases were identified using MEROPS [40]. To reconstruct the metabolic pathways, genome bins were upload to RAST (Rapid Annotation using Subsystem Technology) [41] for gene prediction and annotation, and their CDS were annotated at the KEGG automatic annotation server (KAAS) [42]. The metabolic potential (i.e., hydrolysis, fermentation and syntrophic acetogenesis) of major populations of known lineages were further manually confirmed.

Genome tree phylogeny

A total of 16 universally present and rarely horizontally transferred ribosomal proteins (i.e., L2, L3, L4, L5, L6, L14, L15, L16, L18, L22, L24, S3, S8, S10, S17 and S19) were selected to reconstruct the phylogenetic tree [43]. Protein sequences were extracted separately using AMPHORA2 [44]. Multiple sequence alignments (MSAs) of the individual protein sequence were built using Mafft [45] (iterate 100 times). Poorly aligned regions were filtered by TrimAL [46] to remove columns containing > 95% gap positions, and then the 16 filtered ribosomal protein sequences were concatenated for each genome. The phylogenomic tree was constructed using IQ-TREE (version 1.6.10) [47] with the following parameters: -mset JTT -mrate I,G,I + G. The resulting newick tree file was uploaded to iTOL v4 [48] for visualization and formatting. To confirm their phylogeny, MAGs were further classified with GTDBtk (version 1.0.2, database release 89) [49].

Relative abundance and community composition

Taxonomic classification of contigs was performed using Kraken 2 as described [50] with a database comprising 38,156 bacterial, 11,953 viral, and 535 archaeal complete reference genomes. The sequencing reads were then re-aligned back to the contigs using BMap [32] to calculate their coverages. Taxon read hit counts were obtained by combining the re-aligned reads of all contigs belonging to the taxon, which were further normalized by genome size to calculate taxon abundance. For the relative abundance and diversity analysis, genomes with coverage less than 1 were removed to decrease effect of low abundance misclassification. The abundance of MAGs was calculated based on retrieved genome coverage and normalized as reads per kilobase per million (RPKM) as described previously [51, 52].

Results

The superkingdom community composition of digestion sludge microbiomes

Four metagenomes (144 Gb total raw sequencing reads) from APAD, HS-APAD, Thermal-AD and Control-AD were co-assembled, generating 833,655 contigs with a combined length of 2,395 Mb (N50 = 3,658 bp; Table S1). Community composition investigation based on the relative number of microorganisms showed that Bacteria, Archaea and DNA viruses were the three major taxonomic groups of the digestion sludge microbiome, accounting for an average relative abundance of 92.9%, 2.4% and 4.7%, respectively (Fig. 1A). Notably, the lowest relative abundance of virus in Thermal-AD (i.e., 2.9% in Thermal-AD vs. 5.9% in Control-AD, 4.9% in APAD and 4.9% in HS-APAD) might be due to the efficient removal of bacteriophages in thermal sludge pre-treatment [53]. Also, compared to Control-AD, the higher abundance of methanogenic Archaea in APAD and HS-APAD corroborated their enhanced carbon removal and biogas generation [2, 8]. Principal coordinates analysis (PCoA) showed that thermal and alkaline/acid pre-treatment had very different impacts on the digestion sludge microbiome. By contrast, APAD and HS-APAD shared similar microbial community patterns (Fig. 1B), consistent with their 16S rRNA gene-based community clustering (Fig. S1) [2, 8].

Population genome binning of the co-assembled metagenomes enabled the recovery of 245 bacterial and 9 archaeal population genomes with > 70% completeness and < 5% contamination (Fig. 2; Table S2). Both bacterial and archaeal MAGs represented major populations in digestion sludge microbiomes, i.e., 66.4%, 65.7%, 38.2% and 30.2% of total community in APAD, HS-APAD, Thermal-AD and Control-AD, respectively, based on percentages of reads mapping. These populations were phylogenetically diverse and belonged to 28 different phyla, of which 69 MAGs represented novel lineages at order and even higher taxonomic levels (Fig. S2). Major populations fulfilling different steps in anaerobic sludge digestion (i.e., hydrolysis, fermentation, syntrophic acetogenesis and methanogenesis) had significant difference in their abundance normalized as reads per kilobase per million (RPKM), e.g., 31.4-200.9 for fermentative bacteria, 2.5–12.5 for syntrophic acetogens and 0.75–8.3 for methanogens (Fig. 2), which could be further corroborated by the 16S rRNA gene-based analyses (Fig. S1).

Metabolic networks in sludge digesters

To construct the metabolic networks, metabolic potential of major microorganisms (or MAGs) involving conversion of sludge organic matter into biogas were identified based on the carbon flow from organic macromolecules to methane. The predominant organic macromolecules to be degraded in the sludge digesters were polysaccharides, proteins and lipids. These macromolecules could be converted into methane under mediation of three major groups of functional microorganisms, i.e., hydrolyzing and fermentative bacteria, syntrophic acetogenic bacteria and methanogenic archaea (Table S3).

Pertaining to metabolic networks in the four sludge digesters, presence of multiple populations capable of fulfilling the same function (e.g., hydrolysis, fermentation and methanogenesis) suggested high levels of functional redundancy (Fig. 3; Table S4). The high levels of functional redundancy were contributed by

both core functional populations (major and essential functional microorganisms) and redundant functional populations (minor and non-essential functional populations), and the two different groups of functional populations could be interconvertible when changing operational conditions of sludge digestion. In line with the microbial community composition, the overall metabolic networks in APAD and HS-APAD were highly similar ($p = 0.696$; Fig. 3C and D), but different from them in Thermal-AD and Control-AD (Fig. 3A and B). These distinct metabolic networks indicated different digestion substrates released from alkaline/acid and thermal sludge pre-treatments. Notably, the polysaccharide and protein metabolisms were significantly enriched in APAD and HS-APAD (Fig. 3C and D), suggesting effective destruction of both EPS and microbial cells to release polysaccharides and proteins in the alkaline/acid pre-treatment. Accordingly, bacterial MAGs of Firmicutes (e.g., bin132 and bin166) and Bacteroidetes (e.g., bin58 and bin252) phyla were selectively enriched as core functional populations in APAD and HS-APAD. In addition, the predominant hydrolyzing and fermentative bacterial populations were different in Thermal-AD and Control-AD, e.g., bin169 of Chloroflexi and bin202 of Lentisphaerae in the Thermal-AD, and bin140 of Proteobacteria and bin161 of Actinobacteria in the Control-AD. In syntrophic acetogenic process, propionate and butyrate were major volatile fatty acids (VFAs) as substrates for syntrophic acetogenic bacteria (e.g., bin42 and bin137 of Synergistetes and Firmicutes phyla, respectively, for butyrate oxidation, and bin75 and bin77 of Proteobacteria phylum for propionate oxidation), which were significantly enriched as common and core syntrophic acetogens in sludge digesters fed with thermal or alkaline/acid pre-treated WAS (Fig. 3; Table S4). For the methanogenesis, a variety of acetoclastic and hydrogenotrophic methanogens were observed with a high level of functional redundancy in the four digesters. Particularly, Methanosaeta (bin51) capable of producing methane from acetate, formate and H_2 was enriched as core methanogens in APAD and HS-APAD (Fig. 3C and D), being in good agreement with their high methane production [2, 8].

Microbial dark matter

The "microbial dark matter" as microorganisms evaded cultivation and yet-to-be-characterized physiologically [54] could play important roles in sludge digesters. From the four digestion sludge microbiomes, 30 MAGs belonging to 12 phyla were identified as representative dark matter (Fig. 4 and Table S5), which were phylogenetically distant from their closest known sister lineages and represented novel microorganisms at order and even higher taxonomic levels. Genome-predicted metabolic potential suggested that 24 of 30 MAGs involved in hydrolysis and fermentation of cellular macromolecules in WAS. The abundances of these hydrolyzing and fermentative dark matter were generally < 2.1 RPKM, with an exception of the abundant bin17 and bin106 populations in Thermal-AD. Notably, several populations of candidate phyla with small genomes (< 1 Mb), e.g., Candidatus Stahlbacteria (TM7), Candidatus Shapirobacteria, and Candidatus Dojkabacteria, were identified in the sludge microbiomes without functions directly linked to the major methanogenic digestion procedures (Fig. 4).

Of the 30 MAGs, the two most abundant populations (i.e., bin17 and bin106 with 89.2% and 76.6% genome completeness, respectively) belonging to the Planctomycetes phylum were chosen for subsequent metabolic reconstruction (Fig. 5; Table S6). Both genomes encoded enzymes of glycosyllic

hydrolase (GH) families for hydrolysis of cellulose, hemicellulose and cellobiose, which could convert sludge polysaccharides into simple sugars for their subsequent fermentation and acetogenesis. In contrast to the nearly complete tricarboxylic acid (TCA) cycle in bin106 genome, genes encoding 2-oxoglutarate/2-oxoacid ferredoxin oxidoreductase and succinyl-CoA synthetase were absent in bin17 genome. Consequently, the partial TCA cycle might only provide biosynthetic precursors for anabolism of bin17 population. Notably, in addition to the glycolysis pathway, genes in bin106 genome also encoded enzymes for β -oxidation of fatty acids and phenol degradation, as well as a quinone-dependent electron transport chain containing cbb3-type cytochrome-c oxidase with high O_2 -affinity. These potential metabolic traits implied a versatile lifestyle of bin106 population and its aerobic respiration capability under microaerobic conditions. In addition, cell mobility of bin17 and bin106 populations might be different based on their gene contents. For example, genes encoding proteins for basal body, hook and filament assembly, as well as type IV pilus, were detected in bin17 genome, suggesting its high mobility and capability to move toward favorable growth environment. In contrast, only pilus-encoding genes were detected in bin106 genome, implying limited motility of bin106 population.

Genomic traits of digestion sludge microbiomes

Genomic traits (e.g., genome size, coding sequence or CDS, and GC content) of microorganisms could be defining features of microbial cell growth, and consequently associated with population abundance in a complex microbiome [55, 56, 57]. For instance, populations with small genome size would be expected to grow rapidly and therefore to have high abundance, because small genomes could reduce mutational load and nutrient demands for genome replication [58, 59]. In the four sludge digesters, average genome size and CDS of functional microorganisms (i.e., fermentative bacteria, acetogenic syntrophs and methanogens) decreased with the successive methanogenic digestion procedures (Fig. 6A; Fig. S3), consistent with the Gibbs free energy changes of each redox reactions. Particularly, the genome size was shown to have a significant ($p < 0.01$) correlation with the abundance of overall digestion sludge microbiome (Fig. 6B). Of the three major functional groups of microorganisms, acetogenic syntrophs and methanogens mediated the rate-limiting steps (i.e., acetogenesis and methanogenesis) and their growth was substrate-replete compared to fermentative bacteria. Accordingly, genome size could best predict the cell growth rate of the substrate-competitive fermentative bacteria ($p < 0.01$; Fig. 6B; Fig. S4A and B). Interestingly, correlations between genome sizes of fermentative bacteria and their abundance were significant ($p < 0.01$) in APAD and HS-APAD with high carbon removal efficiencies ($> 50\%$), in contrast to their low correlations in Control-AD ($p = 0.35$) and Thermal-AD ($p = 0.78$; Fig. 6C). The low correlations in Control-AD and Thermal-AD might be due to their low efficiencies (30–40%) in destruction and removal of sludge flocs and associated microbial cells [2]. The low efficiencies further resulted in high retaining of cellular macromolecules (e.g., genomic DNA) of influent sludge cells (Fig. 6A; Fig. S4C) and consequent complication of the digestion sludge microbiome.

Discussion

This study provided the first genome-centric metagenomic insight into how WAS pre-treatments changed the digestion sludge microbiome and subsequent digestion performance, as well as into the associated microbial dark matter. The highly complex digestion sludge microbiome guaranteed efficient conversion of organic polymers into biogas, of which performance could be changed by a variety of parameters, particularly influent sludge composition and operational conditions (e.g., pH, temperature and salinity) [60, 61]. Nonetheless, impacts of these parameters on digestion sludge microbiomes were combined and twisted, and it was consequently challenging to identify the exact impacts of a specific parameter on the sludge microbiome and to differentiate the core functional populations from functionally redundant and to-be-digested microorganisms derived from influent WAS. In view of the microbial community composition and function, prokaryotic populations in digestion sludge microbiomes could be classified into three major groups: (1) Group-I, core functional populations which were abundant and played essential roles in sludge digestion (e.g., bin166, bin252 and bin58 as the predominant hydrolyzing and fermentative bacteria, bin77 as syntrophic acetogens and bin51 as methanogenic archaeal in APAD and HS-APAD); (2) Group-II, redundant functional populations which had parallel functions with the core functional populations but were minor and non-essential populations; (3) Group-III, non-functional populations which entered digesters as to-be-digested substrates (e.g., aerobic bacteria derived from influent WAS). Recently, a study investigating the digestion efficiency in full-scale anaerobic sludge digesters estimated that 82% of microbial populations in feeding WAS were the Group-III non-functional populations and could be digested [10]. Most of the (Group-III) non-functional populations in feeding sludge were facultative aerobes or anaerobes, functions of which could be barely inferred from their 16S rRNA gene-based information and be easily confused with functional populations (Group-I and Group-II microorganisms) in digestion sludge microbiomes. In this study, composition and function of the three groups of microorganisms, particularly Group-I and Group-II populations, were clearly differentiated by employing genome-centric metagenomics and metabolic network reconstruction. Accordingly, several interesting observations were obtained: (1) WAS pre-treatments changed Group-I assemblies (core functional populations) by releasing different digestion substrates from WAS; (2) Group-II microorganisms with high levels of functional redundancy provided a candidate pool for flexible and robust sludge digestion, and the Group-II minor and non-essential functional populations could be converted into Group-I predominant and functionally essential microorganisms when changing operational conditions. These results might guide future optimization of sludge digestion by providing knowledge on specific impacts of WAS pre-treatment on digestion sludge microbiomes.

The methanogenic sludge digesters with prolific nutrients, optimum temperature and long retention time provided excellent habitats for an extremely wide range of anaerobic microorganisms, supporting diverse and complicated microbiomes for efficient sludge digestion. Of the digestion sludge microbiome, many populations were yet-to-be-cultivated due to their inability to grow in standard culture media [62]. Recent advances in single-cell genomic and metagenomic techniques enabled researchers to bypass the complicated microbial cultivation and facilitate the discovery of numerous previously unknown, deep branches of the tree of life without cultivated representatives [63, 64]. For example, a complete genome sequence of *Ca. Cloacamonas Acidaminovorans* belonging to *Ca. Cloacimonetes* (formerly known as

WWE1) was retrieved and suggested that *Ca. C. Acidaminovorans* could derive most of its carbon and energy from the fermentation of amino acids [65]. In our study, several novel bacterial lineages were identified in the sludge microbiomes, including Group-I core functional populations (e.g., bin17 and bin106), Group-II functionally redundant populations (e.g., bin160 and bin214) and functionally unknown microorganisms (e.g., bin180 and bin224). All Group-I and Group-II functional microorganisms of novel lineages were identified to be fermentative bacteria, which suggested the high functional redundancy supported by the thermodynamically favorable fermentation, in contrast to the subsequent rate-limiting steps mediated by both acetogenic syntrophs and methanogens. Interestingly, energy metabolism of the functionally unknown lineages did not directly involved in the sludge digestion procedures. Consequently, these microorganisms might be second-hand metabolizers and their survival might depend on secreta of the Group-I and Group-II populations. Nonetheless, their detailed functional information awaited future in-depth analyses.

Conclusions

This study provided unprecedented genome-centric metagenomic insights into how WAS pre-treatments changed the digestion sludge microbiomes, as well as their metabolic networks. Results suggested that (1) WAS pre-treatments governed the core functional population assemblies by changing digestion substrates, while redundant functional populations provided a candidate pool for flexible and robust sludge digestion; (2) The genome-predicted metabolic potentials of microbial dark matter in digestion sludge microbiomes suggested their predominant roles in fermentation; (3) The genome sizes of the fermentative bacteria and their abundance were significantly correlated; (4) The digestion sludge microbiome could be a unique source for exploring microbial dark matter, as well as bioresource for future biomass industrial implications.

Abbreviations

WAS: waste activated sludge; EPS: extracellular polymeric substances; APAD: alkaline/acid pre-treatment and anaerobic digestion; HS-APAD: sludge pre-treated with 0.8 mol/L alkaline/acid and anaerobic digestion; Thermal-AD: thermal pre-treated sludge and anaerobic digestion; Control-AD: fresh waste activated sludge and anaerobic digestion; CDS: protein coding sequences; MAGs: metagenomic-assembled genomes

Declarations

Author contributions

S.W. conceived the idea and designed the research. Z.L. performed the experiments. Z.L., J.S. and C.W. analyzed the data. S.W. wrote the manuscript with inputs from all authors.

Funding

This study was financially supported by the Natural Science Foundation of Guangdong Province (2018B030314012).

Availability of data and materials

Metagenomic sequencing data for all samples have been deposited in the European Nucleotide Archive (ENA) database with accession number from ERR3454363 to ERR3454366.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Environmental Microbiomics Research Center, School of Environmental Science and Engineering, Guangdong Provincial Key Laboratory of Environmental Pollution Control and Remediation Technology, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-Sen University, Guangzhou, China 510006; ²Center for Ecosystem Science and Society, Northern Arizona University, Flagstaff, AZ, 86011, USA.

References

1. Ju F, Li B, Ma L, Wang Y, Huang D, Zhang T. Antibiotic resistance genes and human bacterial pathogens: Co-occurrence, removal, and enrichment in municipal sewage sludge digesters. *Water Res.* 2016;91:1-10.
2. Wang S, Yu S, Lu Q, Liao Y, Zhang F, Wang X et al. Development of a combined alkaline/acid pretreatment and anaerobic digestion (APAD) process for methane generation from waste activated sludge. *Sci Total Environ.* 2019; doi: 10.1016/j.scitotenv.2019.134564.
3. Mao C, Feng Y, Wang X, Ren G. Review on research achievements of biogas from anaerobic digestion. *Renew Sust Energ Rev.* 2015;45:540-55.
4. Carrere H, Antonopoulou G, Affes R, Passos F, Battimelli A, Lyberatos G et al. Review of feedstock pretreatment strategies for improved anaerobic digestion: From lab-scale research to full-scale application. *Bioresour Technol.* 2016;199:386-97.
5. Wang X, Duan X, Chen J, Fang K, Feng L, Yan Y et al. Enhancing anaerobic digestion of waste activated sludge by pretreatment: effect of volatile to total solids. *Environ Technol.*

- 2016;37(12):1520-29.
6. Li X, Zhao J, Wang D, Yang Q, Xu Q, Deng Y et al. An efficient and green pretreatment to stimulate short-chain fatty acids production from waste activated sludge anaerobic fermentation using free nitrous acid. *Chemosphere*. 2016;144:160-67.
 7. Gonzalez A, Hendriks ATWM, van Lier JB, de Kreuk M. Pre-treatments to enhance the biodegradability of waste activated sludge: Elucidating the rate limiting step. *Biotechnol Adv*. 2018;36(5):1434-69.
 8. Lu Q, Yu Z, Yu S, Liang Z, Li H, Sun L et al. Organic matter rather than salinity as a predominant feature changes performance and microbiome in methanogenic sludge digesters. *J Hazard Mater*. 2019;377:349-56.
 9. Guo J, Peng Y, Ni BJ, Han X, Fan L, Yuan Z. Dissecting microbial community structure and methane-producing pathways of a full-scale anaerobic reactor digesting activated sludge from wastewater treatment by metagenomic sequencing. *Microb Cell Fact*. 2015;14:33.
 10. Mei R, Narihiro T, Nobu MK, Kuroda K, Liu WT. Evaluating digestion efficiency in full-scale anaerobic digesters by identifying active microbial populations through the lens of microbial activity. *Sci Rep*. 2016;6:34090.
 11. Mei R, Nobu MK, Narihiro T, Kuroda K, Munoz Sierra J, Wu Z et al. Operation-driven heterogeneity and overlooked feed-associated populations in global anaerobic digester microbiome. *Water Res*. 2017;124:77-84.
 12. Campanaro S, Treu L, Rodriguez LM, Kovalovszki A, Ziels RM, Maus I et al. The anaerobic digestion microbiome: a collection of 1600 metagenome-assembled genomes shows high species diversity related to methane production. 2019. doi: <https://doi.org/10.1101/680553>.
 13. Yang Y, Yu K, Xia Y, Lau FT, Tang DT, Fung WC et al. Metagenomic analysis of sludge from full-scale anaerobic digesters operated in municipal wastewater treatment plants. *Appl Microbiol Biot* 2014;98(12):5709-18.
 14. Abendroth C, Vilanova C, Gunther T, Luschig O, Porcar M. Eubacteria and archaea communities in seven mesophile anaerobic digester plants in Germany. *Biotechnol Biofuels*. 2015;8:87.
 15. Westerholm M, Crauwels S, Van Geel M, Dewil R, Lievens B, Appels L. Microwave and ultrasound pre-treatments influence microbial community structure and digester performance in anaerobic digestion of waste activated sludge. *Appl Microbiol Biotechnol*. 2016;100(12):5339-52.
 16. Vanwonterghem I, Jensen PD, Rabaey K, Tyson GW. Genome-centric resolution of microbial diversity, metabolism and interactions in anaerobic digestion. *Environ Microbiol*. 2016;18(9):3144-58.
 17. Pervin HM, Dennis PG, Lim HJ, Tyson GW, Batstone DJ, Bond PL. Drivers of microbial community composition in mesophilic and thermophilic temperature-phased anaerobic digestion pre-treatment reactors. *Water Res*. 2013;47(19):7098-108.
 18. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J et al. Best practices for analysing microbiomes. *Nat Rev Microbiol*. 2018;16(7):410-22.
 19. Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM et al. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*.

- 2004;428(6978):37-43.
20. Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA et al. Environmental genome shotgun sequencing of the Sargasso Sea. *Science*. 2004;304(5667):66-74.
 21. Turaev D, Rattei T. High definition for systems biology of microbial communities: metagenomics gets genome-centric and strain-resolved. *Curr Opin Biotech*. 2016;39:174-81.
 22. Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. Computational approaches to predict bacteriophage-host relationships. *FEMS Microbiol Rev*. 2016;40(2):258-72.
 23. Blauwkamp TA, Thair S, Rosen MJ, Blair L, Lindner MS, Vilfan ID et al. Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat Microbiol*. 2019;4(4):663-74.
 24. Xu G, Lu Q, Yu L, Wang S. Tetrachloroethene primes reductive dechlorination of polychlorinated biphenyls in a river sediment microcosm. *Water Res*. 2019;152:87-95.
 25. Joshi NA, Fass JN. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. 2011. Available at <https://github.com/najoshi/sickle>.
 26. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455-77.
 27. Kang DD, Froula J, Egan R, Wang Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*. 2015;3:e1165.
 28. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. 2016;32(4):605-7.
 29. Alneberg J, Bjarnason BS, de Bruijn I, et al. CONCOCT: clustering contigs on coverage and composition. *arXiv*. 2013;312:4038.
 30. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome*. 2018;6(1):158.
 31. Parks DH, Rinke C, Chuvochina M, Chaumeil PA, Woodcroft BJ, Evans PN et al. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat Microbiol*. 2017;2(11):1533-42.
 32. Bushnell B. BBMap: a fast, accurate, splice-aware aligner. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2014.
 33. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25(7):1043-55.
 34. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068-9.
 35. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27-30.
 36. Huerta-Cepas J, Szklarczyk D, Heller D, Hernandez-Plaza A, Forslund SK, Cook H et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090

- organisms and 2502 viruses. *Nucleic Acids Res.* 2019;47(D1):D309-D14.
37. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods.* 2015;12(1):59-60.
 38. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 2018;46(1):95-101.
 39. Berlemont R, Martiny AC. Phylogenetic distribution of potential cellulases in bacteria. *Appl Environ Microbiol.* 2013;79(5):1545-54.
 40. Rawlings ND, Barrett AJ, Finn R. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 2016;44(D1):D343-50.
 41. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics.* 2008;9:75.
 42. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic acids res.* 2007;35:W182-5.
 43. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ et al. A new view of the tree of life. *Nat Microbiol.* 2016;1:16048.
 44. Wu M, Scott AJ. Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics.* 2012;28(7):1033-4.
 45. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772-80.
 46. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25(15):1972-3.
 47. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 2015;32(1):268-74.
 48. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 2019;47(1):256-59.
 49. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol.* 2018;36:996-
 50. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* 2019;20:257.
 51. Bischoff V, Bunk B, Meierkolthoff JP, Sproer C, Poehlein A, Dogs M et al. Cobaviruses – a new globally distributed phage group infecting Rhodobacteraceae in marine ecosystems. *ISME J.* 2019;13(6):1404-1421.
 52. Kruger K., Chafee M, Francis TB, Rio TG, Becher D, Schweder T et al. In marine Bacteroidetes the bulk of glycan degradation during algae blooms is mediated by few clades using a restricted set of genes. *ISME J.* 2019;13(11):2800-2816.

53. Mocellivina L, Muniesa M, Pimentavale H, Lucena F, Jofre J. Survival of Bacterial Indicator Species and Bacteriophages after Thermal Treatment of Sludge and Sewage. *Appl Environ Microbiol.* 2003;69(3):1452-1456.
54. Solden L, Lloyd K, Wrighton K. The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Curr Opin Microbiol.* 2016;31:217-26.
55. Cox RA. Correlation of the rate of protein synthesis and the third power of the RNA : protein ratio in *Escherichia coli* and *Mycobacterium tuberculosis*. *Microbiology.* 2003;149(3):729-37.
56. Scott M, Klumpp S, Mateescu EM, Hwa T. Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol Syst Biol.* 2014;10:747.
57. Li J, Mau RL, Dijkstra P, Koch BJ, Schwartz E, Liu XA et al. Predictive genomic traits for bacterial growth in culture versus actual growth in soil. *ISME 2019.*
58. Lynch M. Streamlining and simplification of microbial genome architecture. *Annu Rev Microbiol.* 2006;60:327-49.
59. Hessen DO, Jeyasingh PD, Neiman M, Weider LJ. Genome streamlining and the elemental costs of growth. *Trends Ecol Evol.* 2010;25(2):75-80.
60. Lefebvre O, Quentin S, Torrijos M, Godon JJ, Delgenes JP, Moletta R. Impact of increasing NaCl concentrations on the performance and community composition of two anaerobic reactors. *Appl Microbiol Biotechnol.* 2007;75(1):61-9.
61. Shin J, Jang HM, Shin SG, Kim YM. Thermophilic anaerobic digestion: Effect of start-up strategies on performance and microbial community. *Sci Total Environ.* 2019;687:87-95.
62. Narihiro T, Nobu MK, Kim NK, Kamagata Y, Liu WT. The nexus of syntrophy-associated microbiota in anaerobic digestion revealed by long-term enrichment and community survey. *Environ Microbiol.* 2015;17(5):1707-20.
63. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature.* 2013;499(7459):431-7.
64. Doud DFR, Woyke T. Novel approaches in function-driven single-cell genomics. *FEMS Microbiol Rev.* 2017;41(4):538-48.
65. Pelletier E, Kreimeyer A, Bocs S, Rouy Z, Gyapay G, Chouari R et al. "Candidatus Cloacamonas acidaminovorans": genome sequence reconstruction provides a first glimpse of a new bacterial division. *J Bacteriol.* 2008;190(7):2572-9.

Figures

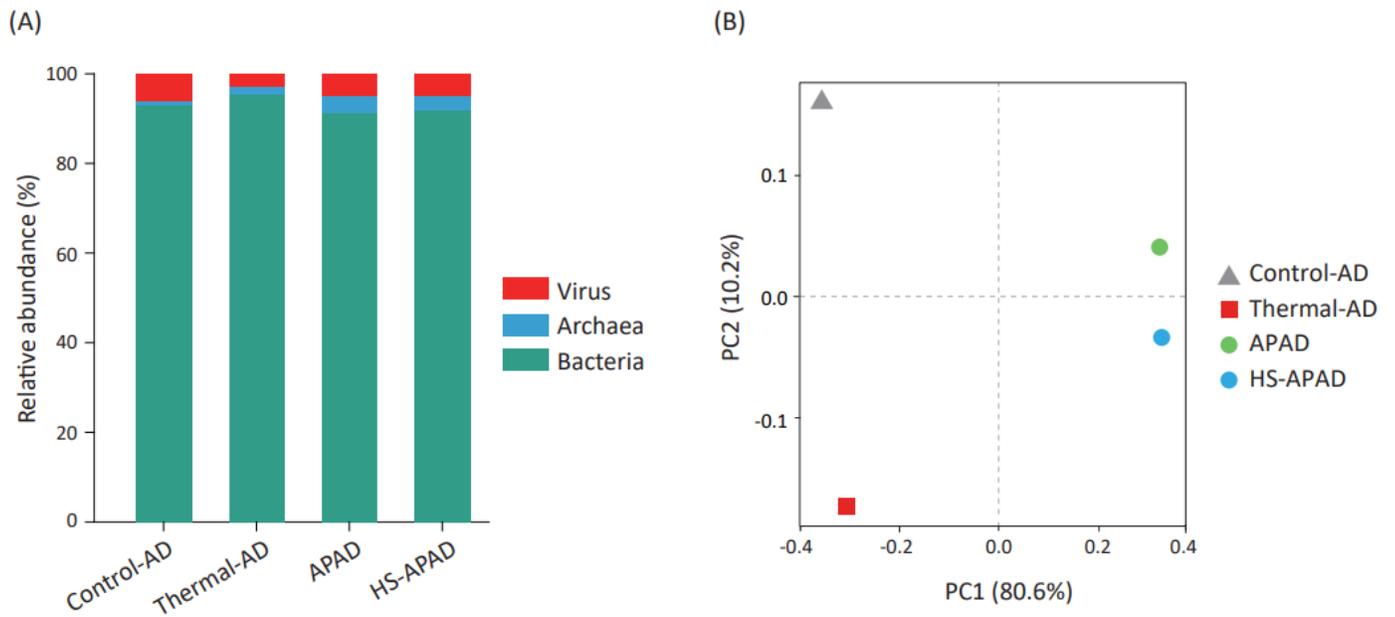


Figure 1

Main taxonomic groups of the digestion sludge microbiome at the domain/kingdom level (A) and their principal coordinates analysis (PCoA).

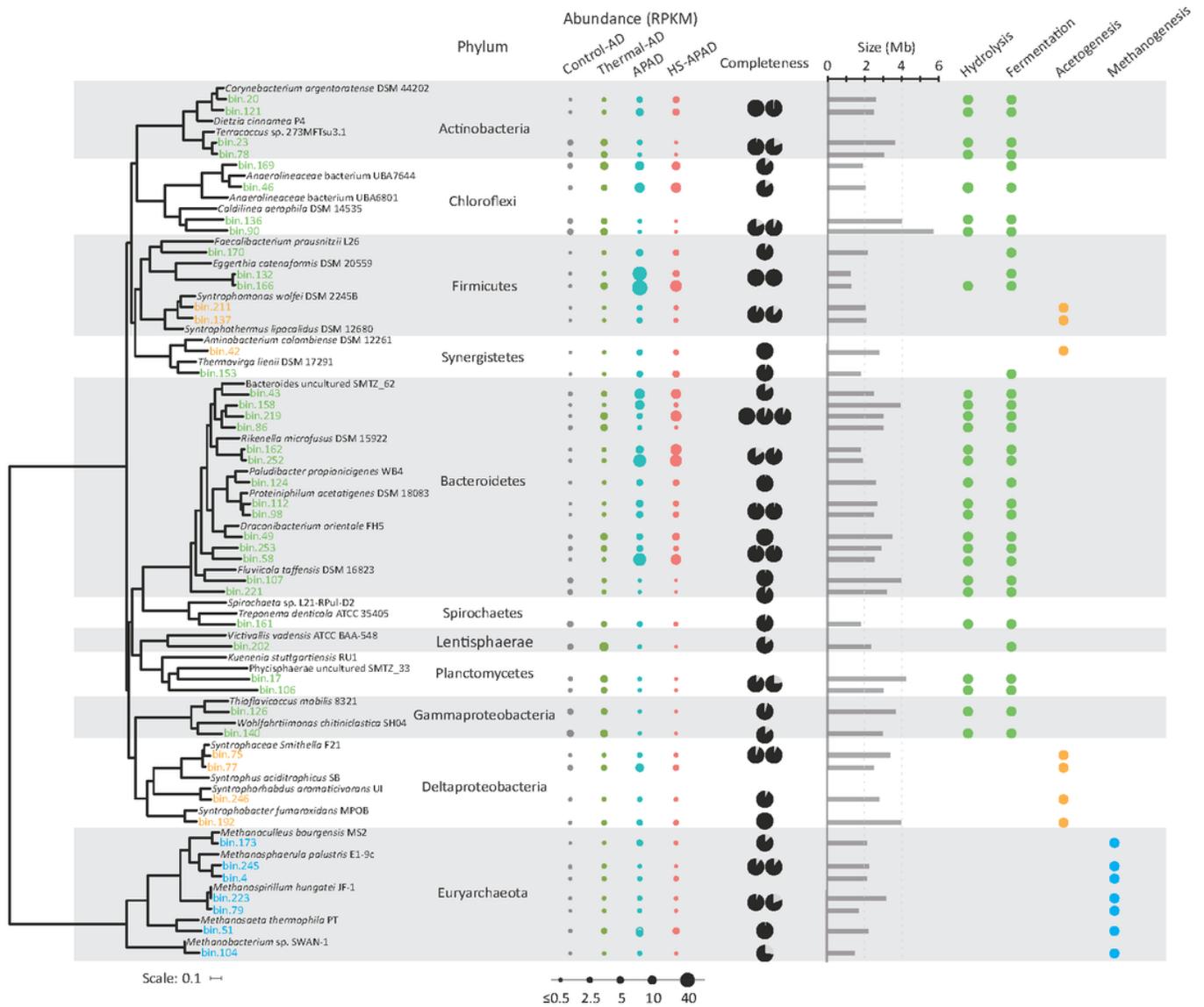


Figure 2

Phylogeny, genome completeness and abundance of retrieved MAGs which were the major populations involved in sludge digestion. The phylogenetic tree was constructed based on an alignment of concatenated ribosomal protein sequences from the MAGs and their related lineages with phyla distinguished by alternating background color. The genome completeness of the MAGs was shown as pie charts. The following right-hand columns indicated the presence (closed circle) and absence (blank) of the MAGs' functions (i.e., hydrolysis, fermentation, syntrophic acetogenesis and methanogenesis) in sludge digestion. See Figure S2 for phylogenetic analysis of all of the 254 MAGs.

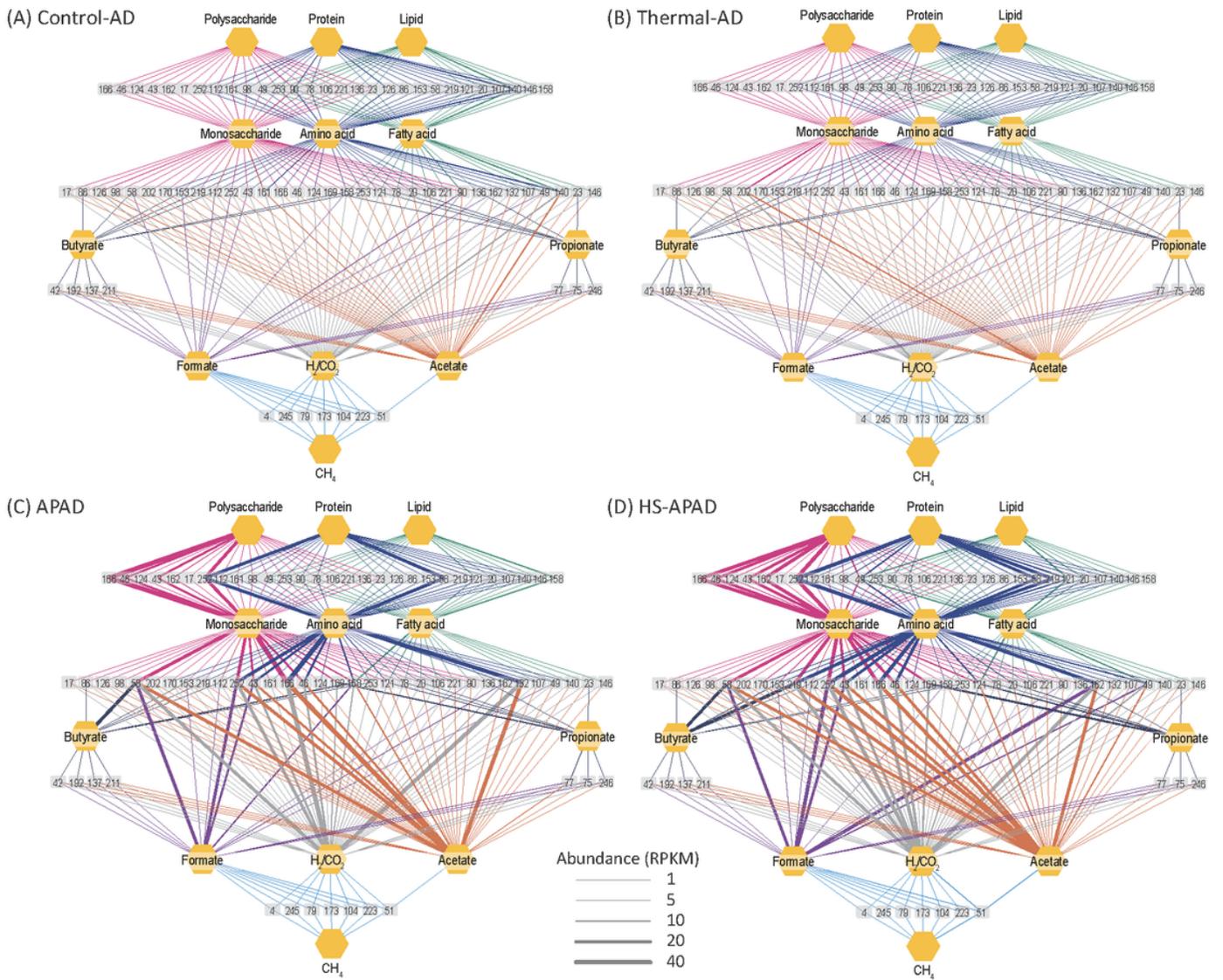


Figure 3

Metabolic networks based on the functional classification of all populations with abundance >33.8 RPKM or top-seven most abundant lineages for methanogenic digestion in Control-AD (A), Thermal-AD (B), APAD (C) and HS-APAD (D). The colors of the edges corresponded to conversion of different substrates, and the thickness of the edges was representative of the MAG abundance. See Table S4 for detailed abundance of the MAGs.

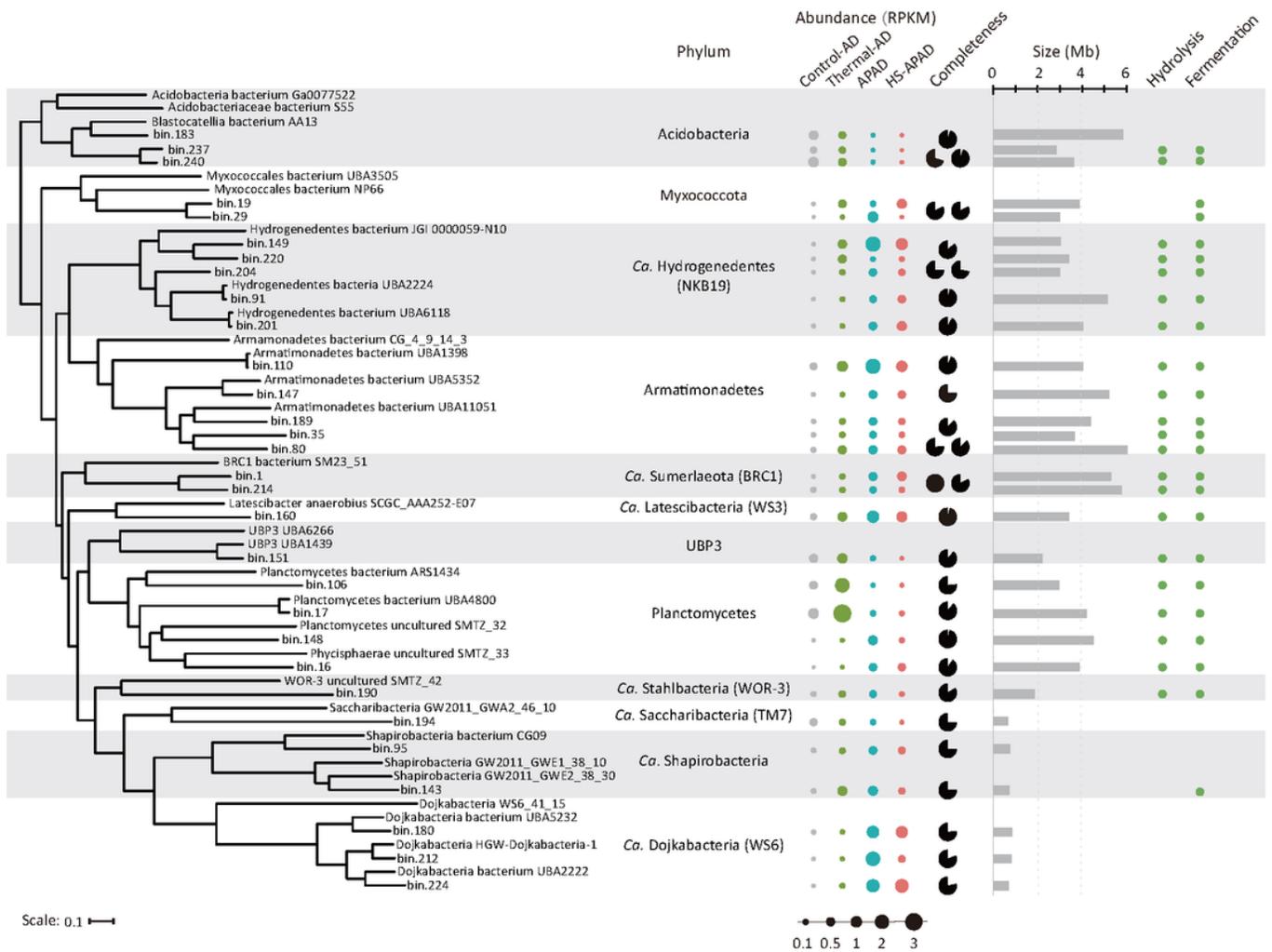


Figure 4

Phylogeny, genome completeness, genome size, and abundance of 30 MAGs which were the representative dark matter in the four digestion sludge microbiome. The phylogenetic tree was constructed based on an alignment of concatenated ribosomal protein sequences from the MAGs and their related lineages with phyla distinguished by alternating background color. The genome completeness of the MAGs was shown as pie charts. The following right-hand columns indicated the presence (closed circle) and absence (blank) of the MAGs' functions (i.e., hydrolysis, fermentation, syntrophic acetogenesis and methanogenesis) in sludge digestion.

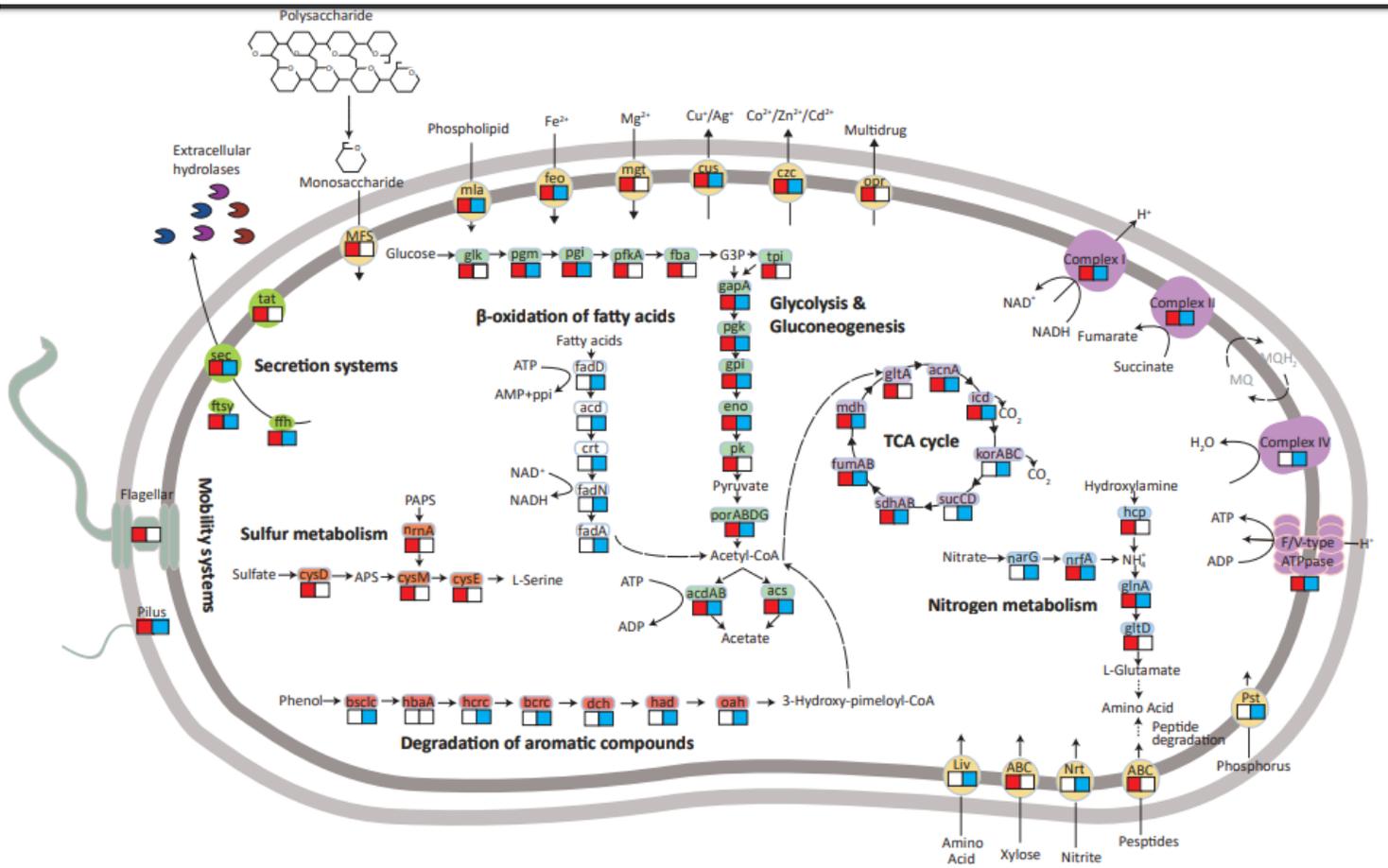


Figure 5

Reconstructed metabolic capability of bin17 and bin106 populations. Pathways with genes detected in bin17 and bin106 genomes were depicted in red and blue colors, respectively. Detailed information on genes assigned to specific metabolic pathways was available in Table S6.

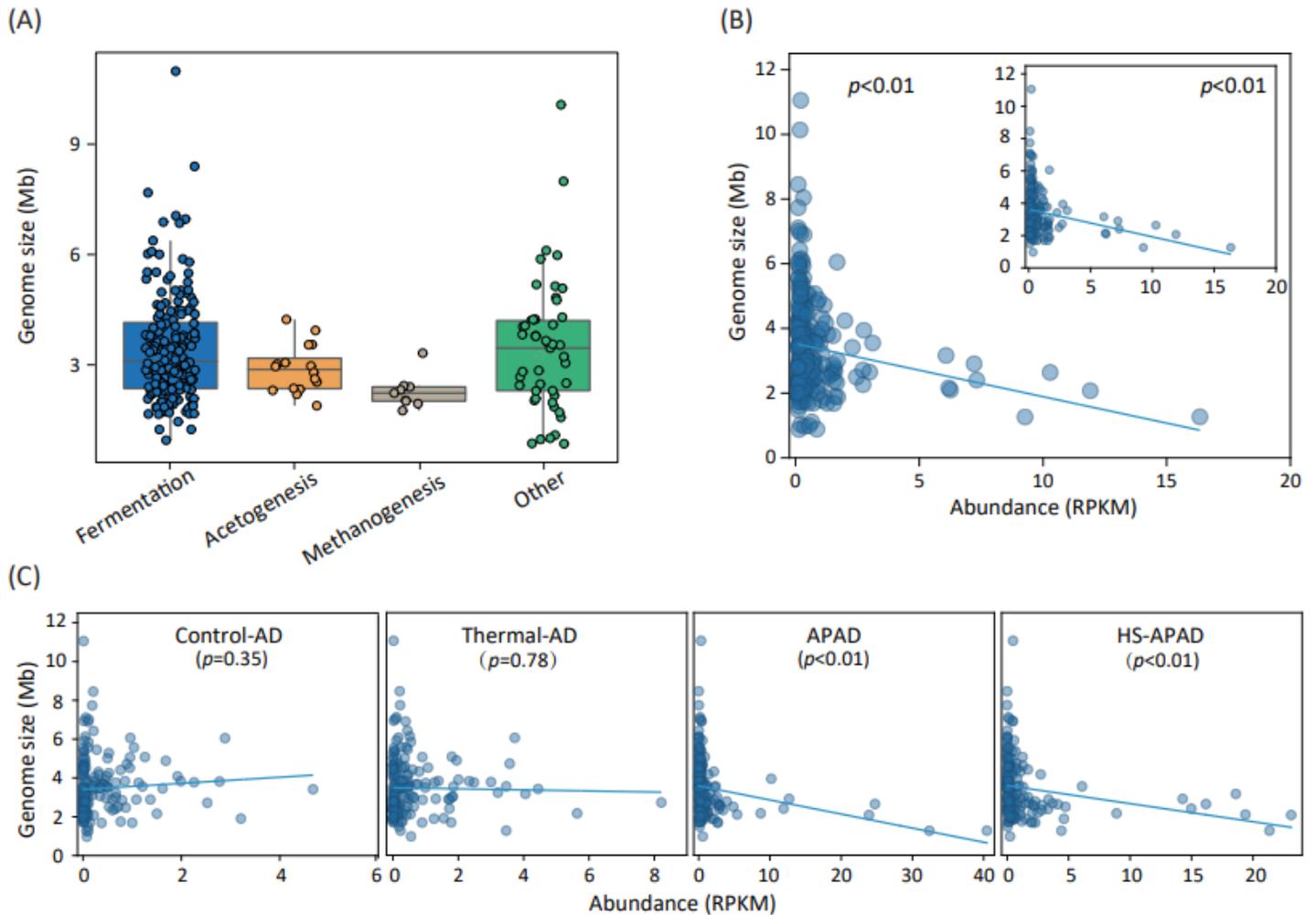


Figure 6

Genome size distribution and relationship between genome size and abundance of MAGs. (A) Genome sizes of functional microbial groups in the co-assembly of the four digestion sludge metagenomes. (B) Relationship between genome sizes and abundance of all retrieved MAGs; inside figure showed the relationship between genome sizes and abundance of fermentative bacterial MAGs. The abundance represented the average of the MAGs in the four digestion sludge metagenomes. (C) Relationships between genome sizes and abundance of all retrieved MAGs in the four sludge digesters.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfiles1.pdf](#)
- [Additionalfiles2.xlsx](#)