MICROBIAL COMMUNITIES IN SEPTIC TANK ANAEROBIC DIGESTERS AND THEIR INTERACTIONS WITH DIGESTER DESIGN AND CHEMICAL ENVIRONMENT

MICROBIAL COMMUNITIES IN SEPTIC TANK ANAEROBIC DIGESTERS AND THEIR INTERACTIONS WITH DIGESTER DESIGN AND CHEMICAL ENVIRONMENT

BY JAMES NAPHTALI, B.Sc

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University

McMaster University MASTER OF SCIENCE (2020) Hamilton, Ontario

(Biology)

TITLE: Microbial Communities in Septic Tank Anaerobic Digesters and Their Interactions with Digester Design and Chemical Environment

AUTHOR: James Naphtali, B. Sc. (Redeemer University College, Ancaster, ON)

SUPERVISOR: Professor Herb E. Schellhorn

NUMBER OF PAGES: xvii, 191

Lay Abstract

Anaerobic digesters are used throughout North America to treat residential sewage. Despite their prevalence, the composition and function of the microbial communities driving sewage degradation in residential digesters has not been studied. We used DNA sequencing to compare the microbial communities and functional genes in different anaerobic digester designs across Southern Ontario. Our findings suggest there are successive microbial groups along the length of septic tanks and that different septic tank designs harbor characteristic sulfidogenic and methanogenic microbes. Characterization of these microbes could inform septic tank bioaugmentation, design and operational optimization strategies to improve sewage treatment performance.

Abstract

Anaerobic digester design and operation influences the biomass degradation efficiency performed by complex and diverse microbial communities. Optimum anaerobic digester design and operational parameters in residential on-site wastewater treatment sites (OWTS) establishes physiochemical environments suitable for the growth and stability of the microbial communities responsible for organic waste degradation. A comparative study of the microbial communities and their functional profiles between different OWTS designs and operational parameters have not been done despite their functional importance in residential organic waste removal. Using whole-metagenome shotgun sequencing, microbial community compositions and functions were compared between two digester designs: conventional box septic tanks and septic tanks equipped with a novel closed-conduit tube called the *InnerTube*™. Wastewater was sampled along the length of each digester to explore the microbial community stratification during the anaerobic digestion treatment process. Additionally, the effect of effluent, aerobic recirculating-lines on the digester microbiome was also explored. Physiochemical characteristics in the form of oxygen demand, nitrogen and solids content was used as endpoints and correlated with microbial community and functional gene abundances to explore the microbes driving anaerobic digestion. Conventional digesters were characterized by syntrophic proprionate-oxidizing microbes and acetoclastic methanogens, while *InnerTube*™ digesters were characterized by syntrophic sulfate-reducing microbes and hydrogenotrophic methanogens.

Recirculating digesters were enriched with denitrifying microbial consortia in syntrophy with hydrogenotrophic methanogens. Microbial communities were organized according to hydrolytic, acidogenic, acetogenic, and methanogenic groups along the digester treatment process. Insight into the core microbiome of OWTS can inform bioaugmentation and digester design and operation optimization strategies to improve the treatment of decentralized residential sewage sources.

Acknowledgements

McMaster University's Biology Master's Program provided me with first-hand experience into the world of corporate scientific research. In my time as a graduate student in Dr. Herb E. Schellhorn's lab, I learned countless practical laboratory skills, conversed with industry partners, greatly developed my computational biology skills and even published my first review article. For all these skills, I would like to first extend my gratitude to Dr. Schellhorn, who patiently tutored me through many unproductive days, improved my scientific writing style, and helped me develop my soft-skills with industry partners. Your resourcefulness, reputation and mentorship as an accomplished scientist provided me many platforms and opportunities for me to develop invaluable people-skills, bioinformatics experience, and to mature as a research scientist. I would like to also extend my sincere thanks to my co-supervisor, Dr. Brian Golding, for your valuable bioinformatics knowledge and insight you provided in my committee meetings and through your BIO-720 class, both of which played a vital role in enabling me to properly analyze sequencing data.

I would also like to extend my thanks to Craig Jowett, Christ Jowett and the employees at Waterloo Biofilter Systems for their extensive cooperation with the MacWater lab. I am thankful for the time you gave to accommodate for our sampling schemes and schedules of your septic tanks. I thank all the undergraduate volunteers who persevered through the hundreds of DNA extractions and endured all the unpleasant aromas that naturally comes with a microbial community study

on sewage. As for the past and present Schellhorn lab members, Ethan, Bansri, Nicole, Meagan and Mahi, I thank you for making lab life survivable through laughter and comradery. I especially would like to express my sincere thanks to my right-hand lab man, Alex Chan, who endured snow in the middle of August while sampling, extracted DNA with me past midnight, provided me valuable second opinions and even co-authored a review article with me. I literally would not have a project without you!

To my family, I thank you for the mental and spiritual support you've provided me throughout my time as a Masters student. Thank you for letting me voice my concerns, anxieties and ideas and giving me invaluable advice on how to handle life's hurdles. Last but not least, I would like to thank my God, who surrounded me with people who love me and most of all gave me His Son, Jesus Christ, the loving Savior of my soul. He has taught me perseverance and taught me where my hope and purpose in life lies, not in my own strength, but in Him. May He get the glory. Soli Deo Gloria.

Contents

List of Figures

List of Tables

List of Supplementary Tables

List of Abbreviations and Symbols

AD	Anaerobic Digestion
ABR	Anaerobic-baffled reactor
CBOD	carbonaceous biological oxygen demand
COD	chemical oxygen demand
CR	Conventional recirculating
CS	Conventional single-pass
HRT	Hydraulic retention time
IR	*InnerTube*™ recirculating
IS	*InnerTube*™ single-pass
N	nitrogen
NMDS	Non-metric multidimensional scaling
OLR	Organic loading rate
OTU	Operational taxonomic units
OWTS	Onsite- Wastewater Treatment system
P	phosphorous
PERMANOVA	Permutational Multivariate Analysis of Variance
PFR	Plug-flow reactor
RDA	Redundancy analysis
SIMPER	Similarity Percentage analysis
TKN	total Kjeldahl nitrogen
TSS	total suspended solids
UASB	Upflow-anaerobic sludge blanket
VFA	volatile fatty acids
WBS	Waterloo Biofilter Systems
WMS	Whole-metagenome sequencing
WWTP	Wastewater treatment plant

1 Chapter 1: The Use of DNA Sequencing in the Microbiological Analysis of Wastewater Treatment Systems

1.1 Wastewater Sources and their Environmental Impact

Anthropogenic activities continue to decrease the quality of ecosystems in watersheds across the globe (Beiras, 2018; Verhougstraete et al., 2015; Wang et al., 2007). Agricultural fertilizer use along with human and animal sewage waste is globally the largest anthropogenic source of water pollution (Dumont et al., 2005; Harrison et al., 2010). Point source and non-point source pollution from centralized municipal (WWTP) and decentralized onsite-wastewater treatment plants (OWTS) respectively, release organic and inorganic compounds commonly containing nitrogen (N), phosphorus (P) and heavy trace metals (Goyette et al., 2018; Kaushal et al., 2011). Nitrogenous and phosphorous inorganic compounds such as ammonia (NH_3), nitrate (NO_3^-), and phosphate (PO_4^{2-}) and carbonaceous compounds in the form of pharmaceuticals, hormones, detergents and human food, fecal, and urinary waste directly contribute to increased nutrient loading in watershed bodies (Frei et al., 2020; Kaushal et al., 2011; Schaider et al., 2017). Nutrient loading in waterbodies results in algal blooms where prokaryotic blue-green algae called cyanobacteria as well as eukaryotic algae utilize the excess phosphorous and nitrogen while simultaneously depleting the oxygen content in water bodies for growth (Lunau et al., 2013; Steven et al., 2012; Wang et al., 2007). Consequently, the ecological effects of algal blooms includes a decrease in aquatic plant populations through the increase of water turbidity leading to a reduction of light

penetration, and a decrease in aquatic vertebrate and invertebrate populations caused by a reduction of dissolved oxygen and an increase in pH due to increased photosynthetic activity of cyanobacteria (Ferrao-Filho Ada & Kozlowsky-Suzuki, 2011; Wiegand & Pflugmacher, 2005). Cyanobacteria also release toxins called microcystins which accumulates in fish tissue causing growth inhibitions and mortality (Drobac et al., 2016; Ferrao-Filho Ada & Kozlowsky-Suzuki, 2011). Microcystins also adversely affects humans by impairing liver, kidney, and muscular function thus preventing recreational water use and seafood consumption (Buratti et al., 2017). Heavy metals as well as pharmaceuticals and other organic compounds such as polyaromatic hydrocarbons (PAH) from septic tanks are also toxic to aquatic fauna (Patel et al., 2019).

Due to the localized nature of point-source WWTP pollution, effluent water quality from municipal and industrial WWTP's have already been extensively studied and tracked by screening for chemical and biological indicators such as carbonaceous biological oxygen demand (CBOD), chemical oxygen demand (COD) (Drobac et al., 2016), total suspended solids (TSS), inorganic compounds such as ammonia (NH_3), nitrate (NO_3^-), and phosphate (PO_4^{2-}) and bacterial counts (Garcia et al., 2013; Tran et al., 2018). As a result, improvements in digester design and operating strategies have been made to increase organic waste degradation rates and more stringent effluent quality regulations have been implemented throughout the world to reduce nutrient loading in regional watersheds (Hendriks & Langeveld, 2017; Holeton et al., 2011; Hvala et al., 2018;

van Lier, 2008). However, despite the fact that decentralized OWTS's serve around 25% of the U.S population (Conn et al., 2006) and around 20% on average across the world (Cookey et al., 2016), non-point source WWTP discharge from OWTS's have not been studied to the same extent as centralized WWTP's in terms of system chemical and microbiological characteristics (Schaider et al., 2017; Wigginton et al., 2020). Consequently, the environmental effects of OWTS's sewage effluent streams on surface waters are currently not fully understood and due to their ubiquity, microbiological degradation mechanics must be studied to remedy nutrient loading from decentralized, non-point source sewage pollution (Diaz-Elsayed et al., 2017; Schaider et al., 2017).

1.2 Microbial Ecology of Anaerobic Digestion

Residential on-site wastewater treatment systems degrade organic waste in the form of black water and grey water consisting of fecal, urine, food waste, and household detergents which are comprised of complex carbohydrates, proteins, fats and other non-biological organic compounds such as aromatics, organophosphorous and organonitrogen (Figure 1.1) (Braguglia et al.; Carballa et al., 2015; Schaider et al., 2017). In OWTS's, the digestion of these complex molecules occurs in an anaerobic digestion (AD) stage where complex macromolecules are degraded into gases such as methane (CH_4), hydrogen sulfide (H_2S), hydrogen gas (H_2), carbon dioxide (CO_2) and ammonia by obligate and facultative anaerobic microbes (Fontana, Campanaro, et al., 2018). Afterwards in an aerobic digestion stage, ammonia produced from anaerobic digestion is

converted into nitrate through a process called nitrification driven by a series of ammonia and nitrite-oxidizing (NO_2^-) aerobic microbes (Alzate Marin et al., 2016; Campanaro et al., 2018; Raszka et al., 2011). Residual simple-fatty acids, monosaccharides, and amino acids produced in the anaerobic digestions stage are also further broken down into simple polyatomic N and P compounds by aerobic microbes in the aerobic stage of digestion (Alzate Marin et al., 2016; Cydzik-Kwiatkowska & Zielinska, 2016; Isazadeh et al., 2016).

High-throughput DNA and RNA sequencing have enabled the analysis of not only taxonomic compositions of the AD microbiome, but also the compositions of functional genes coding for enzymes and proteins driving microbial metabolic pathways (Campanaro et al., 2018). Culture-independent DNA and RNA techniques have revealed that anaerobic digestion occurring within OWTS septic tanks employ several groups of microorganisms to breakdown polymeric carbohydrate, lipid and protein substrates in influent sewage and sludge through a series of syntrophic, microbially-driven reactions. The reactions include hydrolysis, fermentation, classified into acidogenesis and acetogenesis, and methanogenesis in the absence of oxygen (Figure 1.2) (Campanaro et al., 2016).

The degradation of organic waste begins with the hydrolysis of complex carbohydrates, lipids and proteins by hydrolytic microorganisms through the use extracellular hydrolytic enzymes (Fontana, Campanaro, et al., 2018; Vanwonterghem, Evans, et al., 2016). Enzymes that breakdown polysaccharides into monosaccharides are called Carbohydrate active enzymes (CAZymes) and

include glycoside hydrolases, carbohydrate esterases, and polysaccharide lyases and other auxiliary enzymes (Bertucci et al., 2019). Additional hydrolytic enzyme families include proteases and peptidases which digest complex proteins into constituent amino acids, and lipases, which breakdown complex lipids into long-chain fatty acids (LCFA) and glycerol (Fontana, Campanaro, et al., 2018; Verma & Sharma, 2020). The most abundant phyla responsible for hydrolysis include facultative anaerobes and anaerobes within Firmicutes, Bacteriodetes, Fibrobacter, Spirochaetes, Clostridium and Thermotogae phyla. (De Vrieze et al., 2017; Rui et al., 2015).

In acidogenesis, hydrolytic products are catabolized into short-chain volatile fatty acids such as pyruvic, propanoic and butyric acids (VFA), alcohols including methanol, ethanol and butanol, carbon dioxide (CO_2), and hydrogen gas (H_2) (Bengtsson et al., 2008; Campanaro et al., 2016; Vanwonterghem, Jensen, et al., 2016). Monosaccharides and fatty acids are metabolized primarily by Firmicutes, Proteobacteria, and Bacteriodetes through glycolysis producing pyruvate, formate, and acetyl-CoA (Cai et al., 2016). The glycolysis products are then fermented into short-chain fatty acids, alcohols, ketones (acetone) and hydrogen gas (Guo et al., 2015; Yang et al., 2014) by a wide range of enzymes including NADH:ferredoxin oxidoreductase (NFOR) and formate-hydrogen lyase (FHR) (Cai et al., 2016; Sikora et al., 2017). Amino acid deanimation and fermentation into acetic, propionic acids, CO_2 and H_2 is performed by Synergistetes

and Firmicutes (*Clostridia*) (De Francisci et al., 2015; De Vrieze et al., 2016; Sieber et al., 2012).

Acetogenesis is the conversion of VFA, alcohols and the gases CO_2 and H_2 produced from acidogenesis into acetate (Guo et al., 2015; Rui et al., 2015). Acetogens are responsible for the fermentation, oxidation, and reduction of VFAs, alcohols and CO_2 respectively (Guo et al., 2015; Rui et al., 2015). A prominent pathway utilized by homoacetogens primarily in the phylum Firmicutes is the Wood-Ljungdahl pathway which involves the anaerobic reduction of CO_2 with H_2 as the electron donor into acetate (Cai et al., 2016; Campanaro et al., 2016; Rui et al., 2015). Syntrophic acetogens are another type of acetogens called syntrophic acetogens also included in the phylum Firmicutes which oxidize VFAs and alcohols into acetate (Vanwonterghem, Jensen, et al., 2016).

Methanogenesis is the conversion of acetogenic products into methane gas. Methanogenic metabolism is syntrophically tied with acetogenic processes, with the end products of acetogenesis supplying substrates for methanogenesis (Cai et al., 2016; Chen et al., 2018; Rui et al., 2015). Methanogens are classified into three categories based on the substrate utilized in the production of methane. Acetoclastic methanogens break down acetate, hydrogenotrophic methanogens reduce CO_2 with electrons from H_2, and methylotrophic methanogens utilize methylated compounds such as methanol and methylated amines (De Francisci et al., 2015). The majority of methanogen populations in anaerobic digesters is made up of microbes from the Euryarchaeota phylum (Campanaro et al., 2016; Nordgard et al., 2017).

Acetoclastic and methylotrophic methanogens are comprised of archaea in the order *Methanosarcinales*, with the genera *Methanosaeta* being strictly acetoclastic and *Methanosarcina* capable of producing methane from CO_2, H_2 and other methylated compounds in addition to acetate (Cai et al., 2016; Nordgard et al., 2017) .

1.3 Employing Second-Generation High-Throughput Sequencing for OWTS Microbial Analysis

1.3.1 Current State of OWTS Microbiological Analyses

The majority of microbiological studies using traditional culture-assays or high-throughput DNA and RNA sequencing on sewage treatment have been done using lab-scale or centralized municipal and industrial-scale WWTP's (Ai et al., 2019; Campanaro et al., 2018; De Vrieze et al., 2018; Fontana, Campanaro, et al., 2018 2018). In contrast, analyses of OWTS's have been limited to effluent chemical analyses (Kaushal et al., 2011; Schaider et al., 2017), culture-dependent methods (Keegan et al., 2014; Pussayanavin et al., 2015) and low-throughput DNA analyses using quantitative polymerase-chain reaction (qPCR) assays (Keegan et al., 2014; Nie et al., 2011; Robertson et al., 2012; Tomaras et al., 2009). The widespread use of OWTS's in treating residential sewage prompts the need for more in-depth taxonomic and functional microbiome analyses to elucidate core microbes and their roles in sustaining efficient anaerobic digestion of household wastes (De Vrieze et al., 2016; Wigginton et al., 2020).

1.3.2 Whole-Metagenome Shotgun Sequencing: Advantages over 16S Amplicon Sequencing

Whole-Metagenome Sequencing (WMS) is a DNA sequencing method which analyzes nucleotide sequences by randomly shearing all genomes within a sample and subsequently referencing gene and non-redundant protein databases for taxonomic and functional gene annotations (Quince et al., 2017). This is in contrast to other SGS DNA sequencing methods such as 16S Amplicon and 18S amplicon sequencing which relies on the amplification of conserved marker genes targeted by primers for taxonomic annotations (Tessler et al., 2017). Since WMS does not rely on the primer amplification of gene sequences and instead uses all randomly fragmented sequences in a given sample, issues with primer bias towards certain groups of microbes are minimized (Klindworth et al., 2013; Walters et al., 2016). Therefore, without the effect of primer choice on observed microbial community profiles, WMS results in reads which more accurately reflect actual community compositions (Jovel et al., 2016).

Amplicon sequencing can be used to analyze microbial communities based on the divergence of a singular region in bacterial genomes based on primer specificity (Jovel et al., 2016; Ranjan et al., 2016). Accurate classification of microbes down to the species level is dependent on differentiating between conserved sequences that are more than 97% similar (Jovel et al., 2016; Ranjan et al., 2016). As a result, the majority of highly similar 16S gene sequences that belong to different yet closely related species can only be accurately binned and

classified into Operational Taxonomic Units (OTUs) at the genera level rather than at the species level. (Bokulich et al., 2018; Jovel et al., 2016; Ranjan et al., 2016). The use of conserved sequences to differentiate between closely related species leads to a loss of taxonomic information (Bokulich et al., 2018; Jovel et al., 2016; Ranjan et al., 2016). In contrast, since WMS enables the direct annotation of marker and functional gene sequences from complete genomes, WMS provides greater classification accuracy at the genus and species levels than 16S amplicon sequencing (Hillmann et al., 2018; Jovel et al., 2016; Ranjan et al., 2016; Zaheer et al., 2018). Additionally, WMS more accurately detects and quantifies a greater number of species at high (~34%) and low abundances (~ 3.3%) at equivalent read lengths and read depths compared to 16S Amplicon sequencing (Ranjan et al., 2016). The taxonomic profiling advantages that WMS provides over amplicon sequencing is due to the use of sequence classification methods that are not affected by primer bias and the fact that WMS sequences broad regions of the genome compared to a singular region in amplicon sequencing (Jovel et al., 2016; Ranjan et al., 2016). Achieving accurate classifications and abundances of low abundance microbes is pertinent when analyzing the AD microbiome, since functionally important organisms, particularly methanogens, are often found at low abundances (Campanaro et al., 2016). Another advantage that WMS provides over amplicon sequencing is more accurate functional gene profiling since WMS sequences the entirety of genomes in fragments within a sample (Campanaro et al., 2016; Jovel et al., 2016;

Vanwonterghem, Jensen, et al., 2016). This is in contrast to amplicon-based functional profiling, where correlations are drawn to phylogenetic trees to obtain clusters of shared genes between taxa (Langille et al., 2013). With access to curated databases such as SEED (Overbeek et al., 2014) and KEGG (Kanehisa et al., 2014) containing functional genes coding for proteins organized into putative functional roles, WMS enables the direct annotation of DNA fragments to functional genome sequences (Jovel et al., 2016).

1.3.3 Whole-Metagenome Shotgun Sequencing: Optimal Sample Preparation Techniques

WMS and amplicon sequencing are traditionally performed using two sequencing techniques. The first, called pyrosequencing, is a discontinued sequencing technology by 454 Life Sciences characterized by long read lengths (~ 800 bp) and lower read counts, and sequencing-by-synthesis (SBS) which is currently the most widely-used sequencing technique, pioneered by Illumina, Inc. characterized by short read lengths (<300 bp) and substantially higher read counts (Luo et al., 2012; Schirmer et al., 2016; Solonenko et al., 2013). After DNA extraction and isolation, sequencing libraries are prepared by fragmentation, size-selection, labelling and enrichment of the extracted DNA with commercial library preparation kits unique to Illumina MiSeq and HiSeq platforms or Roche 454 pyrosequencers (Solonenko et al., 2013). The majority of recent metagenomic projects employ Illumina sequencing due to the higher read coverage it provides to ensure lower gene annotation error rates for genome assembly compared to

pyrosequencing platforms (Luo et al., 2012; Utturkar et al., 2017). Since this project employed Illumina HiSeq DNA sequencing, only sample preparation techniques specific to Illumina sequencing platforms are discussed.

The choice of library preparation techniques influences downstream bioinformatics analyses and results (Schirmer et al., 2016; Solonenko et al., 2013). For metagenomic library preparation, Illumina's TruSeq Nano and PCR-free sample preparation kits are preferred over the the Nextera XT kits (Schirmer et al., 2016; Solonenko et al., 2013). The Nextera XT kit tags fragmented DNA strands using transposase enzymes in a process called *tagmentation* (Schirmer et al., 2016). Although this library preparation method allows the input DNA masses of as little as 50 ng, tagmentation causes biased read coverage (Jones et al., 2015; Schirmer et al., 2016). Transposase enzymes target specific sequence motifs resulting in biased read-coverage towards high sequences with higher G+C content compared to the TruSeq Nano and PCR-free preparation kits (Schirmer et al., 2016). Despite the bias, Nextera XT kits provide slightly lower read error rates after read trimming and error correction compared to TruSeq kits (Schirmer et al., 2016). In contrast, TruSeq Nano PCR-free kits require higher amounts of DNA input mass (1 to 2 µg) to achieve higher read-coverages on sequences (Solonenko et al., 2013). However, TruSeq PCR-free kits only require around 5 ng of DNA input mass due to the use of PCR-amplification for cluster-generation of fragments(Sanders et al., 2019). Therefore, the choice of sequencing techniques and preparation kits depends on the amount of DNA available after extraction and

isolation (Solonenko et al., 2013). This subsequently influences downstream bioinformatics analyses and results with respect to the number of reads available for aligning and assignment and the proportion of sequence coverage-biases present (Schirmer et al., 2016; Solonenko et al., 2013).

1.3.4 Whole-Metagenome Sequence Processing Methods

WMS sequence reads are outputted in the FASTQ format and are quality trimmed for Illumina adapters and filtered for low quality base-calls based on the "Phred" nucleobase quality scoring system using tools such as BBDuk (Institute, 2019) and Trimmomatic (Bolger et al., 2014). For reference-based sequence analyses, taxonomic classification of whole-metagenome libraries is accomplished by detecting dissimilar clade-specific regions of genomes to differentiate between closely-related species given that sufficient sequencing depth is achieved for each sequence (Jovel et al., 2016; Truong et al., 2015). Taxonomic classification and assignment tools are categorized into alignment-based tools such as BLAST and DIAMOND BLAST (Buchfink et al., 2015), marker-based tools such as MetaPhlAn (Segata et al., 2012), PhyloSift (Darling et al., 2014) and MEGAN6 (Huson et al., 2016) and read-based tools such as as Kraken (Wood & Salzberg, 2014) and CLARK (Ounit et al., 2015). In this project, DIAMOND-BLAST was used to align sequences to NCBI's non-redundant (nr) protein sequence database in conjunction with MEGAN6 to detect and estimate bacterial and archaeal relative abundances from the aligned sequences. The marker-based tool MEGAN6 assigns sequenced reads to a set of clade-specific marker genes from > 7500 bacterial and

archaeal species or reference protein sequences based on previously-sequenced microbial genomes (McIntyre et al., 2017). Alternatively, read-based tools such as Kraken (Wood & Salzberg, 2014) and CLARK (Ounit et al., 2015) assigns reads to complete genomes, often employing unique *k*-mer (sequences of user-defined length *k*) abundances for taxonomic classification. In terms of classification accuracy, read-based assignment tools generate slightly greater taxonomic false positives as read depths increase than marker-based or alignment-based tools in conjunction with marker-based tools (McIntyre et al., 2017). This suggests that reads that are initially falsely identified as one particular species are progressively miscalled as another species at greater read depths (McIntyre et al., 2017). However, for marker-based assignment and alignment tools, any additional miscalled reads are identified as the same falsely identified species as read-depth increases thus resulting in a lower overall false discovery rate (McIntyre et al., 2017). Initial taxonomic false-positive rates produced by marker-based assignment tools for WMS can be mitigated by classifying aligned reads using the lowest-common ancestor (LCA) algorithm as provided by MEGAN (Huson et al., 2016; McIntyre et al., 2017). Thus, the LCA algorithm classification and binning algorithm was used for this project.

1.3.5 Comparison between Reference-Based and Assembly-Based Classification Methods

Assembly-based classification methods aligns raw reads together into contiguous sequences to assemble draft genomes of microbes in metagenomic

samples using tools such as MetaSpades and Megahit (D. Li et al., 2015; Nurk et al., 2017; Quince et al., 2017). Genomic contigs are then *de novo* binned into complete genomes using unique *k*-mer frequencies (Alneberg et al., 2014). Contigs can then be directly aligned to a nucleotide sequence database and classified using either MetaPhlAn or MEGAN6 to quantify taxonomic and functional gene abundances (McIntyre et al., 2017). By detecting open reading frames corresponding to transcription initiation sites using tools such as Prodigal (Hyatt et al., 2010) and PROKKA (Seemann, 2014) genomes are functionally annotated based on protein-coding sequences. However, the *de novo* assembly of genomes, especially those assembled from metagenomic sequence data, requires greater read depths and read-lengths compared to just employing reference-based tools such as MetaPhlAn and Kraken. Higher read depths mitigate the effects of read errors between paired-end reads, increases chances of contig overlap and assembly and increases metagenome assigned genomes (MAG) completeness (Table 1.1) (Hillmann et al., 2018; Olson et al., 2017; Parks et al., 2015; Ranjan et al., 2016). Due to the diversity of genome sizes, types of genes and gene copy numbers in metagenomic samples, a universal *a priori* read-depth range that ensures sufficient genome and compositional completeness for any metagenome does not currently exist (Ni et al., 2013; Ranjan et al., 2016). The program COVER (Tamames et al., 2012) provides *a priori* estimates of read depths based on 16S amplicon sequencing data required to achieve high coverage for dominant genomes in a metagenomic sample. Nonpareil 3 (Rodriguez et al., 2018) has been recently developed to

estimate read coverages of metagenomic shotgun sequencing datasets by comparing a subset of sequencing reads against the entire metagenome using a Turing-Good estimator.

Assembly-based pipelines enables the discovery of novel genomes since reads are combined *de novo* without the use of reference databases by aligning reads based on read overlap and similarity metrics (Nurk et al., 2017). Given enough sequencing depth, complete genomes of both known or novel species can be obtained and subsequently mapped to functional databases (Campanaro et al., 2016; Olson et al., 2017). Metabolic pathways can then be constructed using the functional assignments (Quince et al., 2017). Genome bins of species possessing functional metabolic processes of interest can then be mapped onto the metabolic pathways, forming a comprehensive depiction of microbes directly responsible for metabolic processes of interest (Campanaro et al., 2016). In contrast, since reference-based pipelines forgo the assembly of genomes, determining which microbes possess genes for a functional process can only be inferred from putative roles revealed by previous genome sequencing or culture-based studies on a given microbe. However, by forgoing assembly, reference-based pipelines do not require high read-depths compared to assembly-based pipelines to obtain accurate taxonomic and functional gene abundance profiles for a given metagenome (Table 1.1) (Hillmann et al., 2018). Consequently, the low read-depth requirements of reference-based pipelines allow higher numbers of samples to be studied since increasing the number of samples sequenced on a given Illumina flow-cell

decreases the achievable read depth for each sample (Table 1.1) (Fumagalli, 2013). Therefore, reference-based pipelines are typically used in metagenomic surveys across environments and longitudinal studies where high numbers of samples (> 15) are studied in a single sequencing run (Table 1.1).

1.4 The Use of DNA Sequencing for Bio-augmenting Digester Microbial Communities

Bioaugmentation is the addition of a microbial consortium of defined composition to a microbiological system to enhance the performance of that system (Kouzuma & Watanabe, 2011). Bioaugmentation for anaerobic digesters has been employed to remedy or prevent AD process failure by enhancing the digestion of inhibitory compounds that are not readily degradable or by supplementing existing microbes in digesters overloaded above maximum organic loading rates (OLR) (Y. Li, L. Li, et al., 2018; Y. Li et al., 2017; Venkiteshwaran et al., 2016). Longitudinal 16S Amplicon sequencing and replicated WMS projects have been used to elucidate the core microbiome in digesters during start-up and long-term operation to determine the microbes responsible for jumpstarting and sustaining the AD processes (Campanaro et al., 2018; De Vrieze et al., 2016; Goux et al., 2016). Furthermore, high-throughput DNA sequencing in conjunction with chemical monitoring of digesters has pinpointed chemical compounds which accumulate as microbial by-products during the AD process and inhibit the growth and survivability of downstream syntrophic bacteria (Joyce et al., 2018; L. Li et al., 2015; N. Li et al., 2017). In particular, the accumulation of volatile fatty acids

(VFAs) produced by acetogens have been the cause of AD failures across several types of anaerobic digesters (Fontana, Campanaro, et al., 2018 2018; L. Li et al., 2015; N. Li et al., 2017). Therefore, bioaugmentation studies on anaerobic digesters have focused on the effect of adding VFA-degrading bacterial consortia into digesters under chemical stress (L. Li et al., 2018; N. Li et al., 2017) and adding microbes capable of performing all four steps of AD (Poszytek et al., 2019). Bioaugmentation efforts in anaerobic digesters have so far proven successful, increasing the abundances of core taxa and pertinent functional pathways (Y. Li, L. Li, et al., 2018). Furthermore, bioaugmentation increases biodegradation performance by increasing methane production rates in operational digesters (Y. Li et al., 2017; Poszytek et al., 2019) and has recovered failing digesters overloaded with VFA (Y. Li, G. Yang, et al., 2018; Y. Li et al., 2017).

1.5 Limitations of DNA Sequencing on Deducing the Functionality of the Active Microbiome

DNA-based sequencing methods such as WMS and amplicon sequencing cannot differentiate whether DNA sequences are derived from dead and live cells (De Vrieze et al., 2018; Emerson et al., 2017). This prevents the direct analysis of the active microbiome in a given system and DNA sequencing analyses can therefore only provide taxonomic and functional potentials of a microbiome (De Vrieze et al., 2018; Emerson et al., 2017). Furthermore, functionally important microbes are not always restricted to high abundance taxa (Rivett & Bell, 2018), so a more direct method of revealing the microbial activity is desirable. Meta-

transcriptomic sequencing also referred to as RNA-seq, sequences all RNA strands in a sample and thus profiles genes that are actively expressed, providing insight into the functionality of active microbes (Bashiardes et al., 2016; Singer et al., 2017). A few meta-transcriptomic studies have been performed on anaerobic digesters, but these limited studies demonstrate the high sensitivity of RNA-seq in revealing the *in-situ* activity of low-abundant microbes, particularly methanogens, in response to chemical changes within the digester environment (De Vrieze et al., 2018; Fontana, Campanaro, et al., 2018 2018; Hao et al., 2020; Zhu et al., 2020). By combining WMS to reveal microbial taxonomic and functional potentials with RNA-Seq to provide active gene expression profiles, these studies offer a more holistic picture than DNA-based analyses in detailing the microbes that are not only present but also the microbes that are currently alive and active in an anaerobic digester. Despite the limitations of high-throughput DNA sequencing technologies for directly deducing the functionality of an active microbiome, it has been shown that functional interactions within a community have been correlated with changes in the abundance of phylotypes (Rivett & Bell, 2018), demonstrating that DNA-based sequencing remains a viable method in elucidating large-scale microbial community dynamics.

1.6 Thesis Project Overview

The widespread use of OWTS throughout North America and the detrimental environmental and health effects of untreated sewage runoff into watersheds prompts the need to determine the core microbiome and its functional

attributes responsible for the operation of efficient OWTS anaerobic digesters. In partnership with Waterloo Biofilter Systems Inc. (WBS), I sought to determine the microbial consortia responsible for AD in replicated conventional and modified anaerobic digesters equipped with plug-flow like tube design, called the *InnerTube*™. To do this, I employed whole-metagenome sequencing to 1.) identify characteristic microbial consortiums present in different digester designs, 2.) reveal if there is a progression of taxonomic and functional profiles across the digestion treatment process 3.) determine the response of taxonomic and functional gene counts towards the digester chemical environment and 4.) quantify the taxonomic abundance variability between replicate anaerobic digesters. Ultimately, this project aims to determine pertinent microbial consortiums and their functions that sustain efficient AD and digester designs that optimally support the functioning of microbial consortiums driving AD.

Table 1.1: WMS projects and study designs used in representative water and wastewater monitoring projects.

Project Design and Community Profiling	No. of Samples	Type and no. of Replicates	Read Length (bp)	Depth (avg. reads per sample)
Genome binning and assembly of replicate anaerobic digesters. (Campanaro et al., 2018)	6	Biological, 6	2 x 150	2.5×10^8
Longitudinal study of pathogens and cyanobacteria in drinking water supply. Samples were pooled for *de novo* assembly and reference-based tools were used for community profiling (Otten et al., 2016)	15	0	2 x 101	4.3×10^7
Metagenomic characterization of nitrogen-contaminated groundwater (Ludington et al., 2017)	5	0	2 x 100	5×10^7
Whole-metagenome sequencing of shore water and sand-water samples (Mohiuddin et al., 2017)	32	Biological, 16	2 x 100	9.4×10^6
Reference-based metagenomic survey of antibiotic-resistance genes in water reservoir environments, sewage, and beaches (Fresia et al., 2019)	20	0	2 x 125	3.9×10^7
Taxonomic and functional analysis of anaerobic digesters processing sludge and manure (Luo et al., 2016)	14	Biological, 6	2 x 100	2.8×10^7
Temporal study of cellulose-degrading lab-scale anaerobic digesters over 1 year. Assembled complete genomes (Vanwonterghem, Jensen, et al., 2016)	6	Biological, 3	2 x 150	1×10^8

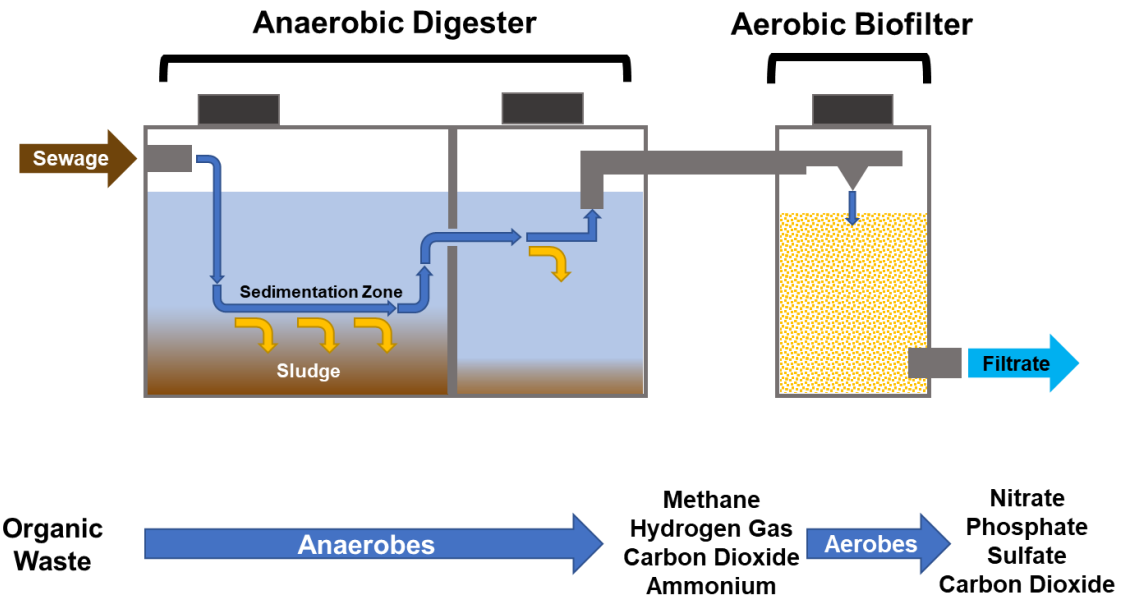


Figure 1.1: Diagram of an Onsite-Wastewater Treatment System (OWTS) with a conventional septic tank anaerobic digester and a downstream aerobic biofilter. Blue arrows represent the flow of sewage and yellow arrows represent the sedimentation of sludge biomass.

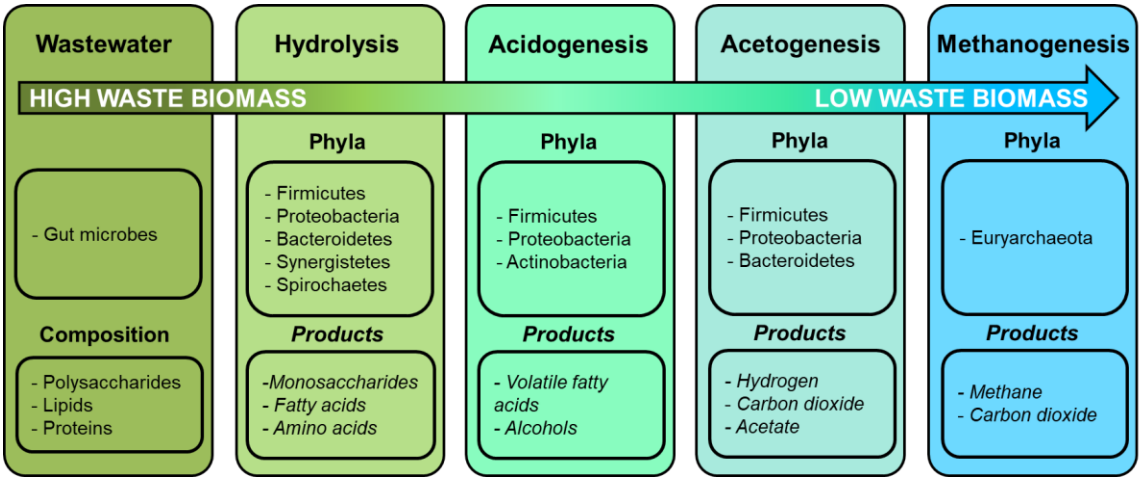


Figure 1.2: Overview of the most abundant Phyla driving and chemical by-products produced during the four stages of anaerobic digestion. Adapted from Chan et al., 2019.

2 Chapter 2: The Impact of Septic Tank Design and Flow on the Microbial Community

2.1 Introduction

The ability of anaerobic digesters to achieve effective degradation of organic waste is dependent on balanced community proportions of hydrolytic, acidogenic, acetogenic, and methanogenic microbial communities (Campanaro et al., 2016; De Vrieze et al., 2016; Vanwonterghem et al., 2014). Biomass degradation is driven by the metabolic pathways expressed by anaerobic microbes (N. Li et al., 2017; Vanwonterghem et al., 2014; Yin et al., 2018). Any degradation or improvement in anaerobic digestion efficiency due to digester design can be quantitatively determined by the physiochemical environment within each digester including the amount of chemical oxygen demand reduced and biogas produced (N. Li et al., 2017; Nascimento et al., 2018). It follows then that differences in physiochemical parameters indicative of degradation performance between digester designs influence the abundance of taxonomic and functional (metabolic) features of the biogas microbiome (Fontana, Campanaro, et al., 2018; Gao et al., 2019; Xu et al., 2018). Conversely, microbial communities participating in AD metabolize organic waste and produce constituent organic byproducts, further impacting the digester chemical environment (Heyer et al., 2019; Maus et al., 2017; Mosbaek et al., 2016).

2.2 The Impact of Digester Design on Microbial Communities

2.2.1 Digester Operational Parameters and the Impact on Microbial Community Stratification

Anaerobic digester designs have evolved from "low-rate" reactor designs such as conventional septic settling digesters (Figure 1.1) classified by the requirement of extended hydraulic retention times (HRT) for efficient organic waste removal to "high-rate" reactor designs requiring short HRT such as the Upflow Anaerobic Sludge Blanket digesters (UASB), Anaerobic Baffled Reactors (ABR) and plug-flow reactors (PFR) (Figure 2.1) (Abbasi et al., 2012; McAteer et al., 2020). HRT refers to the amount of residence time of the liquid portion of influent waste in the septic tank holding compartment. In anaerobic digesters, the settling and buildup of complex organic molecules establish a nutrient-rich environment in the form of semi-solid sludge for anaerobic microbes to grow on (Nascimento et al., 2018; Owusu-Agyeman et al., 2019; Sallis & Uyanik, 2003). Sludge deposited in anaerobic digesters such as UASB, PFR, ABR, and conventional tank digesters as those employed by WBS consists of flocculated particles with the core and surfaces of the granule made up of an active community of hydrolytic, fermentative, acetogenic, and methanogenic microbes in syntrophic relationships surrounded by a matrix of carbohydrates, proteins, and lipids (Figure 2.1, 2.2) (Nascimento et al., 2018; Owusu-Agyeman et al., 2019; Sallis & Uyanik, 2003; Satoh et al., 2007). Microbial groups that share chemical byproducts and are in direct syntrophic relationships are spatially organized close to each other

(Gonzalez-Gil et al., 2001; Satoh et al., 2007). Sludge particles contain dominant populations of hydrogenotrophic and acetoclastic methanogens such as *Methanosaetaeceae, Methanomicrobiales* respectively, in syntrophy and close-proximity with acetate- and hydrogen-producing acetogens such as *Anaerolinaeceae* and *Syntrophus* respectively (Kuroda et al., 2016; Nascimento et al., 2018; Owusu-Agyeman et al., 2019). The retention of these granules within the digestion chamber is crucial for maintaining the activity of slow-growing microbes, particularly acetogens and methanogens, to degrade organic loads into acetate and methane and thus maintain digester efficiency (McAteer et al., 2020).

2.2.2 The Influence of Anaerobic Digester Designs on Sewage and Microbial Biomass Contact

The majority of OWTS septic tanks employed by WBS are low-rate reactors, where a high HRT and a resulting low-rate of sewage flow is required to allow the settling and development of slow-growing acetogenic and methanogenic communities through the sedimentation of sewage sludge (Figure 2.2) (Jowett et al., 2017; McAteer et al., 2020). The septic tank design relegates the active microbial biomass to the bottom and corners of the septic tank while untreated sewage influent flows above (Figure 1.1, 2.2). Excessive influent flow rate will cause untreated to sewage to “short-circuit” the settled microbial biomass which prevents efficient biodegradation due to minimal contact between sewage and microbial biomass (Jowett et al., 2017; Wang et al., 2011). Additionally, excessive flow-rates can also cause the wash-out of the active microbial biomass into the

effluent (Jowett et al., 2017; Masse et al., 2013). Thus, for conventional septic tanks in OWTS, a long HRT (low influent flow-rate) is required for maximal contact between influent sewage and the active microbial biomass (Jowett et al., 2017; McAteer et al., 2020). In contrast, UASB digesters can receive higher influent sewage flow rates while still maximizing contact between the sewage and microbial sludge. UASB reactors accomplish this by pumping influent sewage against gravity up through a tube or tank with a high flow rate such that the active microbial biomass aggregated into granules become suspended in the untreated sewage water (Figure 2.1) (Si et al., 2015). Gravity prevents microbial granules from escaping into the effluent, and the high flow-rate ensures active mixing (Si et al., 2015). PFR and ABR digesters have similar HRT to conventional septic tanks, but these digesters can more readily retain microbial biomass and achieve maximal sewage-sludge mixing by forcing sewage through closed-conduits and multiple, baffled compartments (L. Dong et al., 2019; Jiang et al., 2018; Wang et al., 2011). Since there is no active mixing apparatus or mechanism in PFR and ABR, contact between the sewage and microbial biomass is improved due to greater compartmentalization of microbial biomass compared to conventional septic tanks (Figure 2.1) (L. L. Dong et al., 2019; Jiang et al., 2018; Masse et al., 2013). The high number of treatment compartments established by baffles in PFR and ABR digesters forces the sewage to contact the active microbes multiple times (Figure 2.1) (L. L. Dong et al., 2019; Jiang et al., 2018).

2.2.3 Improvements to WBS Septic Tank Anaerobic Digester Design

To increase the degradation efficiency of conventional OWTS septic tanks, WBS has outfitted conventional septic tanks with an *InnerTube*™. Waterloo Biofilter System's *InnerTube*™ is a corrugated plastic pipe that is 15 feet long with a diameter of around 12 inches (Jowett et al., 2017). It is connected to the inlet of the septic digester tank and directs the flow of sewage into the septic tank by looping around towards the front of the septic tank (Figure 2.2) (Jowett et al., 2017). The *InnerTube*™ -equipped septic tank resembles a PFR digester, where the *InnerTube*™ acts as a closed-conduit compartment which the sewage passes through before entering the greater septic tank volume (Figure 2.2) (Jowett et al., 2017). Internal testing by WBS demonstrated a reduction of COD and TSS in the effluent using *InnerTube*™ - equipped septic tanks compared to conventional digester systems of approximately 34% and 20%, respectively (Jowett et al., 2017). This agrees with other PFR efficiency measures, where PFR digesters achieved higher COD removal rates compared to non-compartmentalized digesters such as continuously-stirred tank reactors (CSTR) (Masse et al., 2013). Chemical oxygen demand is a measure of the total oxygen consumed by a system used to degrade colloidal and soluble organic compounds through oxidation, quantified as the mass of O_2 per volume of sample (Vanwonterghem et al., 2014; Wang et al., 2011). The higher COD removal rates in PFR and PFR-like (*InnerTube*™) digesters demonstrate greater biodegradation efficiencies of organic compounds.

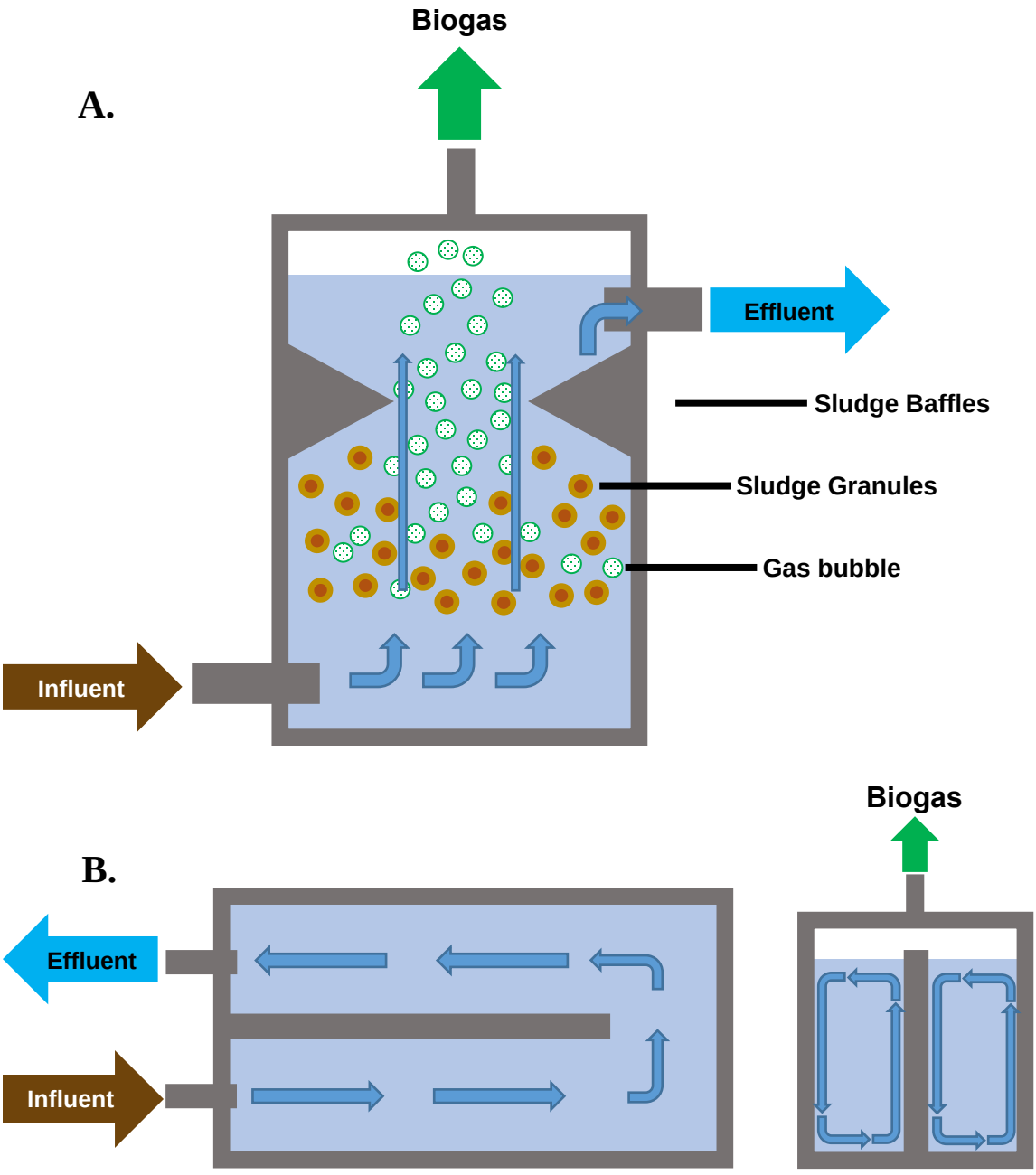
2.3 The Function of Recirculating Sewage Lines in Anaerobic Digesters

Anaerobic digesters have often been coupled with aerobic digesters to further reduce residual COD, TSS and total nitrogen (TN) produced from the anaerobic digestion stage (Akizuki et al., 2016; Novak et al., 2011). Organic waste (COD and TSS) removal is biologically coupled with nitrogen removal. Nitrogen removal occurs during anaerobic digestion primarily through the biological denitrification pathway, where facultatively anaerobic heterotrophic bacteria consuming saccharides as a carbon source utilize NO_3^- as an electron-acceptor to enzymatically produce ATP (Kuypers et al., 2018). The reduction of NO_3^- produces NO_2^- which is further reduced into NH_3 (Kuypers et al., 2018; Wang & Chu, 2016). NH_3 is then oxidized and nitrified back into NO_3^- in the aerobic digester (Akizuki et al., 2016). A recirculating line from the aerobic digester takes a proportion of the nitrified aerobic digester effluent and directs it back into the septic tank to replenish the NO_3^- supply in the anaerobic digester (Figure 2.3) (Akizuki et al., 2016). The breakdown of organic compounds and thus the enhanced removal COD and TSS are ensured by the continuous supply of NO_3^- from the recirculating line, thereby establishing a nitrification-denitrification cycle (Akizuki et al., 2016; Wang & Chu, 2016). The coupling of aerobic digesters to anaerobic digesters is useful not only for increasing nutrient removal rates, but also to reduce the accumulation rates of NH_3, which is inhibitory towards methanogens. Therefore, studying the effects of aerobic recirculating lines in OWTS is crucial for determining the effects of nitrogen cycling on AD process stability.

2.4 Chapter 2 Objectives

The objectives of this chapter were to investigate the characteristic changes in the digester microbial community composition due to differences in digester design and flow-rate. The first objective was to compare two anaerobic digester designs of a conventional septic tank and a PFR-like *InnerTube*™- equipped septic tank at the whole-community level and the level of specific, enriched genera. The second objective was to study the effect of aerobic recirculation in anaerobic digesters on microbial community compositions. The third objective was to determine the relationships between the digester chemical environment with microbial consortia that were enriched in conventional and *InnerTube*™- equipped septic tanks and between recirculating and single-pass digesters. Relating microbial community abundances to chemical parameters can reveal the potential response of the digester microbial community to changes in the tank chemical environment imposed by different digester configurations (Wang et al., 2011; Xu et al., 2018). Any changes in anaerobic digestion efficiency due to digester design can be quantitatively determined by the physiochemical environment within each digester including the amount of chemical oxygen demand reduced and biogas produced (N. Li et al., 2017; Nascimento et al., 2018). It follows then that changes in physicochemical parameters indicative of performance between digester designs influence the abundance of taxonomic and functional (metabolic) features of the biogas microbiome (Gao et al., 2019; Kougias et al., 2016; Xu et al., 2018). Conversely, microbial communities participating in hydrolysis, acidogenesis.

acetogenesis, and methanogenesis metabolize organic waste and produce constituent organic byproducts, further impacting the digester chemical environment (Heyer et al., 2019; Maus et al., 2017; Mosbaek et al., 2016). Therefore, it is hypothesized that distinct microbial populations will be enriched according to their putative metabolic characteristics in response to the differences in digester environments. Between the two flow configurations, aerobic effluent recirculation and single-pass, greater abundances of denitrifying microbes would arise in recirculating systems as a response to the elevated NO_3^- levels provided by the recirculating line. Between the two digester designs, the conventional septic tank and the *InnerTube*™-equipped septic tank, the higher COD and TSS removal rates observed in *InnerTube*™ reported in previous internal studies by WBS (Jowett, 2007) could be attributed to greater abundances of microbial acetogenic and methanogenic syntrophic phylotypes that have been implicated with greater solids and organic waste removal rates (Liu et al., 2016; Rivett & Bell, 2018; Shin et al., 2016).


A.
Biogas
Effluent
Sludge Baffles
Sludge Granules
Gas bubble
Influent
B.
Biogas
Effluent
Influent

C.

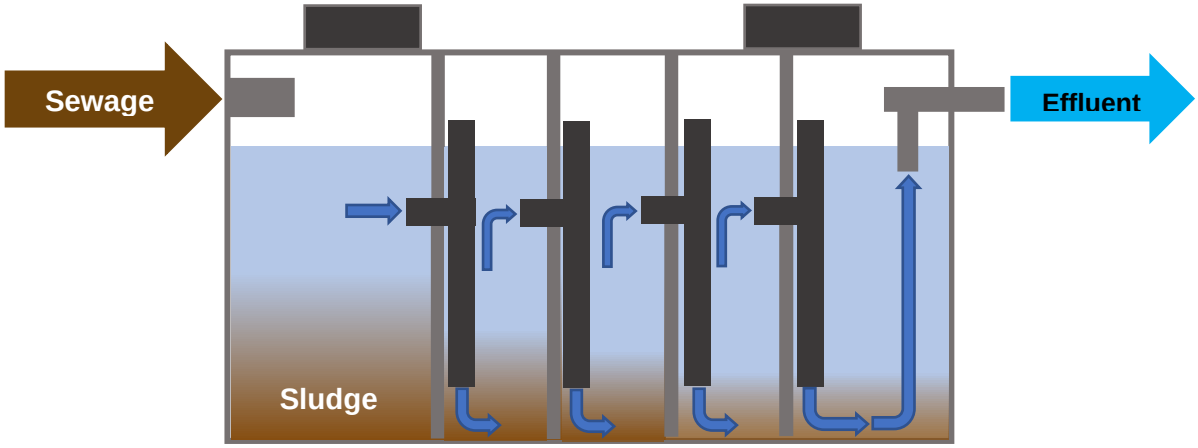


Figure 2.1: A.) Upflow Anaerobic Sludge Blanket reactor (UASB), B.) Anaerobic Plug Flow Reactor and C.) Anaerobic Baffled Reactor. A.) The sludge blanket is comprised of microbial sludge granules suspended by the high-rate of influent sewage flow. Microbial digestion of organic matter produces biogas by-products. Light- blue arrows indicate the flow of sewage. B.) Top-down view of the reactor and its baffle (left), Cross-sectional view of the reactor (right). Light-blue arrows indicate the flow of sewage. Adapted from Dong et al, 2019. C.) Black rectangles represent conduit pipes connecting sewage flow into each baffle. Adapted from Sallis and Uyanik (2003).

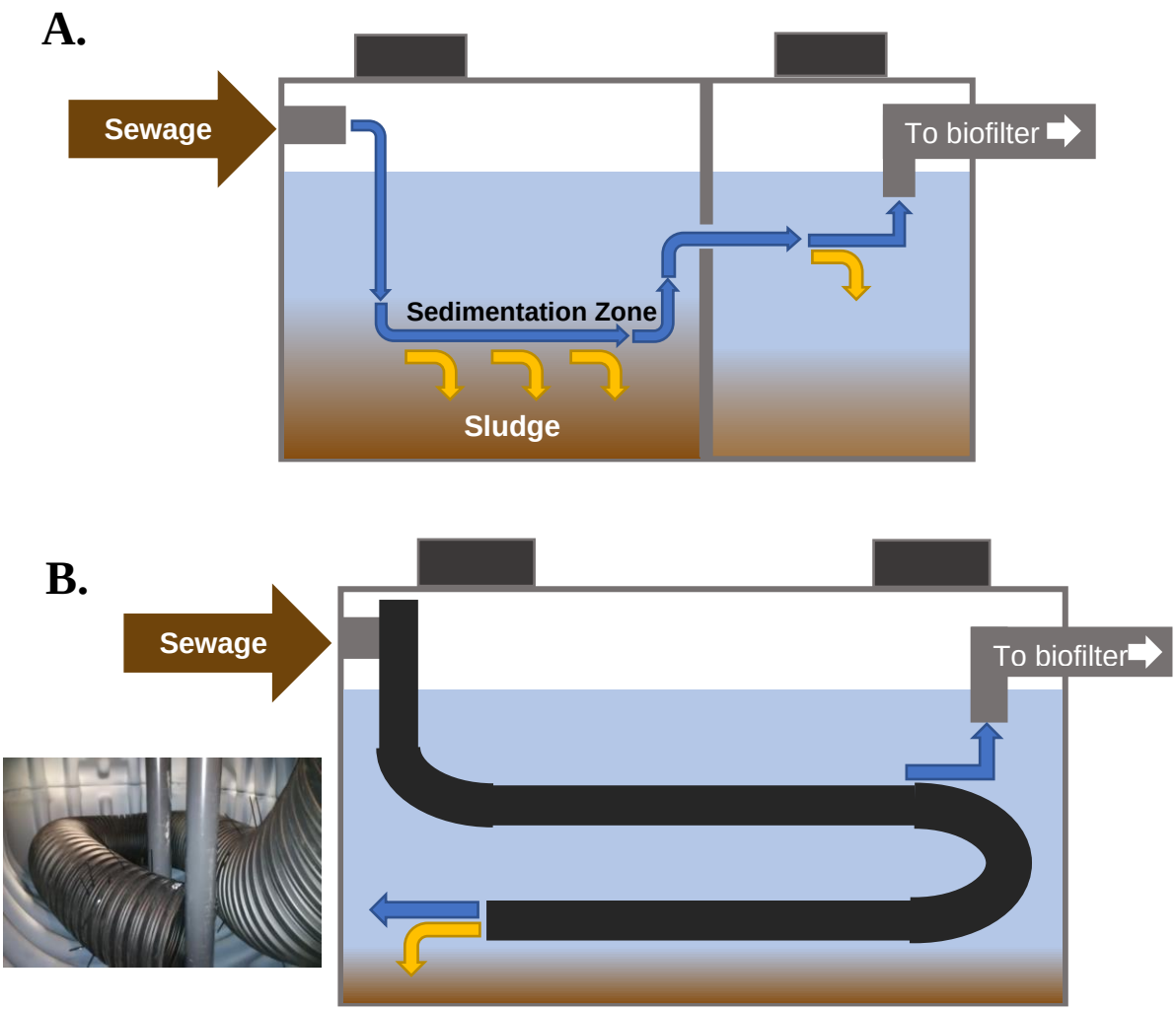


Figure 2.2: Conventional and InnerTube-equipped Septic Tanks. A.) Conventional anaerobic septic tank digester B.) InnerTube-equipped anaerobic septic tank digester. Blue arrows indicate the flow of sewage, yellow arrows indicate sludge sedimentation.

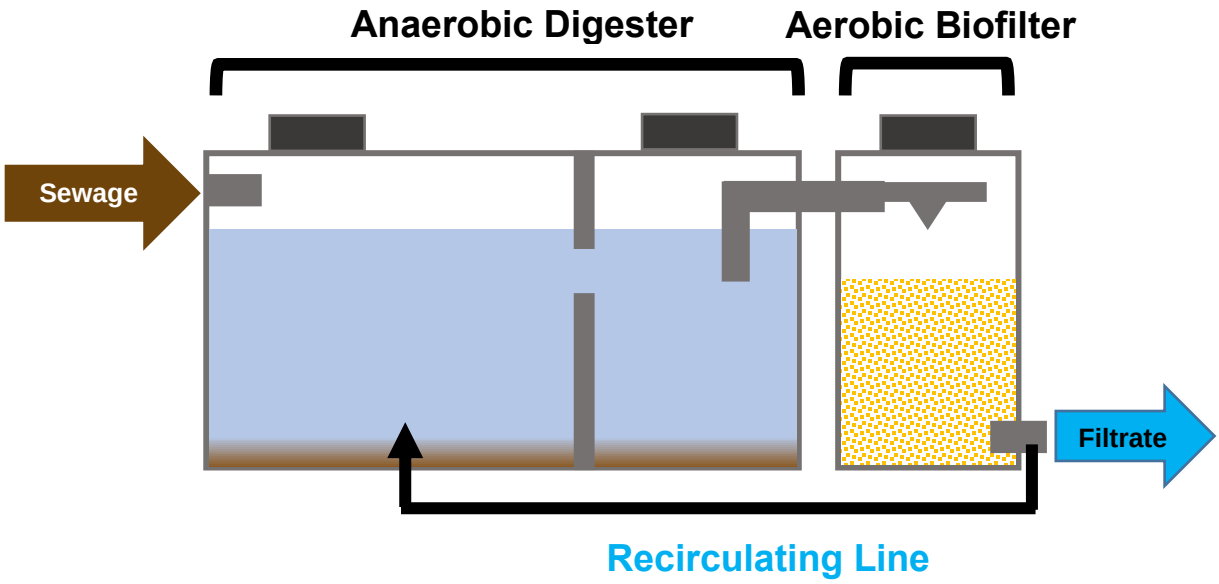


Figure 2.3: Onsite-Wastewater Treatment System equipped with a recirculating-flow line.

2.5 Methods

2.5.1 Digester Site Descriptions

Replicate digester samples were collected by McMaster researchers and WBS employees from *InnerTube*™ and conventional septic tank installations across Southern and Central Ontario (Figure 2.4) from September 2018 to January 2019. All systems were located on residential properties and serviced single households. Each OWTS was comprised of a septic tank connected to a downstream aerobic biofilter unit (Figure 2.5). Four types of systems were sampled:

1. *InnerTube*™ recirculating septic digesters (IR) (*n = 6*),
2. *InnerTube*™ single-pass septic digesters (IS) (*n* = 5),
3. Conventional recirculating septic digesters (CR) (*n = 6*),
4. Conventional single-pass septic digesters (CS) (*n = 6*).

However, only samples from three digester sites (*n*=3) for each type that had the highest average DNA concentrations were shotgun sequenced (Table S 1 and Table S 4). Samples are identified by an alphanumeric code, where numbers prefixing the digester acronym represents a digester biological replicate, and numbers suffixing the digester acronym represents the sampling point, where "1" Influent, "2" Tank, and "3" Effluent. Recirculating digesters are systems that recirculated a portion of aerobic biofilter effluent into the septic tank. Single-pass systems do not recirculate aerobic biofilter effluent. The flow of effluent from the aerobic biofilter recirculated back into the anaerobic digester tank was periodically turned on and off and was not continuous.

2.5.2 Sewage Sampling Techniques and Sample Pre-Treatment

For each digester type and replicate, effluent, tank, and influent samples were taken in that order to minimize chemical cross-contamination between samples that were treated (effluent) being treated (tank) and untreated (influent). The tube sampler and collection vessel were washed with commercial bottled water between sampling attempts to minimize microbial cross-contamination. Sewage was sampled through the riser openings installed in all digester types (Figure 2.5). Conventional septic tank effluent samples were taken from the effluent holding tank by submerging the collection vessel into the effluent sewage (Figure 2.6) directly beneath the septic tank outlet (Figure 2.5). For Conventional septic systems, influent and tank sewage were collected by using the tube sampler (Figure 2.6). The tube sampler was vertically submerged into sewage to a depth marked by the green tape. Subsequently, the sewage inside the tube was collected by plugging off the bottom end of the tube with the foam ball by pulling up and fastening the rope around the hook (Figure 2.6).

InnerTube™ effluent samples were collected from the spray nozzle feeding the aerobic biofilter unit by placing the collection vessel directly beneath the nozzle (Figure 2.5). *InnerTube*™ influent and tank sewage were collected from the inlet opening of the *InnerTube*™ and the approximate outlet of the *InnerTube*™ within the tank, respectively, using the tube sampler (Figure 2.2, 2.5). All samples collected using the tube sampler were then emptied into the collection vessel (Figure 2.6) by releasing the rope and foam ball. Sewage was repeatedly collected

until the 1.5 L collection vessel was filled. The contents of the collection vessel were then emptied into duplicate sterilized 500 mL plastic bottles. One of each of the duplicate bottles was shipped to Fleming College's Centre for Advancement of Wastewater Treatment (CAWT) for chemical testing and another transported to McMaster University for DNA extraction. A field blank was obtained at the end of each sampling day by washing the collection vessel with spring water and collecting the washing into a 500 mL sterile plastic bottle. All sewage samples stored in the 500 mL plastic bottles were immediately placed on ice and transported within ice coolers.

To assess the effects of freezing and alternate sample preservation methods on representative microbial community profiles, 4IR1, and 4RI2 influent and tank samples respectively, (second biological replicate of the *InnerTube*™ recirculating digester, 2RI), was sampled three more times for a total of five 4RI replicates collected (Table S 1, 2). To clarify, five 500 mL influent and tank samples from the 4RI site were collected, one for CAWT chemical analyses, three to serve as technical replicates for DNA extraction, and one more to compare against other frozen digester types (Table S 1, 2). All five 500 mL samples were transported on ice, however, two were frozen in storage at -80ºC before chemical analyses and sequencing while the other three were immediately DNA extracted and sequenced, forgoing freezing (Table S 2). In addition to five 500 mL samples, three 10 mL 4RI influent sewage samples were collected into three 50 mL falcon tubes, each pre-filled with 10 mL of deionized water as a control, 10 mL of 70% ethanol, and 5 mL

of Norgen Biotek's Blood DNA 3x preservative buffer, resulting in sewage-to-preservative mixture ratios of 1:1, 1:1 and 2:1 respectively. The 2:1 ratio of sewage sample to Blood DNA preservative was obtained from Norgen Biotek's Blood DNA preservation protocol. The 10 mL influent were transported on ice and stored at -80ºC before DNA extraction and sequencing (Table S 2).

2.5.3 Chemical Analyses

For each sampling point, two 100 mL bottles both preloaded with 200 µL sulfuric acid (H_2SO_4) for NH_3/TKN and COD measurements were filled with wastewater. All sewage samples (Table S 1) were sent to CAWT on ice for chemical analysis. Chemical measures of total suspended solids (TSS), total Kjeldahl nitrogen (TKN), ammonia (NH_3), and chemical oxygen demand (COD) were performed for all samples except 4RI technical replicates (Table S 2). CAWT protocols are described in the CALA Directory of Laboratories listing (member number: 3628) and are based on Standard Methods for the Examination of Water and Wastewater, 22nd ed. (Rice et al., 2012). Biological Oxygen Demand (BOD) was measured for only influent and effluent samples in McMaster University from Method 5210b in Standard Methods for the Examination of Water and Wastewater, 22nd ed. (Rice et al., 2012). BOD dissolved oxygen was measured with the YSI Pro-BOD probe and onsite pH, temperature, and dissolved oxygen were measured with the YSI Professional Plus multimeter (Xylem Inc.). Chemical oxygen demand is defined as the oxygen consumed in a system by either biological or inorganic reactions, total Kjeldahl nitrogen is the total organic nitrogen content in a system,

total suspended solids refers to the amount of undissolved suspended particles, and dissolved oxygen and BOD are measures of free doublet oxygen present in a system, and the amount of dissolved oxygen consumed by aerobic organisms, respectively (Rice et al., 2012). Tank residence volume and Flow rates of each digester type were provided by Waterloo Biofilters Inc. and the HRT values were calculated by dividing the Tank residence volume (L) by the Flow rate (L/day) (Table S 3).

2.5.4 DNA Extraction

Sewage samples were thawed to room temperature from -80ºC and then vacuum filtered through sterile 0.22 µm cellulose filters to concentrate microbes (Brown et al., 2015; Mohiuddin et al., 2019). Sample volumes were filtered through until the filter clogged and no liquid passed through. The filters were then folded and fit into sterile capped stock tubes pre-filled with 0.25 µL of 0.1 mm zirconium beads (Bag et al., 2016).

The microfuge tubes with filters were then extracted for DNA using the Norgen Biotek Soil DNA Isolation Plus Kit following the included protocols with some changes that are outlined (Norgen Biotek Corp.). All extraction solutions are proprietary unless otherwise noted. Cell lysis buffer was first added to the stock tubes and the tubes were subsequently mechanically agitated and bead-beaten with a Sonibeast Small Sample Cell Disruptor (Bio spec Products, Inc.) for 3 intervals of 40s, for a total of 120 s. After bead-beating, the stock tubes were then centrifuged at 20000 x g for 1 minute and supernatants were collected while

avoiding the zirconium beads. Cell protein content was precipitated by mixing a binding buffer and placing the samples on ice for 5 minutes. Samples were centrifuged again at 20000 x g for 1 minute to pellet protein and residual particles. Humic acid contamination was removed using a proprietary additive followed by centrifugation and supernatant isolation. DNA was then isolated through DNA-binding spin columns through centrifugation. Columns were washed twice with an ethanol solution and eluted from the column by centrifugation and 50 µL of elution buffer. The DNA extracted by the buffer was then quantified with the Qubit 2.0 Fluorometer (Invitrogen™ by Thermofischer Scientific.). Extracted DNA samples were stored in -20ºC before sequencing.

2.5.5 Whole-Metagenome Shotgun Sequencing

DNA concentrations were normalized based on two sample groups due to the wide range of DNA concentrations (Table S 4). The first group being only 4RI pre-treatment technical replicates (bolded samples in Table S 4) was normalized to the sample with the lowest concentration, 4IR1-EtOH 2, and the second group being all other samples, was normalized to the sample with the lowest concentration, 5CR3 (Table S 4). Samples were sent to the Farncombe Sequencing Institute at McMaster University for paired-end HiSeq 2500 sequencing, with a read length of 150 bp and insert size of 500 bp fragments using the NEBNext® Ultra™ II DNA library preparation kit for Illumina (New England Biolabs Inc.). Samples were sequenced on two separate HiSeq lanes.

2.5.6 Raw Read Sequence Pre-processing, Alignment, and Classification

Since the samples were spread across two lanes, two raw FASTQ files were generated from each sample. These FASTQ files were concatenated with the UNIX `cat` command and quality filtered using **Trimmomatic v0.39** (Bolger et al., 2014), using TruSeq3 paired-end adapters, at a Phred score cutoff of 33, using a sliding window model with 3 leading and trailing base pairs, a width of 4 bases, and a minimum quality score of 20. All subsequent downstream analysis steps were performed only on forward reads since forward reads have higher read quality than reverse reads in Illumina sequencing (Luo et al., 2012). Read quality statistics were assessed using the R package 'fastqcr' (Kassambara, 2019) based on the quality-checking tool **FastQC**. The average mean read counts of all samples after trimming was 5.3 M reads (SD ± 1.6 M) (Table S 4).

Sequences were aligned using **DIAMOND-BLASTx** algorithm (Buchfink et al., 2015) against NCBI's non-redundant (nr) protein database, with an e-value of $1x10^{-5}$ and a maximum number of target sequences of 25 per query to report alignments for. All other settings were left to their program default values. The output files were formatted as **DIAMOND** alignment archives (.daa) for **MEGAN6 V6.18.4** to use for sequence classification and binning (Huson et al., 2016). Sequences were binned into taxonomic groups using the Weighted Lowest Common Ancestor Algorithm provided by **MEGAN6** with a minimum quality to assign sequences of 50 and a maximum E-value cut-off of 0.01. The sequences were assigned by referencing the NCBI nr-protein database for taxonomic

classification and SEED Functional Gene Subsystem Database (Overbeek et al., 2014) for functional gene classification. Taxonomic and functional gene counts were exported as tab-delimited .csv files to import into the R statistical software for statistical analyses (R Development Core Team, 2010).

2.5.7 Taxonomic Count Data Pre-Processing and Statistical Analyses

Before statistical analysis, taxonomic levels that were labeled "no ranked" from the NCBI taxonomic database were removed via the R package 'taxonomizr' (Sherrill-Mix, 2019) to remove missing taxonomic ranks and clades and to instead obtain standard taxonomic rank profiles under Kingdom, Phylum, Class, Order, Family, Genus, and Species. This was done to achieve full compatibility for downstream analyses with the 'phyloseq' R package (McMurdie & Holmes, 2013). Additionally, taxa that were unassigned, or with NA values, across all taxonomic ranks with an abundance of < 0.0001% of total reads were removed. Eukaryotic and virus reads were removed using taxonomizr. Subsequently, the remaining reads that were unassigned were labeled with the format "Unclassified_*taxa*" where "*taxa"* represented the lowest taxonomic rank of the organism that was assigned.

Count data were Hellinger-standardized using the R package 'biodiversityR' (Kindt & Coe, 2005) for all downstream statistical analyses except for relative abundance, diversity indices, power curve, rank abundance, rarefaction, and differential abundance graphs. Hellinger-standardization, or the square-root of species relative abundances across samples, was applied as it is recommended for use on species abundance data that characteristically have many species with low

counts and many zeroes (Buttigieg & Ramette, 2014; Ramette, 2007). Hellinger-standardization is suitable for metagenomic count data due to it giving low weights to variables with low counts and many zeroes (Buttigieg & Ramette, 2014; Legendre & Gallagher, 2001). All graphs were generated by ggplot2 (Wickham, 2016) except for correlation heatmaps, hierarchical dendrograms, and rarefaction curves. Relative abundances graphs were generated by agglomerating to either the Phylum level or in the case of Euryarchaeota abundances, at the Family level using phyloseq.

Power curves based on the multivariate dissimilarity-based standard error estimates (multSE) were generated to determine the required replicates to minimize standard errors for each digester type. Power curves were based on the multSE function by (Anderson & Santana-Garcon, 2014). Constrained redundancy analysis (RDA) between microbial community compositions and chemical parameters and non-metric multidimensional scaling (NMDS) ordination methods for all taxonomic and functional samples were generated using phyloseq and ggplot2 using the Bray-Curtis dissimilarity index. RDA ordination was used to reveal the influence of chemical parameters on the differences in variance of the microbiome in the four digester types, Conventional Recirculating (CR), Conventional Single-pass (CS), *InnerTube*™ Recirculating (IR), and *InnerTube*™ Single-pass (IS). The proportions of variance of Bray-Curtis dissimilarities explained by chemical parameters were generated using the varpart() function in the 'vegan' package. Hierarchical agglomerative clustering based on the Bray-Curtis dissimilarity index

was used on taxonomic and functional profiles using the R package 'dendextend' (Galili, 2015). Differential Abundance pairwise comparison graphs between *InnerTube*™ and Conventional digester types, and recirculating and single-pass digester flow configurations were generated using the 'DESeq2' package to identify differentially abundant taxonomic counts (Ferrocino et al., 2018; Jonsson et al., 2016; Love et al., 2014). Raw taxonomic count datasets were used to generate **DESeq2** pairwise comparisons of significantly enriched taxa in log2-fold changes between digester types (Conventional vs. InnerTube™) and flow-configurations (Recirculating vs. Single-pass) according to published **DESeq2** protocol (Love et al., 2014). Additionally, log10-mean normalized abundances were also generated for each factor level in a given comparison. Taxonomic rank counts were agglomerated at the genus level using phyloseq before **DESeq2** comparisons. For digester type and flow configuration comparisons, an abundance cutoff of 0.1% of total agglomerated genera counts and an adjusted p-value cutoff of < 0.001 was applied to include highly significantly differentially enriched rare genera. The order of terms in **DESeq2** design formulas is important, where the final term refers to the variable containing the factors to be compared, and the terms prior to the final term refers to variables to be controlled for as co-variables (Love et al., 2014; Wilkinson & Rogers, 1973). The **DESeq2** formula design for pairwise comparisons of digester type is $\sim Sampling\ point + Flow\ Configuration + Digester\ Type$ and for flow configurations it is $\sim Sampling\ point + Digester\ Type + Flow\ Configuration$ to treat for sampling points (influent, tank and effluent) and

the flow or digester type variance effects depending on which variable was compared (Love et al., 2014). The Wald test with a "local" fit type was used to set up the hypothesis tests for each gene, with the null hypothesis being that a given genera log2-fold change equals to zero (Love et al., 2014). Outlier genera were culled using a Cook's distance cutoff built into the **DESeq2** results() function (Love et al., 2014). Log10-normalized mean abundances of each comparison group were then generated to demonstrate the log2-fold changes in terms of abundances. Differentially abundant genera between digester types and flow configurations were then correlated using Spearman rank-order correlations to chemical parameters using the R function rcorr() in the 'Hmisc' package (Frank E Harrell Jr, 2020) and plotted with the 'corrplot' package (Taiyun Wei, 2017). Rarefaction curves were generated with the rarecurve() function in the vegan package.

The Permutational Multivariate Analysis of Variance (PERMANOVA) significance test was used to determine if digester types differed in overall microbial composition and which chemical parameters had a significant contributing effect on the variance of overall microbiome compositions (Anderson, 2017). PERMANOVA tests across digester types were done by using the vegan function adonis() as it uses Type I Sequential Sums of Squares to obtain the sum of squares of digester types as well as the interaction between sampling point and digester type (Table 2.1). The adonis2() function was used to test the contribution of chemical parameters to the variances of microbial communities as it uses Type III Marginal Sums of Squares to obtain the sum of squares for each chemical

parameter independent of the sequence of terms (Table 2.1) (Jari Oksanen et al., 2019). Pairwise PERMANOVA comparisons were performed among the four digester types, Conventional Recirculating (CR), Conventional Single-pass (CS), *InnerTube*™ Recirculating (IR), and *InnerTube*™ Single-pass (IS), using the package pairwiseAdonis (Arbizu, 2017). Beta dispersion tests provided by the permutest() function in the vegan package was used to determine whether within-digester-group-variances were homogenous throughout all digester types (Jari Oksanen et al., 2019).

Similarity Percentage analysis (**SIMPER**, Clarke and Warwick (2001)) was used to determine those genera that contributed the most to Bray-Curtis dissimilarities between the four digester types. **SIMPER** analysis was done using the simper() function from the vegan package implemented into a custom script created by Steinberger (2018) called simper.pretty(). The script allows custom **SIMPER** group comparisons based on sample group labels (CR, IR, CS, IS) and implements percent relative abundance and percent contribution cutoffs to reduce the number of taxa returned (Steinberger, 2018). Only a **SIMPER** percent contribution cutoff of 0.1% was applied. Subsequently, pairwise Kruskal-Wallis rank-sum significance comparisons between the four digester types, adjusted with the Benjamini & Hochberg Correction for false discovery rate, was performed on the genera selected by **SIMPER** analysis to determine the genera that were significantly different in abundance. Pairwise Kruskal-Wallis rank-sum tests were performed using another custom script created by Steinberger (2018) called

kruskal.pretty(). Only genera showing differential abundances at an adjusted *p-value* < 0.05 were included (Table 2.3).

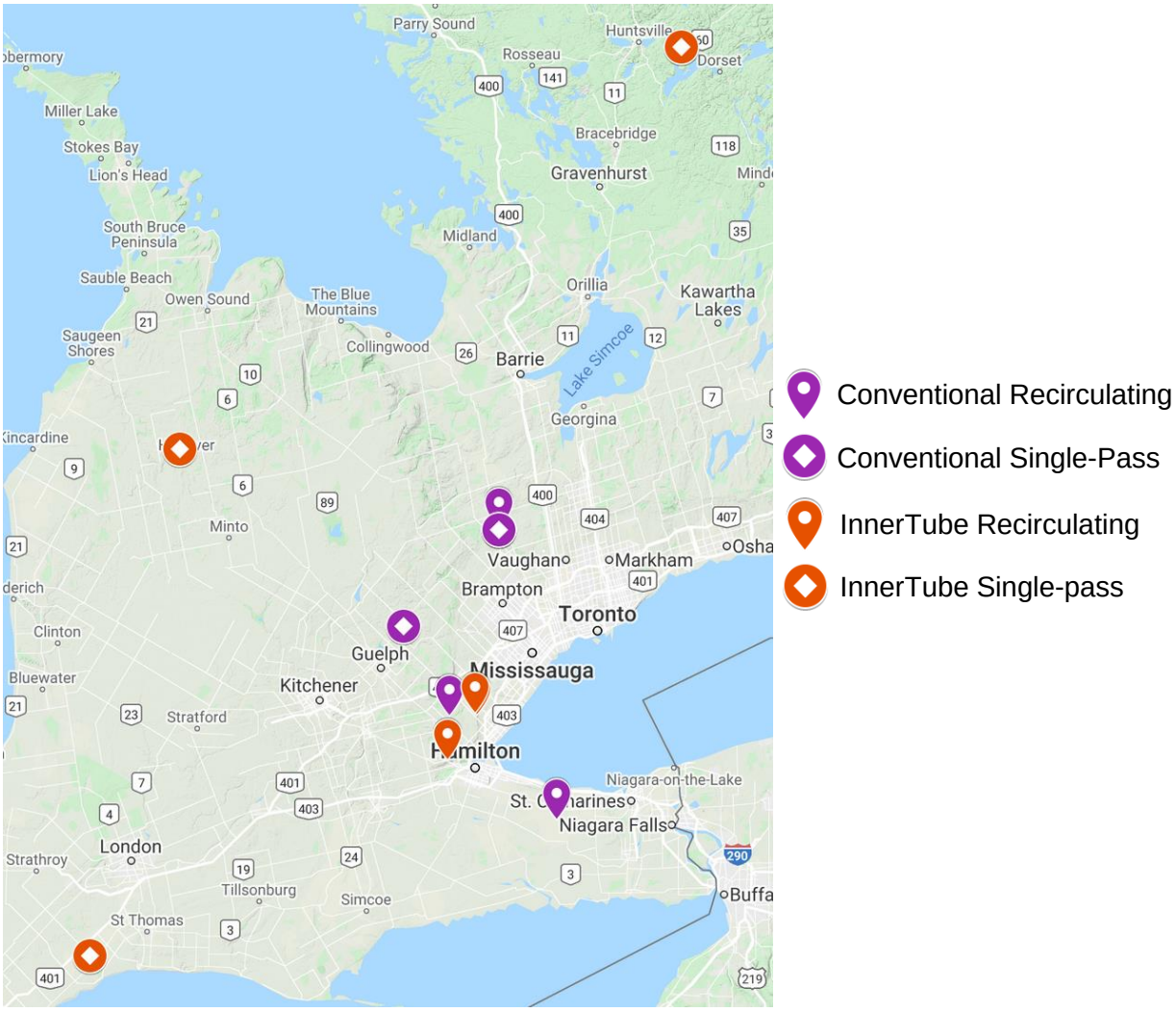


Figure 2.4: Sampling locations of the four digester types.

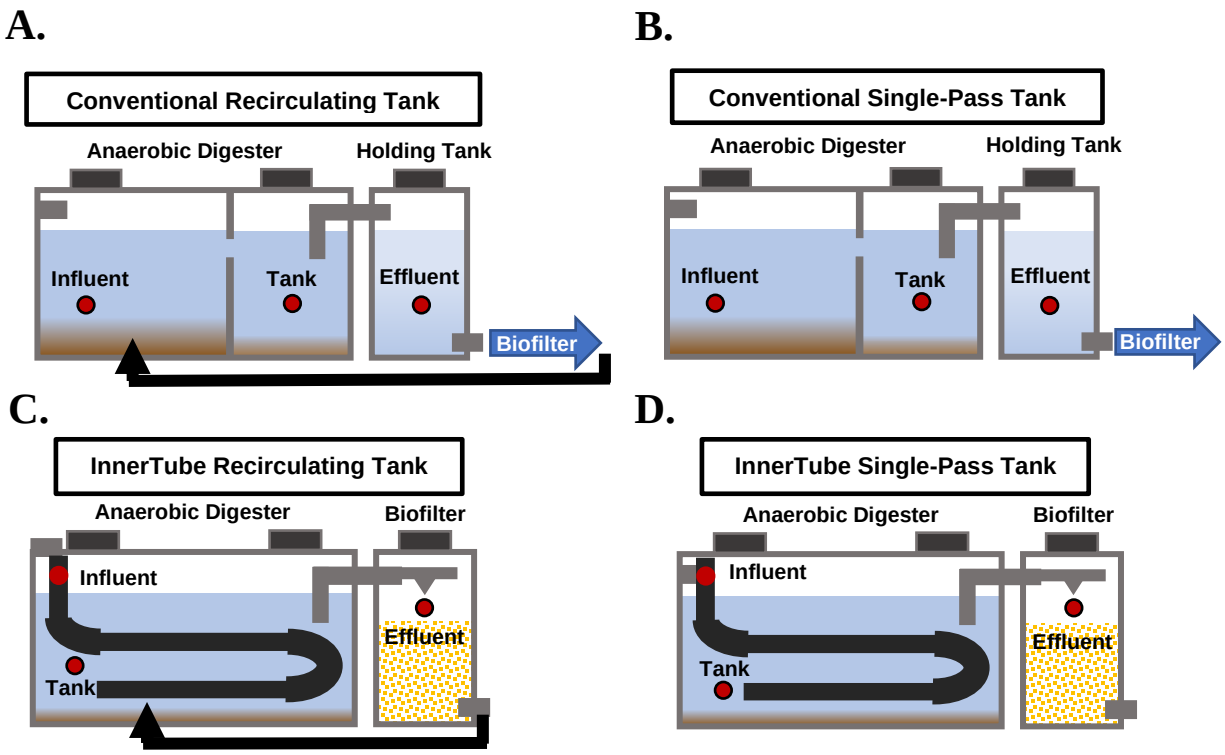


Figure 2.5: Four digester types examined with sampling points. Three replicate digesters were each sampled at Influent, Tank, and Effluent treatment points:

A.) Conventional Recirculating (*n=3*)

B.) Conventional Single-pass (*n=3*)

C.) InnerTube Recirculating (*n=3*)

D.) InnerTube Single-Pass (*n=3*)

A.

B.

Figure 2.6: Septic tank sampling apparatus. A.) Effluent sampler (top) – A PVC container with an approximate volume of 0.75 L attached to the bottom of a ~4 ft. pole with two metallic screws. Collection vessel (bottom) – A PVC container with an approximate volume of 1.25 L attached to the bottom of a ~4 ft. plastic pole with two metallic screws. B.) Tube sampler – A ~4 ft. hollow plastic tube with two open ends. A rope with a hard foam ball on the bottom end (left) runs through the entire length of the tube and exits at the white-tipped end at the top of the tube (right). A screw protrudes near the top of the tube below the white-tipped end to serve as a hook for the rope to be secured on. The green tape designates an approximate 1 L volume mark of the sampler.

2.6 Results

2.6.1 Chemical Analyses

The chemical environment of the four digester types across influent, tank, and effluent treatment points to provide an estimate on the biodegradation performance achieved in terms of COD, BOD, and TSS removal rates and to gain insight into the substrate conditions in pH, temperature, DO, and nitrogen content (TKN and NH_3) that the microbial communities operated in (Figure 2.17). *InnerTube*™ systems demonstrated higher COD, BOD, and TSS removal rates than conventional systems (Figure 2.17). TKN and NH_3 remained similar from influent to effluent in recirculating systems, while decreasing in conventional single-pass systems and increasing in *InnerTube*™ single-pass systems (Figure 2.17). pH and DO increased throughout from influent to effluent in all digester types, while temperature decreased except in *InnerTube*™ recirculating (Figure 2.17).

2.6.2 2.6.2 Relative Taxonomic Abundance Variability Across Digester Types

Anaerobic digester efficiency has been reported to be sustained by a core consortium of microbes (Calusinska et al., 2018). Therefore, relative phyla abundances were calculated for each sample collected across three replicate *InnerTube*™ recirculating and single-pass digesters and conventional recirculating and single-pass digesters to provide a broad overview of differences in community proportions across influent, tank, and effluent sewage samples in the four digester types (Figure 2.7). Additionally, since anaerobic digester function have been associated with keystone methanogenic microbes under the Euryarchaeota phylum,

relative family abundances of the Euryarchaeota phylum were also calculated to identify the compositions of methanogens, which have been reported to exclusively drive the finals step of AD, methanogenesis, across samples (Figure 2.8) (Campanaro et al., 2016). For conventional recirculating and single-pass digesters, the most abundant phyla were Proteobacteria (37.5%), Firmicutes (20.4%), Bacteroidetes (16.4%), Synergistetes (8.6%), Actinobacteria (5.5%), Euryarchaeota (4.1%), with Verrucomicrobia at 1.2% (Figure 2.7, 2.9). For *InnerTube*™ recirculating and single-pass digesters, the most abundant phyla were Proteobacteria (43.3%), Firmicutes (24.2%), Bacteroidetes (13.4%), Actinobacteria (6.5%), Verrucomicrobia (1.8%), and Euryarchaeota (1.0%) (Figure 2.7, 2.9). Synergistetes was not within the top 6 phyla in *InnerTube*™ systems, although Euryarchaeota was more abundant than in conventional systems (Figure 2.7, 2.9). The scatterplot of mean Hellinger-standardized abundances with standard errors reveals that abundances differ within the same digester type and across replicates, with some phyla abundances replicates such as Proteobacteria in *InnerTube*™ single-pass digesters, dispersing far from the mean (Figure 2.9). Abundance trends of the top 6 phyla were observed across influent, tank, and effluent sampling points (Figure 2.9). Proteobacteria increases from the influent to effluent except in IR digesters, Synergistetes increases from influent to effluent in all digester types except in CR digesters and Euryarchaeota increases from influent to effluent while Actinobacteria conversely decreases from influent to effluent in all four digester types (Figure 2.9). For methanogen compositions, the top families

under the Euryarchaeota phylum in conventional digesters were Methanosaetaceae (34.0%), Methanospirillaceae (26.5%), Methanosarcinaeceae (9.8%), Methanobacteriaceae (8.3%), an Unclassified Methanomicrobiales (7.5%), and Methanoregulaceae (5.2%) (Figure 2.8). For *InnerTube*™ digesters, Methanospirillaceae (36.3%), Methanobacteriaceae (16.7%), Methanosarcinaceae (16.5%), Methanosaetaceae (9.7%), Unclassified Methanomicrobiales (6.3%), and Methanoregulaceae (2.3%) dominated the Euryarchaeota proportions (Figure 2.8). Proportions of Euryarchaeota families also differed among replicates, although some trends across influent, tank, and effluent sewage can be seen, where for example, Methanospirillaceae and Methanobacteriaceae proportions decreased across sampling points in *InnerTube*™ Conventional and *InnerTube*™ single-pass digesters (Figure 2.8).

Pairwise Wilcoxon-Rank sum significance tests were also performed to determine significantly different ($p < 0.05$) abundances of dominant phyla between digester types (Figure 2.9). There were no significant rank abundance differences in any of the digesters among the top 2 phyla, Proteobacteria, and Firmicutes (Figure 2.9). However, Bacteroidetes abundances were significantly different among recirculating systems (CR vs. IR), *InnerTube*™ systems (IR vs. IS) and CR vs. IS digesters. Synergistetes abundances were only significantly different between single-pass systems (CS vs. IS). Actinobacteria was significantly different between recirculating systems (CR vs. IR) and *InnerTube*™ systems (IR vs. IS).

Euryarchaeota was significantly different between recirculating systems, single-pass systems, CR vs. IS, and IR vs. CS systems (Figure 2.9).

Power curves of digester types based on multivariate Bray-Curtis dissimilarity standard error estimates (multSE) were generated to determine the minimum number of digester type replicates at which no quantifiable difference in standard error could be observed (Figure 2.10). All digester types required at least 7 replicates, while *InnerTube*™ recirculating digesters only required 6 replicates. Across all replicates, IS systems had the highest standard errors, followed by CS, IR, and finally CR systems with the lowest standard errors (Figure 2.10).

2.6.3 Ordination of Digester Types

Constrained ordination of digester types based on Bray-Curtis dissimilarity measures with chemical parameters was done using Redundancy Analysis (RDA), also known as constrained analysis of principal components (CAP) (Figure 2.11). RDA was used to determine how much of the variance in microbial community compositions between digester types can be explained, or constrained, by changes in chemical parameters (Figure 2.11) (Paliy & Shankar, 2016). RDA of microbial communities grouped into digester type and sampling point showed greater percent variance in the primary CAP (Constrained Analysis of Principal Coordinates) axis compared to the secondary axis (Figure 2.11). Variances of *InnerTube*™ single-pass communities were influenced by ammonia (NH_3) and TKN, while the other three digester types were mostly influenced by temperature, pH, COD, and TSS (Figure 2.11). Specifically, pH influenced tank and effluent sewage from CS,

temperature on IR and CR systems, and COD and TSS on IR, CR, and CS systems (Figure 2.11). Sample community variances were significantly correlated with temperature (PERMANOVA; pseudo-f = 1.775; p = 0.029), NH_3 (PERMANOVA; pseudo-f = 2.220; p = 0.007), and pH (PERMANOVA; pseudo-f = 2.365; p = 0.003) (Table 2.1). Temperature, NH_3, and pH explained 6.5%, 13.0%, and 10.8%, respectively of the variation between digester type microbial communities. Non-metric dimensional scaling (NMDS) ordination is an unconstrained ordination method and does not take into account the effect of chemical parameters on microbial community clustering (Paliy & Shankar, 2016). Similar to RDA ordinations, *InnerTube*™ Single-pass digesters did not cluster with the other three digester types (Figure 2.12). NMDS ordination was also performed on technical replicates of *InnerTube*™ recirculating digester #2 to assess the effect that the sampling and pre-treatment protocol has on microbial community compositions (Figure 2.18, Table S 2). The unpreserved samples that were extracted immediately without storage at -80ºC clustered closely together compared to all other pre-treatment methods (Figure 2.18). Additionally, unpreserved samples were separated similarly to the frozen unpreserved samples for both influent and tank samples along the primary NMDS axis (Figure 2.18).

2.6.4 Significance Tests of Ordinations Between Digester Types and Treatment Points

PERMANOVA was used to assess the significance of microbial community composition variances between digester types and influent, tank, and effluent

sampling points (Table 2.1). Microbial community compositions significantly differed in variance between the four digester types (PERMANOVA; pseudo-f = 2.496; p = 0.001), although sampling points and the interaction between digester types and sampling points were not significantly different between influent, tank and effluent community compositions (Table 2.1). Pairwise PERMANOVA tests further confirm this, where all pairwise digester type comparisons were significantly different whereas pairwise digester type comparisons with respect to sampling points were not Variation partitioning also confirmed that digester type accounts for most of the microbial community variation between samples at 13.6% explained variance whereas sampling points explained only -0.01% variance. BETADISPER tests were then performed to determine if the significant between-digester type variances were caused by within-digester type variance among replicates (Table 2.1). BETADISPER tests revealed that within-digester type variances were significantly different among replicates, while replicates categorized in terms of influent, tank, and effluent sampling points were not (Table 2.1). PERMANOVA and BETADISPER tests were also performed on InnerTube™ recirculating digester #2 technical replicates to determine group dispersions between pre-treatment method and among technical replicates (Table S 5). Both PERMANOVA and BETADISPER returned significant p-values ($p < 0.05$) (Table S 5).

2.6.5 Hierarchical Agglomerative Clustering of Digester Types

Hierarchical Agglomerative Clustering (HAC) reveals the similarity of microbial community compositions across digester samples by joining samples with comparable Bray-Curtis dissimilarity indices into common nodes which are progressively joined until all samples have been incorporated into a single-interconnected structure (Paliy & Shankar, 2016). HAC of digester samples revealed similar Bray-Curtis indices among sampling points within all digester types (Figure 2.13). As for clustering among replicates, most IS and CS digester replicates clustered together in their respective nodes except for 1IS and 3SC (Figure 2.13). However, CR and IR digester replicates did not consistently agglomerate into similar nodes (Figure 2.13). The 3CR replicate clustered under the same node as a 2CS digester, the 2CR replicate clustered under the 2IR node, and 1CR clustered with the majority of CS samples (Figure 2.13).

2.6.6 SIMPER Analysis of Top Dissimilarity-Contributing Taxa Between Digester Types

Similarity Percentage (SIMPER) analysis was used to calculate percent contributions to Bray-Curtis dissimilarities between digester types (Clarke, 1993). However, SIMPER analysis does not provide measures of significance between the abundances of genera. Therefore, the genera that most contribute to Bray-Curtis dissimilarities selected by SIMPER analysis were further tested for differential abundances by Kruskal-Wallis rank-sum tests (Table 2.3). The average percent dissimilarities contributed by differentially abundant genera ($p < 0.05$) were

0.167% between CR and IR, 0.53% between CR and IS, 0.37% between CS and CR, 0.63% between CS and IR, 0.93% between CS and IS, and 0.20% between IS and IR (Table 2.3). Genera of putative functional importance include methanogens such as *Methanoregula spp.* (0.2% contribution between CR and SI), *Methanothrix spp.* (1.5% contribution between CS and IR), *Methanospirillum spp.* (0.65% contribution between CS and IR) and *Methanosphaerula spp.* (0.24% between IS and IR) (Table 2.3). Other genera of interest with the greatest percent contributions throughout all pairwise comparisons include *Desulfomicrobium spp. Desulfobulbus spp.,* and *Desulfomonile spp. Faecalibacterium spp., Proteocatella spp.,* and unclassified genera under the Bacteriodetes and Synergistetes phyla (Table 2.3).

2.6.7 Differentially Abundant Genera Between Digester Type and Flow

DESeq2 analysis (Differential gene expression) analysis was used to reveal significantly enriched genera between pairwise comparisons of the two digester types, conventional and *InnerTube*™ digesters, as well as the two flow configurations, recirculating and single-pass flow ($p < 0.001$) (Figure 2.14). Between conventional and *InnerTube*™ digester types, there were a total of 12 differentially abundant genera (Figure 2.14, Table S 6). Log_2-fold changes range from -5.068 in conventional systems to 2.987 in *InnerTube*™ systems (Table S 6). Genera of putative functional importance include a methanogen enriched in conventional systems, *Methanothrix spp.,* as well as *Desulfovibrio spp.* highly enriched in *InnerTube*™ systems and *Desulfobulbus spp.* and *Desulfomicrobium spp.* enriched in conventional systems (Figure 2.14, Table S 6). Other genera of

functional interest include *Acidovorax spp.*, *Faecalibacterium spp.* and *Paracoccus spp.* all enriched in *InnerTube*™ systems and *Pelolinea spp.* highly enriched in conventional systems (Figure 2.14).

Between recirculating and single-pass flow configurations, there were a total of 13 differentially abundant genera, with Log_2-Fold changes ranging from -5.25 in recirculating digesters and 5.44 in single-pass digesters (Figure 2.14). Genera of putative functional importance include another methanogen, *Methanospaherula spp.*, as well as *Acidithiobacillus spp.*, *Dechloromonas spp.*, *Sulfurimonas spp.*, an unknown genus belonging to the Candidate phyla Cloacimonetes, *Geobacter spp.*, and *Pseudoarcobacter spp.* (Figure 2.14). In single-pass digesters, genera of putative functional importance include *Simplicispira spp.*, *Pleomorphomonas spp.*, *Phenylobacterium spp.*, and an unclassified genus under the order Rhizobiales (Figure 2.14, 2.16).

2.6.8 Spearman Rank Correlations of Enriched Genera to Chemical Parameters

Spearman rank-sum correlations were generated between chemical parameters and differentially abundant genera to elucidate the influence of chemical conditions on genera abundances. In conventional digesters, *Pelolinea spp.*, *Methanothrix spp.*, and *Desulfobulbus spp.* were all significantly correlated with one or more of the biomass degradation indicators, TKN, NH_3, TSS, and COD (Figure 2.15). *Desulfobulbus spp.* and *Methanothrix spp.* were both negatively correlated with TKN and NH_3, with *Desulfobulbus spp.* also negatively correlated with COD while *Pelolinea spp.* was positively correlated with TSS (Figure 2.15).

In *InnerTube*™ systems, *Paracoccus spp.* was positively correlated with COD and both *Acidovorax spp.* and *Faecalibaterium spp.* were negatively correlated with TKN and NH_3 (Figure 2.15). Of note, *Acidovorax spp.* was also positively correlated with pH and temperature (Figure 2.15).

Correlations of digester flow configurations to chemical parameters demonstrated that the majority of enriched genera in recirculating systems were highly significant and negatively correlated to biomass degradation indicators TKN, NH_3, and COD and positively correlated with pH conditions (Figure 2.16). However, only *Pseudoarcobacter spp.,* and *Sulfurimonas spp.,* were negatively correlated to TSS (Figure 2.16). Single-pass enriched genera demonstrated an opposite response to the biomass degradation indicators, where *Phenylobacterium spp.,* and the unclassified Rhizobiales were both positively correlated to TKN, NH_3, and COD, except for *Simplicispira spp.* which was positively correlated with TKN and NH_3 (Figure 2.16).

2.7 Discussion

2.7.1 Sources of Microbial Community Composition Variability Between Digester Types

Anaerobic digester community compositions are impacted by several factors such as influent sewage load (COD, TSS), digester design and configuration, and by interspecies interactions along the anaerobic digestion process (Delforno et al., 2017; Shin et al., 2016; Zhu et al., 2020). By using PERMANOVA to compare between-digester group-dispersions and

BETADISPER analyses to compare within-digester group-dispersions, it was determined that differences in microbial community compositions were brought about by changes in chemical parameters, and different digester design factors (Anderson, 2017; Jari Oksanen et al., 2019). Therefore, a significant PERMANOVA p-value in conjunction with a significant BETADISPER p-value demonstrated that microbial community dispersions across the four digester types, CR, IR, CS, and IS, were significantly different in part due to significantly dissimilar abundances across replicates ($p < 0.05$) (Table 2.1). (Anderson, 2017). In addition to dissimilar communities among replicates, the significantly different microbial community dispersions between digester groups were also caused by temperature, NH_3, and pH chemical parameters, implying that the digester chemical environment partly influenced the differences in community proportions across digester groups (Table 2.1). However, the effect of sampling point, or the influent, tank, and effluent treatment stages sampled from each digester, was not significant in both whole-group PERMANOVA and pairwise-PERMANOVA tests (Table 2.1). This suggests that the significantly different beta-dispersions of digester types were caused by inherent community differences brought about by digester design, flow, and chemical parameters.

Relative abundance, RDA, and HAC analyses also corroborated the fact that microbial community abundances between digester groups highly differed among replicates than among sampling points (Figure 2.7**Error! Reference source not found.**, 2.8, 2.11, 2.13). The primary axis in RDA analysis showed greater percent

variance than the secondary axis, and digester types, particularly *InnerTube*™ single-pass, were more separated along the primary axis than the secondary axis (Figure 2.11). In contrast, replicates and sampling points were more separated along the secondary axis (Figure 2.11). This supports the conclusion that microbial communities differed between digester types revealed by PERMANOVA analysis. NMDS also demonstrated the distinct separation of *InnerTube*™ single-pass digesters from all other digester types (Figure 2.12). Additionally, HAC analysis demonstrated higher degrees of dissimilarities between digester types and replicates compared to sampling points (Figure 2.13). The high degree of dissimilarities among replicate digesters prompts the need for a higher sample size as confirmed by multSE analysis, where at least 6 replicates are required for IR and 7 for CS, IS, and CR digesters to minimize standard errors across replicates (Figure 2.10). Technical replicates of one *InnerTube*™ single-pass digester were also sampled and analyzed to determine if the sampling protocol affects the microbial community variability (Figure 2.18, Table S 5). The pre-treatment method seemed to contribute more to microbial community variabilities compared to the sampling protocol used, as technical replicates clustered more closely together than across pre-treatment methods and the PERMANOVA test across pre-treatment methods was significant (Figure 2.18, Table S 5). However, a significant BETADISPER among technical replicates suggests that although the samples that were not preserved as studied for this project clustered closely together, technical replicates

of preserved samples were significantly different in microbial community composition (Figure 2.18, Table S 5).

Microbial communities in wastewater systems are inherently heterogeneous. Even within a digester system actively mixing microbial biomass and sewage, communities can differ from granule to granule (Kuroda et al., 2016). Therefore, technical and biological replicates are required in metagenomic studies to account for variability from sampling and bioinformatic analysis protocols and microbial community heterogeneity (Knight et al., 2012). Previous replicated metagenomic studies of anaerobic digesters revealed that microbial community proportions were similar across replicates given that influent substrate type, digester design, and operational parameters such as HRT, and organic loading rate (OLR), which is the amount of biodegradable compounds measured in COD/L passing through a digester per day, were consistent across replicates (Vanwonterghem et al., 2014). However, another study decreasing temperature and the solids retention time (SRT), which refers to the residence time of the sludge portion of waste held within the septic tank, among replicated digesters, caused community shifts from acetoclastic to hydrogenotrophic-dominated methanogenesis, as well as an enrichment of Clostridium, Fibrobacter, and Ruminococcus genera (Vanwonterghem et al., 2015). The change in temperature and SRT accounted for at least 50% of the variability in the principal component analysis (PCA) across all samples (Vanwonterghem et al., 2015). Another metagenomic survey of 20 full-scale anaerobic digesters with different designs and

agricultural, animal waste, and wastewater substrate types revealed a core microbial consortium common in all of the digesters that significantly differed in proportions in each digester (Calusinska et al., 2018). Therefore, it is reasonable to conclude that the significant differences in microbial community proportions were caused by different operational and chemical parameters such as HRT, temperature, pH, and NH_3 rather than differences brought about by community changes across the digestion process from influent, tank to effluent sampling points. Microbial community shifts from influent, tank to effluent sampling points are discussed in the following chapters.

2.7.2 Community Differences Between Digester Types Were Characterized by Syntrophic Acetoclastic and Hydrogenotrophic Methanogens

Changes in the proportions of specific microbial members in communities have been previously linked to changes in the chemical levels such as pH and COD imposed by different digester operational parameters (Jo et al., 2018; Zhu et al., 2020). An enrichment of a microbial consortium is indicative of niche functional specializations by the enriched microbes (Rivett & Bell, 2018). Relative abundance graphs and SIMPER/Kruskal-Wallis analysis corroborated the results from differential abundance analysis showing the enrichment of the methanogen *Methanothrix spp.* as well as *Desulfobulbus spp.*, *Desulfomicrobium spp.*, *Pelolinea spp.*, and *Protecatella spp.* in conventional systems compared to *InnerTube*™ systems (Figure 2.14; Order, Family and Genus ranks of enriched microbes are shown in Table S 6). Comparative studies of microbial communities between

replicated PFR digesters and conventional septic tank digesters have yet to be performed. However, in terms of chemical degradation performance, PFR-based digesters exhibited higher COD and TSS removal and methane production rates than other commonly used digester designs such as the CSTR digester (Yue et al., 2011) and based on in-house testing by WBS, conventional septic tanks (Jowett, 2007). From these previous findings, it was hypothesized that *InnerTube*™ digester types would contain higher abundances of methanogenic microbes as higher COD and total solids removal rates have been associated with higher methane production rates from enriched methanogen populations (Liu et al., 2016; Shin et al., 2016). Higher COD and total solids removal rates in *InnerTube*™ digesters compared to conventional digesters were also observed in this study (Figure 2.17). However, genera abundances showed enrichment of *Methanothrix spp.* in conventional systems (Figure 2.14) even though *InnerTube*™ systems achieved higher COD and TSS removal rates (Figure 2.17). *Methanothrix spp.*, under the archaeal family Methanosaetaceae, is an acetoclastic methanogen, utilizing exclusively acetate to produce methane (Ziels et al., 2019) and their enrichment has been found to be syntrophically linked with the presence of propionate-oxidizing acetogens such as *Desulfobulbus spp.* and *Syntrophobacter spp.* (Ying Li et al., 2018; Wang et al., 2019). Differential abundance and SIMPER/Kruskal-Wallis results support this previously reported syntrophic relationship between *Methanothrix spp.* and propionate-oxidizing acetogens since *Desulfobulbus spp.* was enriched alongside *Methanothrix spp.* in conventional systems compared to *InnerTube*™ systems.

Furthermore, Spearman rank-correlations demonstrated a negative relationship between COD, TKN, and NH_3 levels towards *Methanothrix spp.* and *Desulfobulbus spp.* abundances, suggesting that both syntrophs were influential in biomass degradation in the form of COD removal and were potentially sensitive to high nitrogen content in the form of ammonia (Li et al., 2016; C. Zhang et al., 2018). In contrast, although differential abundance of *InnerTube*™ systems did not reveal any significantly enriched methanogens, SIMPER analysis showed an enrichment of *Methanospirillum spp.* under the Methanospirillaceae family, which in contrast to *Methanothrix spp.*, is a hydrogenotrophic methanogen, utilizing hydrogen and formate as sources of electrons for CO_2 reduction to methane (Gunsalus et al., 2016). Other hydrogenotrophic methanogen families, such as Methanobacteriaceae and Methanosarcinaceae were also present in higher relative abundances in *InnerTube*™ systems (Figure 2.8). The enrichment of hydrogenotrophs such as *Methanospirillum spp.* suggests elevated levels of acidic compounds in *InnerTube*™ systems, since hydrogenotrophic methanogens are exclusive H_2-scavengers in acidic environments (Fontana, Campanaro, et al., 2018; Han et al., 2019; Horn et al., 2003). *Desulfovibrio spp.* was also enriched in *InnerTube*™ systems which is a sulfate-reducing bacteria (SRB) that can utilize sulfate (SO_4^{2-}) and using hydrogen gas, formate, lactate, or short-chain alcohols as electron donors to produce hydrogen sulfide (H_2S) (Chen et al., 2019; Keller et al., 2014). SRB's such as *Desulfovibrio spp.* have been syntrophically linked with hydrogenotrophic methanogens, whereby H_2 was reported to be consumed and exchanged between

Desulfovibrio spp. and hydrogenotrophic methanogens such as *Methanospirillum* and *Methanococcus.* (Chen et al., 2019; Meyer et al., 2013; Ozuolmez et al., 2015). Other genera of interest enriched in *InnerTube*™ systems include *Faecalibacterium spp.,* a gut-microbe that consumes acetate to produce the VFA butyrate that has been found in digester sludge treating human waste (Ju et al., 2016; Khan et al., 2012) and *Acidovorax spp.,* an aerobic, denitrifying and nitrate-reducing bacteria that can grow anaerobically using NO_3^- and NO_2^-, which possibly explains its negative correlation with total nitrogen content (TKN) and NH_3 (Heylen et al., 2008; Martins et al., 2020). It was hypothesized that *InnerTube*™ systems would contain higher overall methanogen abundances than conventional systems. However, relative abundances demonstrated lower average Euryarchaeota abundances in *InnerTube*™ digesters. Yet, *InnerTube*™ systems were characterized by higher hydrogenotrophic methanogens abundances along with sulfate-reducing microbes despite containing lower average methanogen abundances, while conventional digesters harbored greater syntrophic propionate-oxidizing acetogen and acetoclastic methanogen populations. Higher COD and TSS removal rates in *InnerTube*™ might not be caused by the abundances of only a single phylotype such as methanogens, but by specific unidentified interspecies interactions across multiple microbial subgroups. Previous metagenomic studies reported, where rare co-occurring phylotypes resulted in greater abundances of niche functional features such as degrading substances (Rivett & Bell, 2018). Additionally, higher degrees of interactions within acidogenic and methanogenic

phylotypes have been reported in anaerobic digesters (Wu et al., 2020). Therefore, although unique consortia were found to be differentially enriched between conventional and *InnerTube*™ systems, the direct reason for higher waste degradation rates in *InnerTube*™ requires further investigation.

The significant influence of chemical parameters on microbial community composition can also be observed since the separation of IS digesters in the RDA ordination was influenced by NH_3, TKN, COD, and TSS levels (Figure 2.11, Table 2.2). The separation across IS replicates could be since IS digesters were the only digesters that were geographically spread the farthest apart to each other and not clustered within a single housing complex (Figure 2.4, Table S 1). This resulted in different influent COD, NH_3, TSS, TKN levels across IS replicates (Table S 1).

2.7.3 Recirculating Systems were Characterized by Sulfur-driven Denitrifying Populations

Implementing downstream aerobic recirculating lines in anaerobic systems facilitates the anaerobic digestion process by replenishing NO_3^- for consumption by nitrate-reducing (denitrifying) microbes, preventing ammonia inhibition through accumulation, and providing an acidic buffering capacity by ensuring the constant use and flow of H_2 by nitrate-reducing microbes (Akizuki et al., 2016; Gulhane et al., 2017; Wang & Chu, 2016; Zhang et al., 2020). It was hypothesized then, that recirculating digesters in this study would contain higher denitrifying populations than single-pass systems. Nitrate was not measured in this project, but organic nitrogen levels (TKN) was recorded, which measures the nitrogen content found in

proteins, amino acids, and other nitrogenous organic compounds (Rice et al., 2012). A removal of TKN thus indicates the degradation of organic nitrogenous compounds such as proteins and amino acids (Novak et al., 2011). Recirculating systems were enriched with a group of microbes that were highly negatively correlated with TKN, NH_3, TSS, and COD while enriched genera in single-pass systems exhibited an opposite trend (Figure 2.16). The majority of microbes enriched in recirculating systems have putative sulfidogenic (sulfur-oxidation and reduction) and denitrifying roles previously discovered to be facilitated by aerobic effluent recirculation (Qiu et al., 2020). *Pseudoarcobacter spp.*, recently renamed from *Arcobacter spp.* (Perez-Cataluna et al., 2018), *and Sulfurimonas spp.* are anaerobic sulfur-oxidizing bacteria (SOB) by coupling nitrate-reduction processes as an electron acceptor (C. Chen et al., 2017; Kobayashi et al., 2012). *Geobacter spp.* is a sulfur-reducer (SRB), metabolizing elemental sulfur to sulfide (Qiu et al., 2020) and have also been found to be syntrophs with hydrogenotrophic methanogens by fermenting acetate to form H_2 for methanogens to consume (Oyiwona et al., 2017; Yee et al., 2019). Members under the candidate phylum Cloacimonetes is prevalent in many anaerobic digesters and has been found to be another syntroph producing H_2 and CO_2 from formate (Solli et al., 2014). *Methanosphaerula spp.,* is a hydrogenotrophic methanogen (Buettner et al., 2019) and *Acidothiobacillus spp.* is another SOB which produces CO_2 (Wang et al., 2018). Members of the Bacteriodales order has been previously reported to participate in carbohydrate hydrolysis in anaerobic digestion (Xia et al., 2018). The

negative relationships these microbes exhibited towards biomass levels in terms of TKN, NH_3, TSS, and COD suggest that these microbes play a significant role in the digestion performance of recirculating digesters. Since removal of organic nitrogen removal (TKN) has been reported to be performed by denitrifying microbes (Novak et al., 2011), the enrichment of sulfur-driven denitrifying microbes in recirculating systems gives credence to the fact that recirculation resulted in greater organic nitrogen metabolic activity by these microbes. possibly due to the function of the recirculating line supplying nitrate into the digester (Novak et al., 2011; Qiu et al., 2020).

Furthermore, SIMPER analysis revealed significantly enriched methanogens in recirculating digesters that were not shown in **DESeq2** analysis such as *Methanoregula spp.,* and *Methanospirllum spp.* (Table 2.3). Both are acid-tolerant hydrogenotrophs that have been found in anaerobic digesters (Jain et al., 2020; Zhang et al., 2019). The prevalence of sulfidogenic and nitrogenic microbes producing H_2 for consumption by hydrogenotrophic methanogens demonstrates that one of the means of digestion in recirculating systems were possibly sulfur-driven, denitrification processes coupled with hydrogenotrophic methanogenesis (Saia et al., 2016). The prevalence of microbes in syntrophy with hydrogenotrophic methanogens prevents the buildup of H_2 and thus a thermodynamically unfavorable environment for acetogenesis and methanogenesis to occur (Regueiro et al., 2015). The only enriched microbe in recirculating digesters that have not been implicated in sulfidogenic or nitrogenic reduction/oxidation functions was *Dechloromonas*

spp. which is a perchlorate reducer and has been found to participate in the degradation of polyhydroxyalkanoates (PHA) in anaerobic digesters (Mei et al., 2016). Overall, the significant dissimilarities between microbial communities were due to digester design differences and flow configurations. As for single-pass enriched microbes, *Simplicispira spp.* is another nitrate-reducing microbe (Chu & Wang, 2017) suggesting that nitrate processing also occurred in single-pass systems. However, the roles of other microbes such as *Phenylobacterium spp.* and *Pusillimonas spp.*, enriched in the single-pass have yet to be studied with regards to their roles in anaerobic digestion.

Table 2.1: A.) Permutational Analysis of Variance (PERMANOVA) and B.) permutation-based test of multivariate homogeneity of group dispersions (BETADISPER) of Bray-Curtis dissimilarities for species and functional genes of digester types and chemical parameters.

A.

	Taxonomic (Species)					Functional Genes				
Terms[2]	d*f*	Sum of Squares	Pseudo *f*	R2	*Adj. p*[1]	d*f*	Sum of Squares	Pseudo *f*	R2	*Adj. p*[1]
Digester	3	0.810	2.496	0.210	**0.001**	3	0.082	2.420	0.203	**0.001**
Sampling Point	2	0.182	0.839	0.047	0.782	2	0.018	0.810	0.045	0.701
Digester:Sampling Point[3]	6	0.264	0.407	0.069	1.000	6	0.033	0.488	0.082	1.000
Temperature	1	0.150	1.775	0.039	**0.029**	1	0.042	4.860	0.103	**0.001**
pH	1	0.199	2.365	0.052	**0.003**	1	0.021	2.397	0.051	**0.013**
DO	1	0.1134	1.346	0.029	0.164	1	0.019	2.257	0.048	**0.031**
TSS	1	0.126	1.489	0.033	0.098	1	0.011	1.257	0.027	0.237
COD	1	0.109	1.291	0.028	0.189	1	0.008	0.892	0.019	0.520
NH3	1	0.187	2.220	0.049	**0.007**	1	0.009	1.087	0.023	0.352
TKN	1	0.074	0.874	0.019	0.527	1	0.007	0.760	0.016	0.633

B.

	Terms	d*f*	Sum of Squares	Pseudo *f*	*Adj. p*[1]
Taxonomic (Species)	Digester	3	0.077	16.577	**0.001**
	Sampling Point	2	0.005	0.523	0.638
Functional Genes	Digester	3	0.009	10.497	**0.001**
	Sampling Point	2	0.001	0.539	0.606

[1]: Significant terms are bolded ($p < 0.05$).
[2]: Digester and Sampling Point comparisons were performed sequentially (Type I sequential Sum of Squares) while chemical parameter comparisons were performed marginally (Type III marginal Sum of Squares).
[3]: Colon represents the interaction between Digester and Sampling Point

Table 2.2: Pairwise Permutational Analysis of Variance (PERMANOVA) comparisons between Digester Types

		Taxonomic (Species)					Functional Genes				
Comparisons	Terms[2]	d*f*	Sum of Squares	Pseudo *f*	R2	*Adj. p*[1]	d*f*	Sum of Squares	Pseudo *f*	R2	*Adj. p*[1]
Conventional Single-pass vs. Conventional Recirculating	Digester	1	0.182	2.074	0.130	**0.022**	1	0.021	2.188	0.138	**0.034**
	Sampling Point	2	0.100	0.571	0.071	0.971	2	0.010	0.531	0.067	0.927
	Digester: Sampling Point	2	0.066	0.379	0.047	1.000	2	0.006	0.293	0.037	0.997
Conventional Single-pass vs. InnerTube Single-Pass	Digester	1	0.333	2.564	0.152	**0.008**	1	0.032	2.178	0.135	**0.029**
	Sampling Point	2	0.190	0.734	0.087	0.853	2	0.018	0.613	0.076	0.915
	Digester: Sampling Point	2	0.112	0.432	0.051	0.999	2	0.011	0.379	0.047	0.997
Conventional Single-Pass vs. InnerTube Recirculating	Digester	1	0.237	2.553	0.151	**0.011**	1	0.024	0.004	0.141	**0.028**
	Sampling Point	2	0.143	0.773	0.091	0.784	2	0.014	0.715	0.086	0.777
	Digester: Sampling Point	2	0.075	0.404	0.048	1.000	2	0.008	0.414	0.050	0.989
Conventional Recirculating vs. InnerTube Single-Pass	Digester	1	0.375	3.026	0.179	**0.001**	1	0.042	3.192	0.190	**0.008**
	Sampling Point	2	0.123	0.496	0.059	0.989	2	0.013	0.496	0.059	0.967
	Digester: Sampling Point	2	0.105	0.423	0.050	0.999	2	0.009	0.330	0.039	0.998
Conventional Recirculating vs. InnerTube Recirculating	Digester	1	0.191	2.200	0.139	**0.020**	1	0.014	1.588	0.103	0.146
	Sampling Point	2	0.083	0.481	0.061	0.985	2	0.008	0.476	0.062	0.958
	Digester: Sampling Point	2	0.060	0.345	0.044	1.000	2	0.007	0.421	0.055	0.987
InnerTube Single-Pass vs. InnerTube Recirculating	Digester	1	0.305	2.364	0.143	**0.014**	1	0.025	1.851	0.116	0.094
	Sampling Point	2	0.169	0.656	0.079	0.930	2	0.017	0.629	0.079	0.847
	Digester: Sampling Point	2	0.110	0.428	0.052	0.999	2	0.011	0.423	0.053	0.992

[1]: Significant terms are bolded ($p < 0.05$)
[2]: Term comparisons were performed with Type I sequential Sum of Squares

Table 2.3: Pairwise Comparisons of differentially abundant genera between digester types identified by SIMPER analysis and Kruskal-Wallis rank-sum significance tests

Comparison	NCBI Taxa ID	Differentially Abundant Genera	% contribution to dissimilarity[1]	Adj. *p*-value[2]	% average abundance (First digester)	% average abundance (Second digester)
Conventional Recirculating vs. InnerTube Recirculating	1168169	Methylomonas	0.204	0.049	0.260	0.028
	899	Desulfomicrobium	0.196	0.048	0.328	0.106
	51642	Nitrosomonas	0.102	0.048	0.147	0.031
Conventional Recirculating vs. InnerTube Single-Pass	1895922	Unclassified_Bacteroidetes	1.739	0.049	2.847	0.343
	894	Desulfobulbus	1.269	0.036	2.027	0.169
	1400053	Unclassified_Bacteroidales	0.947	0.036	1.566	0.184
	511	Alcaligenes	0.308	0.049	0.003	0.455
	2013734	Unclassified_Candidatus Cloacimonetes	0.267	0.049	0.398	0.009
	899	Desulfomicrobium	0.212	0.036	0.328	0.017
	1168169	Methylomonas	0.170	0.036	0.260	0.011
	2049433	Unclassified_Desulfobacteraceae	0.151	0.048	0.235	0.016
	2094242	Unclassified_Victivallales	0.142	0.036	0.243	0.034
	419257	Yersinia	0.138	0.036	0.003	0.206
Conventional Single-Pass vs. Conventional Recirculating	1400053	Unclassified_Bacteroidales	1.092	0.049	0.278	1.566
	80881	Simplicispira	0.204	0.049	0.288	0.041
	882104	Methanoregula	0.200	0.043	0.012	0.260
	1168169	Methylomonas	0.190	0.049	0.027	0.260
	2094242	Unclassified_Victivallales	0.180	0.036	0.019	0.243
Conventional Single-Pass vs. InnerTube Recirculating	181070	Proteocatella	1.580	0.036	2.17	0.079
	2223	Methanothrix	1.534	0.036	2.09	0.041

	853	Faecalibacterium	1.171	0.043	0.35	1.885
	913107	Pelolinea	0.777	0.036	1.04	0.006
	1262952	Ruminococcus	0.728	0.049	0.22	1.165
	668570	Methanospirillum	0.648	0.048	0.99	0.146
	28117	Alistipes	0.587	0.049	0.17	0.909
	648	Aeromonas	0.316	0.048	0.12	0.528
	40520	Blautia	0.204	0.049	0.21	0.462
	81412	Aminomonas	0.180	0.049	0.25	0.013
	2358	Desulfomonile	0.167	0.036	0.23	0.005
	351091	Oscillibacter	0.149	0.049	0.11	0.301
	328301	Syntrophorhabdus	0.133	0.048	0.18	0.007
Conventional Single-Pass vs. InnerTube Single-Pass	2013842	Unclassified_Synergistetes	4.373	0.048	7.212	0.756
	181070	Proteocatella	1.343	0.049	2.175	0.286
	899	Desulfomicrobium	0.957	0.036	1.449	0.017
	913107	Pelolinea	0.692	0.048	1.043	0.008
	1793	Mycolicibacterium	0.612	0.049	0.950	0.061
	1978231	Unclassified_Acidobacteria	0.505	0.049	0.787	0.040
	53461	Nakamurella	0.344	0.048	0.521	0.009
	41950	Candidatus Microthrix	0.162	0.049	0.243	0.000
	2358	Desulfomonile	0.147	0.049	0.227	0.008
	264148	Millisia	0.120	0.048	0.180	0.000
InnerTube Single-Pass vs. InnerTube Recirculating	475088	Methanosphaerula	0.237	0.049	0.022	0.357
	360807	Roseburia	0.225	0.049	0.083	0.400
	419257	Yersinia	0.140	0.036	0.206	0.003

[1]: Bray-Curtis Dissimilarity. Only Genera with greater than 0.1% contribution to dissimilarity were included.
[2]: Genera are ordered from largest to smallest % contribution. Kruskal-Wallis rank-sum test *p*-values were adjusted with the Benjamini & Hochberg Correction for false discovery rate. Only genera showing significant differential abundances at $p < 0.05$ were included.

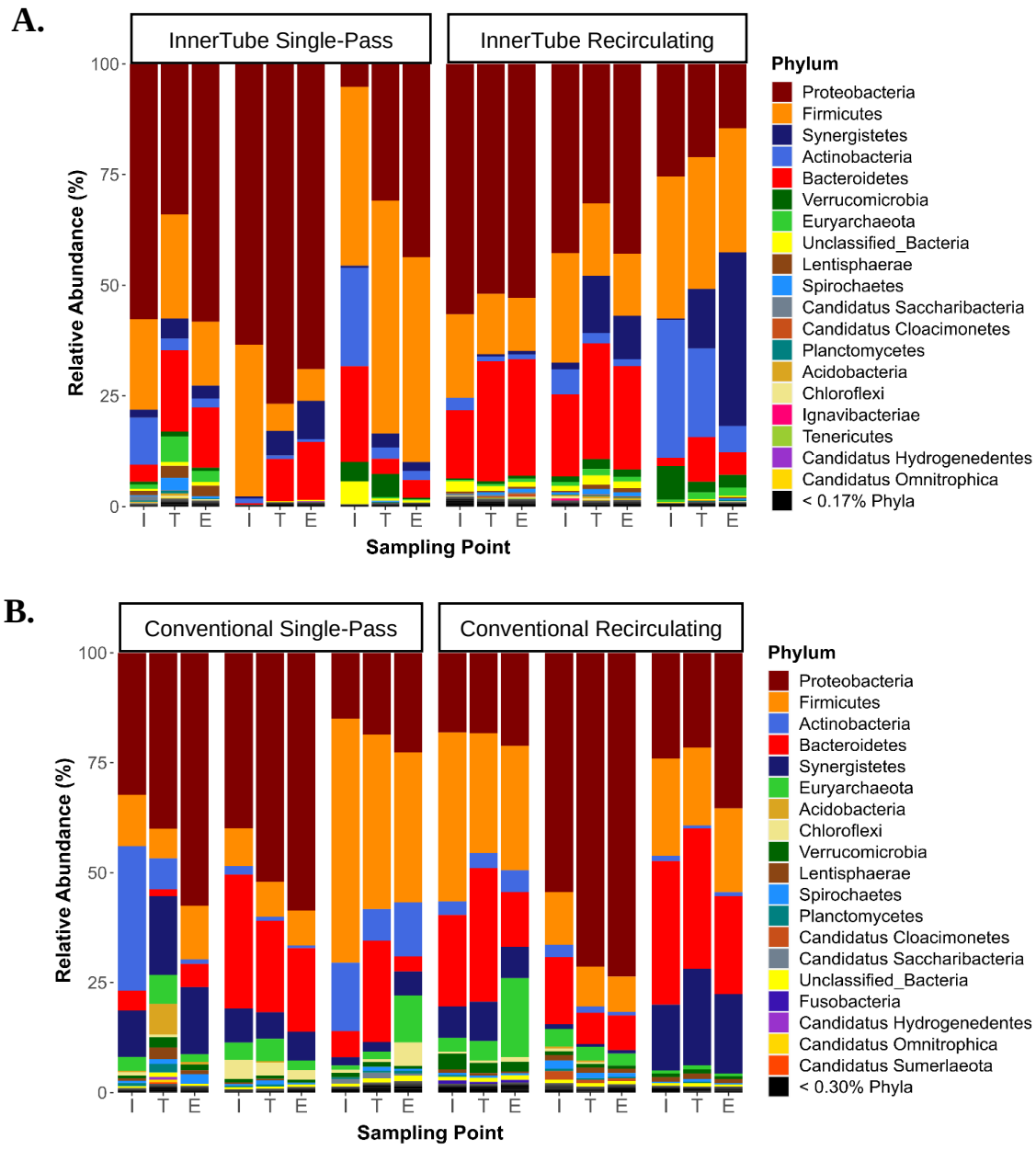


Figure 2.7: Relative abundance bar graphs showing the top 20 phyla for triplicate InnerTube Single-pass, InnerTube Recirculating, Conventional Single-Pass, and Convectional Recirculating across sampling points.

I=Influent, T=Tank, E=Effluent. Lowest abundant phyla were aggregated into a cutoff percentage labeled in black.

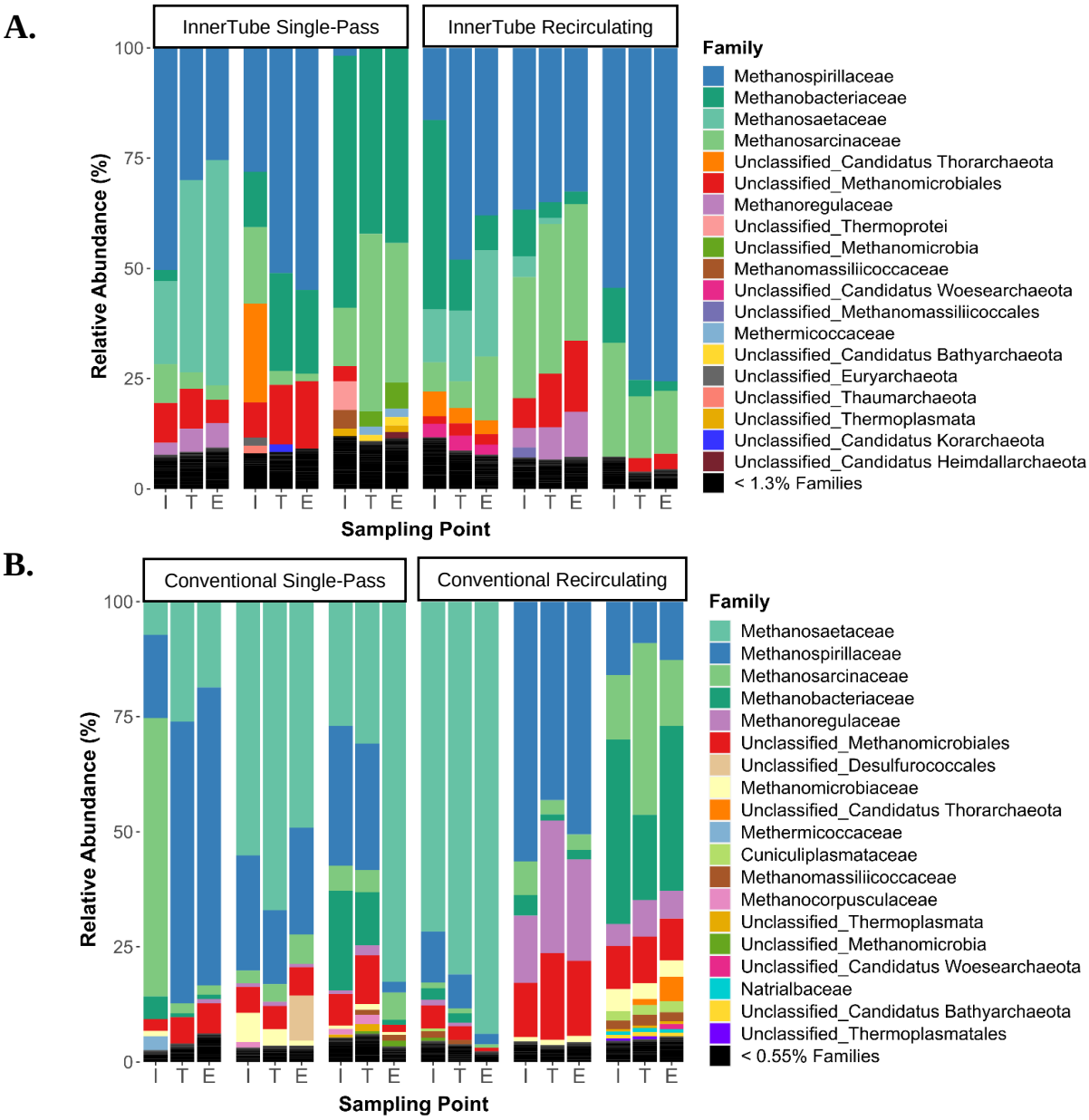


Figure 2.8: Relative abundance bar graphs showing the top 20 families under Euryarchaeota for triplicate InnerTube Single-pass, InnerTube Recirculating, Conventional Single-Pass, and Convectional Recirculating across sampling points. I=Influent, T=Tank, E=Effluent. Lowest abundant phyla were aggregated into a cutoff percentage labeled in black.

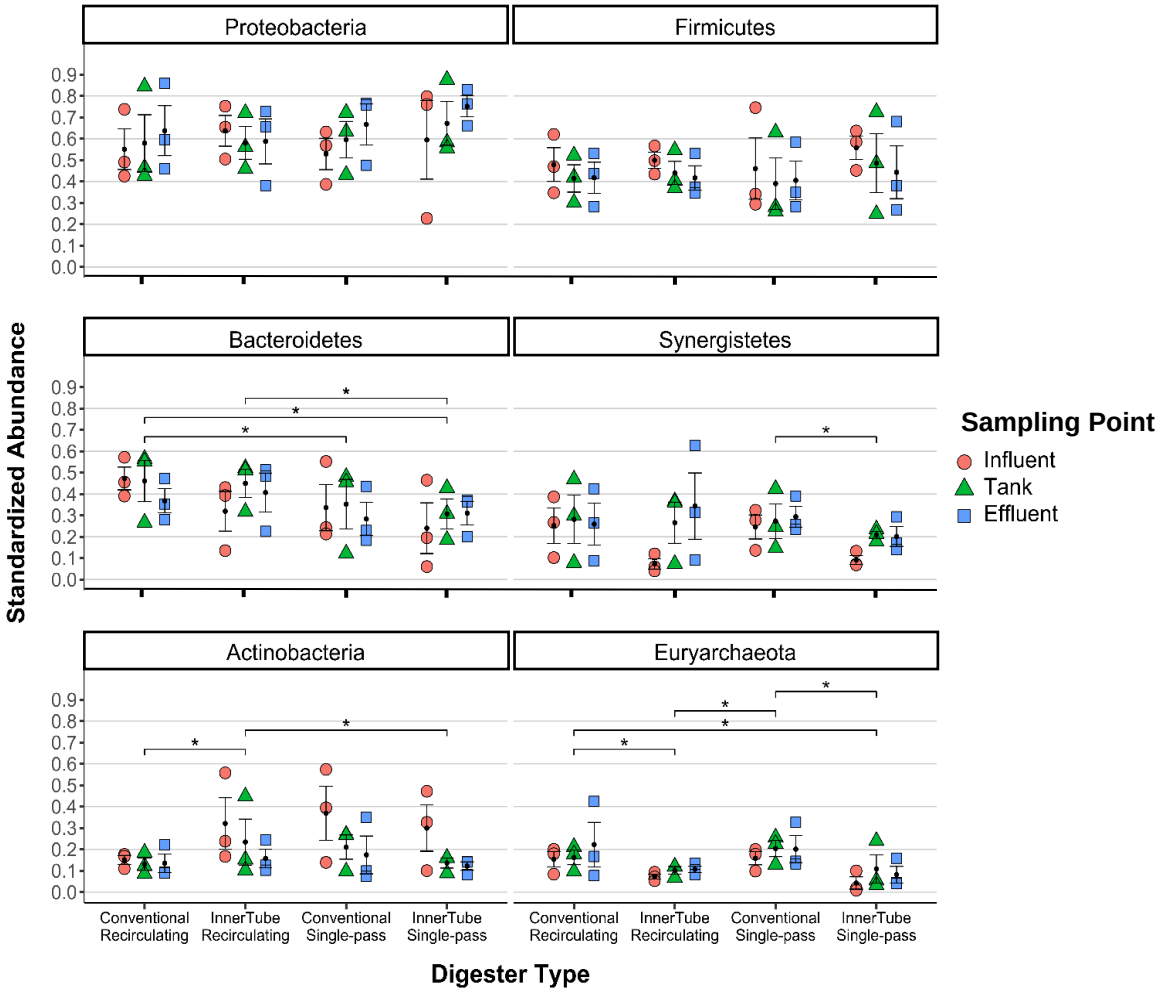


Figure 2.9: Scatterplot of the mean and standard error of the top 6 most abundant phyla across digester types and sampling points. Abundances were Hellinger-standardized. Pairwise Wilcoxon-Rank Sum Tests were performed to test significant rank differences between relative phyla abundances of digesters (*adj*. $p < 0.05$). Only significant pair-wise differences are shown.

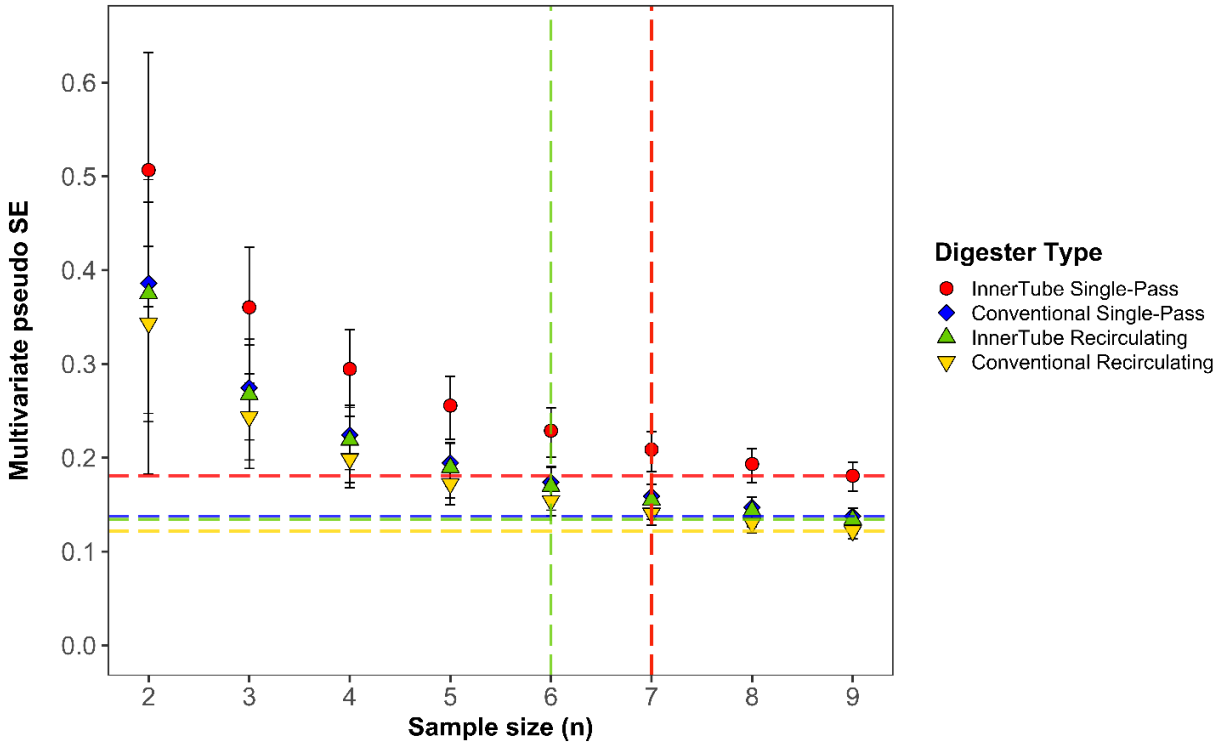


Figure 2.10: Power curve of multivariate dissimilarity-based standard error estimates (multSE) of digester types vs. minimum required sample size. Scatterplot indicates the number of replicates (*n=6* for InnerTube Recirculating, *n=7* for all other digesters) at which there is no quantitative decrease in the standard error of Bray-Curtis dissimilarities within a given digester type.

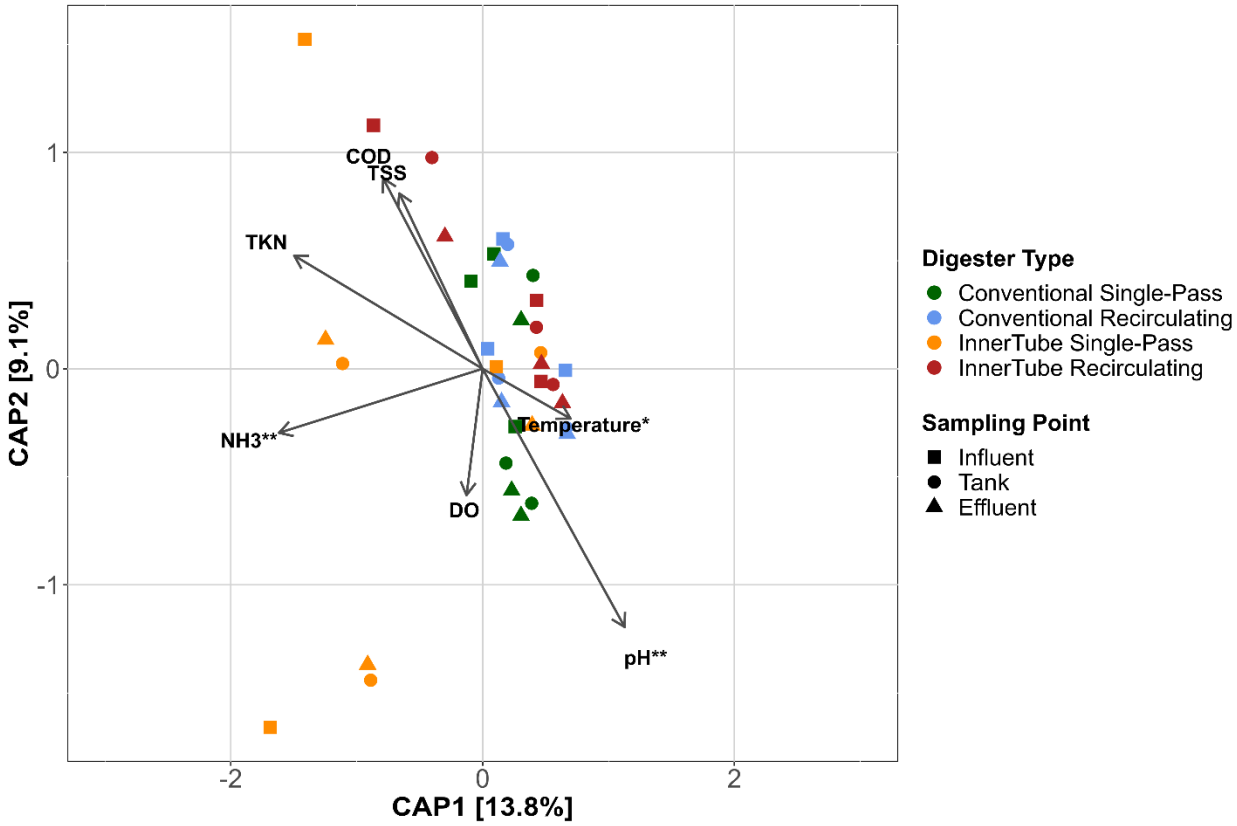


Figure 2.11: Redundancy Analysis of digester type and sampling point against chemical parameters. Chemical parameters that significantly influenced the ordination revealed by the PERMANOVA test are denoted with asterisks (*adj. $p < 0.05$*), where '***' 0.001, '**' 0.01, '*' 0.05.

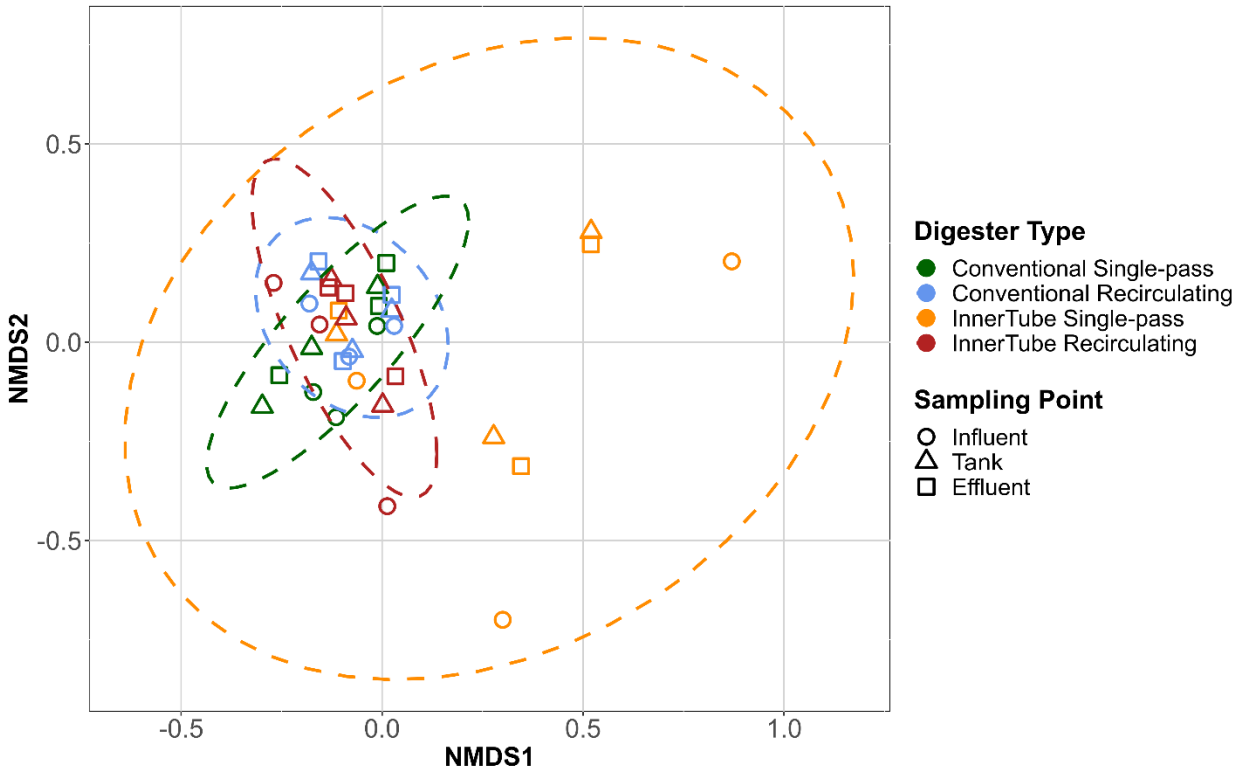


Figure 2.12: Non-metric dimensional scaling ordination of Bray-Curtis dissimilarities between digester types and sampling points.

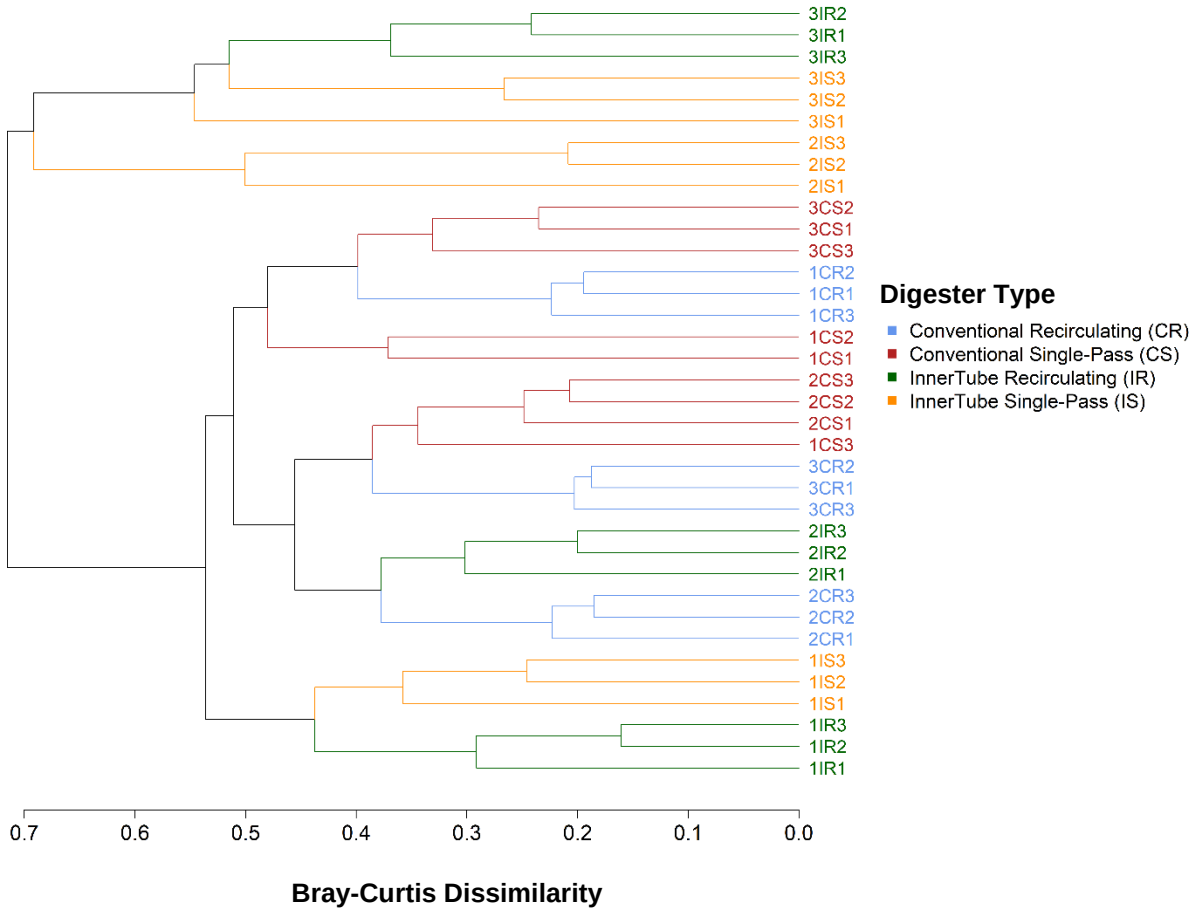


Figure 2.13: Hierarchical agglomerative clustering based on Bray-Curtis Dissimilarities of triplicate digester types across sampling points. Numbers prefixing digester labels indicate a replicate, and numbers suffixing digester labels indicate a sampling point, where 1=Influent, 2=Tank, 3=Effluent.

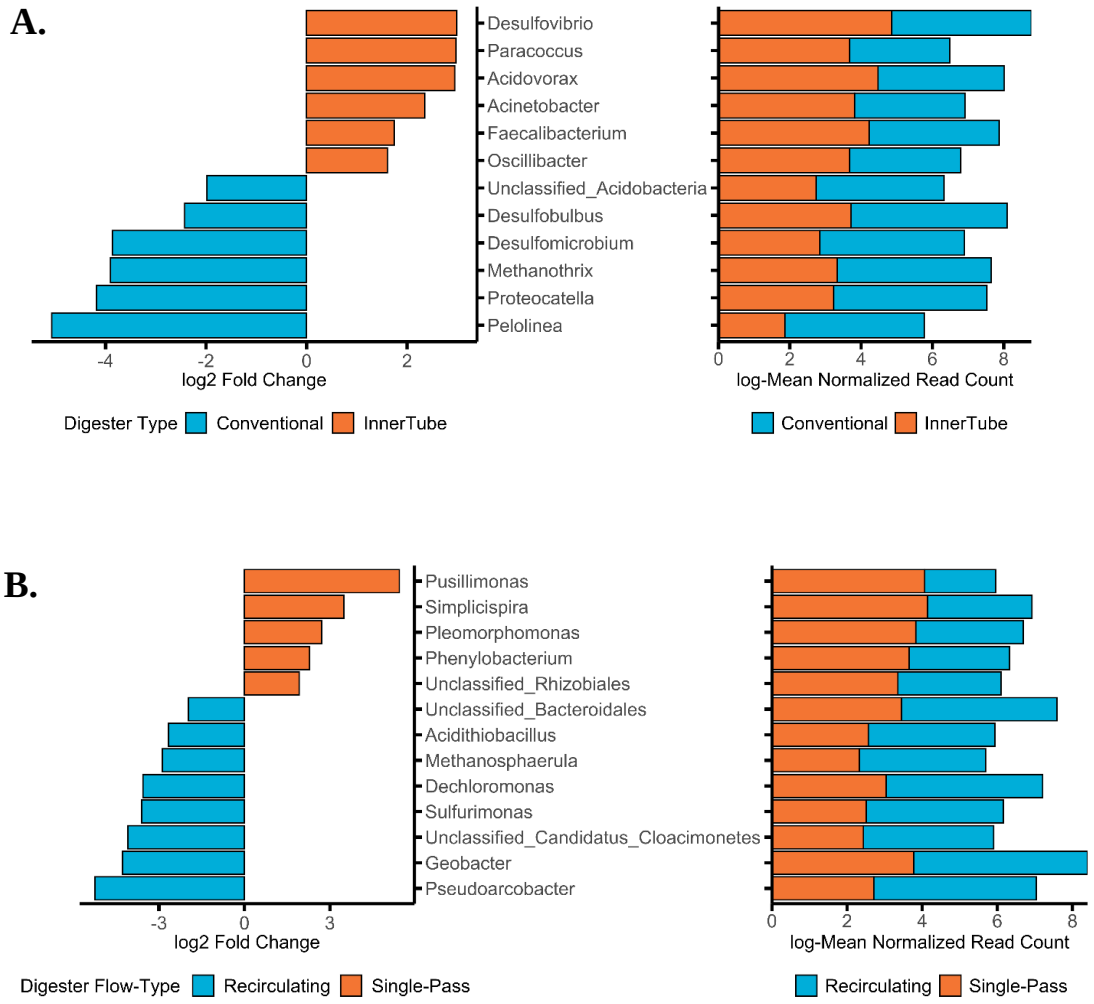


Figure 2.14: Pairwise comparisons of differentially-abundant genera among A.) Conventional and InnerTube and B.) Recirculating and Single-Pass Digester types. Differential abundance analysis was performed using **DESeq2** with an abundance cutoff of 0.1% of total agglomerated genera counts and an adjusted p-value of $p < 0.001$.

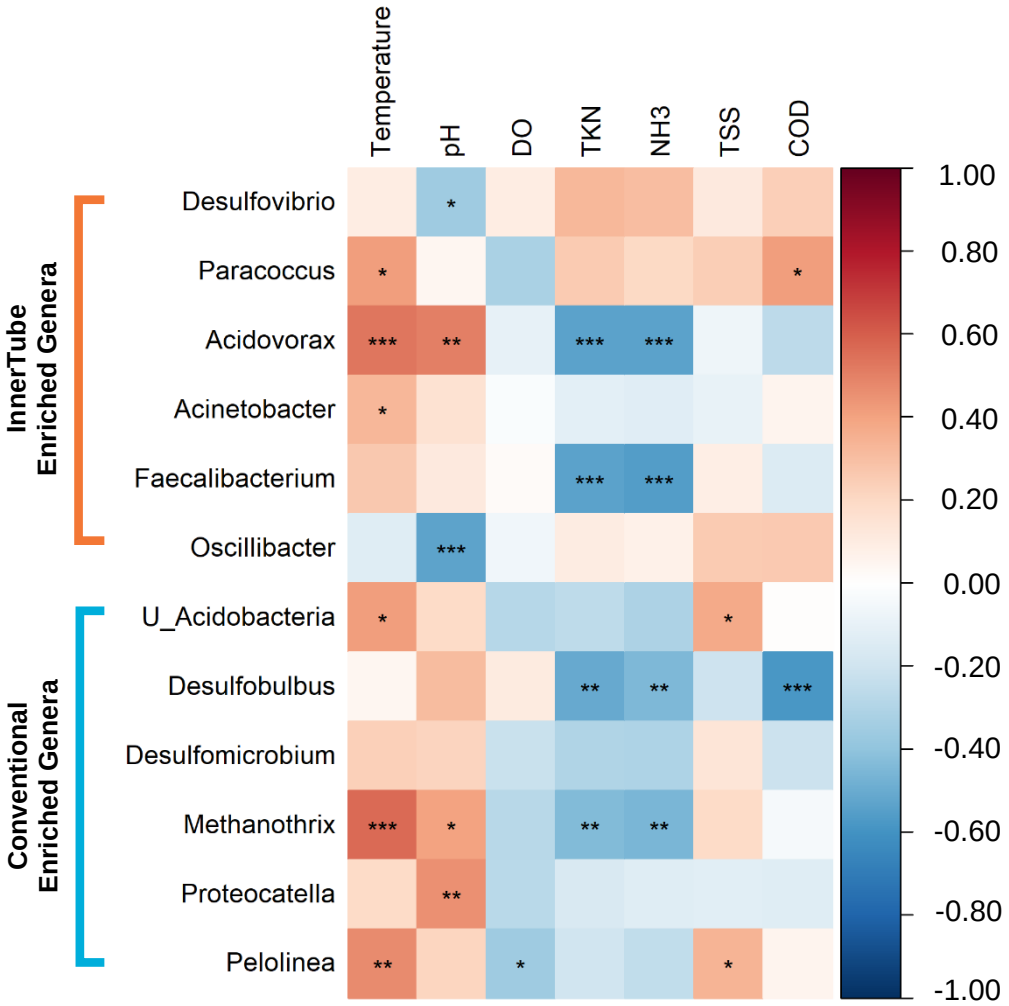


Figure 2.15: Spearman Rank Correlation heatmap of differential enriched genera between conventional and InnerTube digesters to chemical parameters. The prefix "U_" before a genera name represents an unclassified genus. Significant correlations are marked with asterisks ($p < 0.05$), where '***' 0.001, '**' 0.01, '*' 0.05.

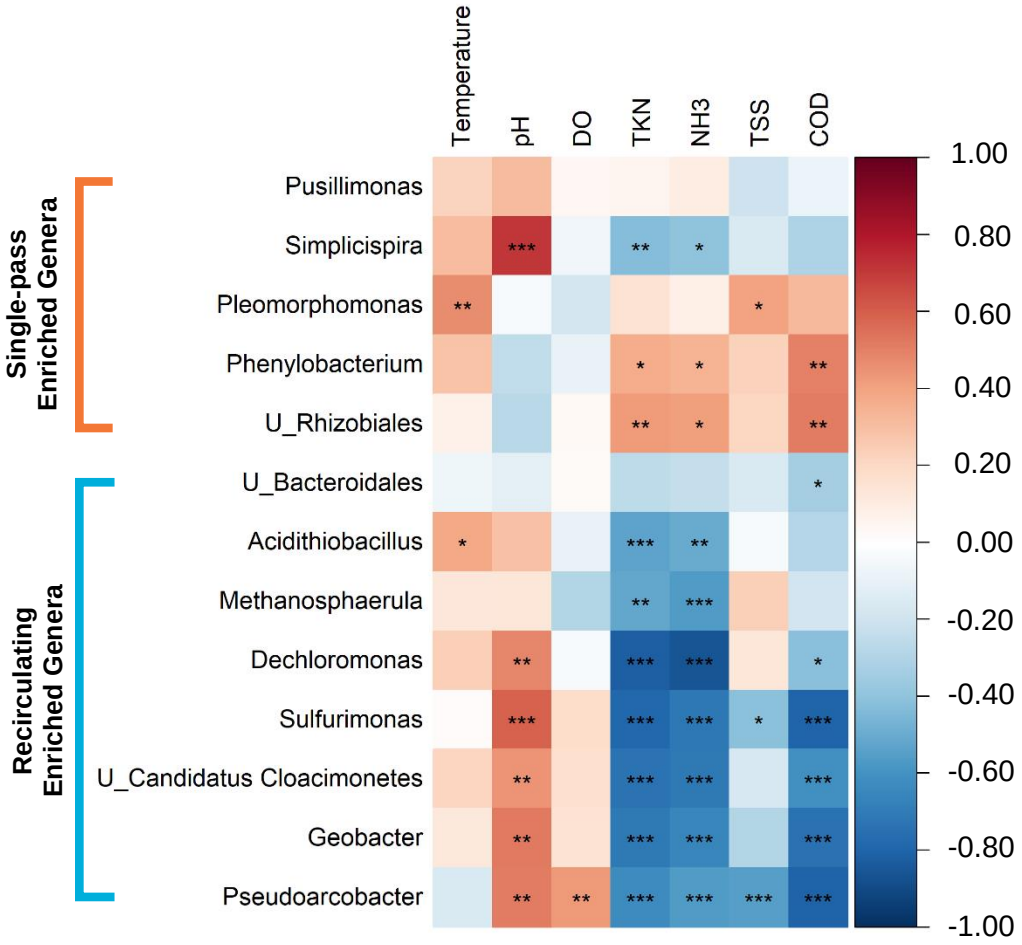


Figure 2.16: Spearman Rank Correlation heatmap of differential enriched genera between single-pass and recirculating digesters to chemical parameters. The prefix "U_" before a genera name represents an unclassified genus. Significant correlations are marked with asterisks ($p < 0.05$) , where '***' 0.001, '**' 0.01, '*' 0.05.

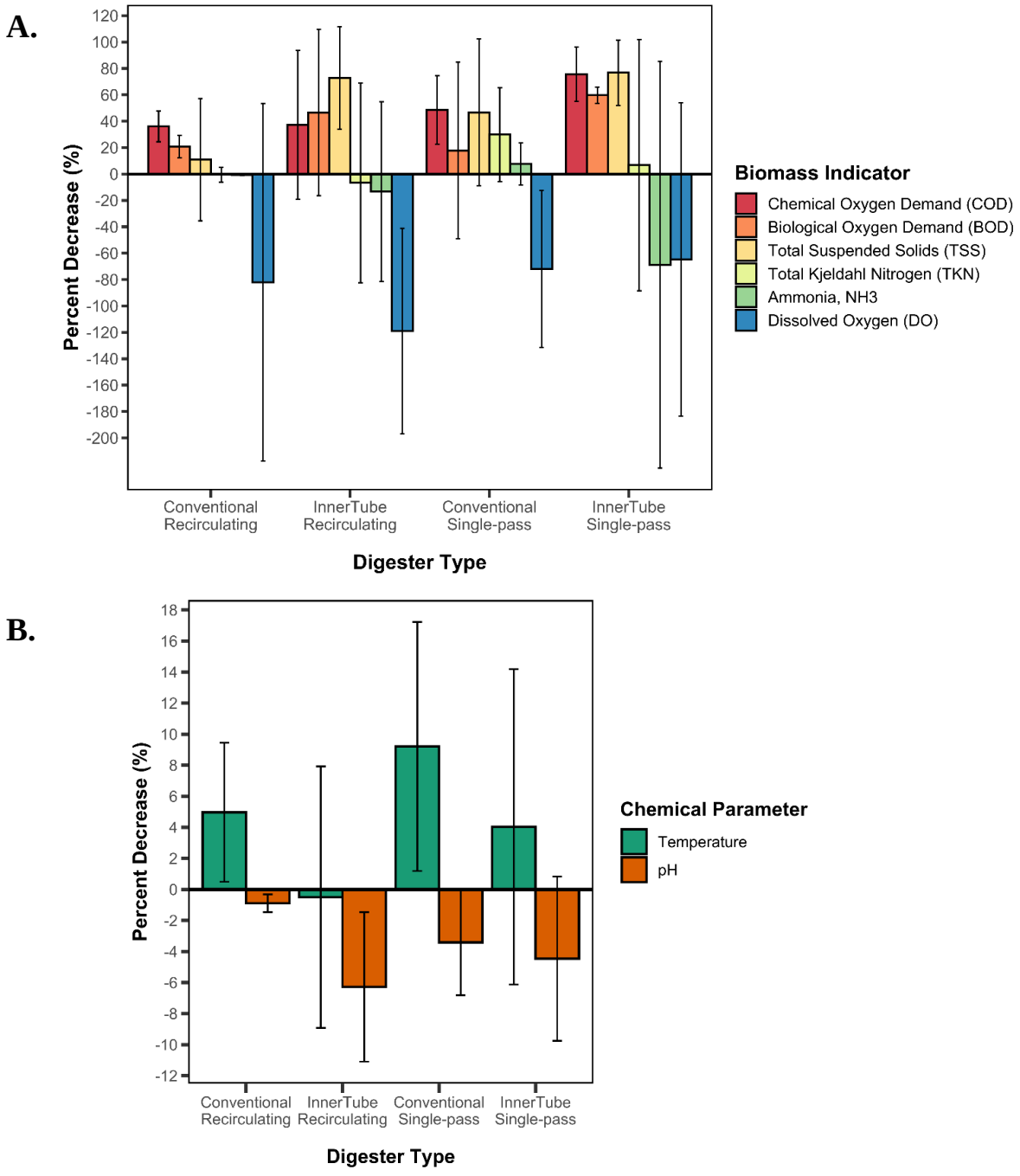


Figure 2.17: Percent decrease in chemical parameters from influent to effluent. A.) Percent decreases of chemical measures of biomass in digester types, B.) Percent decreases in temperature and pH in digester types.

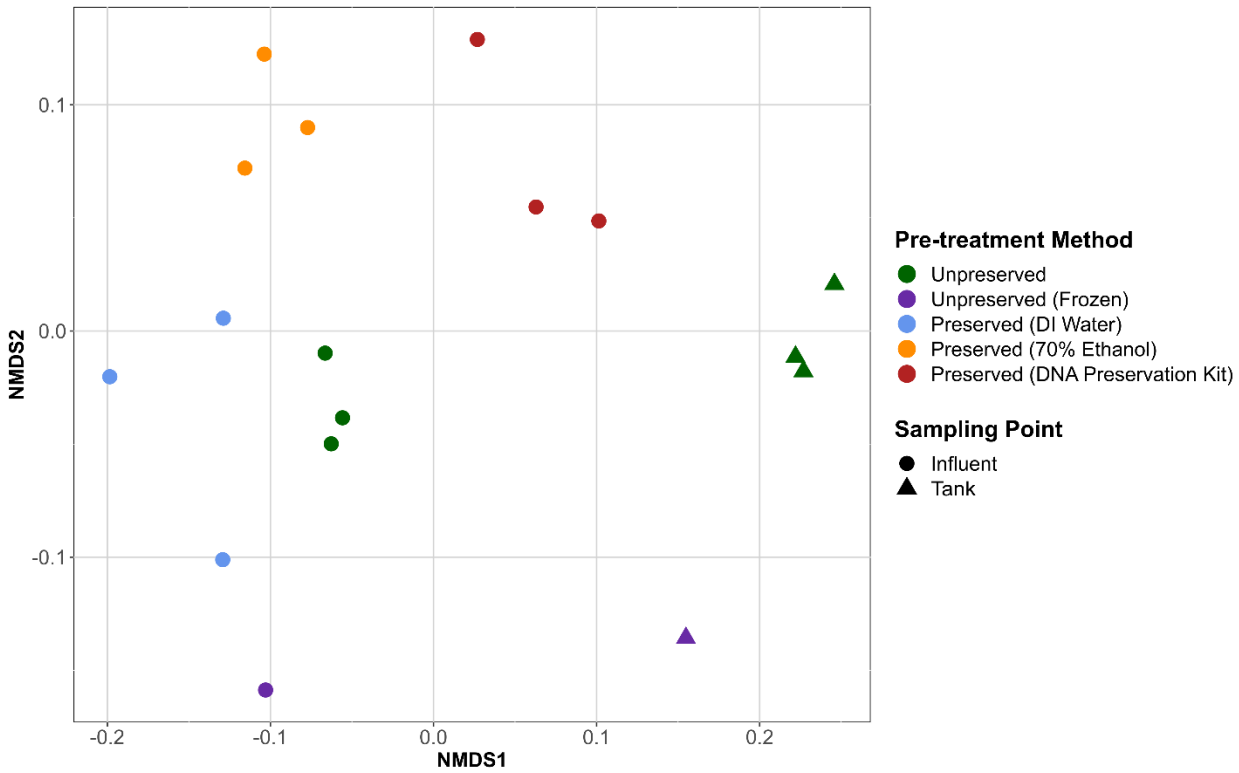


Figure 2.18: NMDS Ordination of InnerTube Recirculating untreated and pre-treated technical replicates. "Unpreserved (Frozen)" are InnerTube Recirculating influent (2RI1) and tank (2RI2) samples used in project analyses (Table S 1 to Table S 3).

3 Chapter 3: Exploring the Microbes Associated with the Sewage Treatment Process

3.1 Introduction

3.1.1 Ecological Succession of Microbes in Plug-Flow Anaerobic Digesters

Since AD is a stepwise process, it has been theorized that hydrolytic, acidogenic, acetogenic, and methanogenic communities are spatially organized longitudinally along the treatment train of anaerobic digesters, responding in growth to substrate composition and availability (De Vrieze et al., 2016; L. Dong et al., 2019; García-Lozano et al., 2019). This longitudinal stratification is especially apparent in PFR and PFR-like digesters such as the ABR due to the highly-compartmentalized flow-path of PFR digesters (Gulhane et al., 2017; Li et al., 2014). It has been observed that hydrolysis occurs at the front of the digester, acidogenesis occurs in the middle, and acetogens and methanogens gradually increase throughout the digester, growing to maximal abundance at the distal end of the digester (L. Dong et al., 2019; Jiang et al., 2018; Ruiz-Sanchez et al., 2018). Only one study has been done on the spatial distribution of microbial communities in a conventional anaerobic box-like tank treating animal waste, showing a decrease in community alpha-diversity in the middle of the tank and an increase near the end (García-Lozano et al., 2019). Other than box-tank or PFR-based digesters, studies on the spatial distribution of microbial communities have only been done a CSTR digester, which expectedly did not exhibit zoning of microbial communities due to continuous mixing (Heyer et al., 2020). There have been no metagenomic studies

done comparing the microbial stratification along the treatment process between conventional anaerobic septic tank digesters and PFR-like digesters thus far.

3.2 Chapter 3 Objectives

The objectives for this chapter was to investigate the microbial community composition across influent, tank, and effluent sewage in the four digester types. The first objective was to investigate the community diversity in each of the treatment points. The second objective was to determine if influent, tank, and effluent treatment points differed in composition at the whole-community level and the genus level. The third objective was to relate enriched microbes to chemical parameters across the treatment point. It is hypothesized that since AD is performed successionally by hydrolytic, acidogenic, acetogenic, and methanogenic microbes, these phylotypes would also be spatially organized accordingly across a digester's influent, tank, and effluent sewage treatment points, where an intermediate population in the tank linking the hydrolytic and methanogenic processes of AD is expected to reside. It has been previously reported that protein breakdown by acidogenic and acetogenic organisms occurs near the effluent outlet of a digester, distal from the influent sewage inlet (Ruiz-Sanchez et al., 2018). Overall, this project sought to elucidate the core consortia of microbes across the anaerobic digestion treatment process while taking into account the effect that the digester environment has on these microbes.

3.3 Methods

3.3.1 Statistical Analyses

Shannon and Chao1 alpha-diversity indices were calculated using phyloseq. Rank abundance graphs were generated using the ampvis2 R package (Andersen et al., 2018). Procrustes analysis on NMDS ordinations between influent, tank, and effluent samples was performed using the ade4 package and the "procuste" function (Dray & Dufour, 2007). M^2 procrustes statistics and its *p*-values were calculated using the vegan package and the "protest" function (Jari Oksanen et al., 2019). An M^2 -statistic was considered significant if a $p < 0.05$ was achieved. **DESeq2** analysis, used to reveal significantly enriched genera between Influent, Tank, and Effluent sewage, was performed with three levels of contrast comparisons (influent, tank, and effluent) for the **DESeq2** object and with a design formula of $\sim Digester\ Type + Flow\ Configuration + Sampling\ point$ (Love et al., 2014). An abundance cutoff at 0.01% of total agglomerated genera counts and an adjusted *p*-value cutoff of $p < 0.05$ was used for all sampling point pairwise comparisons. Otherwise, differential abundance analysis was performed with the same parameters discussed in Section 2.5.7. Spearman rank-sum correlations between enriched genera with chemical parameters were also performed with the same settings listed in Section 2.5.7.

3.4 Results

3.4.1 Alpha-Diversity Along the Sewage Treatment Process

The alpha-diversity Shannon, Inverse Simpson, and Chao1 indices were calculated for each digester sewage sample (Figure 3.1). The Shannon diversity index incorporates both species evenness which refers to the degree of homogeneity of species proportions, and species richness measures, which is the total number of species (Daly et al., 2018). The Chao1 index is categorized as a richness estimator based on abundance data, where richness is estimated based on the number of species that are observed either once or twice in each sample and is therefore weighted towards rare species (Daly et al., 2018). The Inverse Simpson index measures the probability that two randomly sampled individuals from an area will belong to the same species, taking into account richness and evenness and is weighted towards dominant species (Daly et al., 2018). Higher Shannon and Inverse Simpson values signify greater species diversity, while greater Chao1 values signify greater species richness (Kim et al., 2017). Recirculating systems were observed to decrease slightly in species richness and diversity from influent to effluent sampling points based on all three indices while single-pass systems peaked in diversity and richness within the tank and ultimately exhibited greater diversity and richness in the effluent compared to the influent (Figure 3.1).

Archaea and Bacteria rank abundance curves between influent, tank, and effluent sampling points were also generated to visualize archaeal and bacterial species evenness and richness trends across sampling points (Figure 3.2). The

species with the lowest read abundances are ranked at the beginning of the x-axis while the most abundant are ranked towards the tail-end of the graphs (Figure 3.2). The degree of species evenness is indicated by the slope of the line, where a steep slope indicates low community evenness and a shallow slope indicates high community evenness. Species richness is indicated on the x-axis, where the total number of ranked species are shown (Figure 3.2). For bacteria, effluent samples exhibited a slightly steeper gradient than influent samples until around 5000 species, indicating that effluent communities were less even and were dominated by a smaller number of low abundant species than influent communities (Figure 3.2). Bacterial tank samples mimicked influent evenness trends also, with the exception from 1 to 10 species, where it was similar in evenness to effluent communities (Figure 3.2). Archaeal rank abundance curves showed similar trends, where effluent samples also showed a steeper gradient until around 200 species, indicating that archaeal effluent communities were less even and were dominated by a smaller number of low abundant species than influent communities (Figure 3.2). However, archaeal tank communities exhibited a more similar slope and thus degree of evenness with effluent samples at all numbers of species compared to bacterial tank communities (Figure 3.2). At high species richness, however, influent, tank and effluent archaeal and bacterial communities exhibited the same degree of evenness as shown by the plateau in all the slopes (Figure 3.2).

3.4.2 Procrustes Analysis of Influent, Tank, and Effluent Ordinations

Procrustes analysis compares the degree of similarity between two ordination matrices by randomly rotating, translating, and dilating two ordination matrices to minimize the sum of squared residual values (M^2) between each point in at least two ordination matrices (Peres-Neto & Jackson, 2001). Smaller M^2 values of a procrustean comparison indicate more similar ordinations (Peres-Neto & Jackson, 2001). A p-value of less than 0.05 indicates significant concordance between the two ordinations (Peres-Neto & Jackson, 2001). NMDS ordinations were compared between influent and tank, tank and effluent, and influent and effluent digester communities to determine if microbial community compositions changed significantly at each successive treatment point (Figure 3.3-3.5). All three sampling point ordinations exhibited significant concordance with each other ($p < 0.05$). The M^2 value between tank and effluent ordinations ($M^2 = 0.2694$, $p = 0.001$) was the lowest out of the three comparisons (Figure 3.4).

3.4.3 Differential Abundance of Enriched Genera Across the Sewage Treatment Process

Differential abundance analysis was used to perform the following three pairwise-comparisons across the sampling points, influent vs. tank (influent-tank), tank vs. effluent (tank-effluent), and influent vs. effluent (influent-effluent) (Figure 3.6, Table S 7). While Procrustes analysis revealed whole-microbial community differences across sampling points, differential abundant analysis was used to elucidate microbial compositional changes at the genera level at each sampling

point. Differential abundance analysis revealed 18 differentially abundant genera between influent and tank samples, with log_2-fold changes ranging from -3.52 to 2.01 (Figure 3.6, Table S 7). Only two differentially abundant genera were found enriched in effluent compared to tank samples, *Halothiobacillus spp.* and *Desulfamplus spp.* (Figure 3.6, Table S 7). No genera were found to be enriched in the tank when compared to effluent sewage. (Figure 3.6, Table S 7). Influent samples were enriched with Actinobacteria including genera from *Thiothrix spp.* to *Collinsella spp.* and Firmicutes with several genera of relevance to anaerobic digestion under the Clostridiales order such as *Anaerostipes spp.*, *Ruminococcus spp.*, *Erysipelatoclostridium spp.*, and *Blautia spp.* (Figure 3.6, Table S 7). There were only three enriched genera in tank samples, *Desulfonatronum spp.*, and unclassified Verrucomicrobia and Deltaproteobacteria members (Figure 3.6, Table S 7).

In contrast to influent vs. tank and tank vs. effluent comparisons, differential abundance analysis between influent and effluent samples revealed a much higher number of enriched genera (Figure 3.7, Table S 7). A total of 58 differentially abundant genera were found, with log_2-fold changes ranging from -3.59 to 5.08 (Figure 3.7, Table S 7). Similar to influent to tank comparisons, influent-effluent comparisons a group of genera under the Clostridiales Order, which include Ruminococceae and Lachnospiraceae genera members such as *Ruminococcus spp.* and *Blautia spp.*, respectively, enriched in the (Figure 3.7, Table S 7). Additionally, a group of Rhizobiales members and a potential pathogen, *Yersinia spp.* were also

enriched in the influent (Figure 3.7, Table S 7). For effluent enriched genera, there was an abundance of notable genera as part of the Desulfobacterales, Desulfovibrionales and Desulfuromonadales and Nitrosomonadales Orders (Figure 3.7, Table S 7). There was also an enriched methanogen, *Methanomethylvorans spp.* in the effluent (Figure 3.7, Table S 7).

3.4.4 Spearman Rank Correlations between Chemical Parameters and Genera Enriched Across the Treatment Process

The aforementioned effluent enriched genera under the orders Desulfobacterales, Desulfovibrionales and Desulfuromonadales and Nitrosomonadales were all significantly negatively correlated with TKN, NH_3 and COD and positively correlated with pH chemical parameters (Figure 3.8). In influent samples, Clostridiales genera such as *Ruminococcus, spp., Lachnospira spp., Anaerostipes spp., Faecalibacterium spp.* were also significantly negatively correlated with TKN and NH_3 levels (Figure 3.8). Additionally, there was a group of influent enriched genera significantly positively correlated with TSS and COD levels (Figure 3.8).

3.5 Discussion

3.5.1 Effluent Sewage Was Enriched with Sulfur-Reducing and Methanogenic Microbes

Multiple **DESeq2** pairwise-comparisons between influent vs. tank, tank vs. effluent, and influent vs. effluent sampling points were used to identify the microbes that significantly changed in abundance across sampling points in all digesters. Differential abundance analysis revealed that influent sewage contained higher abundances of putative Clostridiales gut bacteria under the orders Lachnospiraceae and Ruminococcaceae (El Hage et al., 2019) which have also been found to perform hydrolytic and acidogenic functions in digesters (Černý et al., 2018). Differential abundance tests also revealed a group of effluent sulfur-reducing microbes, nitrifying bacteria, and a methanogen significantly correlated with biomass indicators. Effluent enriched sulfate-reducing bacteria include *Desulfamplus spp.* (Lefèvre et al., 2013), *Desulfovibrio spp.* (Keller et al., 2014), *Desulfobacterium spp* (Liu et al., 2018)., and *Desulfobacter spp.* (Liu et al., 2018). Effluent enriched elemental sulfur-reducing bacteria producing H_2S coupled with acetate and/or propionate-oxidation include *Geobacter spp.* (Qiu et al., 2020), *Desulfonatronum spp.* (Zhilina et al., 2005), *Desulfuromonas spp.* (Badalamenti et al., 2016) and *Desulfobulbus spp.* (Liu et al., 2018). The obligately methylotrophic methanogen, *Methanomethylvorans spp.* which utilizes methanol as an electron-source to produce methane, was also enriched in the effluent (Lomans et al., 1999). *Yersinia spp.*, a potential pathogen, was enriched in the influent yet reduced to low

abundances at the effluent, indicating that the digesters could remove this particular pathogen by relying on the elevated levels of toxic chemicals such as ammonia (Zhao & Liu, 2019).

PERMANOVA beta-diversity significance tests (Table 2.1), RDA and NMDS ordinations, and Procrustes analyses demonstrated that at the whole-community level, microbial compositions did not differ significantly between influent, tank, and effluent communities ($p < 0.05$). The number of differentially abundant genera enriched from influent-tank and tank-effluent sampling points corroborated the low degree of community compositional change from sampling point to sampling point (Figure 3.6, Table S 7). Influent-tank differential abundance comparisons only yielded 3 enriched genera in the tank and tank-effluent comparisons only yielded two enriched genera in the effluent (Figure 3.6). However, differential abundance analyses between influent-effluent comparisons revealed several enriched genera at each sampling point. The disparity of results between whole-community beta-diversity significance tests and differential abundance genera tests suggests that whole-community analyses were not sufficient to resolve community compositional changes across sampling points at the genus level. The notably higher number of differentially abundant genera between influent-effluent treatment points at 58 genera compared to influent to tank (18 genera) and tank to effluent enriched genera (3 genera) implies that the majority of the microbes that had significantly different abundances may have arisen at the effluent end of digesters.

These results corroborate previous digester stratification studies showing the accumulation of hydrolytic and fermentative bacteria near the beginning of a digester which converts complex polysaccharides, proteins, and LCFAs to VFA products such as propionate and butyrate (L. Dong et al., 2019; García-Lozano et al., 2019; Li et al., 2014). These VFAs and SCFA's were then oxidized by acetogens near the end of the digester, which in this case were sulfate and sulfur-reducing bacteria, to produce H_2 gas for syntrophic methanogens such as *Methanomethylvorans spp.* to consume at the effluent (L. L. Dong et al., 2019; García-Lozano et al., 2019; Li et al., 2014). Ultimately this results in a decrease in COD as well as total Kjeldahl nitrogen levels in the form of ammonia, which perhaps explains the strong negative correlation these genera exhibited towards COD, TKN, and NH_3 levels (Ju et al., 2017; Ruiz-Sanchez et al., 2018). It is worth noting that most of the effluent enriched genera such as the sulfidogenic and methanogenic microbes were low abundance organisms of less than < 5% mean abundance, corroborating the fact that though rare, low abundant organisms still influence narrow functional niches in a microbial community system (Rivett & Bell, 2018).

3.5.2 Diversity of Microbial Communities Across Sampling Points

With regards to diversity throughout the sampling points, the disparity in richness and diversity trends between recirculating and single-pass systems could perhaps be due to the increased specialization towards denitrifying microbes in

syntrophy with hydrogenotrophic methanogens as covered in Section 2.7.3 brought about by the increase in NO_3^- levels replenished by the recirculating line. There are currently no metagenomic studies analyzing the spatial stratification of microbial communities in recirculating digesters. The increase in diversity and richness at the tank (middle of the digester) and effluent (end of the digester) points in single-pass systems could be due to a more diverse set of acetogenic and methanogenic specialists present compared to recirculating systems since single-pass systems were not subject to pulses of nitrate which would otherwise induce the growth of exclusively hydrogenotrophic methanogenesis via syntrophy with denitrifying and sulfidogenic microbes (C. Chen et al., 2017; Li et al., 2020; Saia et al., 2016). Additionally, the difference in diversity trends between single-pass and recirculating systems could be due to differences in OLR, and substrate type being treated, giving rise to a more diverse community (Xu et al., 2018). Further chemical measurements of VFA and methane levels must be performed to confirm the types of microbial degradation processes occurring (Ju et al., 2017; Li et al., 2014; Peces et al., 2018). Additionally, microbial interaction networks are also required to elucidate specialized microbial phylotypes and their relationship to functional profiles and chemical parameters responsible for the observed community diversity shifts across sampling points and digester types (Rivett & Bell, 2018; Wu et al., 2020).

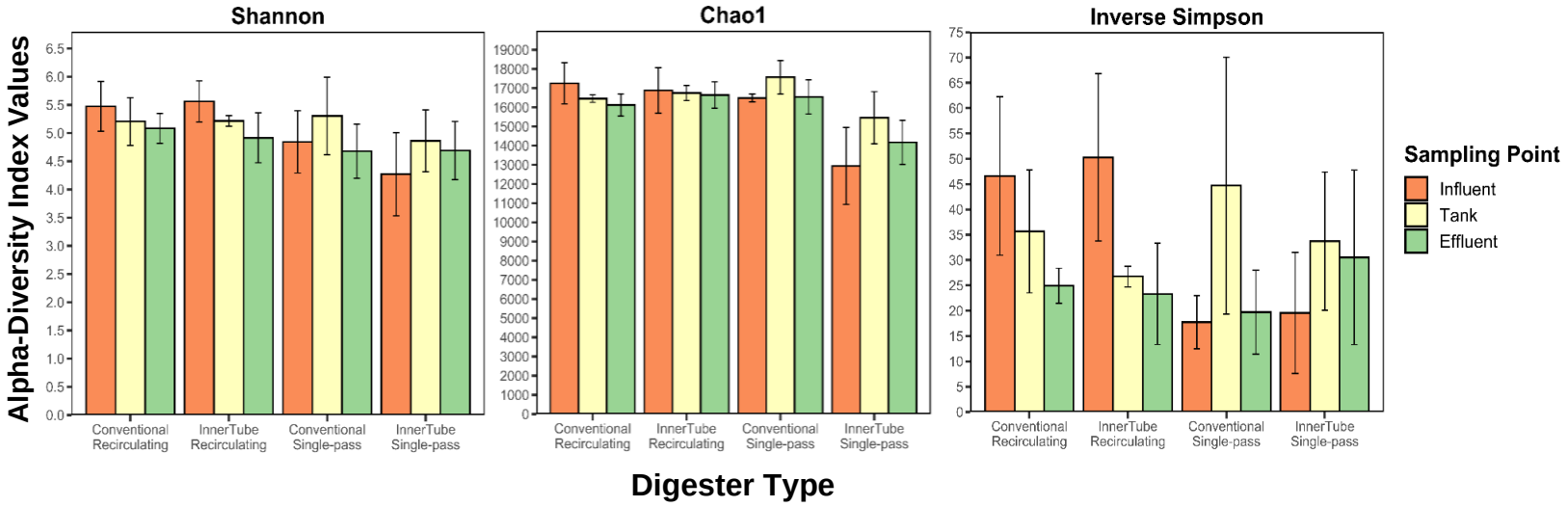


Figure 3.1: Shannon, Chao1, and Inverse Simpson mean alpha diversity indices of digester types across sampling points. Standard errors are shown (*n=3*).

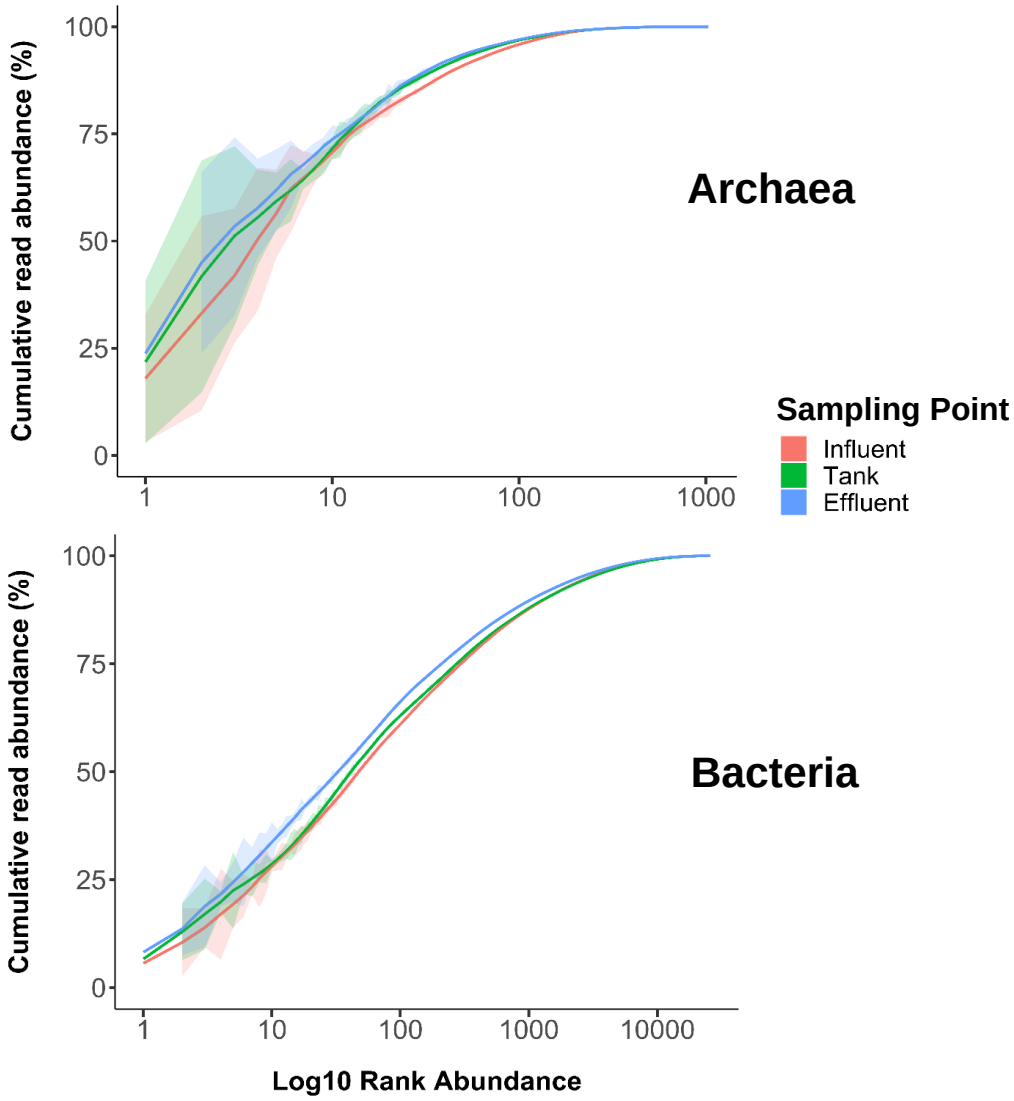


Figure 3.2: Rank Abundance Curves of Archaeal and Bacterial species among sampling points. Standard deviations of mean percent cumulative read abundances at each rank abundance are shown in colored shadings according to sampling points.

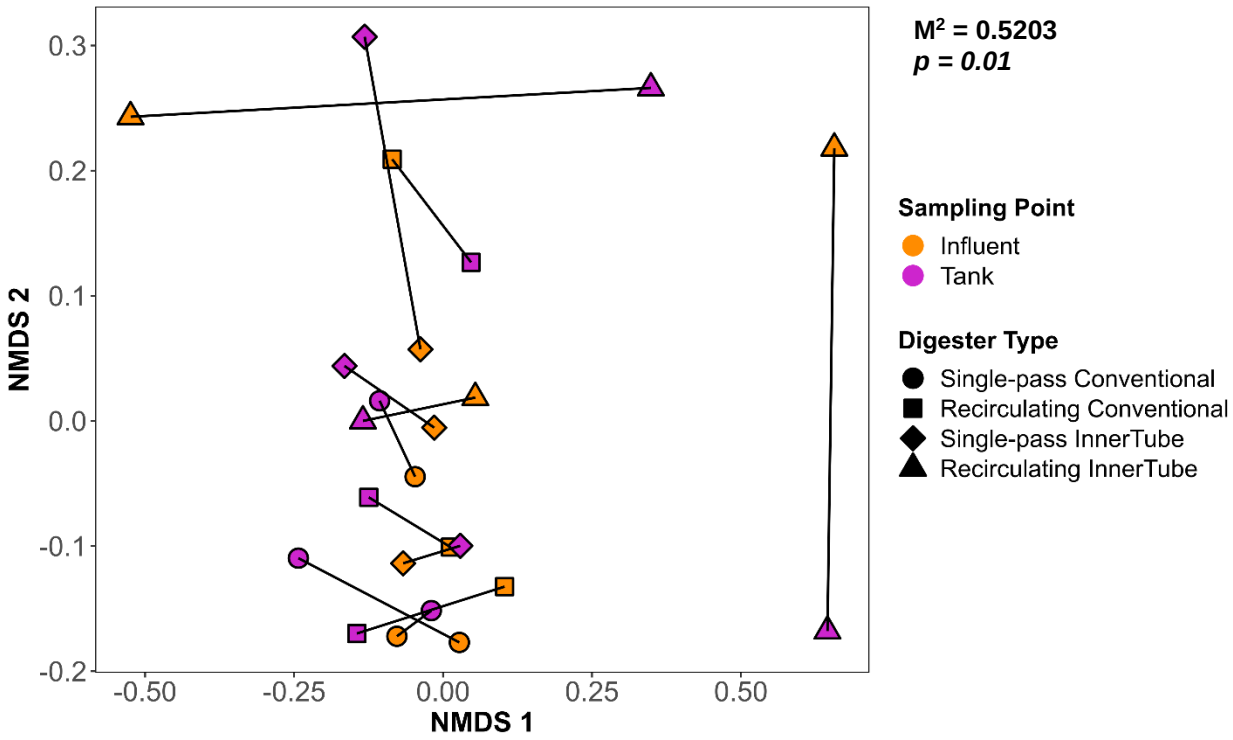


Figure 3.3: Procrustes analysis of NMDS ordinations between Influent and Tank. Significance test of the M^2 Procrustes statistic was performed by the R vegan function 'protest' ($p < 0.05$).

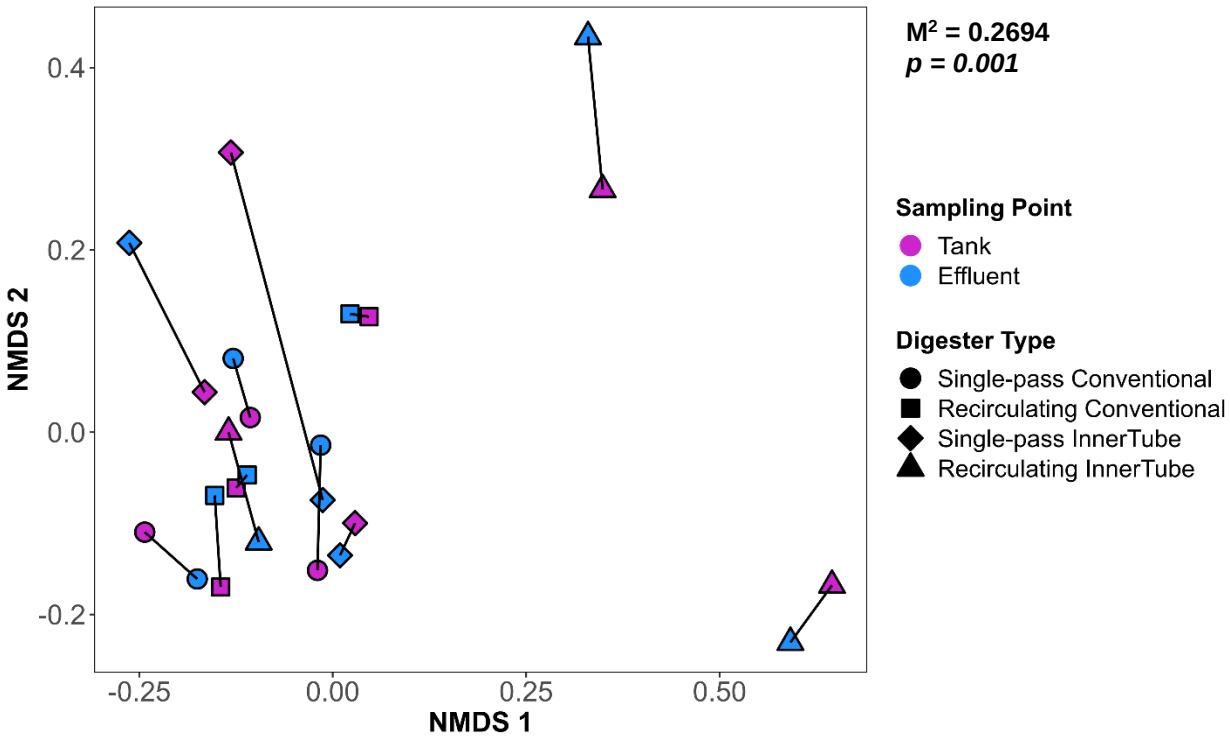


Figure 3.4: Procrustes analysis on NMDS ordinations between Tank and Effluent. Significance test of the M^2 Procrustes statistic was performed by the R vegan function 'protest' ($p < 0.05$).

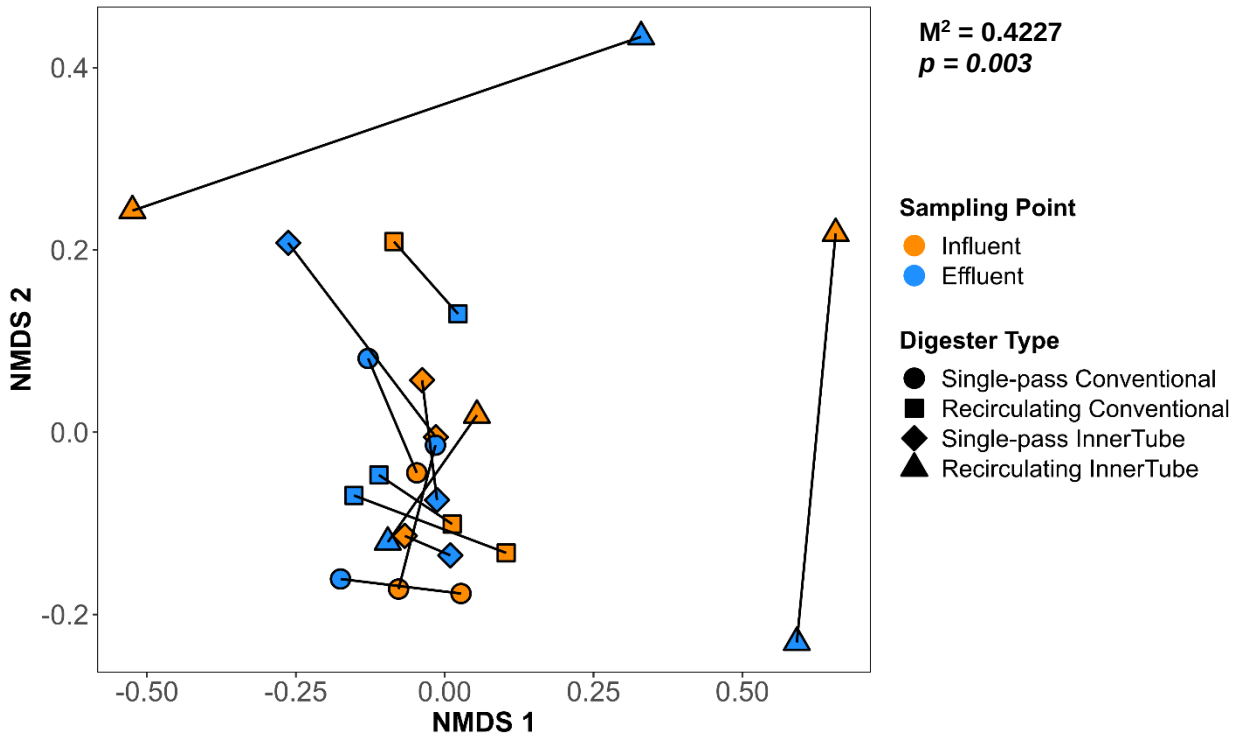


Figure 3.5: Procrustes analysis of NMDS ordinations between Influent and Effluent. Significance test of the M^2 Procrustes statistic was performed by the R vegan function 'protest' ($p < 0.05$).

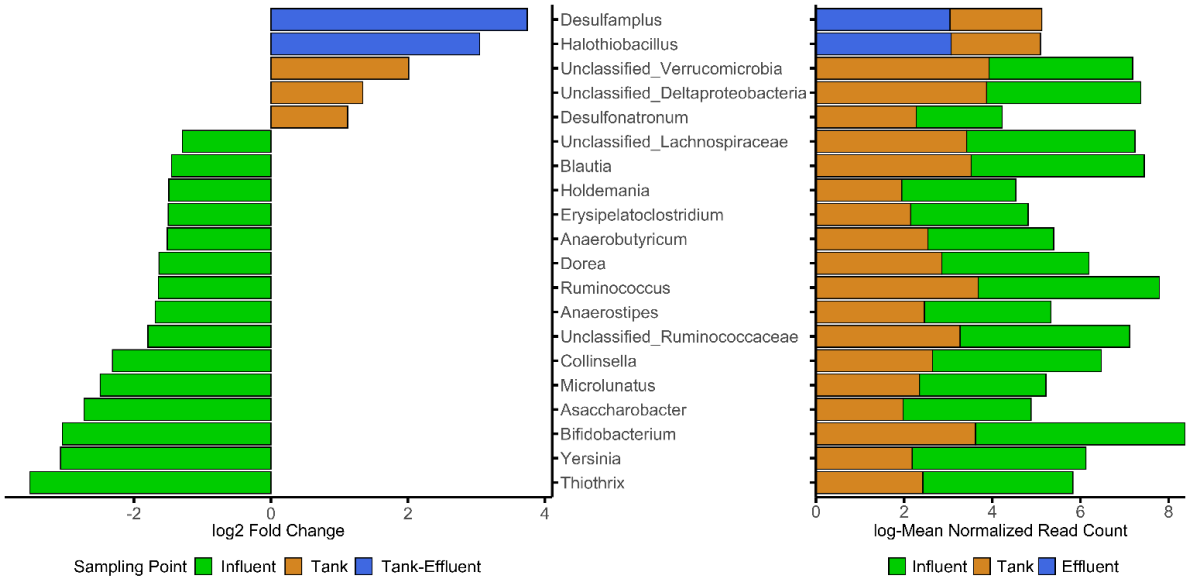


Figure 3.6: Pairwise comparisons of differentially-abundant genera from Influent, Tank, and Tank to Effluent sewage. Differential abundance analysis was done using **DESeq2** with an abundance cutoff of 0.01% of total agglomerated genera counts and an adjusted p-value of $p < 0.05$.

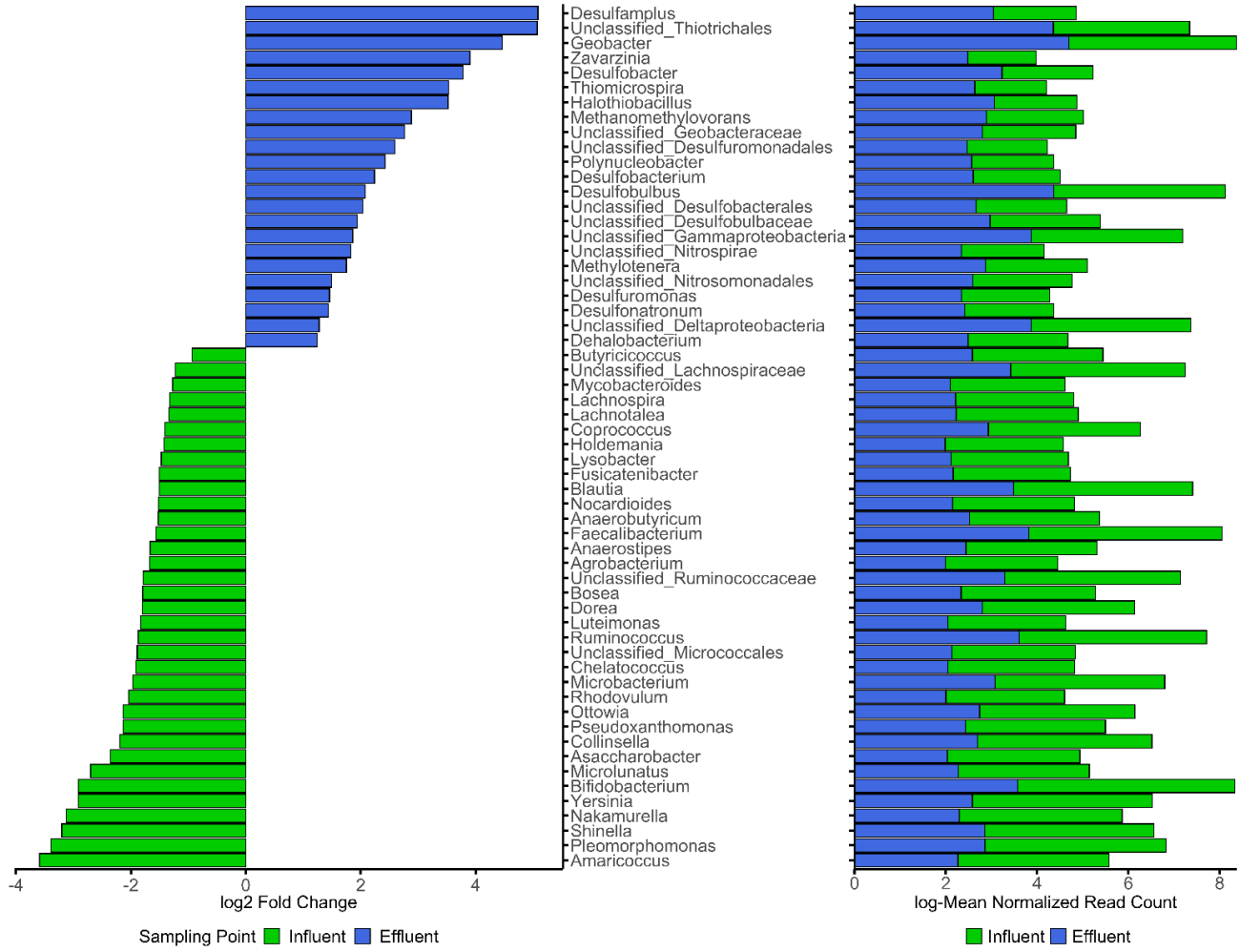


Figure 3.7: Pairwise comparisons of differentially-abundant genera between Influent and Effluent sewage. Differential abundance analysis was done using **DESeq2** with an abundance cutoff at 0.01% of total agglomerated genera counts and an adjusted p-value cutoff at $p < 0.05$.

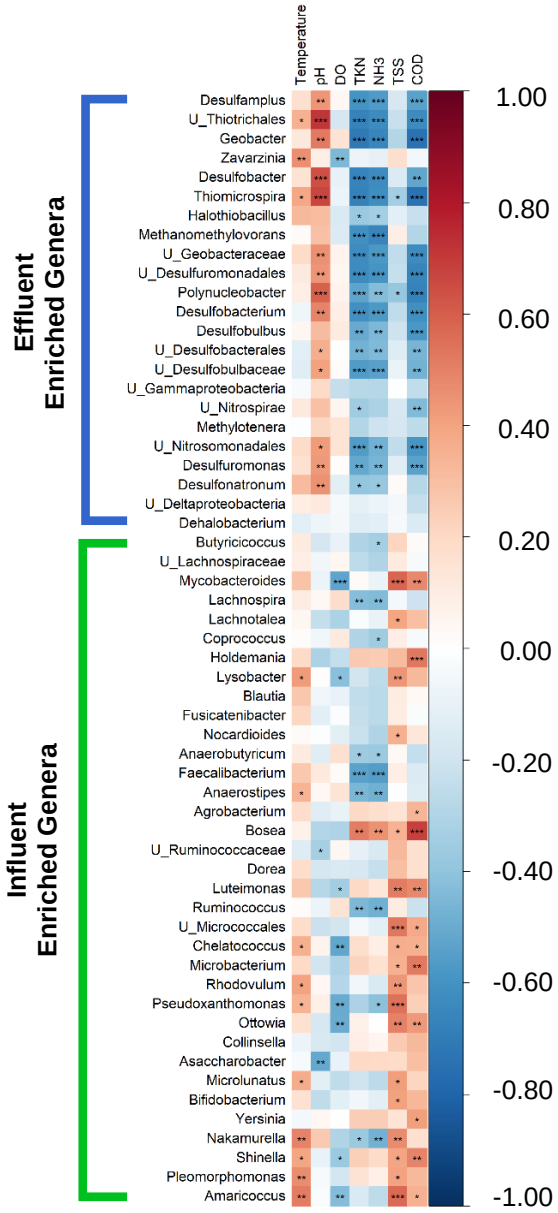


Figure 3.8: Spearman Rank Correlation heatmap of enriched genera between Influent and Effluent sewage to chemical parameters. The prefix "U_" before a genera name represents an unclassified genus. Significant correlations are marked with asterisks ($p < 0.05$) , where '***' 0.001, '**' 0.01, '*' 0.05.

4 Chapter 4: Functional Profiles of Digester Types, Flow, and Treatment Process

4.1 Introduction

4.1.1 Impact of Digester Design on Microbial Functional Potential

Analysis of metagenomic functional genes allows direct implications to be made of the roles that microbes perform in the anaerobic digester environment, providing a sort of 'snapshot' of a digester's microbial activity at a moment in time (Vanwonterghem et al., 2014). Microbial functions range from genes encoding for enzymes driving carbohydrate and protein metabolism to genes regulating cell motility and chemotaxis (Cai et al., 2016). There have been many replicated shotgun metagenomic sequencing analyses done on digester types such as CSTR (Fontana, Campanaro, et al., 2018), UASB (Delforno et al., 2017; Guo et al., 2016), recirculating digesters (Zhang et al., 2020) and conventional biogas septic tanks (Campanaro et al., 2016; Kougias et al., 2018; Luo et al., 2016). Depending on the digester design studied, different functional attributes of the microbiome were enriched. For example, metabolism of amino acids and derivatives gene subsystems were enriched in UASB reactors (Delforno et al., 2017) and cell motility functions were enriched in a conventional biogas reactor treating grass waste, as microbes attach to the surface of lignocellulose particles (Kougias et al., 2018) as opposed to the surface of granules in UASB reactors (Guo et al., 2016). However, there have been no comparative metagenomic studies performed assessing the effect of

different digester designs and operation on the functional microbiome, particularly between recirculating and conventional anaerobic digesters.

4.2 Chapter 4 Objectives

In this chapter, comparative analysis of functional gene profiles based on SEED Gene Subsystems between digester types, conventional and *InnerTube*™ and flow types, recirculating and single-pass, was done. Additionally, functional profiles of influent, tank, and effluent treatment points were also analyzed. By analyzing functional gene profiles, this project sought to confirm the putative functions of the microbes that were implicated for driving anaerobic digestion in different digester designs. It is hypothesized that functional gene subsystems corresponding to the putative function of the enriched genera that were observed in previous chapters would similarly be enriched within digester types, flow configurations, and sampling points.

4.3 Methods

4.3.1 Statistical Analyses

The same statistical analysis methods as discussed in Section 2.5.7 were performed on functional count datasets generated by **MEGAN6** based on the SEED Functional Gene Subsystems Database (Overbeek et al., 2014). Functional genes in SEED Subsystems are organized into three categories based on the putative gene roles, Level I, II, and III Subsystems and individual genes-products (Overbeek et al., 2014). **MEGAN6** outputted only Level I and Level III Subsystems along with

individual functional genes. Functional counts were Hellinger-standardized and averaged across digester replicates for the Top 50 Level 3 Subsystems graph. Functional counts were agglomerated at Gene Subsystem Level III for **DESeq2** and Spearman rank-sum correlation analyses. Differential abundance analysis of functional gene counts for digester type and flow configurations were done with an abundance cutoff at 0.1% total agglomerated Level III Subsystem counts and an adjusted p-value cutoff of $p < 0.05$. Differential abundance analysis between influent and effluent subsystems counts was done with an abundance cutoff at 0.01% of total agglomerated Subsystem Level III counts and an adjusted p-value cutoff at $p < 0.05$. All other statistical analyses were done at the level of gene products.

4.4 Results

4.4.1 Top 50 Abundant Functional Gene Subsystems

The top 50 most abundant Level I SEED Subsystems included Amino Acid and Derivatives, Cofactors, Vitamins, Prosthetic Groups, Pigments, Iron Acquisition and Metabolism, Protein Metabolism, DNA Metabolism, RNA Metabolism, and Carbohydrates (Figure 4.1). Within Amino Acid Derivatives, "Branched-Chain Amino Acid Biosynthesis and Glutamine", "Glutamate, Aspartate, and Asparagine Biosynthesis Subsystems" and "Methionine Biosynthesis" were equally abundant across all digester samples. However, "Isoleucine Degradation" was more abundant in the IR influent, the influent and tank of CS, and effluent of SI (Figure 4.1). "Lysine and threonine metabolism in

plants" was also most abundant in the influent of SI (Figure 4.1). Under Cofactors, Vitamins, Prosthetic Groups, Pigments, Iron Acquisition, and Metabolism, "5-FCL-like-OF" was highly abundant across all digesters, while "Biotin Biosynthesis" was most abundant in IR and CS influent and tank sewage (Figure 4.1). Under Protein Metabolism, "EC 6.1.1.- Ligases forming aminoacyl-tRNA and related compounds" was highly abundant across all digesters. Under DNA Metabolism, "Restriction-Modification Systems" was highly abundant across all digesters, and under RNA metabolism, only the "RNA polymerase bacterial" subsystem was abundant across all digester types (Figure 4.1). Under Carbohydrates, the "Entner-Doudoroff Pathway" was abundant throughout all digesters, while "Lactose and Galactose Uptake and Utilization" exhibited decreasing abundances from influent to effluent in all digesters (Figure 4.1).

4.4.2 Ordination of Digester Functional Profiles with Chemical Parameters

Similar to taxonomic analysis, RDA was used to determine how much of the variance in microbial functional gene compositions between digester types can be explained, or constrained, by changes in chemical parameters. RDA ordination of functional gene counts showed greater percent variance in the primary CAP axis compared to the secondary axis (Figure 4.2). Clustering across the primary axis was mostly influenced by TSS, COD, and pH parameters, while variance along the secondary axis was mostly influenced by NH_3, TKN, Temperature, and DO (Figure 4.2). Similar to the RDA ordination of species counts, gene profiles of *InnerTube*™ single-pass systems separated away from the main cluster of IR, CR, and CS

samples (Figure 4.2). Specifically, replicates of IR, IS, and CR digester types separated along the primary CAP axis while IS replicates separated along the secondary CAP axis. Digester Types also separated along the secondary CAP axis (Figure 4.2). Sample functional gene count variances were significantly correlated with temperature (PERMANOVA; pseudo-f = 4.860; p = 0.001), pH (PERMANOVA; pseudo-f = 2.397; p = 0.013), and DO (PERMANOVA; pseudo-f = 2.257; p = 0.031) (Table 2.1). Separation of digester types and replicates were significantly influenced by temperature, pH, and DO, similar to species ordination with the exception that DO did not significantly influence microbial community variances (Figure 2.11, 4.2). Temperature, pH, and DO explained 5.6%, 11.4%, and 1.2%, of the variation between digester type functional gene profiles, respectively.

4.4.3 Significance Tests of Ordinations Between Digester Types Functional Profiles

PERMANOVA tests revealed that functional gene compositions between digester types were significantly different (PERMANOVA; pseudo-f = 2.420; p = 0.001) (Table 2.1). Sampling points were not significantly different between influent, tank, and effluent functional gene profiles (Table 2.1). Pairwise PERMANOVA tests somewhat confirm this, as all pairwise digester type comparisons returned significant except between conventional and *InnerTube*™ recirculating systems and *InnerTube*™ single-pass and *InnerTube*™ recirculating systems (Table 2.2). However, sampling points and interactions between digester type and sampling points did not differ significantly throughout all pairwise

comparisons (Table 2.2). Variation partitioning also confirmed that digester type accounts for most of the microbial community variation between samples at 12.0% explained variance whereas sampling points explained -1.2% variance. BETADISPER tests revealed similar results as taxonomic profiles, where within-digester type variances were significantly different among replicates, but replicates categorized in terms of influent, tank, and effluent sampling points were not (Table 2.1).

4.4.4 Hierarchical Agglomerative Clustering of Functional Profiles of Digester Types

HAC analysis of functional genes showed similar results to the hierarchical agglomeration of taxonomic counts (Figure 2.13, 4.3). Bray-Curtis dissimilarities of sampling points agglomerated within the same digester type and replicate except for the first and second replicate of CS digesters which were outliers from their respective digester replicate (Figure 4.3). Replicates mostly agglomerated with the same digester types, except for IR, clustering with CR, and 1IS digesters (Figure 4.3). The 1SI replicate were also outliers from SI digesters, agglomerating at completely different nodes than the 2SI and 3SI replicates (Figure 4.3).

4.4.5 Differential enrichment of Functional Genes – Digester Types and Flow and Chemical Correlations

Between conventional and *InnerTube*™ digesters, a total of 19 differential abundant gene at Subsystem Level III were observed with a $\log_2$-fold change ranging from -1.04 to 0.54 (Figure 4.4, Table S 8). Enriched genes subsystems in

conventional systems putatively pertinent to anaerobic digestion include "Methanogenesis" (Carbohydrates), "Hydrogenases" (Respiration), and "Galactose Degradation in Plants" (Carbohydrates) (Figure 4.4, Table S 8). However, the only significantly correlated subsystems with chemical parameters implicated in anaerobic digestion were 'Methanogenesis" and "Galactose degradation in plants" (Figure 4.6, Table S 8). Two pertinent functional subsystems, the "Flagellum" (Motility and Chemotaxis) and "Potassium Homeostasis" (Potassium Metabolism) were enriched in *InnerTube*™ digesters (Figure 4.4, Table S 8).

Differentially abundant gene subsystems enriched in recirculating systems include a group of subsystems relevant to anaerobic digestion that were significantly negatively correlated TKN, NH_3, and COD and positively correlated with pH (Figure 4.7, Table S 8). These subsystems were "Propionyl-CoA to Succinyl-CoA Module" (Carbohydrates), "PII Superfamily" (Amino Acids and Derivatives), " "FOL Commensurate regulon activation" (Stress Response), "Nitrate and nitrite ammonification" (Nitrogen Metabolism), "Respiratory Complex I" (Respiration), "Multidrug Resistance Efflux Pumps" (Virulence), "Flagellar motility" (Motility and Chemotaxis) (Figure 4.7, Table S 8). In single-pass systems, relevant digestion subsystems exhibited the opposite trend, where most were significantly positively correlated to TKN, NH_3, and COD levels (Figure 4.7, Table S 8). The majority of single-pass-enriched subsystems positively correlated with biomass indicators were under the Carbohydrates subsystem, such as "Acetone Butanol Ethanol Synthesis", "Fermentations in Streptococci", and "D-

galactarate, D-glucarate, and D-glycerate catabolism", "Acetyl-CoA biosynthesis in plants", "D-ribose utilization", "Glycerol fermentation to 1,3-propanediol" (Figure 4.7, Table S 8). Additionally, other enriched subsystems of note include "Cinnamic Acid Degradation" (Amino Acids and Derivatives), "Alkanesulfonate assimilation" (Sulfur Metabolism), "Choline and Betaine Uptake and Betaine Biosynthesis" (Stress Response) (Figure 4.7, Table S 8).

4.4.6 Differential enrichment of Functional Genes – Influent, Tank, and Effluent and Chemical Correlations

Differential abundance tests between influent vs. tank and tank vs. effluent comparisons did not result in any significantly enriched subsystems. However, there was a total of 11 differentially abundant subsystems between the influent and effluent, ranging from $\log_2$-fold changes of -1.49 to 1.725 (Figure 4.5, Table S 9). There were only two enriched subsystems in the influent, "Ferrous iron transporter EfeUOB, low-pH-induced" (Iron acquisition and metabolism) and "2-O-alpha-mannosyl-D-glycerate utilization" (Carbohydrates) (Figure 4.5, Table S 9). Both were significantly positively correlated with TKN, and COD and negatively correlated with pH (Figure 4.8). "2-O-alpha-mannosyl-D-glycerate utilization" (Carbohydrates) was also positively correlated with NH_3 (Figure 4.8). In the effluent, subsystems of relevance to anaerobic digestion included "Selenocysteine Metabolism" (Protein Metabolism), which was significantly negatively correlated with COD levels (Figure 4.8). "Pyruvate metabolism I: anaplerotic reactions, PEP in Mycobacteria" (Carbohydrates) and "Hydrogenases" (Respiration) are also

relevant to anaerobic digestion but were not significantly negatively correlated with any chemical parameter (Figure 4.8).

4.5 Discussion

4.5.1 Whole-Community Functional Comparisons between Digester Types

To test whether significant variations in microbial community compositions between digester types (conventional and *InnerTube*™) and across digester replicates extend similarly at the functional gene level, PERMANOVA and BETADISPER tests were used on functional profiles. Functional profiles were significantly different across digester types and replicates as revealed by significant PERMANOVA between-group dispersion and BETADISPER within-group dispersion tests. This corroborates the results from taxonomic compositional analyses, demonstrating that digester design and chemical parameters affected both the presence of the microbes themselves and the functional potential of those microbes. Temperature and pH remained significant in influencing both functional profiles and taxonomic profiles, while dissolved oxygen (DO) significantly affected functional profiles. All three parameters have been previously reported to significantly alter digester community and functional profiles (Fontana, Kougias, et al., 2018; Lin et al., 2016b; Musa et al., 2018). Psychrophilic, mesophilic, or thermophilic microbes would dominate depending on the digester operating temperature (Lin et al., 2016a), and both temperature and pH levels would cause certain methanogenic syntrophic processes such as hydrogenotrophic methanogenesis at low pH levels, to dominate over others (Fontana, Kougias, et al.,

2018; Lin et al., 2016a). The biological oxygen demand of a system (BOD) is measured based on dissolved oxygen and along with COD, is a measurement of the quantity of biodegradable compounds and are indicators of the organic loading rate which is the amount biodegradable compounds in the influent (Gerba & Pepper, 2009; Musa et al., 2018). The significant influence that dissolved oxygen (DO) levels had on the variation of functional profiles across digester types indicates that each digester received significantly influent organic loading rates which thus, significantly affected microbial functional composition within digesters (Shin et al., 2019). Similar to taxonomic profiles, SI digesters did not agglomerate with its replicates in Bray-Curtis dissimilarities, demonstrating the effect that different influent chemical parameters have on both the taxonomic and functional profiles of anaerobic digesters (Calusinska et al., 2018).

4.5.2 Enriched Metabolic and Cellular Processes Between Digester Types

As shown in the Top 50 most abundant Level III Subsystems (Figure 4.1), the gene subsystem "5-FCL-like-OF" (Cofactors, Vitamins, Prosthetic Groups, Pigments, Iron Acquisition and Metabolism) was expected to be equally and highly abundant across all samples since this protein is ubiquitous in all domains of life and has been associated to folate metabolism used for basic cell function and environmental adaptation (Morgan et al., 2018; Pribat et al., 2011). Other subsystems that were equally abundant throughout all sampling points and digester types include the "Entner-Doudoroff Pathway" (Carbohydrates) and within Amino Acid Derivatives, "Branched-Chain Amino Acid Biosynthesis and Glutamine",

"Glutamate, Aspartate, and Asparagine Biosynthesis Subsystems" and "Methionine Biosynthesis" which are gene subsystems that include metabolic processes common in prokaryotes to use substrates such as glucose, glutamate, aspartate, and asparagine, for the Entner-Doudoroff pathway and Citric Acid Cycle and methionine biosynthesis for basic cell metabolism and function (Ferla & Patrick, 2014; Reitzer, 2004).

Differential abundance analysis between conventional and *InnerTube*™ digesters revealed that there was an enrichment of the "Methanogenesis" subsystem in conventional digesters over *InnerTube*™ digesters. This could be due to the fact that conventional systems were enriched over *InnerTube*™ systems with the highly-abundant acetoclastic methanogen *Methanothrix spp.* However, it is worth noting that the $\log_2$-fold change of all subsystems between conventional and *InnerTube*™ systems including the "Methanogenesis" subsystem only ranged from -1.05 to 0.54 fold-change, suggesting that the functional methanogenic processes in *InnerTube*™ systems were still abundant, as indicated by the fact that hydrogenotrophic methanogens were enriched in *InnerTube*™ digesters instead of acetoclastic methanogens. A positive correlation of TSS with the "Methanogenesis" subsystem was observed and can be explained by the fact that higher solids content facilitated by the retention of degradable solids has been directly correlated with increased methanogenic microbes as the increased solids content provides higher levels of substrate available for digestion (Vanwonterghem et al., 2015). The slight enrichment of "Hydrogenases" (Respiration) and

"Galactose Degradation in Plants" subsystems in conventional systems is indicative that conventional systems also contained a fermentative, acidogenic population producing bio-hydrogen from galactose, even though acetoclastic methanogens, not hydrogenotrophic methanogens, were dominant in conventional systems (Xia et al., 2016).

In *InnerTube*™ digesters, the "Flagellum" (Motility and Chemotaxis) subsystem was enriched, suggesting that the increased motility function in a plug-flow environment within the *InnerTube*™ might have been caused by a greater number of microbes seeking out areas with higher levels of substrate and adhering to the greater available surface area (Kougias et al., 2018). The enrichment of "Potassium Metabolism" (Potassium Metabolism) could be perhaps associated with the homeostatic potassium systems of enriched hydrogenotrophic methanogens (Sprott et al., 1984) as observed in *InnerTube*™ systems. Coupled with the activity of the SRB, *Desulfovibrio spp.*, which produce ammonia from sulfate, elevated ammonia levels could induce hydrogenotrophic methanogens such as *Methanospirllum spp.,* to adapt by utilizing potassium/ammonium homeostatic exchange processes (Sprott et al., 1984), further supporting the enrichment of hydrogenotrophic methanogenic activity in *InnerTube*™ systems.

4.5.3 4.5.3 Enriched Metabolic and Cellular Processes Between Flow Configurations

The significantly negatively correlated subsystems towards the biomass indicators TKN, NH_3 and COD in recirculating systems were all processes related

to carbohydrate, nitrogen, and amino acid metabolism. Most notably, “Nitrate and Nitrite ammonification” (Nitrogen Metabolism) was enriched, signifying elevated denitrification processes possibly performed by *Pseudoarcobacter spp.,* and *Sulfurimonas spp.,* due to the increase in NO_3^- delivered by the recirculating line (Chuan Chen et al., 2017; Perez-Cataluna et al., 2018; K. Zhang et al., 2018). The enriched “Propionyl-CoA to Succinyl-CoA Module” (Carbohydrates) subsystem is consistent with the increased abundance of propionate-oxidizing microbes such as *Geobacter spp.* (Aklujkar et al., 2009) in recirculating systems. This subsystem contains gene products driving the methylcitrate cycle, which is responsible for the oxidation of propionate to pyruvate in bacteria (Simonte et al., 2017). Other relevant subsystems include “Iron acquisition in Vibrio” (Iron acquisition and Metabolism) (Rakin et al., 2012) and “Multidrug Resistance Efflux Pumps” (Virulence) (Raczkowska et al., 2015), which indicate the presence of potential pathogens and might be attributed to the presence of *Yersinia spp.* in the influent. The enrichment of “Flagellar motility” (Motility and Chemotaxis) in recirculating systems is contrary to previous studies analyzing the metagenome of recirculating digesters, as it was shown that recirculation increased the flow and thus mixing of microbial biomass with sewage, removing the need for flagellar activity to obtain nutrients (Zhang et al., 2020). The discrepancy could be due to the fact that recirculating flow in WBS digesters were done in pulses, and therefore, there were periods when mixing does not occur.

The majority of enriched subsystems in single-pass systems were classified under Carbohydrate metabolism. In general, single-pass systems were characterized by an enrichment of hydrolytic and fermentative processes. For example, the enrichment of "Acetone Butanol Ethanol Synthesis" (Carbohydrate) suggests the presence of *Clostridium spp.* members participating in the fermentation of glucose into acetone, butanol, and ethanol (ABE), although no members from the genus *Clostridium spp.* were found enriched in single-pass systems (Zhang et al., 2011). ABE synthesis has been previously detected in ABR digesters, and *Clostridium spp.* was reported to be syntrophically-linked with methylotrophic methanogens that were able to consume alcohols to produce methane (Zhang et al., 2011). The enrichment of "Glycerol fermentation to 1,3-propanediol" (Carbohydrate) may suggest increased levels of glycerol and subsequent fermentation into 1,3-propanediol in single-pass systems, which have been observed in other anaerobic bioreactors to also give rise to members from the Clostridiaceae family (Zhou et al., 2017). In general, further studies tracking the metabolic activity of microbes that were found enriched in single-pass such as *Simplicispira spp., Phenylobacterium spp.* and *Pusillimonas spp.* using genome-centric metagenomics or isotope-based labeling of chemicals (DNA-SIP) are required as these genera and their functions in AD have yet to be elucidated (Campanaro et al., 2018; Wang et al., 2019).

4.5.4 Enriched Metabolic and Cellular Processes Driving the Treatment Process

In the Top 50 most abundant SEED subsystems (Figure 4.1), the subsystems for “Isoleucine degradation” (Amino Acids and Derivatives), “Lysine and threonine metabolism in plants” (Amino Acids and Derivatives), and “Lactose and Galactose Uptake and Utilization” (Carbohydrates) were most abundant in the influent or tank of digester types, demonstrating the stratification of hydrolytic and acidogenic, or fermentative, steps of anaerobic digestion near the beginning of digesters (García-Lozano et al., 2019; Li et al., 2014; Park et al., 2014).

To test whether functional subsystems were also spatially distributed according to successive putative hydrolytic, acidogenic, acetogenic, and methanogenic consortia roles along the three sampling points, influent, tank and effluent, differential abundance pairwise-comparisons of Level III Subsystems between influent vs. tank, tank vs. effluent and influent vs. effluent were performed (Figure 4.5).There were no differentially enriched subsystems in the influent vs. tank, and tank vs. effluent pairwise-companions, while influent vs. effluent comparisons yielded 11 differentially abundant subsystems. Similar to the fewer differentially abundant genera found between influent-tank and tank-effluent as opposed to influent-effluent comparisons, functional profiles must have changed significantly at the effluent side rather than gradually into the tank from the influent. This is corroborated by the fact that only 2 subsystems were enriched in the influent, while 9 were enriched in the effluent. “2-O-alpha-mannosyl-D-glycerate utilization” (Carbohydrates) enriched in influent has not been implicated in AD

function, but has been reported to be a carbon source for mesophilic bacteria (Sampaio et al., 2004). The "Ferrous iron transporter EfeUOB, low-pH-induced" (Iron acquisition and metabolism) subsystem contains gene-products associated with the regulation of iron uptake and is ubiquitous in all bacteria (Lau et al., 2015). This subsystem is activated in anaerobic conditions (Lau et al., 2015), such as the one found in anaerobic digesters and its enrichment in the influent could perhaps be indicative of microbes generating the enzymes required for AD with high iron requirements(Venkata Mohan et al., 2011), such as the Fe-Fe hydrogenase enzyme family which catalyzes the formation of H_2 in H_2-producing bacteria (Ziganshin et al., 2016).

The enrichment of "Selenocysteine metabolism" (Protein Metabolism) and "Hydrogenases" (Respiration) suggests the enrichment of acetogenic and methanogenic microbes frequently found at the end of an anaerobic digester (García-Lozano et al., 2019; Peters et al., 2004). Sulfate-reducing bacteria such as *Desulfococcus spp.* have been previously reported to utilize selenocysteine-containing proteins to anaerobically degrade the PHA, benzoate, which corroborates with the increased abundance of sulfidogenic bacteria in the effluent. The enrichment of "Hydrogenases" (Respiration) indicates higher abundances of hydrogen-producing enzymes, such as the Fe-Fe hydrogenase enzyme family (Venkata Mohan et al., 2011; Ziganshin et al., 2016), which corresponds to an increase in acetogenic and methanogenic microbes that were enriched in the effluent. Furthermore, the use of pyruvate as a substrate as shown in the enrichment

of “Pyruvate metabolism I: anaplerotic reactions, PEP” (Carbohydrates) has been reported to be performed by acetogens into acetate before being used by methanogens (Stams et al., 2005; Yang et al., 2015). Therefore, the enriched functional gene profiles between influent and effluent similarly reflect the microbial stratification of hydrolytic, acidogenic, acetogenic, and methanogenic microbes throughout the AD treatment process. Although WMS enabled the profiling of metabolic and cellular genes present in a system, biologically-relevant conclusions based on the enrichment of metabolic processes can only be inferred using WMS functional profiles (Emerson et al., 2017). In contrast, the addition of chemical substrates and inhibitors putatively known to affect AD *ex-situ* of digesters, as in culture-enriched functional metagenomics, or *in situ* of operationally stable digesters provides direct experimental evidence into the syntrophic metabolic processes occurring within a digester type (Jia et al., 2016; Treu et al., 2016).

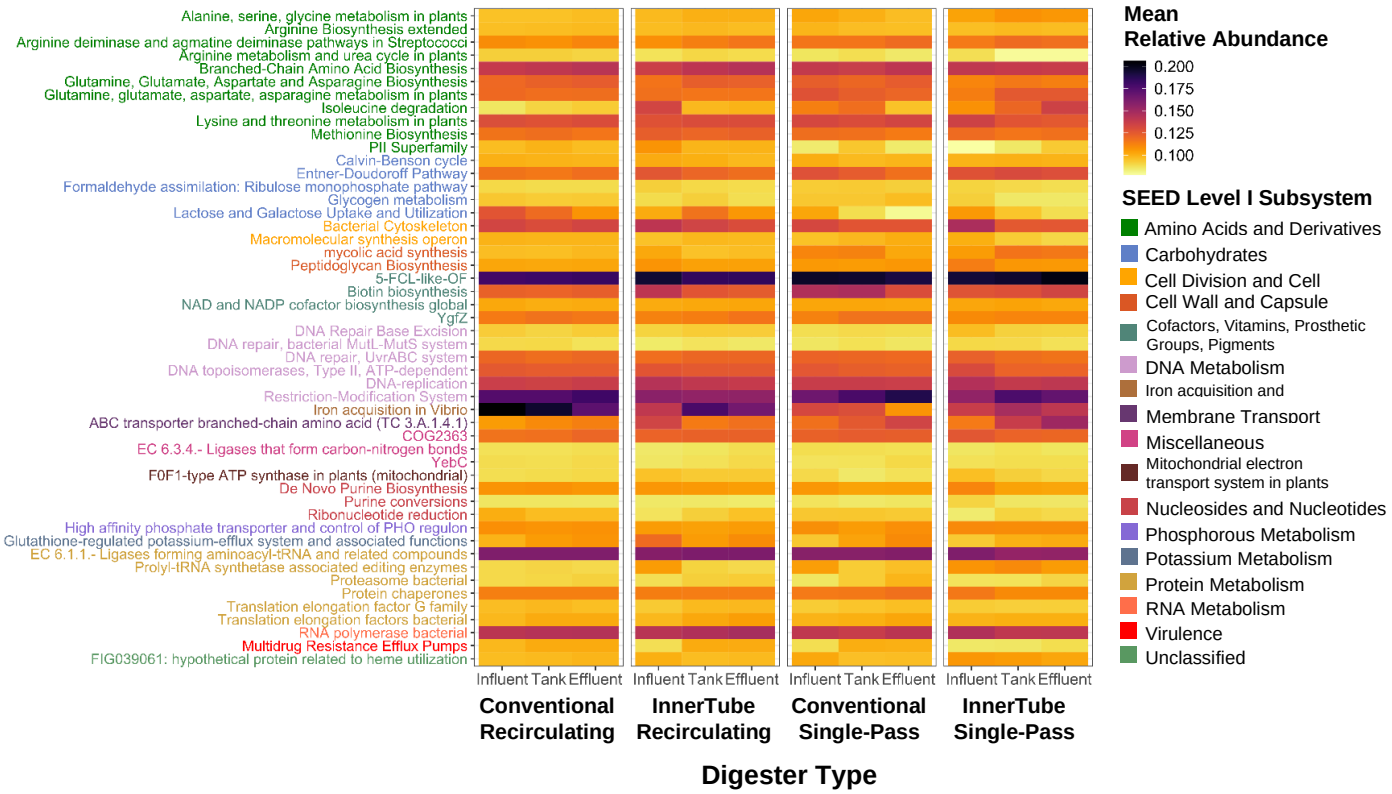


Figure 4.1: Top 50 most abundant functional Gene Subsystems (SEED Level III). Level III Gene Subsystems categorized alphabetically by Level I Classifications are indicated in color.

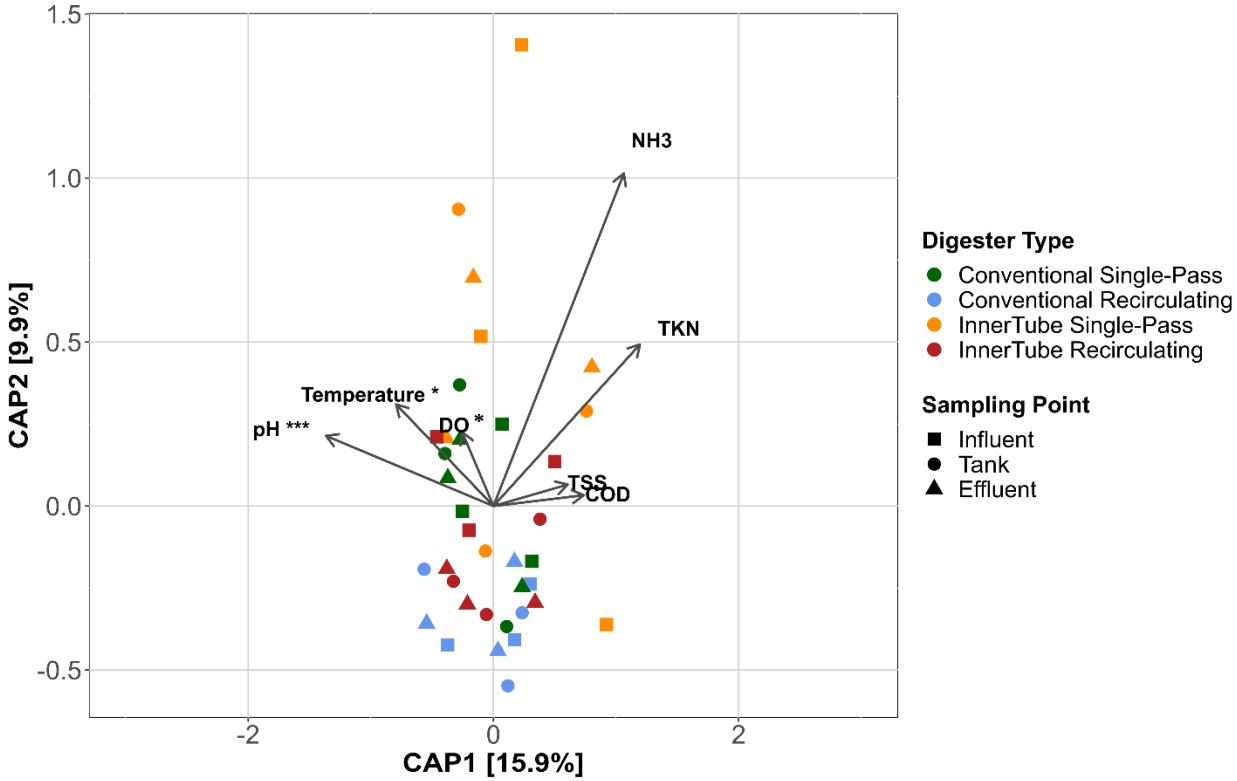


Figure 4.2: Redundancy Analysis of the Functional profiles of digester type and sampling points against chemical parameters. Chemical parameters that significantly influenced the ordination revealed by the PERMANOVA test are denoted with asterisks (*adj. $p < 0.05$*), where '***' 0.001, '**' 0.01, '*' 0.05.

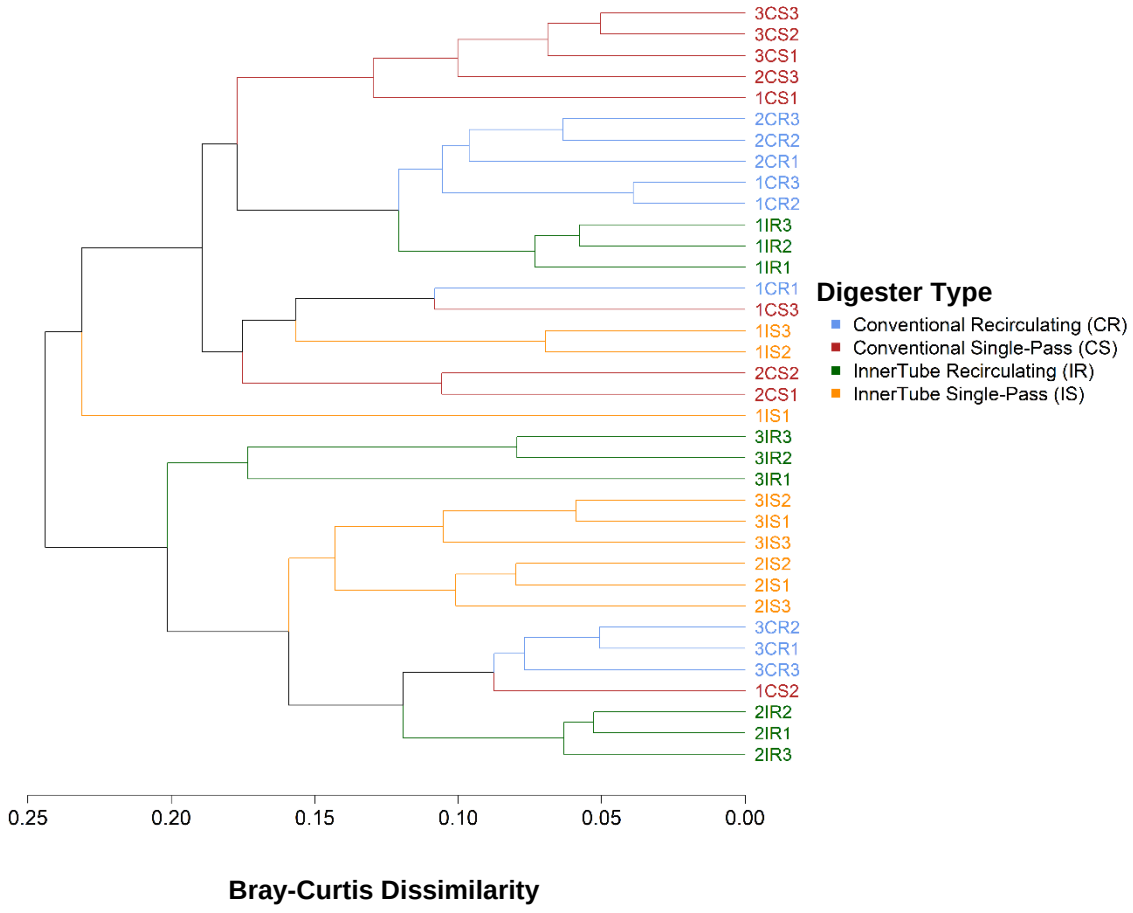
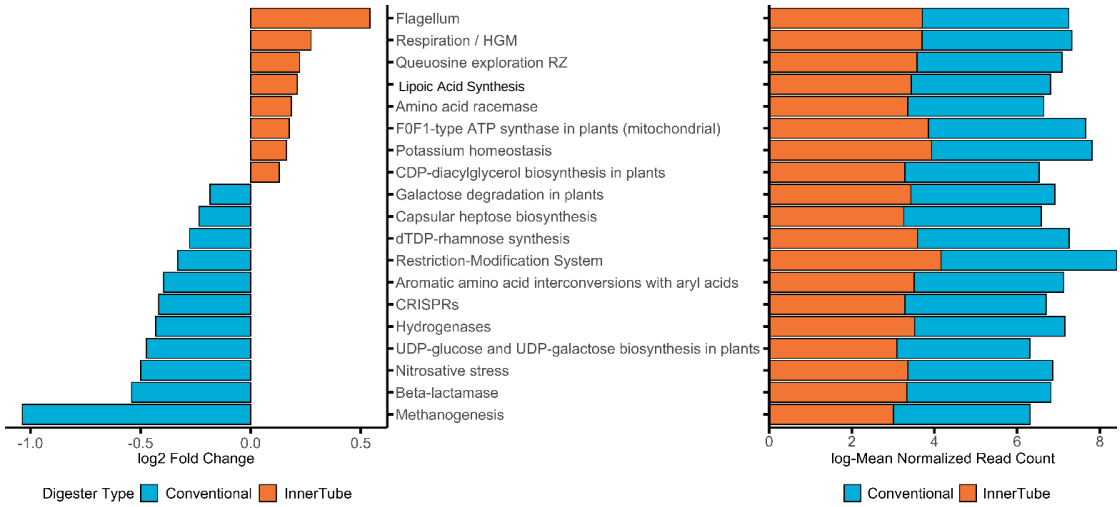


Figure 4.3: Hierarchical agglomerative clustering of functional genes based on Bray-Curtis Dissimilarities of triplicate digester types across sampling points. Numbers prefixing digester labels indicate a replicate, and numbers suffixing digester labels indicate a sampling point, where 1=Influent, 2=Tank, 3=Effluent.

A.


Flagellum
Respiration / HGM
Queuosine exploration RZ
Lipoic Acid Synthesis
Amino acid racemase
F0F1-type ATP synthase in plants (mitochondrial)
Potassium homeostasis
CDP-diacylglycerol biosynthesis in plants
Galactose degradation in plants
Capsular heptose biosynthesis
dTDP-rhamnose synthesis
Restriction-Modification System
Aromatic amino acid interconversions with aryl acids
CRISPRs
Hydrogenases
UDP-glucose and UDP-galactose biosynthesis in plants
Nitrosative stress
Beta-lactamase
Methanogenesis
-1.0
-0.5
0.0
0.5
log2 Fold Change
0
2
4
6
8
log-Mean Normalized Read Count
Digester Type
Conventional
InnerTube
Conventional
InnerTube

B.

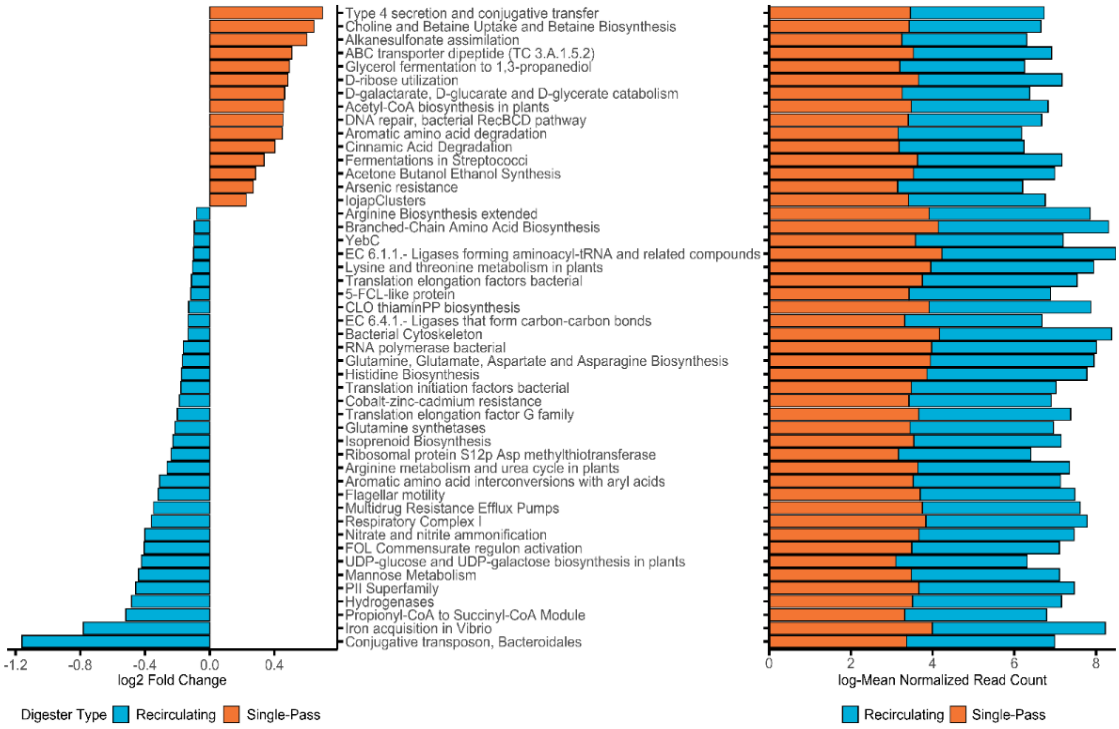


Figure 4.4: Pairwise comparisons of differentially-abundant Gene Subsystems between A.) Conventional and InnerTube and B.) Recirculating and Single-Pass Digester types. Differential abundance analysis was done using **DESeq2** with an abundance cutoff at 0.1% of total agglomerated Subsystem Level III counts and an adjusted p-value cutoff at $p < 0.05$.

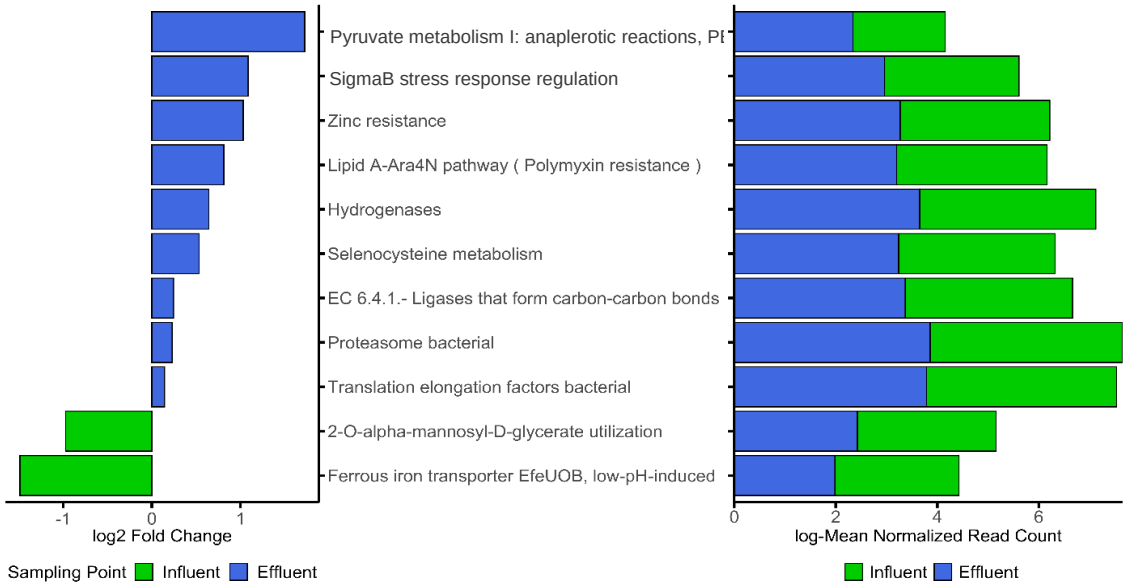


Figure 4.5: Pairwise comparisons of differentially-abundant Gene Subsystems between Influent and Effluent sewage. Differential abundance analysis was done using **DESeq2** with an abundance cutoff at 0.01% of total agglomerated Subsystem Level III counts and an adjusted p-value cutoff at $p < 0.05$.

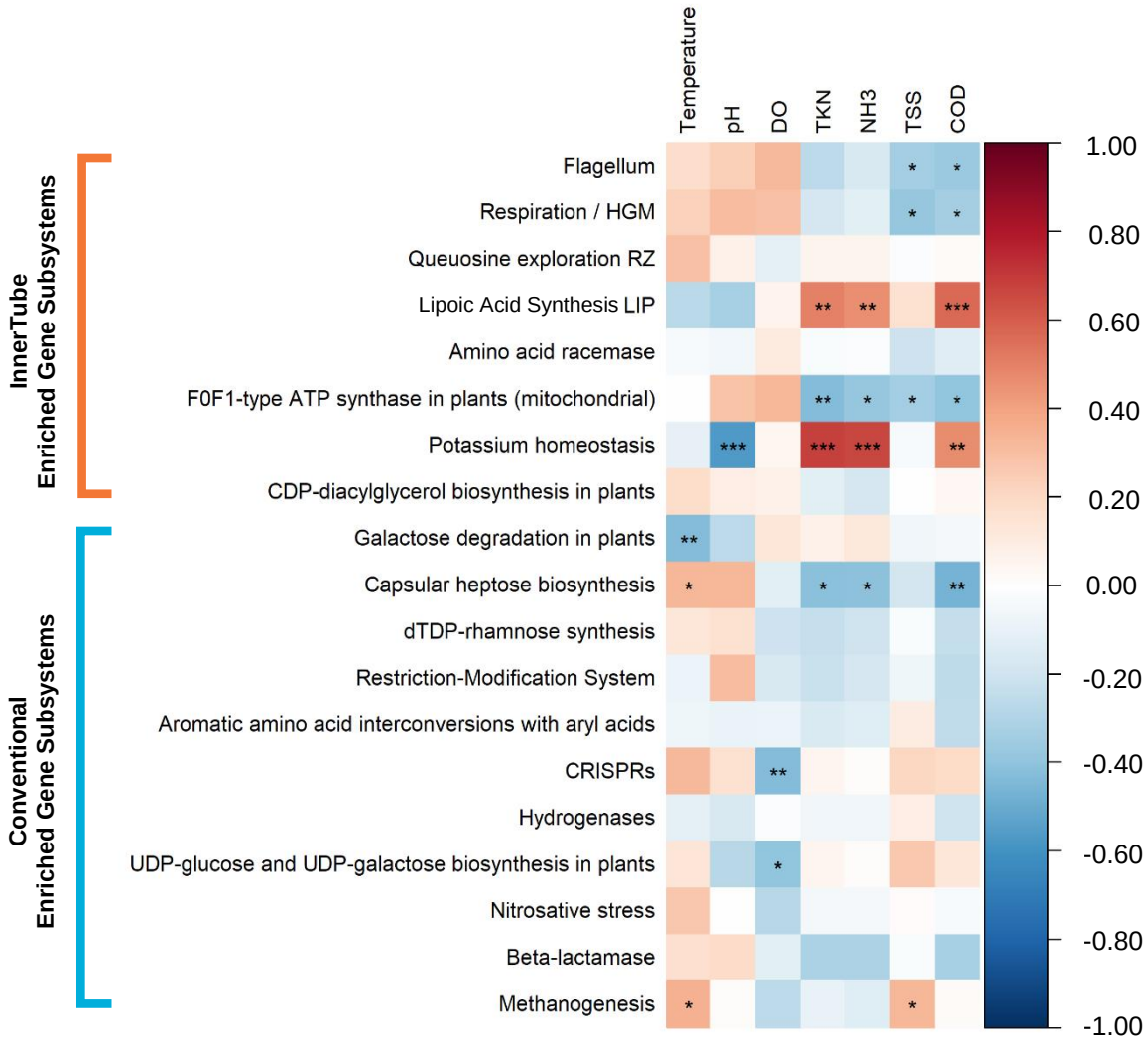


Figure 4.6: Spearman Rank Correlation heatmap of differential enriched SEED Level III Gene Subsystems between Conventional and InnerTube digesters to chemical parameters. Significant correlations are marked with asterisks ($p < 0.05$) , where '***' 0.001, '**' 0.01, '*' 0.05.

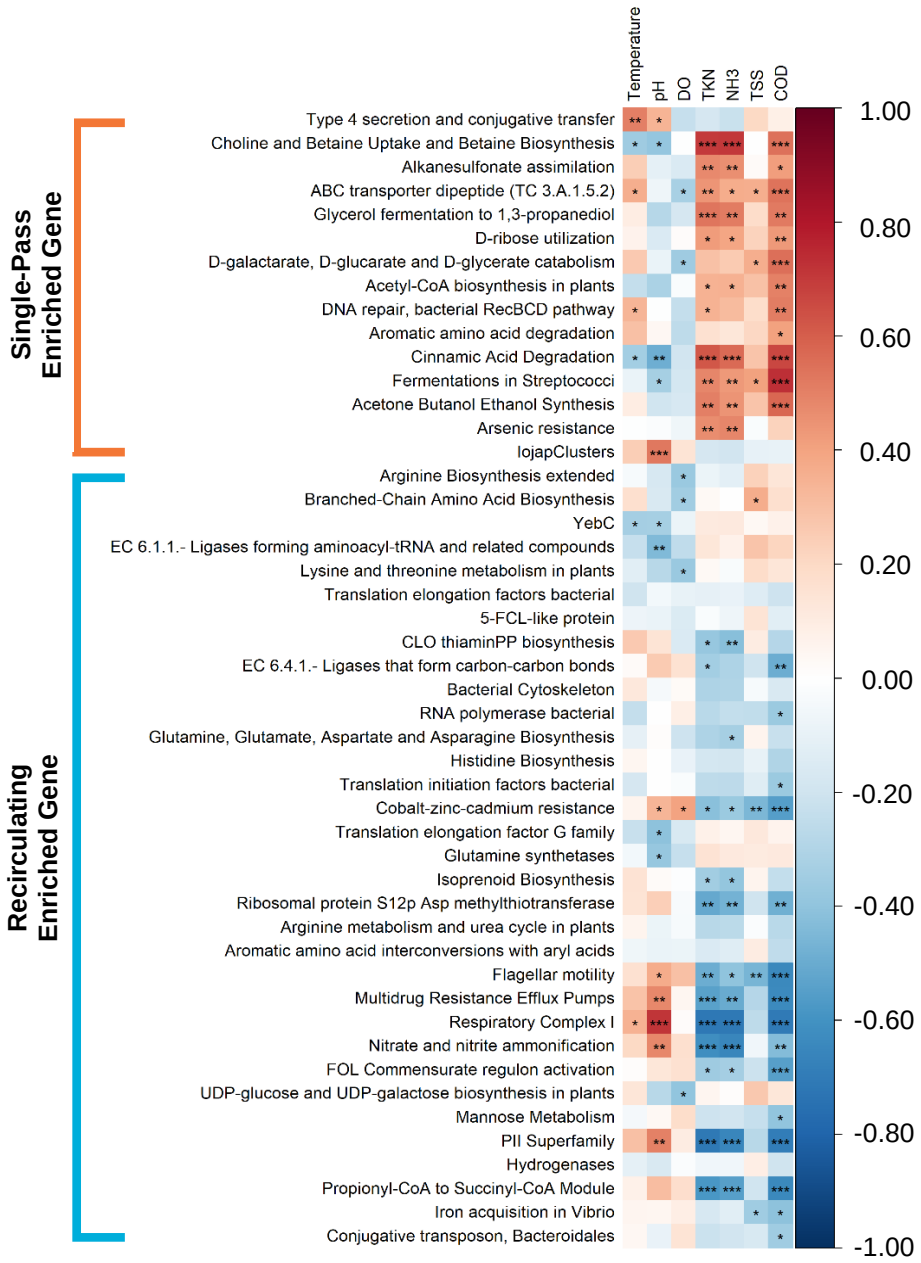


Figure 4.7: Spearman Rank Correlation heatmap of differential enriched SEED Level III Gene Subsystems between Recirculating and Single-pass digesters to chemical parameters. Significant correlations are marked with asterisks ($p < 0.05$) , where '***' 0.001, '**' 0.01, '*' 0.05.

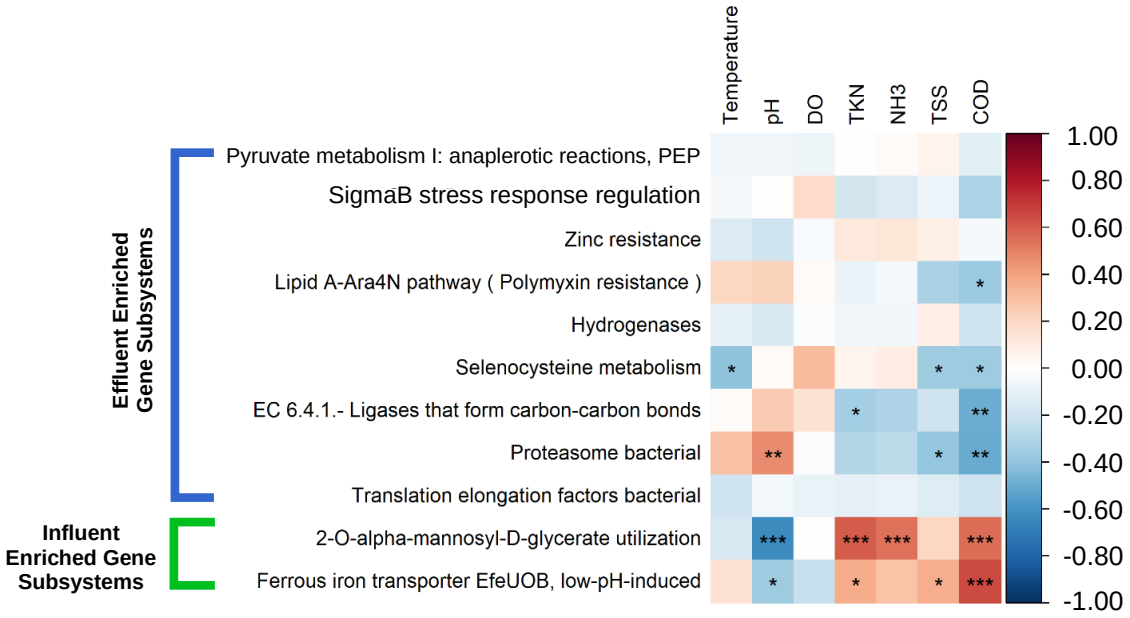


Figure 4.8: Spearman Rank Correlation heatmap of differential enriched SEED Level III Gene Subsystems between Influent and Effluent sewage to chemical parameters. Significant correlations are marked with asterisks ($p < 0.05$). Significant correlations are marked with asterisks ($p < 0.05$) , where '***' 0.001, '**' 0.01, '*' 0.05.

5 Conclusion and Future Steps

To date, a comparative metagenomic analysis of the microbial community and its functions in onsite-wastewater treatment systems (OWTS) has not been done. The widespread implementation of OWTS with anaerobic digesters throughout North America and the world for decentralized sewage treatment prompts the need for accurate characterization of core microbiomes driving anaerobic digestion in OWTS. The use of whole-metagenome shotgun sequencing enabled the analyses of low-abundant microbes that were functionally important to anaerobic digestion. The first objective of this project to characterize the microbial community composition and functional potential of two commonly used OWTS designs, conventional and *InnerTube*™-equipped septic tanks, throughout Southern Ontario. Conventional digesters were enriched with propionate-oxidizing microbes in syntrophic relationships with acetoclastic methanogens, while *InnerTube*™ systems had higher abundances of hydrogenotrophic methanogens in syntrophy with sulfate-reducing microbes. Implemented within the septic tanks' designs were aerobic effluent recirculating lines, which was observed to significantly impact microbial community composition and function since recirculating systems were enriched with sulfur-driven, denitrifying microbial consortia in syntrophy with hydrogenotrophic methanogens. The second objective of this project was to determine a core microbiome across digestion treatment influent, septic tank, and effluent points. Overall, influent sewage was enriched with microbes and functional genes associated with the hydrolysis and acidogenic

process, while the effluent contained microbes and functional genes associated with the acetogenic and methanogenic process to complete anaerobic digestion. The third objective of this project was to relate changes in microbial community compositions with the digester chemical parameters. All enriched microbes and functional genes that are putatively known to drive anaerobic digestion were correlated with waste biomass indicators, confirming their roles in organic waste degradation. Finally, this project implemented a replicated design to assess the level of variability across sequenced microbial communities. It was determined that digester replicates of the same design and flow type significantly differed in microbial community and functional compositions. Overall, microbial communities and their functional profiles were demonstrated to be influenced by changes in chemical parameters brought about by differences in design between conventional and *InnerTube*™-equipped digesters and between aerobic-effluent-recirculating and single-pass digesters.

One of the challenges of metagenomic sequencing analyses is to reduce the size and complexity of data to elucidate biologically pertinent microbes and their function in a given system (Mulcahy-O'Grady & Workentine, 2016). Due to the size and expense of WMS analyses, this project had an insufficient number of replicates available to minimize standard errors as revealed by multivariate-based power analysis. The inherent heterogeneity of microbial communities limits the applicability of this project's findings towards anaerobic digesters in general (Knight et al., 2012). Furthermore, since the genome-centric analysis was not

employed, direct implications of enriched microbes to enriched functional gene groups could not be made (Campanaro et al., 2018). Additional pertinent chemical indicators of digestion performance such as volatile fatty acids, methane and in particular, nitrate to determine the denitrification capabilities in recirculating systems, were not measured, which also limited the insight that could be gained on the functional processes of pertinent microbes. Furthermore, functional metagenomic and metatranscriptomic analyses of *in situ* and *ex-situ* addition of chemical byproducts of anaerobic digestion into stable full-scale digesters and cultures, respectively, enables direct experimental conclusions to be made of the effect of chemical parameters on microbial community dynamics (Jia et al., 2016; Treu et al., 2016). Finally, using microbial network correlation analyses can more clearly demonstrate syntrophic relationships between microbial phylotypes primarily driving stable anaerobic digestion (Wu et al., 2020; Zhang et al., 2017). Mitigating these experimental design limitations can provide greater insight into the composition and function of the core microbiome in OWTS.

Ultimately, this project was a survey of the microbial communities of decentralized OWTS anaerobic digesters across Ontario. Specialized microbes present in digester type, flow, and treatment points were found and these microbes have been implicated in driving anaerobic digestion based on several previous studies analyzing the anaerobic digester microbiome. Characterization of these microbes in onsite-wastewater treatment systems could serve as a starting point in

informing bioaugmentation and system design and operational optimization strategies to improve septic tank performance.

6 Appendices

6.1 Appendix I: Replicated Digester Field Survey

Table S 1: Field Survey sample information and chemical measurements.

Sample Code[1]	Location	Sample Type	pH	DO (mg/L)	Temp. (ºC)	DO (mg/L)	COD (mg/L)	BOD[2] (mg/L)	TSS (mg/L)	TKN (mg/L)	NH3 (mg/L)	Recirculating Valve
1IR1	75 Firelane 2	Inlet	7.42	1.86	16.7	1.86	326	428	138	62.8	54.3	50% open
1IR2	75 Firelane 2	Tank	7.47	1.86	17.1	1.86	227	N/A	52	72.0	65.2	
1IR3	75 Firelane 2	Effluent	7.61	5.58	16.9	5.58	243	242	44	72.8	67.0	
2IR1	1552 Concession 2	Inlet	7.27	4.84	19.2	4.84	979	945	460	107	84.1	50% open
2IR2	1552 Concession 2	Tank	7.13	1.64	18.9	1.64	463	N/A	258	87.7	80.0	
2IR3	1552 Concession 2	Effluent	7.55	1.27	18.6	1.27	191	132.66	28	82.6	76.0	
3IR1	**28 Highway 8**	**Inlet**	**7.53**	**2.25**	**14.2**	**2.25**	**173**	**247**	**46**	**14.6**	**12.3**	**Unknown**
3IR2	**28 Highway 8**	**Tank**	**7.59**	**0.65**	**15.4**	**0.65**	**208**	**N/A**	**24**	**24.4**	**19.6**	
3IR3	**28 Highway 8**	**Effluent**	**7.67**	**2.93**	**15**	**2.93**	**199**	**49**	**33**	**24.0**	**18.9**	
4IR1	**2649 No. 2 Sideroad**	**Inlet**	**7.18**	**1.82**	**11.8**	**1.82**	**219**	**292**	**202**	**30.6**	**24.1**	**50% open**
4IR2	**2649 No. 2 Sideroad**	**Tank**	**7.35**	**1.68**	**12.6**	**1.68**	**174**	**N/A**	**17**	**42.7**	**39.3**	
4IR3	**2649 No. 2 Sideroad**	**Effluent**	**7.58**	**4.57**	**12.4**	**4.57**	**154**	**368**	**18**	**41.2**	**36.5**	
5IR1	**4483 Escarpment Dr.**	**Inlet**	**6.31**	**0.29**	**14.1**	**0.29**	**14070**	**1584**	**7480**	**266**	**131**	**10% open**
5IR2	**4483 Escarpment Dr.**	**Tank**	**6.73**	**0.29**	**13.5**	**0.29**	**1133**	**N/A**	**705**	**78.7**	**59.1**	
5IR3	**4483 Escarpment Dr.**	**Effluent**	**7.03**	**0.80**	**12.8**	**0.8**	**396**	**229**	**79**	**55.8**	**45.5**	
6IR1	8465 Canyon RI	Inlet	7.62	0.78	13.4	0.78	935	818	431	177	146	50% open
6IR2	8465 Canyon RI	Tank	7.45	1.18	13.5	1.18	696	N/A	60	141	122	
6IR3	8465 Canyon RI	Effluent	7.56	4.21	13.3	4.21	682	120	52	139	122	
1IS1	**10091 Iona RD.**	**Inlet**	**7.69**	**1.19**	**22.7**	**1.19**	**1110**	**748**	**248**	**24.9**	**13.5**	**None**
1IS2	**10091 Iona RD.**	**Tank**	**7.58**	**0.37**	**19.5**	**0.37**	**263**	**N/A**	**34**	**47.6**	**42.6**	
1IS3	**10091 Iona RD.**	**Effluent**	**7.59**	**0.35**	**19.3**	**0.35**	**259**	**310**	**46**	**50.6**	**45.7**	
3IS1	1128 Matthiasville RD.	Inlet	6.51	5.10	11.5	5.1	21080	1721	5470	399	81.1	None
3IS2	1128 Matthiasville RD.	Tank	7.76	0.79	8.9	0.79	718	N/A	137	64.8	48.9	
3IS3	1128 Matthiasville RD.	Effluent	7.29	1.36	9.7	1.36	424	535	21	50.1	48.5	
4IS1	**493 Roaslind Lake Road**	**Inlet**	**9.40**	**4.28**	**7.8**	**1.82**	**580**	**727**	**30**	**228**	**207**	**None**
4IS2	**493 Roaslind Lake Road**	**Tank**	**8.86**	**1.71**	**4.2**	**1.68**	**279**	**N/A**	**24**	**85.5**	**77.2**	
4IS3	**493 Roaslind Lake Road**	**Effluent**	**8.82**	**3.51**	**4.3**	**4.57**	**263**	**245**	**15**	**86.8**	**78.2**	

5IS1	1002 Golden Point Road	Inlet	5.61	1.10	8.7	1.1	11890	1715	4020	331	68.8	None
5IS2	1002 Golden Point Road	Tank	6.98	1.94	8.7	1.94	116	N/A	8	54.7	48.7	
5IS3	1002 Golden Point Road	Effluent	6.95	2.35	8.9	2.35	109	156	5	55	49.2	
6IS1	**1206 Charlie Thompson RD.**	**Inlet**	**6.37**	**1.10**	**9.1**	**1.1**	**11310**	**1751**	**3800**	**360**	**92.2**	**None**
6IS2	**1206 Charlie Thompson RD.**	**Tank**	**6.98**	**1.94**	**8.9**	**1.94**	**571**	**N/A**	**74**	**155**	**132**	
6IS3	**1206 Charlie Thompson RD.**	**Effluent**	**6.95**	**2.35**	**8.9**	**2.35**	**507**	**801**	**44**	**139**	**120**	
1CR1	44 Autumn Circle	Influent	7.22	1.76	11.6	1.76	514	387	70	75.7	65.6	10% open
1CR2	44 Autumn Circle	Tank	7.21	1.34	11.6	1.34	489	N/A	33	75.4	71.4	
1CR3	44 Autumn Circle	Effluent	8.25	0.57	10.8	0.57	264	130	39	73.0	62.6	
2CR1	7 Diamondwood Drive	Influent	6.72	0.61	16.9	0.61	950	541	650	98.9	86.1	25% open
2CR2	7 Diamondwood Drive	Tank	7.03	0.75	16	0.75	308	N/A	44	82.8	79.4	
2CR3	7 Diamondwood Drive	Effluent	7.38	1.25	15.5	1.25	255	210	54	74.2	68.4	
3CR1	10 Diamondwood Drive	Influent	7.09	0.65	17.7	0.65	739	243	424	75.8	50.2	Unknown
3CR2	10 Diamondwood Drive	Tank	6.99	0.73	17.1	0.73	308	N/A	150	62.0	59.1	
3CR3	10 Diamondwood Drive	Effluent	7.16	1.95	16.6	1.95	113	47	21	54.1	51.7	
4CR1	**17 Flamborough Hills Drive**	**Influent**	**7.37**	**1.23**	**13.2**	**1.23**	**615**	**821**	**48**	**67.9**	**59.5**	**50% open**
4CR2	**17 Flamborough Hills Drive**	**Tank**	**7.42**	**1.62**	**12.7**	**1.62**	**504**	**N/A**	**83**	**70.8**	**60.3**	
4CR3	**17 Flamborough Hills Drive**	**Effluent**	**7.39**	**3.20**	**13.2**	**3.2**	**431**	**633**	**66**	**71.6**	**60.1**	
5CR1	**40 Diamondwood Drive**	**Influent**	**7.26**	**0.65**	**17.0**	**1.23**	**342**	**220**	**124**	**44.8**	**37.6**	**25% open**
5CR2	**40 Diamondwood Drive**	**Tank**	**7.12**	**0.70**	**16.7**	**1.62**	**216**	**N/A**	**79**	**42.5**	**38.5**	
5CR3	**40 Diamondwood Drive**	**Effluent**	**7.38**	**1.29**	**16.5**	**3.2**	**173**	**195**	**56**	**42.3**	**37.9**	
6CR1	**45 Autumn Circle**	**Influent**	**7.11**	**1.63**	**11.5**	**1.63**	**380**	**297**	**52**	**78.5**	**67.7**	**50% open**
6CR2	**45 Autumn Circle**	**Tank**	**7.09**	**1.41**	**11.1**	**1.41**	**323**	**N/A**	**48**	**81.9**	**72.0**	
6CR3	**45 Autumn Circle**	**Effluent**	**7.21**	**0.42**	**10.5**	**0.42**	**271**	**216**	**44**	**79.8**	**67.8**	
1CS1	**2 Diamondwood Drive**	**Influent**	**7.22**	**1.76**	**11.6**	**0.69**	**739**	**890**	**282**	**190**	**64.8**	**Closed**
1CS2	**2 Diamondwood Drive**	**Tank**	**7.21**	**1.34**	**11.6**	**0.55**	**5864**	**N/A**	**2960**	**64.4**	**57.9**	
1CS3	**2 Diamondwood Drive**	**Effluent**	**8.25**	**0.57**	**10.8**	**1.39**	**193**	**217**	**35**	**61.4**	**56.9**	
2CS1	**22 Diamondwood Drive**	**Influent**	**6.72**	**0.61**	**16.9**	**0.55**	**505**	**222**	**163**	**52.7**	**44.0**	**None**

2CS2	**22 Diamondwood Drive**	**Tank**	**7.03**	**0.75**	**16**	**0.66**	**381**	**N/A**	**125**	**44.9**	**37.9**	
2CS3	**22 Diamondwood Drive**	**Effluent**	**7.38**	**1.25**	**15.5**	**1.16**	**255**	**148**	**50**	**39.5**	**34.8**	
3CS1	925 Longfellow Ave.	Influent	7.09	0.65	17.7	0.54	256	258	37	68.1	62.1	None
3CS2	925 Longfellow Ave.	Tank	6.99	0.73	17.1	0.56	211	N/A	42	70.7	67.2	
3CS3	925 Longfellow Ave.	Effluent	7.16	1.95	16.6	1.14	127	51	26	64.4	61.2	
4CS1	**362 Evert Street**	**Influent**	**7.37**	**1.23**	**13.2**	**0.86**	**558**	**632**	**131**	**51.7**	**41.2**	**None**
4CS2	**362 Evert Street**	**Tank**	**7.42**	**1.62**	**12.7**	**0.57**	**394**	**N/A**	**68**	**52.5**	**43.2**	
4CS3	**362 Evert Street**	**Effluent**	**7.39**	**3.20**	**13.2**	**0.89**	**435**	**1701**	**153**	**53.3**	**45.4**	
5CS1	3105 Dundas St.	Influent	7.26	0.65	17.0	1.42	441	604	432	68.4	55.8	None
5CS2	3105 Dundas St.	Tank	7.12	0.70	16.7	1.61	255	N/A	27	69.9	49.9	
5CS3	3105 Dundas St.	Effluent	7.38	1.29	16.5	2.1	246	373	19	60.8	56.3	
6CS1	2850 Victoria Street	Influent	7.11	1.63	11.5	1.83	486	919	200	47.1	35.0	None
6CS2	2850 Victoria Street	Tank	7.09	1.41	11.1	1.08	218	N/A	35	39.1	30.5	
6CS3	2850 Victoria Street	Effluent	7.21	0.42	10.5	0.71	375	411	200	47.1	32.2	

[1]: Bolded samples represent the ones sequenced for this project
[1]: Numbers prefixing digester labels indicate a specific digester site replicate, and numbers suffixing digester labels indicate a sampling point, where 1=Influent, 2=Tank, 3=Effluent. Digester types are labeled as InnerTube Recirculating (IR), InnerTube Single-pass (IS), Conventional Recirculating (CR), Conventional Single-pass (CS).
[2]: BOD values labeled "N/A" were not measured

Table S 2: Pre-treatment sample information of InnerTube Recirculating (4RI) technical replicates

Sample Code[1]	Region	Sampling Point	Pre-Treatment method
4IR1-C1	2649 No. 2 Sideroad	Influent	-80ºC Frozen, 10 mL Sample; 10 mL of H_2O, Control
4IR1-C2	2649 No. 2 Sideroad	Influent	
4IR1-C3	2649 No. 2 Sideroad	Influent	
4IR1-E1	2649 No. 2 Sideroad	Influent	-80ºC Frozen, 10 mL Sample, 10 mL 70% EtOH
4IR1-E2	2649 No. 2 Sideroad	Influent	
4IR1-E3	2649 No. 2 Sideroad	Influent	
4IR1-P1	2649 No. 2 Sideroad	Influent	-80ºC Frozen, 10 Sample, 5 mL of Preservative Buffer
4IR1-P2	2649 No. 2 Sideroad	Influent	
4IR1-P3	2649 No. 2 Sideroad	Influent	
4IR1-1	2649 No. 2 Sideroad	Influent	DNA extracted immediately
4IR1-2	2649 No. 2 Sideroad	Influent	
4IR1-3	2649 No. 2 Sideroad	Influent	
4IR2-1	2649 No. 2 Sideroad	Tank	DNA extracted immediately
4IR2-2	2649 No. 2 Sideroad	Tank	
4IR2-3	2649 No. 2 Sideroad	Tank	
4RI1	2649 No. 2 Sideroad	Influent	-80ºC Frozen
4RI2	2649 No. 2 Sideroad	Tank	
4RI3	2649 No. 2 Sideroad	Effluent	

[1]: Numbers prefixing digester labels indicate a specific digester site replicate, and numbers suffixing digester labels indicate a sampling point, where 1=Influent, 2=Tank. Digester type is labeled as InnerTube Recirculating (IR). Letters following the hyphen represent the pre-treatment method, Control "C", Ethanol "E", Preserved "P", and numbers following the hyphen represent a technical replicate. The last three samples were not considered as technical replicates and were used to directly compare to other digester types.

Table S 3: Anaerobic septic tank volumes, flow-rates, and hydraulic retention (residence) times

Digester ID[1]	Sample ID[1]	Tank Residence Volume (L)	Flow Rate (L/day)	Hydraulic Residence Time (days)
1IS	1IS	4830	800	6.0
2IS	2IS	6060	800	7.6
3IS	3IS	3600	1000	3.6
3IR	1IR	24000	4000	6.0
4IR	2IR	13000	2625	5.0
5IR	3IR	17280	2888	6.0
1CS	1CS	9170	2000	4.6
2CS	2CS	9170	2000	4.6
4CS	3CS	5400	1000	5.4
4CR	1CR	5400	1250	4.3
5CR	2CR	9173	1750	5.2
6CR	3CR	9170	2250	4.1

[1]: Numbers prefixing digester labels indicate a specific digester site replicate. Digester types are labeled as InnerTube Recirculating (IR), InnerTube Single-pass (IS), Conventional Recirculating (CR), Conventional Single-pass (CS).

6.2 Appendix II: Sequencing Statistics

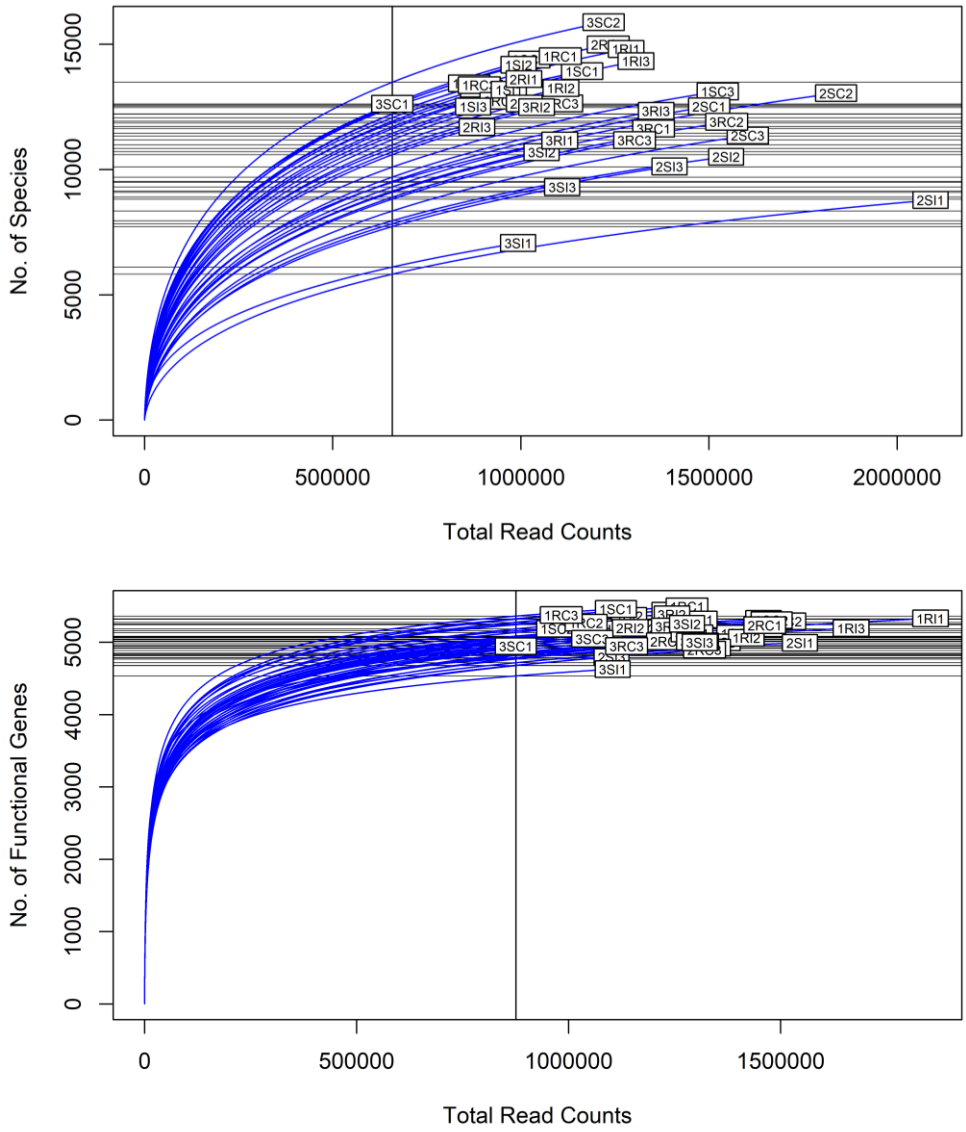


Figure 6.1: Rarefaction curve of A.) unique species vs. total read counts and B.) functional genes vs. total read counts. Vertical lines indicate the sample with the lowest total read counts. Horizontal lines indicate the number of species or genes obtained if rarefied to the sample with the lowest read count.

Table S 4: Qubit DNA concentrations and sample sequencing statistics

Digester ID[1]	Sample ID[2]	Qubit DNA Concentration (ng/uL)	Total Read Count	% GC Content	% Duplicate Reads
3IR1	1IR1	32.36	7927394	53	6.95
3IR2	1IR2	24.56	5860571	46	5.07
3IR3	1IR3	21.16	7129347	46	4.9
4IR1	2IR1	51.60	6164751	52	3.41
4IR2	2IR2	5.16	5067283	50	3.21
4IR3	2IR3	6.68	4505234	49	4.27
5IR1	3IR1	40.40	5465333	54	3.49
5IR2	3IR2	61.60	5408314	52	3.06
5IR3	3IR3	24.32	6135332	50	5.73
4IR1-1	4IR1-1	146.00	6298084	52	2.9
4IR1-2	4IR1-2	131.20	7207749	51	3.47
4IR1-3	4IR1-3	106.40	7665923	52	3.53
4IR2-1	4IR2-1	91.60	5650413	46	3.42
4IR2-2	4IR2-2	101.20	5638008	46	2.84
4IR2-3	4IR2-3	42.40	5241021	46	3.12
1IS1	1IS1	8.20	5071232	54	4.32
1IS2	1IS2	13.76	6317098	50	4.54
1IS3	1IS3	19.88	5249390	53	4.27
4IS1	2IS1	7.32	5813013	49	5.8
4IS2	2IS2	14.12	6031819	51	6.7
4IS3	2IS3	13.38	5441899	50	7.55
6IS1	3IS1	34.48	4630674	49	3.69
6IS2	3IS2	33.72	6253923	50	9.48
6IS3	3IS3	13.68	6035461	50	8.58
4CR1	1CR1	30.40	7120059	48	3.7
4CR2	1CR2	24.60	5463739	48	3.43

4CR3	1CR3	33.20	5133494	51	3.83
5CR1	2CR1	59.80	7615328	48	5.7
5CR2	2CR2	14.14	5550975	54	4.2
5CR3*	2CR3*	5.56	5890315	50	5.63
6CR1	3CR1	19.20	5750570	47	4.13
6CR2	3CR2	10.22	6144909	47	4.44
6CR3	3CR3	60.00	5312626	46	4.03
1CS1	1CS1	102.00	5759504	57	3.34
1CS2	1CS2	26.20	5556880	61	3.86
1CS3	1CS3	13.72	7259837	49	7
2CS1	2CS1	30.20	5118587	49	4.21
2CS2	2CS2	19.80	6593147	50	5.71
2CS3	2CS3	30.20	5595596	49	5.81
4CS1	3CS1	55.20	4564793	51	4.38
4CS2	3CS2	39.40	8381243	47	5.06
4CS3	3CS3	54.20	6115431	52	4.82
4IR1-C1	**4IR1-C1**	**1.12**	**2580272**	**52**	**3.57**
4IR1-C2	**4IR1-C2**	**0.97**	**2481820**	**49**	**3.3**
4IR1-C3	**4IR1-C3**	**0.99**	**2335608**	**48**	**3.27**
4IR1-EtOH1	**4IR2- EtOH 1**	**0.884**	**2475045**	**50**	**3.18**
4IR1-EtOH2	**4IR2- EtOH 2**	**0.464**	**2390762**	**48**	**3**
4IR1-EtOH3	**4IR2- EtOH 3**	**1.108**	**2373709**	**48**	**3.37**
4IR1-P1	**4IR2- P1**	**2.21**	**2589753**	**47**	**2.98**
4IR1-P2	**4IR2- P2**	**1.88**	**3503446**	**46**	**6.35**
4IR1-P3	**4IR2- P3**	**2.07**	**2612062**	**46**	**3.05**

[1*]: Asterisk on 5CR3 notes that this is the sample with the lowest concentration that unbolded samples were normalized to.
[2]: Sample ID, as opposed to the Digester Sample ID, refers to the IDs used in the thesis to represent the replicates of digesters chosen which are bolded in Table S 1. Bolded samples represent the sample group that was normalized to the concentration of 4IR1-EtOH 2 before sequencing. Samples that are not bolded were normalized to the concentration of 5CR3 (denoted with an asterisk).

6.3 Appendix III: Raw Statistical Test Values

Table S 5: PERMANOVA and BETADISPER Significance tests between Pre-treatment methods.

	Terms[1]	d*f*	Sum of Squares	Pseudo *f*	*p*-value
PERMANOVA	Pre-Treatment Method	3	0.142	2.609	0.001
BETADISPER	Pre-Treatment Method	3	0.006	15.118	0.001

[1]: Tank samples for Unpreserved samples are not included in the analysis to maintain a balanced test design

Table S 6: Genera differential abundance (DESeq2) statistics for A.) Conventional vs. InnerTube and B.) Recirculating vs. Single-pass

A.

Order	Family	Genus	Taxon_ID	Conventional Mean Abundance	InnerTube Mean Abundance	log-2 fold change	log-2 fold change SE	Adj. *p*
Desulfovibrionales	Desulfovibrionaceae	Desulfovibrio	184917	8259.06	72550.24	2.987	0.520	0.000001
Rhodobacterales	Rhodobacteraceae	Paracoccus	34003	656.12	4708.28	2.974	0.587	0.000019
Burkholderiales	Comamonadaceae	Acidovorax	1276755	3378.55	30136.98	2.949	0.572	0.000014
Pseudomonadales	Moraxellaceae	Acinetobacter	470	1256.61	6551.63	2.350	0.410	0.000001
Clostridiales	Ruminococcaceae	Faecalibacterium	853	4458.68	16775.72	1.741	0.435	0.000911
Clostridiales	Oscillospiraceae	Oscillibacter	351091	1328.56	4704.68	1.608	0.373	0.000341
Unclassified_Acidobacteria	Unclassified_Acidobacteria	Unclassified_Acidobacteria	1978231	3934.45	543.59	-1.980	0.452	0.000265
Desulfobacterales	Desulfobulbaceae	Desulfobulbus	894	23995.99	5194.96	-2.424	0.562	0.000338
Desulfovibrionales	Desulfomicrobiaceae	Desulfomicrobium	899	11342.24	704.76	-3.866	0.550	0.000000
Methanosarcinales	Methanosaetaceae	Methanothrix	2223	20897.36	2132.61	-3.899	0.917	0.000405
Clostridiales	Peptostreptococcaceae	Proteocatella	181070	20030.62	1703.37	-4.177	0.781	0.000006
Anaerolineales	Anaerolineaceae	Pelolinea	913107	8101.25	72.34	-5.068	0.743	0.000000

B.

Order	Family	Genus	Taxon_ID	Recirculating Mean Abundance	Single_Pass Mean Abundance	log2 fold-change	log2 fold-change SE	Adj. *p*
Burkholderiales	Alcaligenaceae	Pusillimonas	1548123	76.43	11698.20	5.44	0.66	0.000000
Burkholderiales	Comamonadaceae	Simplicispira	80881	610.23	13729.36	3.49	0.61	0.000002
Rhizobiales	Methylocystaceae	Pleomorphomonas	257440	729.97	6715.52	2.71	0.63	0.000368
Caulobacterales	Caulobacteraceae	Phenylobacterium	1736591	480.98	4426.47	2.29	0.53	0.000311
Rhizobiales	Unclassified_Rhizobiales	Unclassified_Rhizobiales	1909294	556.38	2260.47	1.92	0.40	0.000069
Bacteroidales	Unclassified_Bacteroidales	Unclassified_Bacteroidales	1400053	13833.65	2812.20	-1.97	0.45	0.000279
Acidithiobacillales	Acidithiobacillaceae	Acidithiobacillus	930	2364.33	365.51	-2.66	0.63	0.000493
Methanomicrobiales	Methanospirillaceae	Methanosphaerula	475088	2307.15	210.06	-2.88	0.64	0.000219
Rhodocyclales	Azonexaceae	Dechloromonas	73030	14762.60	1095.21	-3.56	0.67	0.000008
Campylobacterales	Helicobacteraceae	Sulfurimonas	39766	4477.83	322.82	-3.60	0.67	0.000007
Unclassified_Candidatus Cloacimonetes	Unclassified_Candidatus Cloacimonetes	Unclassified_Candidatus Cloacimonetes	2013734	2927.01	273.69	-4.10	0.66	0.000000
Desulfuromonadales	Geobacteraceae	Geobacter	115783	42044.47	5902.90	-4.28	0.71	0.000001
Campylobacterales	Campylobacteraceae	Pseudoarcobacter	1912877	21593.94	510.36	-5.25	0.82	0.000000

Table S 7: Genera differential abundance (DESeq2) statistics for A.) Influent vs. Tank and Tank vs. Effluent comparisons and B.) Influent vs. Effluent comparisons

A.

Order	Family	Genus	Taxon_ID	Influent Mean Abundance	Tank Mean Abundance	Effluent Mean Abundance	log2 fold-change	log2 fold-change SE	Adj. *p*
Desulfobacterales	Desulfobacteraceae	Desulfamplus	1246637		120.78	1100.02	3.74	0.87	0.038
Chromatiales	Halothiobacillaceae	Halothiobacillus	1970383		105.71	1168.74	3.05	0.73	0.038
Unclassified_Verrucomicrobia	Unclassified_Verrucomicrobia	Unclassified_Verrucomicrobia	2026799	1800.93	8469.27		2.01	0.58	0.037
Unclassified_Deltaproteobacteria	Unclassified_Deltaproteobacteria	Unclassified_Deltaproteobacteria	34034	3139.86	7401.22		1.34	0.40	0.040
Desulfovibrionales	Desulfonatronaceae	Desulfonatronum	617001	87.94	189.17		1.12	0.34	0.050
Clostridiales	Lachnospiraceae	Unclassified_Lachnospiraceae	39491	6556.00	2635.23		-1.29	0.37	0.034
Clostridiales	Lachnospiraceae	Blautia	40520	8457.20	3309.80		-1.45	0.38	0.025
Erysipelotrichales	Erysipelotrichaceae	Holdemania	61171	387.04	88.48		-1.49	0.45	0.044
Erysipelotrichales	Erysipelotrichaceae	Erysipelatoclostridium	29348	460.98	140.71		-1.50	0.42	0.034
Clostridiales	Lachnospiraceae	Anaerobutyricum	39488	708.71	347.51		-1.51	0.42	0.034
Clostridiales	Lachnospiraceae	Dorea	88431	2148.43	721.14		-1.63	0.43	0.025
Clostridiales	Ruminococcaceae	Ruminococcus	1262952	12706.34	4811.29		-1.64	0.46	0.034
Clostridiales	Lachnospiraceae	Anaerostipes	649756	736.97	288.06		-1.69	0.45	0.025
Clostridiales	Ruminococcaceae	Unclassified_Ruminococcaceae	2048667	7018.97	1861.86		-1.80	0.48	0.028
Coriobacteriales	Coriobacteriaceae	Collinsella	74426	6693.16	442.26		-2.31	0.58	0.019
Propionibacteriales	Propionibacteriaceae	Microlunatus	29405	738.67	223.60		-2.49	0.74	0.040
Eggerthellales	Eggerthellaceae	Asaccharobacter	394340	798.62	94.95		-2.73	0.66	0.019
Bifidobacteriales	Bifidobacteriaceae	Bifidobacterium	216816	56560.44	4173.37		-3.05	0.76	0.019
Enterobacterales	Yersiniaceae	Yersinia	419257	8732.83	152.40		-3.07	0.82	0.026
Thiotrichales	Thiotrichaceae	Thiothrix	525917	2524.77	266.32		-3.52	1.03	0.040

B.

Order	Family	Genus	Taxon_ID	Influent Mean Abundance	Effluent Mean Abundance	log2-fold change	log2-fold change SE	Adj. *p*
Desulfobacterales	Desulfobacteraceae	Desulfamplus	1246637	64.80	1100.02	5.08	0.87	0.00001
Thiotrichales	Unclassified_Thiotrichales	Unclassified_Thiotrichales	1970606	968.23	22629.14	5.07	1.01	0.00020
Desulfuromonadales	Geobacteraceae	Geobacter	115783	4805.67	49406.87	4.46	0.88	0.00019
Rhodospirillales	Acetobacteraceae	Zavarzinia	1264899	31.46	301.75	3.90	0.85	0.00085
Desulfobacterales	Desulfobacteraceae	Desulfobacter	2293	98.66	1705.75	3.78	0.76	0.00024
Thiotrichales	Piscirickettsiaceae	Thiomicrospira	92245	37.15	432.32	3.53	0.74	0.00038
Chromatiales	Halothiobacillaceae	Halothiobacillus	1970383	63.42	1168.74	3.52	0.73	0.00038

Methanosarcinales	Methanosarcinaceae	Methanomethylovorans	101192	133.83	780.35	2.88	0.99	0.04686
Desulfuromonadales	Geobacteraceae	Unclassified_Geobacteraceae	1798316	110.73	636.07	2.76	0.62	0.00086
Desulfuromonadales	Unclassified_Desulfuromonadales	Unclassified_Desulfuromonadales	2099678	56.73	295.75	2.59	0.63	0.00319
Burkholderiales	Burkholderiaceae	Polynucleobacter	1970427	62.21	367.92	2.42	0.54	0.00086
Desulfobacterales	Desulfobacteraceae	Desulfobacterium	201089	80.96	398.35	2.24	0.57	0.00446
Desulfobacterales	Desulfobulbaceae	Desulfobulbus	894	5708.05	23025.73	2.08	0.69	0.03785
Desulfobacterales	Unclassified_Desulfobacterales	Unclassified_Desulfobacterales	1973983	96.70	462.29	2.04	0.52	0.00452
Desulfobacterales	Desulfobulbaceae	Unclassified_Desulfobulbaceae	1629713	258.61	932.36	1.94	0.57	0.01575
Unclassified_Gammaproteobacteria	Unclassified_Gammaproteobacteria	Unclassified_Gammaproteobacteria	1970513	2080.47	7486.02	1.86	0.53	0.01039
Unclassified_Nitrospirae	Unclassified_Nitrospirae	Unclassified_Nitrospirae	1801696	64.68	218.80	1.83	0.41	0.00086
Nitrosomonadales	Methylophilaceae	Methylotenera	2051956	168.88	741.00	1.75	0.54	0.02376
Nitrosomonadales	Unclassified_Nitrosomonadales	Unclassified_Nitrosomonadales	1970535	150.45	391.57	1.49	0.47	0.02466
Desulfuromonadales	Desulfuromonadaceae	Desulfuromonas	892	84.94	222.24	1.46	0.40	0.00816
Desulfovibrionales	Desulfonatronaceae	Desulfonatronum	617001	87.94	260.49	1.43	0.34	0.00251
Unclassified_Deltaproteobacteria	Unclassified_Deltaproteobacteria	Unclassified_Deltaproteobacteria	34034	3139.86	7376.25	1.28	0.40	0.02376
Clostridiales	Peptococcaceae	Dehalobacterium	51515	152.85	308.96	1.24	0.42	0.04321
Clostridiales	Clostridiaceae	Butyricicoccus	2292297	726.01	382.39	-0.94	0.30	0.02761
Clostridiales	Lachnospiraceae	Unclassified_Lachnospiraceae	39491	6556.00	2643.94	-1.23	0.37	0.01637
Corynebacteriales	Mycobacteriaceae	Mycobacteroides	36809	326.88	125.85	-1.27	0.44	0.04964
Clostridiales	Lachnospiraceae	Lachnospira	446043	385.88	163.85	-1.33	0.44	0.03785
Clostridiales	Lachnospiraceae	Lachnotalea	1763509	472.22	169.01	-1.34	0.43	0.03001
Clostridiales	Lachnospiraceae	Coprococcus	410072	2128.39	854.18	-1.41	0.46	0.03183
Erysipelotrichales	Erysipelotrichaceae	Holdemania	61171	387.04	96.80	-1.42	0.45	0.02498
Xanthomonadales	Xanthomonadaceae	Lysobacter	1604334	369.39	131.59	-1.47	0.43	0.01597
Clostridiales	Lachnospiraceae	Fusicatenibacter	1150298	368.46	144.99	-1.50	0.51	0.04648
Clostridiales	Lachnospiraceae	Blautia	40520	8457.20	3028.91	-1.50	0.38	0.00453
Propionibacteriales	Nocardioidaceae	Nocardioides	2049295	467.29	139.78	-1.52	0.50	0.03496
Clostridiales	Lachnospiraceae	Anaerobutyricum	39488	708.71	330.32	-1.53	0.42	0.00831
Clostridiales	Ruminococcaceae	Faecalibacterium	853	17144.48	6606.29	-1.56	0.53	0.04530
Clostridiales	Lachnospiraceae	Anaerostipes	649756	736.97	277.33	-1.67	0.45	0.00706
Rhizobiales	Rhizobiaceae	Agrobacterium	358	287.97	97.51	-1.68	0.51	0.02070
Clostridiales	Ruminococcaceae	Unclassified_Ruminococcaceae	2048667	7018.97	1945.48	-1.79	0.48	0.00798
Rhizobiales	Bradyrhizobiaceae	Bosea	406341	867.65	216.91	-1.79	0.54	0.01665
Clostridiales	Lachnospiraceae	Dorea	88431	2148.43	632.50	-1.80	0.43	0.00195
Xanthomonadales	Xanthomonadaceae	Luteimonas	2172536	380.10	111.09	-1.83	0.54	0.01618
Clostridiales	Ruminococcaceae	Ruminococcus	1262952	12706.34	4053.27	-1.87	0.46	0.00346
Micrococcales	Unclassified_Micrococcales	Unclassified_Micrococcales	1895792	506.41	134.41	-1.89	0.59	0.02376
Rhizobiales	Chelatococcaceae	Chelatococcus	1953771	604.65	110.52	-1.91	0.64	0.03986
Micrococcales	Microbacteriaceae	Microbacterium	51671	5236.07	1192.08	-1.96	0.58	0.01623
Rhodobacterales	Rhodobacteraceae	Rhodovulum	35806	398.27	100.57	-2.04	0.59	0.01371
Burkholderiales	Comamonadaceae	Ottowia	2109914	2530.51	550.12	-2.14	0.59	0.00850
Xanthomonadales	Xanthomonadaceae	Pseudoxanthomonas	314722	1168.36	268.73	-2.14	0.59	0.00831
Coriobacteriales	Coriobacteriaceae	Collinsella	74426	6693.16	491.21	-2.19	0.58	0.00598
Eggerthellales	Eggerthellaceae	Asaccharobacter	394340	798.62	108.10	-2.36	0.66	0.01016
Propionibacteriales	Propionibacteriaceae	Microlunatus	29405	738.67	188.81	-2.70	0.74	0.00816

Bifidobacteriales	Bifidobacteriaceae	Bifidobacterium	216816	56560.44	3787.55	-2.91	0.76	0.00570
Enterobacterales	Yersiniaceae	Yersinia	419257	8732.83	379.76	-2.92	0.82	0.00997
Nakamurellales	Nakamurellaceae	Nakamurella	53461	3690.16	198.07	-3.12	0.83	0.00679
Rhizobiales	Rhizobiaceae	Shinella	1870904	5056.36	708.99	-3.20	0.82	0.00474
Rhizobiales	Methylocystaceae	Pleomorphomonas	257440	9278.89	723.64	-3.39	0.77	0.00107
Rhodobacterales	Rhodobacteraceae	Amaricoccus	1985299	2017.84	184.88	-3.59	0.80	0.00086

Table S 8: Functional Gene subsystem differential abundance (DESeq2) of A.) Conventional and InnerTube and B.) Recirculating and Single-pass digesters

A.

L1 Subsystem	L3 Subsystem	Conventional Mean Abundance	InnerTube Mean Abundance	log2 fold-change	log2 fold-change SE	Adj. *p*
Motility and Chemotaxis	Flagellum	3520.85	5077.21	0.54	0.14	0.0022
Respiration	Respiration / HGM	4179.01	5037.86	0.27	0.10	0.0436
RNA Metabolism	Queuosine exploration RZ	3243.42	3783.92	0.22	0.08	0.0246
Cofactors, Vitamins, Prosthetic Groups, Pigments	Lipoic Acid Synthesis	2365.21	2743.82	0.21	0.06	0.0063
Amino Acids and Derivatives	Amino acid racemase	1973.70	2243.63	0.18	0.06	0.0175
Mitochondrial electron transport system in plants	F0F1-type ATP synthase in plants (mitochondrial)	6364.90	7190.06	0.18	0.06	0.0135
Potassium metabolism	Potassium homeostasis	7595.02	8518.92	0.16	0.05	0.0178
Fatty Acids, Lipids, and Isoprenoids	CDP-diacylglycerol biosynthesis in plants	1762.56	1930.87	0.13	0.05	0.0337
Carbohydrates	Galactose degradation in plants	3055.90	2687.65	-0.19	0.07	0.0402
Cell Wall and Capsule	Capsular heptose biosynthesis	2125.33	1813.67	-0.23	0.09	0.0428
Cell Wall and Capsule	dTDP-rhamnose synthesis	4684.28	3871.64	-0.28	0.08	0.0041
DNA Metabolism	Restriction-Modification System	18116.42	14415.88	-0.33	0.07	0.0001
Amino Acids and Derivatives	Aromatic amino acid interconversions with aryl acids	4168.15	3218.75	-0.39	0.09	0.0003
DNA Metabolism	CRISPRs	2593.65	1940.32	-0.42	0.14	0.0192
Respiration	Hydrogenases	4323.58	3313.25	-0.43	0.16	0.0377
Nucleotide sugars	UDP-glucose and UDP-galactose biosynthesis in plants	1661.81	1230.47	-0.47	0.15	0.0178
Nitrogen Metabolism	Nitrosative stress	3196.68	2283.03	-0.50	0.14	0.0068
Virulence	Beta-lactamase	3076.31	2141.88	-0.54	0.11	0.0000
Carbohydrates	Methanogenesis	2023.35	1017.69	-1.04	0.28	0.0032

B.

L1 Subsystem	L3 Subsystem	Recirculating Mean Abundance	Single Pass Mean Abundance	log2 fold-change	log2 fold-change SE	Adj. *p*
Virulence	Type 4 secretion and conjugative transfer	1858.8	2876.31	0.7	0.26	0.036
Stress Response	Choline and Betaine Uptake and Betaine Biosynthesis	1688.64	2665.59	0.64	0.2	0.012
Sulfur Metabolism	Alkanesulfonate assimilation	1139.2	1766.74	0.6	0.19	0.013
Membrane Transport	ABC transporter dipeptide (TC 3.A.1.5.2)	2400.65	3396.14	0.51	0.12	0.001
Carbohydrates	Glycerol fermentation to 1,3-propanediol	1124.49	1585.97	0.49	0.15	0.011
Carbohydrates	D-ribose utilization	3256.73	4507.8	0.48	0.14	0.006
Carbohydrates	Acetyl-CoA biosynthesis in plants	2209.6	3030.69	0.46	0.15	0.016
Carbohydrates	D-galactarate, D-glucarate, and D-glycerate catabolism	1318.69	1817.15	0.46	0.11	0.001
Amino Acids and Derivatives	Aromatic amino acid degradation	1065.13	1425.1	0.45	0.15	0.019

DNA Metabolism	DNA repair, bacterial RecBCD pathway	1842.65	2528.38	0.45	0.11	0.002
Amino Acids and Derivatives	Cinnamic Acid Degradation	1142.04	1518.92	0.4	0.13	0.015
Carbohydrates	Fermentations in Streptococci	3408.06	4295.71	0.34	0.09	0.004
Carbohydrates	Acetone Butanol Ethanol Synthesis	2826.19	3438.15	0.28	0.06	0.004
Virulence	Arsenic resistance	1150.5	1394.6	0.27	0.1	0.032
Miscellaneous	IojapClusters	2204.03	2579.28	0.23	0.08	0.02
Amino Acids and Derivatives	Arginine Biosynthesis extended	8706.97	8237.05	-0.08	0.03	0.034
Amino Acids and Derivatives	Branched-Chain Amino Acid Biosynthesis	14693.58	13748.07	-0.1	0.03	0.008
Miscellaneous	YebC	4082.6	3815.42	-0.1	0.04	0.041
Protein Metabolism	EC 6.1.1.- Ligases forming aminoacyl-tRNA and related compounds	18145.05	16936.56	-0.1	0.03	0.025
Amino Acids and Derivatives	Lysine and threonine metabolism in plants	9689.48	9006.22	-0.11	0.04	0.026
Protein Metabolism	Translation elongation factors bacterial	6095.46	5638.34	-0.11	0.03	0.006
Cofactors, Vitamins, Prosthetic Groups, Pigments	5-FCL-like protein	2882.78	2660.49	-0.12	0.05	0.049
Cell Division and Cell Cycle	Bacterial Cytoskeleton	16207.9	14806.36	-0.13	0.05	0.025
Central metabolism	EC 6.4.1.- Ligases that form carbon-carbon bonds	2274.9	2083.37	-0.13	0.05	0.033
Cofactors, Vitamins, Prosthetic Groups, Pigments	CLO thiaminPP biosynthesis	9030.91	8246.85	-0.13	0.05	0.046
RNA Metabolism	RNA polymerase bacterial	10653.39	9512.53	-0.16	0.04	0.002
Amino Acids and Derivatives	Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis	10009.13	8908.96	-0.17	0.04	0.002
Amino Acids and Derivatives	Histidine Biosynthesis	8197.05	7261.99	-0.17	0.05	0.009
Protein Metabolism	Translation initiation factors bacterial	3448.98	3045.95	-0.18	0.05	0.003
Virulence	Cobalt-zinc-cadmium resistance	3011.25	2639.14	-0.19	0.07	0.043
Protein Metabolism	Translation elongation factor G family	5239.33	4561.9	-0.2	0.06	0.008
Amino Acids and Derivatives	Glutamine synthetases	3247.82	2797.21	-0.21	0.07	0.011
Fatty Acids, Lipids, and Isoprenoids	Isoprenoid Biosynthesis	4047.26	3461.07	-0.23	0.06	0.005
Protein Metabolism	Ribosomal protein S12p Asp methylthiotransferase	1725.39	1464.52	-0.24	0.07	0.01
Amino Acids and Derivatives	Arginine metabolism and urea cycle in plants	5177.32	4321.89	-0.26	0.06	0.001
Amino Acids and Derivatives	Aromatic amino acid interconversions with aryl acids	4050.37	3336.54	-0.31	0.09	0.006
Motility and Chemotaxis	Flagellar motility	6079.68	4953.67	-0.32	0.12	0.037
Virulence	Multidrug Resistance Efflux Pumps	7163.9	5650.76	-0.35	0.12	0.027
Respiration	Respiratory Complex I	8796.12	6869.29	-0.36	0.13	0.034
Nitrogen Metabolism	Nitrate and nitrite ammonification	6178.46	4668.26	-0.4	0.15	0.043
Stress Response	FOL Commensurate regulon activation	4082.12	3094.2	-0.41	0.13	0.012
Nucleotide sugars	UDP-glucose and UDP-galactose biosynthesis in plants	1636.04	1256.24	-0.42	0.15	0.034
Carbohydrates	Mannose Metabolism	4176.56	3053.34	-0.44	0.13	0.009
Amino Acids and Derivatives	PII Superfamily	6336.34	4629.47	-0.46	0.13	0.006
Respiration	Hydrogenases	4393.07	3243.76	-0.48	0.16	0.017
Carbohydrates	Propionyl-CoA to Succinyl-CoA Module	2969.75	2075.06	-0.52	0.16	0.011
Iron acquisition and metabolism	Iron acquisition in Vibrio	17059.1	9875.04	-0.78	0.28	0.027
Phages, Prophages, Transposable elements, Plasmids	Conjugative transposon, Bacteroidales	4196.02	2310.04	-1.16	0.4	0.025

Table S 9: Functional Gene subsystem differential abundance (DESeq2) of Influent and Effluent comparison.

Category	Group	Influent Mean Abundance	Effluent Mean Abundance	log2 fold-change	log2 fold-change SE	Adj. *p*
Carbohydrates	Pyruvate metabolism I: anaplerotic reactions, PEP in Mycobacteria	65.288	217.960	1.725	0.432	0.019
Stress Response	SigmaB stress responce regulation	443.355	916.860	1.087	0.278	0.020
Virulence	Zinc resistance	883.641	1855.887	1.030	0.308	0.049
Cell Wall and Capsule	Lipid A-Ara4N pathway (Polymyxin resistance)	911.053	1585.135	0.812	0.221	0.038
Respiration	Hydrogenases	2942.025	4501.380	0.643	0.194	0.049
Protein Metabolism	Selenocysteine metabolism	1203.529	1735.457	0.532	0.150	0.040
Central metabolism	EC 6.4.1.- Ligases that form carbon-carbon bonds	1982.259	2341.176	0.244	0.059	0.015
Protein Metabolism	Proteasome bacterial	6144.470	7202.743	0.229	0.051	0.005
Protein Metabolism	Translation elongation factors bacterial	5555.174	6134.552	0.143	0.039	0.038
Carbohydrates	2-O-alpha-mannosyl-D-glycerate utilization	541.945	266.084	-0.973	0.281	0.040
Iron acquisition and metabolism	Ferrous iron transporter EfeUOB, low-pH-induced	277.697	96.327	-1.490	0.434	0.040

7 References

Abbasi, T., Tauseef, S. M., & Abbasi, S. A. (2012). Low-Rate and High-Rate Anaerobic Reactors/Digesters/Fermenters. In T. Abbasi, S. M. Tauseef, & S. A. Abbasi (Eds.), *Biogas Energy* (pp. 35-39). Springer New York. https://doi.org/10.1007/978-1-4614-1040-9_4

Ai, C., Yan, Z., Zhou, H., Hou, S., Chai, L., Qiu, G., & Zeng, W. (2019). Metagenomic Insights into the Effects of Seasonal Temperature Variation on the Activities of Activated Sludge. *Microorganisms, 7*(12), 713. https://www.mdpi.com/2076-2607/7/12/713

Akizuki, S., Matsuyama, T., & Toda, T. (2016, 2016/09/01/). An anaerobic-aerobic sequential batch system using simultaneous organic and nitrogen removal to treat intermittently discharged organic solid wastes. *Process Biochemistry, 51*(9), 1264-1273. https://doi.org/https://doi.org/10.1016/j.procbio.2016.05.011

Aklujkar, M., Krushkal, J., DiBartolo, G., Lapidus, A., Land, M. L., & Lovley, D. R. (2009). The genome sequence of Geobacter metallireducens: features of metabolism, physiology and regulation common and dissimilar to Geobacter sulfurreducens. *BMC microbiology, 9*, 109-109. https://doi.org/10.1186/1471-2180-9-109

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, Nov). Binning metagenomic contigs by coverage and composition. *Nat Methods, 11*(11), 1144-1146. https://doi.org/10.1038/nmeth.3103

Alzate Marin, J. C., Caravelli, A. H., & Zaritzky, N. E. (2016, Jan). Nitrification and aerobic denitrification in anoxic-aerobic sequencing batch reactor. *Bioresour Technol, 200*, 380-387. https://doi.org/10.1016/j.biortech.2015.10.024

Andersen, K. S., Kirkegaard, R. H., Karst, S. M., & Albertsen, M. (2018). ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. *bioRxiv*, 299537. https://doi.org/10.1101/299537

Anderson, M., & Santana-Garcon, J. (2014, 10/01). Measures of precision for dissimilarity-based multivariate analysis of ecological communities. *Ecology Letters, 18*. https://doi.org/10.1111/ele.12385

Anderson, M. J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA). *Wiley StatsRef: Statistics Reference Online*, 1-15. https://doi.org/10.1002/9781118445112.stat07841

Arbizu, P. M. (2017). *pairwiseAdonis: Pairwise Multilevel Comparison using Adonis.* In (Version 0.01) https://github.com/pmartinezarbizu/pairwiseAdonis

Badalamenti, J. P., Summers, Z. M., Chan, C. H., Gralnick, J. A., & Bond, D. R. (2016). Isolation and Genomic Characterization of 'Desulfuromonas soudanensis WTL', a Metal- and Electrode-Respiring Bacterium from Anoxic Deep Subsurface Brine. *Frontiers in Microbiology, 7*, 913-913. https://doi.org/10.3389/fmicb.2016.00913

Bag, S., Saha, B., Mehta, O., Anbumani, D., Kumar, N., Dayal, M., Pant, A., Kumar, P., Saxena, S., Allin, K. H., Hansen, T., Arumugam, M., Vestergaard, H., Pedersen, O., Pereira, V., Abraham, P., Tripathi, R., Wadhwa, N., Bhatnagar, S., Prakash, V. G., Radha, V., Anjana, R. M., Mohan, V., Takeda, K., Kurakawa, T., Nair, G. B., & Das, B. (2016, May 31). An Improved Method for High Quality Metagenomics DNA Extraction from Human and Environmental Samples. *Sci Rep, 6*, 26775. https://doi.org/10.1038/srep26775

Bashiardes, S., Zilberman-Schapira, G., & Elinav, E. (2016). Use of Metatranscriptomics in Microbiome Research. *Bioinformatics and Biology Insights, 10*, 19-25. https://doi.org/10.4137/Bbi.S34610

Beiras, R. (2018). Chapter 1 - Basic Concepts. In R. Beiras (Ed.), *Marine Pollution* (pp. 3-20). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-12-813736-9.00001-5

Bengtsson, S., Hallquist, J., Werker, A., & Welander, T. (2008). Acidogenic fermentation of industrial wastewaters: Effects of chemostat retention time

and pH on volatile fatty acids production. *Biochemical Engineering Journal, 40*(3), 492-499. https://doi.org/10.1016/j.bej.2008.02.004

Bertucci, M., Calusinska, M., Goux, X., Rouland-Lefevre, C., Untereiner, B., Ferrer, P., Gerin, P. A., & Delfosse, P. (2019, Aug 1). Carbohydrate Hydrolytic Potential and Redundancy of an Anaerobic Digestion Microbiome Exposed to Acidosis, as Uncovered by Metagenomics. *Appl Environ Microbiol, 85*(15). https://doi.org/10.1128/AEM.00895-19

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., Huttley, G. A., & Gregory Caporaso, J. (2018, 2018/05/17). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome, 6*(1), 90. https://doi.org/10.1186/s40168-018-0470-z

Bolger, A. M., Lohse, M., & Usadel, B. (2014, Aug 1). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Braguglia, C. M., Gallipoli, A., Gianico, A., & Pagliaccia, P. (2018, Jan). Anaerobic bioconversion of food waste into energy: A critical review. *Bioresour Technol, 248*(Pt A), 37-56. https://doi.org/10.1016/j.biortech.2017.06.145

Brown, C. T., Hug, L. A., Thomas, B. C., Sharon, I., Castelle, C. J., Singh, A., Wilkins, M. J., Wrighton, K. C., Williams, K. H., & Banfield, J. F. (2015, Jul 9). Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature, 523*(7559), 208-211. https://doi.org/10.1038/nature14486

Buchfink, B., Xie, C., & Huson, D. H. (2015, Jan). Fast and sensitive protein alignment using DIAMOND. *Nat Methods, 12*(1), 59-60. https://doi.org/10.1038/nmeth.3176

Buettner, C., von Bergen, M., Jehmlich, N., & Noll, M. (2019, 2019/09/06). Pseudomonas spp. are key players in agricultural biogas substrate degradation. *Scientific Reports, 9*(1), 12871. https://doi.org/10.1038/s41598-019-49313-8

Buratti, F. M., Manganelli, M., Vichi, S., Stefanelli, M., Scardala, S., Testai, E., & Funari, E. (2017, Mar). Cyanotoxins: producing organisms, occurrence, toxicity, mechanism of action and human health toxicological risk evaluation. *Arch Toxicol, 91*(3), 1049-1130. https://doi.org/10.1007/s00204-016-1913-6

Buttigieg, P. L., & Ramette, A. (2014, Dec). A guide to statistical analysis in microbial ecology: a community-focused, living review of multivariate data analyses. *FEMS Microbiol Ecol, 90*(3), 543-550. https://doi.org/10.1111/1574-6941.12437

Cai, M., Wilkins, D., Chen, J., Ng, S. K., Lu, H., Jia, Y., & Lee, P. K. (2016). Metagenomic Reconstruction of Key Anaerobic Digestion Pathways in Municipal Sludge and Industrial Wastewater Biogas-Producing Systems. *Front Microbiol, 7*, 778. https://doi.org/10.3389/fmicb.2016.00778

Calusinska, M., Goux, X., Fossépré, M., Muller, E. E. L., Wilmes, P., & Delfosse, P. (2018, 2018/07/19). A year of monitoring 20 mesophilic full-scale bioreactors reveals the existence of stable but different core microbiomes in bio-waste and wastewater anaerobic digestion systems. *Biotechnology for Biofuels, 11*(1), 196. https://doi.org/10.1186/s13068-018-1195-8

Campanaro, S., Treu, L., Kougias, P. G., De Francisci, D., Valle, G., & Angelidaki, I. (2016). Metagenomic analysis and functional characterization of the biogas microbiome using high throughput shotgun sequencing and a novel binning strategy. *Biotechnol Biofuels, 9*, 26. https://doi.org/10.1186/s13068-016-0441-1

Campanaro, S., Treu, L., Kougias, P. G., Luo, G., & Angelidaki, I. (2018, Sep 1). Metagenomic binning reveals the functional roles of core abundant microorganisms in twelve full-scale biogas plants. *Water Res, 140*, 123-134. https://doi.org/10.1016/j.watres.2018.04.043

Carballa, M., Regueiro, L., & Lema, J. M. (2015, Jun). Microbial management of anaerobic digestion: exploiting the microbiome-functionality nexus. *Curr Opin Biotechnol, 33*, 103-111. https://doi.org/10.1016/j.copbio.2015.01.008

Černý, M., Vítězová, M., Vítěz, T., Bartoš, M., & Kushkevych, I. (2018). Variation in the Distribution of Hydrogen Producers from the Clostridiales

Order in Biogas Reactors Depending on Different Input Substrates. *Energies, 11*(12), 3270. https://www.mdpi.com/1996-1073/11/12/3270

Chen, C., Xu, X.-J., Xie, P., Yuan, Y., Zhou, X., Wang, A.-J., Lee, D.-J., & Ren, N.-Q. (2017, 2017/03/01/). Pyrosequencing reveals microbial community dynamics in integrated simultaneous desulfurization and denitrification process at different influent nitrate concentrations. *Chemosphere, 171*, 294-301. https://doi.org/https://doi.org/10.1016/j.chemosphere.2016.11.159

Chen, C., Xu, X. J., Xie, P., Yuan, Y., Zhou, X., Wang, A. J., Lee, D. J., & Ren, N. Q. (2017, Mar). Pyrosequencing reveals microbial community dynamics in integrated simultaneous desulfurization and denitrification process at different influent nitrate concentrations. *Chemosphere, 171*, 294-301. https://doi.org/10.1016/j.chemosphere.2016.11.159

Chen, J., Wade, M. J., Dolfing, J., & Soyer, O. S. (2019, 2019/05/31). Increasing sulfate levels show a differential impact on synthetic communities comprising different methanogens and a sulfate reducer. *Journal of The Royal Society Interface, 16*(154), 20190129. https://doi.org/10.1098/rsif.2019.0129

Chen, S. S., Li, N., Dong, B., Zhao, W. T., Dai, L. L., & Dai, X. H. (2018, Jan 15). New insights into the enhanced performance of high solid anaerobic digestion with dewatered sludge by thermal hydrolysis: Organic matter degradation and methanogenic pathways. *Journal of Hazardous Materials, 342*, 1-9. https://doi.org/10.1016/j.jhazmat.2017.08.012

Chu, L., & Wang, J. (2017). Denitrification of groundwater using a biodegradable polymer as a carbon source: long-term performance and microbial diversity [10.1039/C7RA11151G]. *RSC Advances, 7*(84), 53454-53462. https://doi.org/10.1039/C7RA11151G

Clarke, K., & Warwick, R. (2001). Clarke KR, Warwick RM.Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. Primer-E Ltd: Plymouth, UK. In.

Clarke, K. R. (1993, 1993/03/01). Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology, 18*(1), 117-143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

Conn, K. E., Barber, L. B., Brown, G. K., & Siegrist, R. L. (2006, Dec 1). Occurrence and fate of organic contaminants during onsite wastewater treatment. *Environ Sci Technol*, *40*(23), 7358-7366. https://doi.org/10.1021/es0605117

Cookey, P. E., Koottatep, T., van der Steen, P., & Lens, P. N. L. (2016). Public health risk assessment tool: strategy to improve public policy framework for onsite wastewater treatment systems (OWTS). *Journal of Water, Sanitation and Hygiene for Development*, *6*(1), 74-88. https://doi.org/10.2166/washdev.2016.081

Cydzik-Kwiatkowska, A., & Zielinska, M. (2016, Apr). Bacterial communities in full-scale wastewater treatment systems. *World J Microbiol Biotechnol*, *32*(4), 66. https://doi.org/10.1007/s11274-016-2012-9

Daly, A. J., Baetens, J. M., & De Baets, B. (2018). Ecological Diversity: Measuring the Unmeasurable. *Mathematics*, *6*(7), 119. https://www.mdpi.com/2227-7390/6/7/119

Darling, A. E., Jospin, G., Lowe, E., Matsen, F. A. t., Bik, H. M., & Eisen, J. A. (2014). PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ*, *2*, e243. https://doi.org/10.7717/peerj.243

De Francisci, D., Kougias, P. G., Treu, L., Campanaro, S., & Angelidaki, I. (2015, Jan). Microbial diversity and dynamicity of biogas reactors due to radical changes of feedstock composition. *Bioresour Technol*, *176*, 56-64. https://doi.org/10.1016/j.biortech.2014.10.126

De Vrieze, J., Christiaens, M. E. R., Walraedt, D., Devooght, A., Ijaz, U. Z., & Boon, N. (2017, Mar 15). Microbial community redundancy in anaerobic digestion drives process recovery after salinity exposure. *Water Res*, *111*, 109-117. https://doi.org/10.1016/j.watres.2016.12.042

De Vrieze, J., Pinto, A. J., Sloan, W. T., & Ijaz, U. Z. (2018, Apr 2). The active microbial community more accurately reflects the anaerobic digestion process: 16S rRNA (gene) sequencing as a predictive tool. *Microbiome*, *6*(1), 63. https://doi.org/10.1186/s40168-018-0449-9

De Vrieze, J., Raport, L., Roume, H., Vilchez-Vargas, R., Jauregui, R., Pieper, D. H., & Boon, N. (2016, Nov 1). The full-scale anaerobic digestion microbiome is represented by specific marker populations. *Water Res, 104*, 101-110. https://doi.org/10.1016/j.watres.2016.08.008

Delforno, T. P., Lacerda Júnior, G. V., Noronha, M. F., Sakamoto, I. K., Varesche, M. B. A., & Oliveira, V. M. (2017). Microbial diversity of a full-scale UASB reactor applied to poultry slaughterhouse wastewater treatment: integration of 16S rRNA gene amplicon and shotgun metagenomic sequencing. *Microbiologyopen, 6*(3), e00443. https://doi.org/10.1002/mbo3.443

Diaz-Elsayed, N., Xu, X., Balaguer-Barbosa, M., & Zhang, Q. (2017, Sep 15). An evaluation of the sustainability of onsite wastewater treatment systems for nutrient management. *Water Res, 121*, 186-196. https://doi.org/10.1016/j.watres.2017.05.005

Dong, L., Cao, G., Wu, J., Yang, S., & Ren, N. (2019, Jul). Reflux of acidizing fluid for enhancing biomethane production from cattle manure in plug flow reactor. *Bioresour Technol, 284*, 248-255. https://doi.org/10.1016/j.biortech.2019.03.092

Dong, L. L., Cao, G. L., Guo, X. Z., Liu, T. S., Wu, J. W., & Ren, N. Q. (2019, Apr). Efficient biogas production from cattle manure in a plug flow reactor: A large scale long term study. *Bioresource Technology, 278*, 450-455. https://doi.org/10.1016/j.biortech.2019.01.100

Dray, S., & Dufour, A.-B. (2007, 2007-09-30). The ade4 Package: Implementing the Duality Diagram for Ecologists. *2007, 22*(4), 20. https://doi.org/10.18637/jss.v022.i04

Drobac, D., Tokodi, N., Lujic, J., Marinovic, Z., Subakov-Simic, G., Dulic, T., Vazic, T., Nybom, S., Meriluoto, J., Codd, G. A., & Svircev, Z. (2016, May). Cyanobacteria and cyanotoxins in fishponds and their effects on fish tissue. *Harmful Algae, 55*, 66-76. https://doi.org/10.1016/j.hal.2016.02.007

Dumont, E., Harrison, J. A., Kroeze, C., Bakker, E. J., & Seitzinger, S. P. (2005). Global distribution and sources of dissolved inorganic nitrogen export to

the coastal zone: Results from a spatially explicit, global model. *Global Biogeochemical Cycles, 19*(4). https://doi.org/10.1029/2005gb002488

El Hage, R., Hernandez-Sanabria, E., Calatayud Arroyo, M., Props, R., & Van de Wiele, T. (2019, 2019-May-31). Propionate-Producing Consortium Restores Antibiotic-Induced Dysbiosis in a Dynamic in vitro Model of the Human Intestinal Microbial Ecosystem [Original Research]. *Frontiers in Microbiology, 10*(1206). https://doi.org/10.3389/fmicb.2019.01206

Emerson, J. B., Adams, R. I., Roman, C. M. B., Brooks, B., Coil, D. A., Dahlhausen, K., Ganz, H. H., Hartmann, E. M., Hsu, T., Justice, N. B., Paulino-Lima, I. G., Luongo, J. C., Lymperopoulou, D. S., Gomez-Silvan, C., Rothschild-Mancinelli, B., Balk, M., Huttenhower, C., Nocker, A., Vaishampayan, P., & Rothschild, L. J. (2017, Aug 16). Schrodinger's microbes: Tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome, 5*(1), 86. https://doi.org/10.1186/s40168-017-0285-3

Ferla, M. P., & Patrick, W. M. (2014, Aug). Bacterial methionine biosynthesis. *Microbiology (Reading), 160*(Pt 8), 1571-1584. https://doi.org/10.1099/mic.0.077826-0

Ferrao-Filho Ada, S., & Kozlowsky-Suzuki, B. (2011, Dec). Cyanotoxins: bioaccumulation and effects on aquatic animals. *Mar Drugs, 9*(12), 2729-2772. https://doi.org/10.3390/md9122729

Ferrocino, I., Bellio, A., Giordano, M., Macori, G., Romano, A., Rantsiou, K., Decastelli, L., & Cocolin, L. (2018). Shotgun Metagenomics and Volatilome Profile of the Microbiota of Fermented Sausages. *Applied and Environmental Microbiology, 84*(3), e02120-02117. https://doi.org/10.1128/AEM.02120-17

Fontana, A., Campanaro, S., Treu, L., Kougias, P. G., Cappa, F., Morelli, L., & Angelidaki, I. (2018, May 1). Performance and genome-centric metagenomics of thermophilic single and two-stage anaerobic digesters treating cheese wastes. *Water Res, 134*, 181-191. https://doi.org/10.1016/j.watres.2018.02.001

Fontana, A., Kougias, P. G., Treu, L., Kovalovszki, A., Valle, G., Cappa, F., Morelli, L., Angelidaki, I., & Campanaro, S. (2018, Oct 27). Microbial

activity response to hydrogen injection in thermophilic anaerobic digesters revealed by genome-centric metatranscriptomics. *Microbiome, 6*(1), 194. https://doi.org/10.1186/s40168-018-0583-4

Frank E Harrell Jr, C. D., et al. (2020). *Hmisc: Harrell Miscellaneous.* In (Version 4.4-0) https://CRAN.R-project.org/package=Hmisc

Frei, R. J., Abbott, B. W., Dupas, R., Gu, S., Gruau, G., Thomas, Z., Kolbe, T., Aquilina, L., Labasque, T., Laverman, A., Fovet, O., Moatar, F., & Pinay, G. (2020, 2020-January-14). Predicting Nutrient Incontinence in the Anthropocene at Watershed Scales [Original Research]. *Frontiers in Environmental Science, 7*(200). https://doi.org/10.3389/fenvs.2019.00200

Fresia, P., Antelo, V., Salazar, C., Giménez, M., D'Alessandro, B., Afshinnekoo, E., Mason, C., Gonnet, G. H., & Iraola, G. (2019, 2019/02/28). Urban metagenomics uncover antibiotic resistance reservoirs in coastal beach and sewage waters. *Microbiome, 7*(1), 35. https://doi.org/10.1186/s40168-019-0648-z

Fumagalli, M. (2013). Assessing the effect of sequencing depth and sample size in population genetics inferences. *PLoS One, 8*(11), e79667. https://doi.org/10.1371/journal.pone.0079667

Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics, 31*(22), 3718-3720. https://doi.org/10.1093/bioinformatics/btv428

Gao, M. J., Guo, B., Zhang, L., Zhang, Y. D., & Liu, Y. (2019, Sep 1). Microbial community dynamics in anaerobic digesters treating conventional and vacuum toilet flushed blackwater. *Water Research, 160*, 249-258. https://doi.org/10.1016/j.watres.2019.05.077

García-Lozano, M., Hernández-De Lira, I. O., Huber, D. H., & Balagurusamy, N. (2019). Spatial Variations of Bacterial Communities of an Anaerobic Lagoon-Type Biodigester Fed with Dairy Manure. *Processes, 7*(7), 408. https://www.mdpi.com/2227-9717/7/7/408

Garcia, S. N., Clubbs, R. L., Stanley, J. K., Scheffe, B., Yelderman, J. C., Jr., & Brooks, B. W. (2013, Jun). Comparative analysis of effluent water quality from a municipal treatment plant and two on-site wastewater treatment systems. *Chemosphere*, *92*(1), 38-44. https://doi.org/10.1016/j.chemosphere.2013.03.007

Gerba, C. P., & Pepper, I. L. (2009). Wastewater Treatment and Biosolids Reuse. *Environmental Microbiology*, 503-530. https://doi.org/10.1016/B978-0-12-370519-8.00024-9

Gonzalez-Gil, G., Lens, P. N., Van Aelst, A., Van As, H., Versprille, A. I., & Lettinga, G. (2001, Aug). Cluster structure of anaerobic aggregates of an expanded granular sludge bed reactor. *Appl Environ Microbiol*, *67*(8), 3683-3692. https://doi.org/10.1128/AEM.67.8.3683-3692.2001

Goux, X., Calusinska, M., Fossepre, M., Benizri, E., & Delfosse, P. (2016, Jul). Start-up phase of an anaerobic full-scale farm reactor - Appearance of mesophilic anaerobic conditions and establishment of the methanogenic microbial community. *Bioresour Technol*, *212*, 217-226. https://doi.org/10.1016/j.biortech.2016.04.040

Goyette, J. O., Bennett, E. M., & Maranger, R. (2018, 2018/12/01). Low buffering capacity and slow recovery of anthropogenic phosphorus pollution in watersheds. *Nature Geoscience*, *11*(12), 921-925. https://doi.org/10.1038/s41561-018-0238-x

Gulhane, M., Pandit, P., Khardenavis, A., Singh, D., & Purohit, H. (2017, 2017/02/01/). Study of microbial community plasticity for anaerobic digestion of vegetable waste in Anaerobic Baffled Reactor. *Renewable Energy*, *101*, 59-66. https://doi.org/https://doi.org/10.1016/j.renene.2016.08.021

Gunsalus, R. P., Cook, L. E., Crable, B., Rohlin, L., McDonald, E., Mouttaki, H., Sieber, J. R., Poweleit, N., Zhou, H., Lapidus, A. L., Daligault, H. E., Land, M., Gilna, P., Ivanova, N., Kyrpides, N., Culley, D. E., & McInerney, M. J. (2016). Complete genome sequence of Methanospirillum hungatei type strain JF1. *Stand Genomic Sci*, *11*, 2. https://doi.org/10.1186/s40793-015-0124-8

Guo, J., Peng, Y., Fan, L., Zhang, L., Ni, B.-J., Kartal, B., Feng, X., Jetten, M. S. M., & Yuan, Z. (2016). Metagenomic analysis of anammox communities in three different microbial aggregates. *Environmental Microbiology, 18*(9), 2979-2993. https://doi.org/10.1111/1462-2920.13132

Guo, J., Peng, Y., Ni, B. J., Han, X., Fan, L., & Yuan, Z. (2015, Mar 14). 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, 14*, 33. https://doi.org/10.1186/s12934-015-0218-4

Han, W., He, P., Lin, Y., Shao, L., & Lü, F. (2019, 2019-November-26). A Methanogenic Consortium Was Active and Exhibited Long-Term Survival in an Extremely Acidified Thermophilic Bioreactor [Original Research]. *Frontiers in Microbiology, 10*(2757). https://doi.org/10.3389/fmicb.2019.02757

Hao, L., Michaelsen, T. Y., Singleton, C. M., Dottorini, G., Kirkegaard, R. H., Albertsen, M., Nielsen, P. H., & Dueholm, M. S. (2020, Apr). Novel syntrophic bacteria in full-scale anaerobic digesters revealed by genome-centric metatranscriptomics. *ISME J, 14*(4), 906-918. https://doi.org/10.1038/s41396-019-0571-0

Harrison, J. A., Bouwman, A. F., Mayorga, E., & Seitzinger, S. (2010). Magnitudes and sources of dissolved inorganic phosphorus inputs to surface fresh waters and the coastal zone: A new global model. *Global Biogeochemical Cycles, 24*(1). https://doi.org/10.1029/2009gb003590

Hendriks, A., & Langeveld, J. G. (2017, May 2). Rethinking Wastewater Treatment Plant Effluent Standards: Nutrient Reduction or Nutrient Control? *Environ Sci Technol, 51*(9), 4735-4737. https://doi.org/10.1021/acs.est.7b01186

Heyer, R., Klang, J., Hellwig, P., Schallert, K., Kress, P., Huelsemann, B., Theuerl, S., Reichl, U., & Benndorf, D. (2020, 2020/10/01/). Impact of feeding and stirring regimes on the internal stratification of microbial communities in the fermenter of anaerobic digestion plants. *Bioresource Technology, 314*, 123679. https://doi.org/https://doi.org/10.1016/j.biortech.2020.123679

Heyer, R., Schallert, K., Siewert, C., Kohrs, F., Greve, J., Maus, I., Klang, J., Klocke, M., Heiermann, M., Hoffmann, M., Püttker, S., Calusinska, M., Zoun, R., Saake, G., Benndorf, D., & Reichl, U. (2019). Metaproteome analysis reveals that syntrophy, competition, and phage-host interaction shape microbial communities in biogas plants. *Microbiome, 7*(1), 69-69. https://doi.org/10.1186/s40168-019-0673-y

Heylen, K., Lebbe, L., & De Vos, P. (2008, Jan). Acidovorax caeni sp. nov., a denitrifying species with genetically diverse isolates from activated sludge. *Int J Syst Evol Microbiol, 58*(Pt 1), 73-77. https://doi.org/10.1099/ijs.0.65387-0

Hillmann, B., Al-Ghalith, G. A., Shields-Cutler, R. R., Zhu, Q., Gohl, D. M., Beckman, K. B., Knight, R., & Knights, D. (2018). Evaluating the Information Content of Shallow Shotgun Metagenomics. *mSystems, 3*(6), e00069-00018. https://doi.org/10.1128/mSystems.00069-18

Holeton, C., Chambers, P., & Grace, L. (2011, 10/01). Wastewater release and its impacts on Canadian waters. *Canadian Journal of Fisheries and Aquatic Sciences, 68*, 1836-1859. https://doi.org/10.1139/f2011-096

Horn, M. A., Matthies, C., Küsel, K., Schramm, A., & Drake, H. L. (2003). Hydrogenotrophic methanogenesis by moderately acid-tolerant methanogens of a methane-emitting acidic peat. *Applied and Environmental Microbiology, 69*(1), 74-83. https://doi.org/10.1128/aem.69.1.74-83.2003

Huson, D. H., Beier, S., Flade, I., Gorska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H. J., & Tappu, R. (2016, Jun). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput Biol, 12*(6), e1004957. https://doi.org/10.1371/journal.pcbi.1004957

Hvala, N., Vrečko, D., & Bordon, C. (2018). Plant-wide modelling for assessment and optimization of upgraded full-scale wastewater treatment plant performance. *Water Practice and Technology, 13*(3), 566-582. https://doi.org/10.2166/wpt.2018.070

Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010, Mar 8). Prodigal: prokaryotic gene recognition and translation

initiation site identification. *BMC Bioinformatics*, *11*, 119. https://doi.org/10.1186/1471-2105-11-119

Institute, J. G. (2019). *BBDuk Guide*. United States Depeartment of Energy Office of Science User Facility. https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbduk-guide/

Isazadeh, S., Jauffur, S., & Frigon, D. (2016, Dec). Bacterial community assembly in activated sludge: mapping beta diversity across environmental variables. *Microbiologyopen*, *5*(6), 1050-1060. https://doi.org/10.1002/mbo3.388

Jain, K. A., Suryawanshi, P. C., & Chaudhari, A. B. (2020, 2020/07/08). Hydrogenotrophic methanogen strain of Methanospirillum from anaerobic digester fed with agro-industrial waste. *Biologia*. https://doi.org/10.2478/s11756-020-00559-y

Jari Oksanen, F. G. B., Michael Friendly, Roeland, Kindt, P. L., Dan McGlinn, Peter R. Minchin, R. B. O'Hara,, Gavin L. Simpson, P. S., M. Henry H. Stevens, Eduard Szoecs, & Wagner, a. H. (2019). *vegan: Community Ecology Package.* In (Version 2.5-6) https://CRAN.R-project.org/package=vegan

Jia, Y., Wilkins, D., Lu, H., Cai, M., & Lee, P. K. H. (2016). Long-Term Enrichment on Cellulose or Xylan Causes Functional and Taxonomic Convergence of Microbial Communities from Anaerobic Digesters. *Applied and Environmental Microbiology*, *82*(5), 1519-1529. https://doi.org/10.1128/aem.03360-15

Jiang, H., Nie, H., Ding, J., Stinner, W., Sun, K., & Zhou, H. (2018, Jan 2). The startup performance and microbial distribution of an anaerobic baffled reactor (ABR) treating medium-strength synthetic industrial wastewater. *J Environ Sci Health A Tox Hazard Subst Environ Eng*, *53*(1), 46-54. https://doi.org/10.1080/10934529.2017.1368297

Jo, Y., Kim, J., Hwang, K., & Lee, C. (2018, 2018/08/01/). A comparative study of single- and two-phase anaerobic digestion of food waste under

uncontrolled pH conditions. *Waste Management*, *78*, 509-520. https://doi.org/https://doi.org/10.1016/j.wasman.2018.06.017

Jones, M. B., Highlander, S. K., Anderson, E. L., Li, W., Dayrit, M., Klitgord, N., Fabani, M. M., Seguritan, V., Green, J., Pride, D. T., Yooseph, S., Biggs, W., Nelson, K. E., & Venter, J. C. (2015, Nov 10). Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proc Natl Acad Sci U S A*, *112*(45), 14024-14029. https://doi.org/10.1073/pnas.1519288112

Jonsson, V., Österlund, T., Nerman, O., & Kristiansson, E. (2016, 2016/01/25). Statistical evaluation of methods for identification of differentially abundant genes in comparative metagenomics. *BMC Genomics*, *17*(1), 78. https://doi.org/10.1186/s12864-016-2386-y

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., Perry, T., Kao, D., Mason, A. L., Madsen, K. L., & Wong, G. K. (2016). Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Front Microbiol*, *7*, 459. https://doi.org/10.3389/fmicb.2016.00459

Jowett, C., Kraemer, J., James, C., & Jowett, E. C. (2017). *IMPROVING SEPTIC TANK PERFORMANCE BY ENHANCING ANAEROBIC DIGESTION*.

Jowett, E. C. (2007). Comparing the Performance of Prescribed Septic Tanks to Long, Narrow, Flooded Designs. *Proceedings of the Water Environment Federation*, *2007*(18), 1027-1038. https://doi.org/10.2175/193864707787452804

Joyce, A., Ijaz, U. Z., Nzeteu, C., Vaughan, A., Shirran, S. L., Botting, C. H., Quince, C., O'Flaherty, V., & Abram, F. (2018). Linking Microbial Community Structure and Function During the Acidified Anaerobic Digestion of Grass. *Front Microbiol*, *9*, 540. https://doi.org/10.3389/fmicb.2018.00540

Ju, F., Lau, F., & Zhang, T. (2017, Apr 4). Linking Microbial Community, Environmental Variables, and Methanogenesis in Anaerobic Biogas Digesters of Chemically Enhanced Primary Treatment Sludge. *Environ Sci Technol*, *51*(7), 3982-3992. https://doi.org/10.1021/acs.est.6b06344

Ju, F., Li, B., Ma, L., Wang, Y., Huang, D., & Zhang, T. (2016, Mar 15). Antibiotic resistance genes and human bacterial pathogens: Co-occurrence, removal, and enrichment in municipal sewage sludge digesters. *Water Res, 91*, 1-10. https://doi.org/10.1016/j.watres.2015.11.071

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2014, Jan). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res, 42*(Database issue), D199-205. https://doi.org/10.1093/nar/gkt1076

Kassambara, A. (2019). *fastqcr: Quality Control of Sequencing Data.* In (Version 0.1.2) https://CRAN.R-project.org/package=fastqcr

Kaushal, S. S., Groffman, P. M., Band, L. E., Elliott, E. M., Shields, C. A., & Kendall, C. (2011, Oct 1). Tracking nonpoint source nitrogen pollution in human-impacted watersheds. *Environ Sci Technol, 45*(19), 8225-8232. https://doi.org/10.1021/es200779e

Keegan, M., Kilroy, K., Nolan, D., Dubber, D., Johnston, P. M., Misstear, B. D., O'Flaherty, V., Barrett, M., & Gill, L. W. (2014). Assessment of the impact of traditional septic tank soakaway systems on water quality in Ireland. *Water Sci Technol, 70*(4), 634-641. https://doi.org/10.2166/wst.2014.227

Keller, K. L., Rapp-Giles, B. J., Semkiw, E. S., Porat, I., Brown, S. D., & Wall, J. D. (2014, Feb). New model for electron flow for sulfate reduction in Desulfovibrio alaskensis G20. *Appl Environ Microbiol, 80*(3), 855-868. https://doi.org/10.1128/AEM.02963-13

Khan, M. T., Duncan, S. H., Stams, A. J., van Dijl, J. M., Flint, H. J., & Harmsen, H. J. (2012, Aug). The gut anaerobe Faecalibacterium prausnitzii uses an extracellular electron shuttle to grow at oxic-anoxic interphases. *ISME J, 6*(8), 1578-1585. https://doi.org/10.1038/ismej.2012.5

Kim, B. R., Shin, J., Guevarra, R., Lee, J. H., Kim, D. W., Seol, K. H., Lee, J. H., Kim, H. B., & Isaacson, R. (2017, Dec 28). Deciphering Diversity Indices

for a Better Understanding of Microbial Communities. *J Microbiol Biotechnol, 27*(12), 2089-2093. https://doi.org/10.4014/jmb.1709.09027

Kindt, R., & Coe, R. (2005). *Tree Diversity Analysis. A Manual and Software for Common Statistical Methods for Ecological and Biodiversity Studies*.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glockner, F. O. (2013, Jan 7). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res, 41*(1), e1. https://doi.org/10.1093/nar/gks808

Knight, R., Jansson, J., Field, D., Fierer, N., Desai, N., Fuhrman, J. A., Hugenholtz, P., van der Lelie, D., Meyer, F., Stevens, R., Bailey, M. J., Gordon, J. I., Kowalchuk, G. A., & Gilbert, J. A. (2012, Jun 7). Unlocking the potential of metagenomics through replicated experimental design. *Nat Biotechnol, 30*(6), 513-520. https://doi.org/10.1038/nbt.2235

Kobayashi, T., Li, Y.-Y., Kubota, K., Harada, H., Maeda, T., & Yu, H.-Q. (2012, 2012/01/01). Characterization of sulfide-oxidizing microbial mats developed inside a full-scale anaerobic digester employing biological desulfurization. *Applied Microbiology and Biotechnology, 93*(2), 847-857. https://doi.org/10.1007/s00253-011-3445-6

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 Genome-Centric Metagenomics. *Applied and Environmental Microbiology, 84*(18), e01244-01218. https://doi.org/10.1128/aem.01244-18

Kougias, P. G., Treu, L., Campanaro, S., Zhu, X., & Angelidaki, I. (2016, Jun 29). Dynamic functional characterization and phylogenetic changes due to Long Chain Fatty Acids pulses in biogas reactors. *Sci Rep, 6*, 28810. https://doi.org/10.1038/srep28810

Kouzuma, A., & Watanabe, K. (2011). 6.04 - Molecular Approaches for the Analysis of Natural Attenuation and Bioremediation. In M. Moo-Young (Ed.), *Comprehensive Biotechnology (Second Edition)* (pp. 25-36). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-08-088504-9.00375-5

Kuroda, K., Nobu, M. K., Mei, R., Narihiro, T., Bocher, B. T., Yamaguchi, T., & Liu, W. T. (2016). A Single-Granule-Level Approach Reveals Ecological Heterogeneity in an Upflow Anaerobic Sludge Blanket Reactor. *PLoS One*, *11*(12), e0167788. https://doi.org/10.1371/journal.pone.0167788

Kuypers, M. M. M., Marchant, H. K., & Kartal, B. (2018, May). The microbial nitrogen-cycling network. *Nat Rev Microbiol*, *16*(5), 263-276. https://doi.org/10.1038/nrmicro.2018.9

Langille, M. G., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G., & Huttenhower, C. (2013, Sep). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol*, *31*(9), 814-821. https://doi.org/10.1038/nbt.2676

Lau, C. K. Y., Krewulak, K. D., & Vogel, H. J. (2015). Bacterial ferrous iron transport: the Feo system. *FEMS Microbiology Reviews*, *40*(2), 273-298. https://doi.org/10.1093/femsre/fuv049

Lefèvre, C. T., Trubitsyn, D., Abreu, F., Kolinko, S., Jogler, C., de Almeida, L. G. P., de Vasconcelos, A. T. R., Kube, M., Reinhardt, R., Lins, U., Pignol, D., Schüler, D., Bazylinski, D. A., & Ginet, N. (2013). Comparative genomic analysis of magnetotactic bacteria from the Deltaproteobacteria provides new insights into magnetite and greigite magnetosome genes required for magnetotaxis. *Environmental Microbiology*, *15*(10), 2712-2735. https://doi.org/10.1111/1462-2920.12128

Legendre, P., & Gallagher, E. D. (2001, Oct). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*(2), 271-280. https://doi.org/10.1007/s004420100716

Li, C., Tao, Y., Fang, J., Li, Q., & Lu, W. (2020, 2020/07/01/). Impact of continuous leachate recirculation during solid state anaerobic digestion of Miscanthus. *Renewable Energy*, *154*, 38-45. https://doi.org/https://doi.org/10.1016/j.renene.2020.02.104

Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015, May 15). 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

Li, L., He, Q., Ma, Y., Wang, X., & Peng, X. (2015). Dynamics of microbial community in a mesophilic anaerobic digester treating food waste: Relationship between community structure and process stability. *Bioresour Technol, 189*, 113-120. https://doi.org/10.1016/j.biortech.2015.04.015

Li, L., He, Q., Ma, Y., Wang, X., & Peng, X. (2016, Apr 25). A mesophilic anaerobic digester for treating food waste: process stability and microbial community analysis using pyrosequencing. *Microb Cell Fact, 15*, 65. https://doi.org/10.1186/s12934-016-0466-y

Li, L., Peng, X., Wang, X., & Wu, D. (2018, Jan). Anaerobic digestion of food waste: A review focusing on process stability. *Bioresour Technol, 248*(Pt A), 20-28. https://doi.org/10.1016/j.biortech.2017.07.012

Li, N., He, J., Yan, H., Chen, S., & Dai, X. (2017, Oct). Pathways in bacterial and archaeal communities dictated by ammonium stress in a high solid anaerobic digester with dewatered sludge. *Bioresour Technol, 241*, 95-102. https://doi.org/10.1016/j.biortech.2017.05.094

Li, Y., Li, L., Sun, Y., & Yuan, Z. (2018, Aug). Bioaugmentation strategy for enhancing anaerobic digestion of high C/N ratio feedstock with methanogenic enrichment culture. *Bioresour Technol, 261*, 188-195. https://doi.org/10.1016/j.biortech.2018.02.069

Li, Y., Sun, Y., Li, L., & Yuan, Z. (2018, 2018/02/01/). Acclimation of acid-tolerant methanogenic propionate-utilizing culture and microbial community dissecting. *Bioresource Technology, 250*, 117-123. https://doi.org/https://doi.org/10.1016/j.biortech.2017.11.034

Li, Y., Yang, G., Li, L., & Sun, Y. (2018, Sep). Bioaugmentation for overloaded anaerobic digestion recovery with acid-tolerant methanogenic enrichment. *Waste Manag, 79*, 744-751. https://doi.org/10.1016/j.wasman.2018.08.043

Li, Y., Zhang, Y., Sun, Y., Wu, S., Kong, X., Yuan, Z., & Dong, R. (2017, May). The performance efficiency of bioaugmentation to prevent anaerobic digestion failure from ammonia and propionate inhibition. *Bioresour Technol, 231*, 94-100. https://doi.org/10.1016/j.biortech.2017.01.068

Li, Y. F., Chen, P. H., & Yu, Z. (2014, Jul). Spatial and temporal variations of microbial community in a mixed plug-flow loop reactor fed with dairy manure. *Microb Biotechnol, 7*(4), 332-346. https://doi.org/10.1111/1751-7915.12125

Lin, Q., De Vrieze, J., He, G., Li, X., & Li, J. (2016a, Sep). Temperature regulates methane production through the function centralization of microbial community in anaerobic digestion. *Bioresour Technol, 216*, 150-158. https://doi.org/10.1016/j.biortech.2016.05.046

Lin, Q., De Vrieze, J., He, G. H., Li, X. Z., & Li, J. B. (2016b, Sep). Temperature regulates methane production through the function centralization of microbial community in anaerobic digestion. *Bioresource Technology, 216*, 150-158. https://doi.org/10.1016/j.biortech.2016.05.046

Liu, C., Li, H., Zhang, Y. Y., Si, D. D., & Chen, Q. W. (2016, Sep). Evolution of microbial community along with increasing solid concentration during high-solids anaerobic digestion of sewage sludge. *Bioresource Technology, 216*, 87-94. https://doi.org/10.1016/j.biortech.2016.05.048

Liu, Z.-h., Yin, H., Lin, Z., & Dang, Z. (2018, 2018/01/01). Sulfate-reducing bacteria in anaerobic bioprocesses: basic properties of pure isolates, molecular quantification, and controlling strategies. *Environmental Technology Reviews, 7*(1), 46-72. https://doi.org/10.1080/21622515.2018.1437783

Lomans, B. P., Maas, R., Luderer, R., Op den Camp, H. J., Pol, A., van der Drift, C., & Vogels, G. D. (1999). Isolation and characterization of Methanomethylovorans hollandica gen. nov., sp. nov., isolated from freshwater sediment, a methylotrophic methanogen able to grow on dimethyl sulfide and methanethiol. *Applied and Environmental Microbiology, 65*(8), 3641-3650. https://doi.org/10.1128/AEM.65.8.3641-3650.1999

Love, M. I., Huber, W., & Anders, S. (2014, 2014/12/05). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology, 15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Ludington, W. B., Seher, T. D., Applegate, O., Li, X., Kliegman, J. I., Langelier, C., Atwill, E. R., Harter, T., & DeRisi, J. L. (2017). Assessing biosynthetic potential of agricultural groundwater through metagenomic sequencing: A diverse anammox community dominates nitrate-rich groundwater. *PLoS One, 12*(4), e0174930. https://doi.org/10.1371/journal.pone.0174930

Lunau, M., Voss, M., Erickson, M., Dziallas, C., Casciotti, K., & Ducklow, H. (2013, May). Excess nitrate loads to coastal waters reduces nitrate removal efficiency: mechanism and implications for coastal eutrophication. *Environ Microbiol, 15*(5), 1492-1504. https://doi.org/10.1111/j.1462-2920.2012.02773.x

Luo, C., Tsementzi, D., Kyrpides, N., Read, T., & Konstantinidis, K. T. (2012). Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. *PLoS One, 7*(2), e30087. https://doi.org/10.1371/journal.pone.0030087

Luo, G., Fotidis, I. A., & Angelidaki, I. (2016). Comparative analysis of taxonomic, functional, and metabolic patterns of microbiomes from 14 full-scale biogas reactors by metagenomic sequencing and radioisotopic analysis. *Biotechnol Biofuels, 9*, 51. https://doi.org/10.1186/s13068-016-0465-6

Martins, C. L., Velho, V. F., Magnus, B. S., Xavier, J. A., Guimarães, L. B., Leite, W. R., & Costa, R. H. R. (2020, 2020/08/01/). Assessment of sludge reduction and microbial dynamics in an OSA process with short anaerobic retention time. *Environmental Technology & Innovation, 19*, 101025. https://doi.org/https://doi.org/10.1016/j.eti.2020.101025

Masse, D. I., Gilbert, Y., Saady, N. M., & Liu, C. (2013, Sep-Oct). Low-temperature anaerobic digestion of swine manure in a plug-flow reactor. *Environ Technol, 34*(17-20), 2617-2624. https://doi.org/10.1080/09593330.2013.781229

Maus, I., Bremges, A., Stolze, Y., Hahnke, S., Cibis, K. G., Koeck, D. E., Kim, Y. S., Kreubel, J., Hassa, J., Wibberg, D., Weimann, A., Off, S., Stantscheff, R., Zverlov, V. V., Schwarz, W. H., König, H., Liebl, W., Scherer, P., McHardy, A. C., Sczyrba, A., Klocke, M., Pühler, A., & Schlüter, A. (2017). Genomics and prevalence of bacterial and archaeal isolates from biogas-producing microbiomes. *Biotechnology for Biofuels*, *10*, 264-264. https://doi.org/10.1186/s13068-017-0947-1

McAteer, P. G., Christine Trego, A., Thorn, C., Mahony, T., Abram, F., & O'Flaherty, V. (2020, Jul). Reactor configuration influences microbial community structure during high-rate, low-temperature anaerobic treatment of dairy wastewater. *Bioresour Technol*, *307*, 123221. https://doi.org/10.1016/j.biortech.2020.123221

McIntyre, A. B. R., Ounit, R., Afshinnekoo, E., Prill, R. J., Henaff, E., Alexander, N., Minot, S. S., Danko, D., Foox, J., Ahsanuddin, S., Tighe, S., Hasan, N. A., Subramanian, P., Moffat, K., Levy, S., Lonardi, S., Greenfield, N., Colwell, R. R., Rosen, G. L., & Mason, C. E. (2017, Sep 21). Comprehensive benchmarking and ensemble approaches for metagenomic classifiers. *Genome Biol*, *18*(1), 182. https://doi.org/10.1186/s13059-017-1299-7

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Mei, R., Narihiro, T., Nobu, M. K., Kuroda, K., & Liu, W.-T. (2016, 2016/09/26). Evaluating digestion efficiency in full-scale anaerobic digesters by identifying active microbial populations through the lens of microbial activity. *Scientific Reports*, *6*(1), 34090. https://doi.org/10.1038/srep34090

Meyer, B., Kuehl, J. V., Deutschbauer, A. M., Arkin, A. P., & Stahl, D. A. (2013). Flexibility of syntrophic enzyme systems in Desulfovibrio species ensures their adaptation capability to environmental changes. *Journal of bacteriology*, *195*(21), 4900-4914. https://doi.org/10.1128/JB.00504-13

Mohiuddin, M. M., Botts, S. R., Paschos, A., & Schellhorn, H. E. (2019, Feb). Temporal and spatial changes in bacterial diversity in mixed use

watersheds of the Great Lakes region. *Journal of Great Lakes Research, 45*(1), 109-118. https://doi.org/10.1016/j.jglr.2018.10.007

Mohiuddin, M. M., Salama, Y., Schellhorn, H. E., & Golding, G. B. (2017, May 15). Shotgun metagenomic sequencing reveals freshwater beach sands as reservoir of bacterial pathogens. *Water Res, 115*, 360-369. https://doi.org/10.1016/j.watres.2017.02.057

Morgan, J., Smith, M., Mc Auley, M. T., & Enrique Salcedo-Sora, J. (2018). Disrupting folate metabolism reduces the capacity of bacteria in exponential growth to develop persisters to antibiotics. *Microbiology, 164*(11), 1432-1445. https://doi.org/https://doi.org/10.1099/mic.0.000722

Mosbaek, F., Kjeldal, H., Mulat, D. G., Albertsen, M., Ward, A. J., Feilberg, A., & Nielsen, J. L. (2016, Oct). Identification of syntrophic acetate-oxidizing bacteria in anaerobic digesters by combined protein-based stable isotope probing and metagenomics. *Isme Journal, 10*(10), 2405-2418. https://doi.org/10.1038/ismej.2016.39

Mulcahy-O'Grady, H., & Workentine, M. L. (2016, 2016-February-03). The Challenge and Potential of Metagenomics in the Clinic [Review]. *Frontiers in Immunology, 7*(29). https://doi.org/10.3389/fimmu.2016.00029

Musa, M. A., Idrus, S., Hasfalina, C. M., & Daud, N. N. N. (2018). Effect of Organic Loading Rate on Anaerobic Digestion Performance of Mesophilic (UASB) Reactor Using Cattle Slaughterhouse Wastewater as Substrate. *International Journal of Environmental Research and Public Health, 15*(10), 2220. https://doi.org/10.3390/ijerph15102220

Nascimento, A. L., Souza, A. J., Andrade, P. A. M., Andreote, F. D., Coscione, A. R., Oliveira, F. C., & Regitano, J. B. (2018, Jul 3). Sewage Sludge Microbial Structures and Relations to Their Sources, Treatments, and Chemical Attributes. *Frontiers in Microbiology, 9*. https://doi.org/ARTN 1462

10.3389/fmicb.2018.01462

Ni, J., Yan, Q., & Yu, Y. (2013). How much metagenomic sequencing is enough to achieve a given goal? *Sci Rep, 3*, 1968. https://doi.org/10.1038/srep01968

Nie, J. Y., Zhu, N. W., Zhao, K., Wu, L., & Hu, Y. H. (2011). Analysis of the bacterial community changes in soil for septic tank effluent treatment in response to bio-clogging. *Water Sci Technol, 63*(7), 1412-1417. https://doi.org/10.2166/wst.2011.319

Nordgard, A. S. R., Bergland, W. H., Vadstein, O., Mironov, V., Bakke, R., Ostgaard, K., & Bakke, I. (2017, Nov 8). Anaerobic digestion of pig manure supernatant at high ammonia concentrations characterized by high abundances of Methanosaeta and non-euryarchaeotal archaea. *Scientific Reports, 7*. https://doi.org/ARTN 15077

10.1038/s41598-017-14527-1

Novak, J. T., Banjade, S., & Murthy, S. N. (2011, Jan). Combined anaerobic and aerobic digestion for increased solids reduction and nitrogen removal. *Water Res, 45*(2), 618-624. https://doi.org/10.1016/j.watres.2010.08.014

Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017, May). metaSPAdes: a new versatile metagenomic assembler. *Genome Res, 27*(5), 824-834. https://doi.org/10.1101/gr.213959.116

Olson, N. D., Treangen, T. J., Hill, C. M., Cepeda-Espinoza, V., Ghurye, J., Koren, S., & Pop, M. (2017, Aug 7). Metagenomic assembly through the lens of validation: recent advances in assessing and improving the quality of genomes assembled from metagenomes. *Brief Bioinform*. https://doi.org/10.1093/bib/bbx098

Otten, T. G., Graham, J. L., Harris, T. D., & Dreher, T. W. (2016, Sep). Elucidation of Taste- and Odor-Producing Bacteria and Toxigenic Cyanobacteria in a Midwestern Drinking Water Supply Reservoir by Shotgun Metagenomic Analysis. *Applied and Environmental Microbiology, 82*(17), 5410-5420. https://doi.org/10.1128/Aem.01334-16

Ounit, R., Wanamaker, S., Close, T. J., & Lonardi, S. (2015, Mar 25). CLARK: fast and accurate classification of metagenomic and genomic sequences

using discriminative k-mers. *BMC Genomics, 16*, 236. https://doi.org/10.1186/s12864-015-1419-2

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F., & Stevens, R. (2014, Jan). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res, 42*(Database issue), D206-214. https://doi.org/10.1093/nar/gkt1226

Owusu-Agyeman, I., Eyice, Ö., Cetecioglu, Z., & Plaza, E. (2019, 2019/10//). The study of structure of anaerobic granules and methane producing pathways of pilot-scale UASB reactors treating municipal wastewater under sub-mesophilic conditions. *Bioresource Technology, 290*, 121733. https://doi.org/10.1016/j.biortech.2019.121733

Oyiwona, G. E., Ogbonna, J., Anyanwu, C. U., Ishizaki, S., Kimura, Z.-i., & Okabe, S. (2017, 2017/02/01). Oxidation of glucose by syntrophic association between Geobacter and hydrogenotrophic methanogens in microbial fuel cell. *Biotechnology Letters, 39*(2), 253-259. https://doi.org/10.1007/s10529-016-2247-4

Ozuolmez, D., Na, H., Lever, M., Kjeldsen, K., Jørgensen, B., & Plugge, C. (2015, 2015-May-27). Methanogenic archaea and sulfate reducing bacteria co-cultured on acetate: teamwork or coexistence? [Original Research]. *Frontiers in Microbiology, 6*(492). https://doi.org/10.3389/fmicb.2015.00492

Paliy, O., & Shankar, V. (2016, Mar). Application of multivariate statistical techniques in microbial ecology. *Mol Ecol, 25*(5), 1032-1057. https://doi.org/10.1111/mec.13536

Park, J., Park, S., & Kim, M. (2014, 2014/05/03). Anaerobic degradation of amino acids generated from the hydrolysis of sewage sludge. *Environmental Technology, 35*(9), 1133-1139. https://doi.org/10.1080/09593330.2013.863951

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015, Jul). CheckM: assessing the quality of microbial genomes

recovered from isolates, single cells, and metagenomes. *Genome Res, 25*(7), 1043-1055. https://doi.org/10.1101/gr.186072.114

Patel, M., Kumar, R., Kishor, K., Mlsna, T., Pittman, C. U., & Mohan, D. (2019, 2019/03/27). Pharmaceuticals of Emerging Concern in Aquatic Systems: Chemistry, Occurrence, Effects, and Removal Methods. *Chemical Reviews, 119*(6), 3510-3673. https://doi.org/10.1021/acs.chemrev.8b00299

Peces, M., Astals, S., Jensen, P. D., & Clarke, W. P. (2018, Sep 15). Deterministic mechanisms define the long-term anaerobic digestion microbiome and its functionality regardless of the initial microbial community. *Water Res, 141*, 366-376. https://doi.org/10.1016/j.watres.2018.05.028

Peres-Neto, P. R., & Jackson, D. A. (2001, Oct). How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia, 129*(2), 169-178. https://doi.org/10.1007/s004420100720

Perez-Cataluna, A., Salas-Masso, N., Dieguez, A. L., Balboa, S., Lema, A., Romalde, J. L., & Figueras, M. J. (2018). Revisiting the Taxonomy of the Genus Arcobacter: Getting Order From the Chaos. *Front Microbiol, 9*, 2077. https://doi.org/10.3389/fmicb.2018.02077

Peters, F., Rother, M., & Boll, M. (2004). Selenocysteine-containing proteins in anaerobic benzoate metabolism of Desulfococcus multivorans. *Journal of bacteriology, 186*(7), 2156-2163. https://doi.org/10.1128/jb.186.7.2156-2163.2004

Poszytek, K., Karczewska-Golec, J., Dziurzynski, M., Stepkowska-Kowalska, O., Gorecki, A., Decewicz, P., Dziewit, L., & Drewniak, L. (2019, Jul 18). Genome-Wide and Functional View of Proteolytic and Lipolytic Bacteria for Efficient Biogas Production through Enhanced Sewage Sludge Hydrolysis. *Molecules, 24*(14). https://doi.org/10.3390/molecules24142624

Pribat, A., Blaby, I. K., Lara-Núñez, A., Jeanguenin, L., Fouquet, R., Frelin, O., Gregory, J. F., 3rd, Philmus, B., Begley, T. P., de Crécy-Lagard, V., & Hanson, A. D. (2011). A 5-formyltetrahydrofolate cycloligase paralog from all domains of life: comparative genomic and experimental evidence

for a cryptic role in thiamin metabolism. *Functional & integrative genomics, 11*(3), 467-478. https://doi.org/10.1007/s10142-011-0224-5

Pussayanavin, T., Koottatep, T., Eamrat, R., & Polprasert, C. (2015). Enhanced sludge reduction in septic tanks by increasing temperature. *J Environ Sci Health A Tox Hazard Subst Environ Eng, 50*(1), 81-89. https://doi.org/10.1080/10934529.2015.964633

Qiu, Y. Y., Zhang, L., Mu, X., Li, G., Guan, X., Hong, J., & Jiang, F. (2020, Feb 1). Overlooked pathways of denitrification in a sulfur-based denitrification system with organic supplementation. *Water Res, 169*, 115084. https://doi.org/10.1016/j.watres.2019.115084

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017, Sep 12). Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol, 35*(9), 833-844. https://doi.org/10.1038/nbt.3935

R Development Core Team. (2010). *R. A language and environment for statistical computing.* In (Version 3.6.3) R Foundation for Statistical Computing. http://www.R-project.org

Raczkowska, A., Trzos, J., Lewandowska, O., Nieckarz, M., & Brzostek, K. (2015). Expression of the AcrAB Components of the AcrAB-TolC Multidrug Efflux Pump of Yersinia enterocolitica Is Subject to Dual Regulation by OmpR. *PLoS One, 10*(4), e0124248-e0124248. https://doi.org/10.1371/journal.pone.0124248

Rakin, A., Schneider, L., & Podladchikova, O. (2012). Hunger for iron: the alternative siderophore iron scavenging systems in highly virulent Yersinia. *Frontiers in cellular and infection microbiology, 2*, 151-151. https://doi.org/10.3389/fcimb.2012.00151

Ramette, A. (2007, Nov). Multivariate analyses in microbial ecology. *FEMS Microbiol Ecol, 62*(2), 142-160. https://doi.org/10.1111/j.1574-6941.2007.00375.x

Ranjan, R., Rani, A., Metwally, A., McGee, H. S., & Perkins, D. L. (2016, Jan 22). Analysis of the microbiome: Advantages of whole genome shotgun

versus 16S amplicon sequencing. *Biochem Biophys Res Commun, 469*(4), 967-977. https://doi.org/10.1016/j.bbrc.2015.12.083

Raszka, A., Surmacz-Gorska, J., Zabczynski, S., & Miksch, K. (2011, Dec). The population dynamics of nitrifiers in ammonium-rich systems. *Water Environ Res, 83*(12), 2159-2169. https://doi.org/10.2175/106143011x12989211841331

Regueiro, L., Lema, J. M., & Carballa, M. (2015, 2015/12/01/). Key microbial communities steering the functioning of anaerobic digesters during hydraulic and organic overloading shocks. *Bioresource Technology, 197*, 208-216. https://doi.org/https://doi.org/10.1016/j.biortech.2015.08.076

Reitzer, L. (2004). Biosynthesis of Glutamate, Aspartate, Asparagine, L-Alanine, and D-Alanine. *EcoSal Plus*. https://doi.org/doi:10.1128/ecosalplus.3.6.1.3

Rice, E. W., American Public Health, A., American Water Works, A., & Water Environment, F. (2012). *Standard methods for the examination of water and wastewater*. American Public Health Association.

Rivett, D. W., & Bell, T. (2018, Jul). Abundance determines the functional role of bacterial phylotypes in complex communities. *Nat Microbiol, 3*(7), 767-772. https://doi.org/10.1038/s41564-018-0180-0

Robertson, W. D., Moore, T. A., Spoelstra, J., Li, L., Elgood, R. J., Clark, I. D., Schiff, S. L., Aravena, R., & Neufeld, J. D. (2012, Jul-Aug). Natural attenuation of septic system nitrogen by anammox. *Ground Water, 50*(4), 541-553. https://doi.org/10.1111/j.1745-6584.2011.00857.x

Rodriguez, R. L., Gunturu, S., Tiedje, J. M., Cole, J. R., & Konstantinidis, K. T. (2018, May-Jun). Nonpareil 3: Fast Estimation of Metagenomic Coverage and Sequence Diversity. *mSystems, 3*(3). https://doi.org/10.1128/mSystems.00039-18

Rui, J., Li, J., Zhang, S., Yan, X., Wang, Y., & Li, X. (2015). The core populations and co-occurrence patterns of prokaryotic communities in household biogas digesters. *Biotechnol Biofuels, 8*, 158. https://doi.org/10.1186/s13068-015-0339-3

Ruiz-Sanchez, J., Campanaro, S., Guivernau, M., Fernandez, B., & Prenafeta-Boldu, F. X. (2018, Feb). Effect of ammonia on the active microbiome and metagenome from stable full-scale digesters. *Bioresour Technol*, *250*, 513-522. https://doi.org/10.1016/j.biortech.2017.11.068

Saia, F. T., Souza, T. S., Duarte, R. T., Pozzi, E., Fonseca, D., & Foresti, E. (2016, Feb). Microbial community in a pilot-scale bioreactor promoting anaerobic digestion and sulfur-driven denitrification for domestic sewage treatment. *Bioprocess Biosyst Eng*, *39*(2), 341-352. https://doi.org/10.1007/s00449-015-1520-6

Sallis, P. J., & Uyanik, S. (2003, Sep). Granule development in a split-feed anaerobic baffled reactor. *Bioresour Technol*, *89*(3), 255-265. https://doi.org/10.1016/s0960-8524(03)00071-3

Sampaio, M.-M., Chevance, F., Dippel, R., Eppler, T., Schlegel, A., Boos, W., Lu, Y.-J., & Rock, C. O. (2004, February 13, 2004). Phosphotransferase-mediated Transport of the Osmolyte 2-O-α-Mannosyl-d-glycerate in Escherichia coli Occurs by the Product of the mngA (hrsA) Gene and Is Regulated by the mngR (farR) Gene Product Acting as Repressor. *Journal of Biological Chemistry*, *279*(7), 5537-5548. https://doi.org/10.1074/jbc.M310980200

Sanders, J. G., Nurk, S., Salido, R. A., Minich, J., Xu, Z. Z., Zhu, Q., Martino, C., Fedarko, M., Arthur, T. D., Chen, F., Boland, B. S., Humphrey, G. C., Brennan, C., Sanders, K., Gaffney, J., Jepsen, K., Khosroheidari, M., Green, C., Liyanage, M., Dang, J. W., Phelan, V. V., Quinn, R. A., Bankevich, A., Chang, J. T., Rana, T. M., Conrad, D. J., Sandborn, W. J., Smarr, L., Dorrestein, P. C., Pevzner, P. A., & Knight, R. (2019, Oct 31). Optimizing sequencing protocols for leaderboard metagenomics by combining long and short reads. *Genome Biol*, *20*(1), 226. https://doi.org/10.1186/s13059-019-1834-9

Satoh, H., Miura, Y., Tsushima, I., & Okabe, S. (2007, Nov). Layered structure of bacterial and archaeal communities and their in situ activities in anaerobic granules. *Appl Environ Microbiol*, *73*(22), 7300-7307. https://doi.org/10.1128/AEM.01426-07

Schaider, L. A., Rodgers, K. M., & Rudel, R. A. (2017, Jul 5). Review of Organic Wastewater Compound Concentrations and Removal in Onsite Wastewater Treatment Systems. *Environ Sci Technol*, *51*(13), 7304-7317. https://doi.org/10.1021/acs.est.6b04778

Schirmer, M., D'Amore, R., Ijaz, U. Z., Hall, N., & Quince, C. (2016, Mar 11). Illumina error profiles: resolving fine-scale variation in metagenomic sequencing data. *BMC Bioinformatics*, *17*, 125. https://doi.org/10.1186/s12859-016-0976-y

Seemann, T. (2014, Jul 15). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068-2069. https://doi.org/10.1093/bioinformatics/btu153

Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., & Huttenhower, C. (2012, Jun 10). Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods*, *9*(8), 811-814. https://doi.org/10.1038/nmeth.2066

Sherrill-Mix, S. (2019). *taxonomizr: Functions to Work with NCBI Accessions and Taxonomy*. In (Version 0.5.3) https://CRAN.R-project.org/package=taxonomizr

Shin, J., Cho, S.-K., Lee, J., Hwang, K., Chung, J. W., Jang, H.-N., & Shin, S. G. (2019). Performance and Microbial Community Dynamics in Anaerobic Digestion of Waste Activated Sludge: Impact of Immigration. *Energies*, *12*(3), 573. https://www.mdpi.com/1996-1073/12/3/573

Shin, S. G., Koo, T., Lee, J., Han, G., Cho, K., Kim, W., & Hwang, S. (2016, Aug). Correlations between bacterial populations and process parameters in four full-scale anaerobic digesters treating sewage sludge. *Bioresource Technology*, *214*, 711-721. https://doi.org/10.1016/j.biortech.2016.05.021

Si, B., Li, J., Li, B., Zhu, Z., Shen, R., Zhang, Y., & Liu, Z. (2015, 2015/09/21/). The role of hydraulic retention time on controlling methanogenesis and homoacetogenesis in biohydrogen production using upflow anaerobic sludge blanket (UASB) reactor and packed bed reactor (PBR).

International Journal of Hydrogen Energy, 40(35), 11414-11421. https://doi.org/https://doi.org/10.1016/j.ijhydene.2015.04.035

Sieber, J. R., McInerney, M. J., & Gunsalus, R. P. (2012). Genomic insights into syntrophy: the paradigm for anaerobic metabolic cooperation. *Annu Rev Microbiol, 66*, 429-452. https://doi.org/10.1146/annurev-micro-090110-102844

Sikora, A., Detman, A., Chojnacka, A., & Blaszczyk, M. K. (2017). Anaerobic Digestion: I. A Common Process Ensuring Energy Flow and the Circulation of Matter in Ecosystems. II. A Tool for the Production of Gaseous Biofuels. In *Fermentation Processes*. https://doi.org/10.5772/64645

Simonte, F. M., Dötsch, A., Galego, L., Arraiano, C., & Gescher, J. (2017, 2017/01/01). Investigation on the anaerobic propionate degradation by Escherichia coli K12. *Molecular Microbiology, 103*(1), 55-66. https://doi.org/10.1111/mmi.13541

Singer, E., Wagner, M., & Woyke, T. (2017, Sep). Capturing the genetic makeup of the active microbiome in situ. *ISME J, 11*(9), 1949-1963. https://doi.org/10.1038/ismej.2017.59

Solli, L., Håvelsrud, O. E., Horn, S. J., & Rike, A. G. (2014, 2014/10/14). A metagenomic study of the microbial communities in four parallel biogas reactors. *Biotechnology for Biofuels, 7*(1), 146. https://doi.org/10.1186/s13068-014-0146-2

Solonenko, S. A., Ignacio-Espinoza, J. C., Alberti, A., Cruaud, C., Hallam, S., Konstantinidis, K., Tyson, G., Wincker, P., & Sullivan, M. B. (2013, May 10). Sequencing platform and library preparation choices impact viral metagenomes. *BMC Genomics, 14*, 320. https://doi.org/10.1186/1471-2164-14-320

Sprott, G. D., Shaw, K. M., & Jarrell, K. F. (1984, 1984/10//). Ammonia/potassium exchange in methanogenic bacteria. *The Journal of biological chemistry, 259*(20), 12602-12608. http://europepmc.org/abstract/MED/6490632

http://intl.jbc.org/cgi/content/abstract/259/20/12602

Stams, A. J. M., Plugge, C. M., de Bok, F. A. M., van Houten, B. H. G. W., Lens, P., Dijkman, H., & Weijma, J. (2005). Metabolic interactions in methanogenic and sulfate-reducing bioreactors. *Water Science and Technology, 52*(1-2), 13-20. https://doi.org/10.2166/wst.2005.0493

Steinberger, A. J. (2018). *seq_scripts.* In (Version 1) http://doi.org/10.5281/zenodo.1458243

Steven, B., McCann, S., & Ward, N. L. (2012, Dec). Pyrosequencing of plastid 23S rRNA genes reveals diverse and dynamic cyanobacterial and algal populations in two eutrophic lakes. *FEMS Microbiol Ecol, 82*(3), 607-615. https://doi.org/10.1111/j.1574-6941.2012.01429.x

Taiyun Wei, V. S. (2017). *R package "corrplot": Visualization of a Correlation Matrix.* In (Version 0.84)) https://github.com/taiyun/corrplot

Tamames, J., de la Pena, S., & de Lorenzo, V. (2012, Jun). COVER: a priori estimation of coverage for metagenomic sequencing. *Environ Microbiol Rep, 4*(3), 335-341. https://doi.org/10.1111/j.1758-2229.2012.00338.x

Tessler, M., Neumann, J. S., Afshinnekoo, E., Pineda, M., Hersch, R., Velho, L. F. M., Segovia, B. T., Lansac-Toha, F. A., Lemke, M., DeSalle, R., Mason, C. E., & Brugler, M. R. (2017, Jul 31). Large-scale differences in microbial biodiversity discovery between 16S amplicon and shotgun sequencing. *Sci Rep, 7*(1), 6589. https://doi.org/10.1038/s41598-017-06665-3

Tomaras, J., Sahl, J. W., Siegrist, R. L., & Spear, J. R. (2009, May). Microbial diversity of septic tank effluent and a soil biomat. *Appl Environ Microbiol, 75*(10), 3348-3351. https://doi.org/10.1128/AEM.00560-08

Tran, N. H., Reinhard, M., & Gin, K. Y. (2018, Apr 15). Occurrence and fate of emerging contaminants in municipal wastewater treatment plants from different geographical regions-a review. *Water Res, 133*, 182-207. https://doi.org/10.1016/j.watres.2017.12.029

Treu, L., Campanaro, S., Kougias, P. G., Zhu, X., & Angelidaki, I. (2016, Jun 7). Untangling the Effect of Fatty Acid Addition at Species Level Revealed Different Transcriptional Responses of the Biogas Microbial Community Members. *Environ Sci Technol, 50*(11), 6079-6090. https://doi.org/10.1021/acs.est.6b00296

Truong, D. T., Franzosa, E. A., Tickle, T. L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C., & Segata, N. (2015, Oct). MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods, 12*(10), 902-903. https://doi.org/10.1038/nmeth.3589

Utturkar, S. M., Klingeman, D. M., Hurt, R. A., & Brown, S. D. (2017, 2017-July-18). A Case Study into Microbial Genome Assembly Gap Sequences and Finishing Strategies [Original Research]. *Frontiers in Microbiology, 8*(1272). https://doi.org/10.3389/fmicb.2017.01272

van Lier, J. B. (2008). High-rate anaerobic wastewater treatment: diversifying from end-of-the-pipe treatment to resource-oriented conversion techniques. *Water Sci Technol, 57*(8), 1137-1148. https://doi.org/10.2166/wst.2008.040

Vanwonterghem, I., Evans, P. N., Parks, D. H., Jensen, P. D., Woodcroft, B. J., Hugenholtz, P., & Tyson, G. W. (2016, Oct 3). Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nat Microbiol, 1*, 16170. https://doi.org/10.1038/nmicrobiol.2016.170

Vanwonterghem, I., Jensen, P. D., Dennis, P. G., Hugenholtz, P., Rabaey, K., & Tyson, G. W. (2014, Oct). Deterministic processes guide long-term synchronised population dynamics in replicate anaerobic digesters. *ISME J, 8*(10), 2015-2028. https://doi.org/10.1038/ismej.2014.50

Vanwonterghem, I., Jensen, P. D., Rabaey, K., & Tyson, G. W. (2015, Feb 16). Temperature and solids retention time control microbial population dynamics and volatile fatty acid production in replicated anaerobic digesters. *Sci Rep, 5*, 8496. https://doi.org/10.1038/srep08496

Vanwonterghem, I., Jensen, P. D., Rabaey, K., & Tyson, G. W. (2016, Sep). Genome-centric resolution of microbial diversity, metabolism and

interactions in anaerobic digestion. *Environ Microbiol*, *18*(9), 3144-3158. https://doi.org/10.1111/1462-2920.13382

Venkata Mohan, S., Agarwal, L., Mohanakrishna, G., Srikanth, S., Kapley, A., Purohit, H. J., & Sarma, P. N. (2011, 2011/07/01/). Firmicutes with iron dependent hydrogenase drive hydrogen production in anaerobic bioreactor using distillery wastewater. *International Journal of Hydrogen Energy*, *36*(14), 8234-8242. https://doi.org/https://doi.org/10.1016/j.ijhydene.2011.04.021

Venkiteshwaran, K., Milferstedt, K., Hamelin, J., & Zitomer, D. H. (2016, Nov 1). Anaerobic digester bioaugmentation influences quasi steady state performance and microbial community. *Water Res*, *104*, 128-136. https://doi.org/10.1016/j.watres.2016.08.012

Verhougstraete, M. P., Martin, S. L., Kendall, A. D., Hyndman, D. W., & Rose, J. B. (2015, Aug 18). Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at the basin scale. *Proc Natl Acad Sci U S A*, *112*(33), 10419-10424. https://doi.org/10.1073/pnas.1415836112

Verma, S. K., & Sharma, P. C. (2020, Jul). NGS-based characterization of microbial diversity and functional profiling of solid tannery waste metagenomes. *Genomics*, *112*(4), 2903-2913. https://doi.org/10.1016/j.ygeno.2020.04.002

Walters, W., Hyde, E. R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J. A., Jansson, J. K., Caporaso, J. G., Fuhrman, J. A., Apprill, A., & Knight, R. (2016, Jan-Feb). Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems*, *1*(1). https://doi.org/10.1128/mSystems.00009-15

Wang, H.-Z., Lv, X.-M., Yi, Y., Zheng, D., Gou, M., Nie, Y., Hu, B., Nobu, M. K., Narihiro, T., & Tang, Y.-Q. (2019, 2019/11/22). Using DNA-based stable isotope probing to reveal novel propionate- and acetate-oxidizing bacteria in propionate-fed mesophilic anaerobic chemostats. *Scientific Reports*, *9*(1), 17396. https://doi.org/10.1038/s41598-019-53849-0

Wang, J., & Chu, L. (2016, Nov 1). Biological nitrate removal from water and wastewater by solid-phase denitrification process. *Biotechnol Adv, 34*(6), 1103-1112. https://doi.org/10.1016/j.biotechadv.2016.07.001

Wang, R., Lin, J. Q., Liu, X. M., Pang, X., Zhang, C. J., Yang, C. L., Gao, X. Y., Lin, C. M., Li, Y. Q., Li, Y., Lin, J. Q., & Chen, L. X. (2018). Sulfur Oxidation in the Acidophilic Autotrophic Acidithiobacillus spp. *Front Microbiol, 9*, 3290. https://doi.org/10.3389/fmicb.2018.03290

Wang, X. L., Lu, Y. L., Han, J. Y., He, G. Z., & Wang, T. Y. (2007). Identification of anthropogenic influences on water quality of rivers in Taihu watershed. *J Environ Sci (China), 19*(4), 475-481. https://doi.org/10.1016/s1001-0742(07)60080-1

Wang, Z. W., Ma, J., & Chen, S. (2011, Jan). Bipolar effects of settling time on active biomass retention in anaerobic sequencing batch reactors digesting flushed dairy manure. *Bioresour Technol, 102*(2), 697-702. https://doi.org/10.1016/j.biortech.2010.08.045

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wiegand, C., & Pflugmacher, S. (2005, Mar 15). Ecotoxicological effects of selected cyanobacterial secondary metabolites: a short review. *Toxicol Appl Pharmacol, 203*(3), 201-218. https://doi.org/10.1016/j.taap.2004.11.002

Wigginton, S. K., Brannon, E. Q., Kearns, P. J., Lancellotti, B. V., Cox, A., Moseman-Valtierra, S., Loomis, G. W., & Amador, J. A. (2020). Nitrifying and Denitrifying Microbial Communities in Centralized and Decentralized Biological Nitrogen Removing Wastewater Treatment Systems. *Water, 12*(6), 1688. https://www.mdpi.com/2073-4441/12/6/1688

Wilkinson, G. N., & Rogers, C. E. (1973). Symbolic Description of Factorial Models for Analysis of Variance. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 22*(3), 392-399. https://doi.org/10.2307/2346786

Wood, D. E., & Salzberg, S. L. (2014, Mar 3). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol*, *15*(3), R46. https://doi.org/10.1186/gb-2014-15-3-r46

Wu, L., Shan, X., Chen, S., Zhang, Q., Qi, Q., Qin, Z., Yin, H., Zhou, J., He, Q., & Yang, Y. (2020). Progressive Microbial Community Networks with Incremental Organic Loading Rates Underlie Higher Anaerobic Digestion Performance. *mSystems*, *5*(1), e00357-00319. https://doi.org/10.1128/mSystems.00357-19

Xia, A., Jacob, A., Herrmann, C., & Murphy, J. D. (2016, 2016/02/01/). Fermentative bio-hydrogen production from galactose. *Energy*, *96*, 346-354. https://doi.org/https://doi.org/10.1016/j.energy.2015.12.087

Xia, Y., Yang, C., & Zhang, T. (2018, 2018/02/21). Microbial effects of part-stream low-frequency ultrasonic pretreatment on sludge anaerobic digestion as revealed by high-throughput sequencing-based metagenomics and metatranscriptomics. *Biotechnology for Biofuels*, *11*(1), 47. https://doi.org/10.1186/s13068-018-1042-y

Xu, R., Yang, Z. H., Zheng, Y., Liu, J. B., Xiong, W. P., Zhang, Y. R., Lu, Y., Xue, W. J., & Fan, C. Z. (2018, Aug). Organic loading rate and hydraulic retention time shape distinct ecological networks of anaerobic digestion related microbiome. *Bioresource Technology*, *262*, 184-193. https://doi.org/10.1016/j.biortech.2018.04.083

Yang, Y., Chen, Q., Guo, J., & Hu, Z. (2015, 2015/12/15/). Kinetics and methane gas yields of selected C1 to C5 organic acids in anaerobic digestion. *Water Research*, *87*, 112-118. https://doi.org/https://doi.org/10.1016/j.watres.2015.09.012

Yang, Y., Yu, K., Xia, Y., Lau, F. T., Tang, D. T., Fung, W. C., Fang, H. H., & Zhang, T. (2014, Jun). Metagenomic analysis of sludge from full-scale anaerobic digesters operated in municipal wastewater treatment plants. *Appl Microbiol Biotechnol*, *98*(12), 5709-5718. https://doi.org/10.1007/s00253-014-5648-0

Yee, M. O., Snoeyenbos-West, O. L., Thamdrup, B., Ottosen, L. D. M., & Rotaru, A.-E. (2019, 2019-April-02). Extracellular Electron Uptake by Two

Methanosarcina Species [Original Research]. *Frontiers in Energy Research, 7*(29). https://doi.org/10.3389/fenrg.2019.00029

Yin, D. M., Westerholm, M., Qiao, W., Bi, S. J., Wandera, S. M., Fan, R., Jiang, M. M., & Dong, R. J. (2018, Sep). An explanation of the methanogenic pathway for methane production in anaerobic digestion of nitrogen-rich materials under mesophilic and thermophilic conditions. *Bioresource Technology, 264*, 42-50. https://doi.org/10.1016/j.biortech.2018.05.062

Yue, Z., Teater, C., MacLellan, J., Liu, Y., & Liao, W. (2011, 2011/05/01/). Development of a new bioethanol feedstock – Anaerobically digested fiber from confined dairy operations using different digestion configurations. *Biomass and Bioenergy, 35*(5), 1946-1953. https://doi.org/https://doi.org/10.1016/j.biombioe.2011.01.035

Zaheer, R., Noyes, N., Ortega Polo, R., Cook, S. R., Marinier, E., Van Domselaar, G., Belk, K. E., Morley, P. S., & McAllister, T. A. (2018, Apr 12). Impact of sequencing depth on the characterization of the microbiome and resistome. *Sci Rep, 8*(1), 5890. https://doi.org/10.1038/s41598-018-24280-8

Zhang, C., Yuan, Q., & Lu, Y. (2018, Dec 1). Inhibitory effects of ammonia on syntrophic propionate oxidation in anaerobic digester sludge. *Water Res, 146*, 275-287. https://doi.org/10.1016/j.watres.2018.09.046

Zhang, J., Wei, Y., Xiao, W., Zhou, Z., & Yan, X. (2011, 2011/08/01/). Performance and spatial community succession of an anaerobic baffled reactor treating acetone–butanol–ethanol fermentation wastewater. *Bioresource Technology*, *102*(16), 7407-7414. https://doi.org/https://doi.org/10.1016/j.biortech.2011.05.035

Zhang, J. X., Mao, L. W., Zhang, L., Loh, K. C., Dai, Y. J., & Tong, Y. W. (2017, Sep 12). Metagenomic insight into the microbial networks and metabolic mechanism in anaerobic digesters for food waste by incorporating activated carbon. *Scientific Reports*, *7*. https://doi.org/ARTN 11293

10.1038/s41598-017-11826-5

Zhang, K., Gu, J., Wang, X., Yin, Y., Zhang, X., Zhang, R., Tuo, X., & Zhang, L. (2018, 2018/09/01/). Variations in the denitrifying microbial community

and functional genes during mesophilic and thermophilic anaerobic digestion of cattle manure. *Science of The Total Environment, 634*, 501-508. https://doi.org/https://doi.org/10.1016/j.scitotenv.2018.03.377

Zhang, L., Guo, B., Mou, A., Li, R., & Liu, Y. (2020, 2020/11/15/). Blackwater biomethane recovery using a thermophilic upflow anaerobic sludge blanket reactor: Impacts of effluent recirculation on reactor performance. *Journal of Environmental Management, 274*, 111157. https://doi.org/https://doi.org/10.1016/j.jenvman.2020.111157

Zhang, Q., Wang, M., Ma, X., Gao, Q., Wang, T., Shi, X., Zhou, J., Zuo, J., & Yang, Y. (2019, May). High variations of methanogenic microorganisms drive full-scale anaerobic digestion process. *Environ Int, 126*, 543-551. https://doi.org/10.1016/j.envint.2019.03.005

Zhao, Q., & Liu, Y. (2019). Is anaerobic digestion a reliable barrier for deactivation of pathogens in biosludge? *The Science of the total environment, 668*, 893-902. https://doi.org/10.1016/j.scitotenv.2019.03.063

Zhilina, T. N., Zavarzina, D. G., Kuever, J., Lysenko, A. M., & Zavarzin, G. A. (2005, May). Desulfonatronum cooperativum sp. nov., a novel hydrogenotrophic, alkaliphilic, sulfate-reducing bacterium, from a syntrophic culture growing on acetate. *Int J Syst Evol Microbiol, 55*(Pt 3), 1001-1006. https://doi.org/10.1099/ijs.0.63490-0

Zhou, J.-J., Shen, J.-T., Jiang, L.-L., Sun, Y.-Q., Mu, Y., & Xiu, Z.-L. (2017, 2017/08/01). Selection and characterization of an anaerobic microbial consortium with high adaptation to crude glycerol for 1,3-propanediol production. *Applied Microbiology and Biotechnology, 101*(15), 5985-5996. https://doi.org/10.1007/s00253-017-8311-8

Zhu, X., Campanaro, S., Treu, L., Seshadri, R., Ivanova, N., Kougias, P. G., Kyrpides, N., & Angelidaki, I. (2020, Feb 15). Metabolic dependencies govern microbial syntrophies during methanogenesis in an anaerobic digestion ecosystem. *Microbiome, 8*(1), 22. https://doi.org/10.1186/s40168-019-0780-9

Ziels, R. M., Nobu, M. K., & Sousa, D. Z. (2019). Elucidating Syntrophic Butyrate-Degrading Populations in Anaerobic Digesters Using Stable-

Isotope-Informed Genome-Resolved Metagenomics. *mSystems*, *4*(4), e00159-00119. https://doi.org/10.1128/mSystems.00159-19

Ziganshin, A. M., Schmidt, T., Lv, Z., Liebetrau, J., Richnow, H. H., Kleinsteuber, S., & Nikolausz, M. (2016, 2016/10/01/). Reduction of the hydraulic retention time at constant high organic loading rate to reach the microbial limits of anaerobic digestion in various reactor systems. *Bioresource Technology*, *217*, 62-71. https://doi.org/https://doi.org/10.1016/j.biortech.2016.01.096