5-2020

Anaerobic Digestion and a Core Microbiome

Andrii Vatulin
avatulin@hotmail.com

Anaerobic Digestion and a Core Microbiome

by

Andrii Vatulin

A Thesis

Submitted to the Graduate Faculty of

St Cloud State University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science in

Biology: Ecology and Evolution

May, 2020

Thesis Committee
Ryan C. Fink, Chairperson
Matthew P. Davis
Omar Al-Azzam

Abstract

Anaerobic Digestion is a microbially mediated process turning organic matter into biogas and biofertilizer. This kind of waste decomposition is advantageous over traditional waste management for its low energy requirements, potential energy recovery, reduction of greenhouse gas released into the atmosphere, and production of environmentally friendly fertilizers. However, lack of information about the establishment and stability of the core microbial community composition needed to sustain this process and to make it economically viable has hampered its deployment. Decrease of the biogas production caused by a fatal microbial community collapse is one of the major issues encountered in small-scaled and commercial enterprise using the technology. This study focuses on the commonalities in microbial community compositions of infeed and digestate present in four anaerobic digesters different in their designs, infeeds, sizes, and operational temperatures to determine a shared microbial community. Anaerobic digesters situated on a farm, at a wastewater treatment facility, University of Minnesota laboratories, and a bench fermentation set up at St Cloud State University were sampled: in-feed, digestate, outfeed. These digesters operate respectively on manure, wastewater and high strength waste from breweries, manure and food waste mix from the campus cafeteria, and a strictly food waste (calculated ingredients proportions). All digesters operate in mesophilic conditions; the sizes were from two liters to 1.6 million liters; and hydraulic retention times were from 9 to 58 days. Samples were collected from all points where organic matter was hypothesized to be changing composition of its microbial community. The microbial communities were characterized using bacterial and archaeal specific 16S rRNA primers and high throughput sequencing with Illumina Miseq to the genus level. Our study determined 14 genera that was high abundant and overlapping at least with two anaerobic digestion systems. Also, our analysis showed that the three out of four sites shared *Methanobrevibacter* as the dominating methanogenic genus; *Lactobacillus* and *Clostridium* (Ruminococcaceae family) were highly abundant (>1%) and shared between all anaerobic digesters. A repeated sampling of the same sites over time would give an even more reliable list of core microorganisms. A furthermore accurate determination of a core microbial “recipe” is a valuable instrument that allows for the establishment of a stable yet diverse community and at the same time will assist an operator in cases when a microbial community is struggling due to the changes in infeed physical or chemical composition.

Table of Contents

Chapter Page

List of Tables 5

List of Figures 6

1. Introduction 11

Traditional Energy Sources: Economic Impacts 11

Traditional Energy Sources: Environmental Impacts 13

Available Alternative Energy Sources 16

Biogas production from anaerobic digestion 18

Metabolic process of anaerobic digestion 19

Microbial Composition: a Cross-sectional Meta-analysis 23

Benchtop Fermentor Study at SCSU: Set up 25

Benchtop Fermentor Study at SCSU: 16S Profiling 27

2. Comparison of Microbial Representation in Dissimilar Anaerobic Digesters 29

Introduction 29

Hypothesis 31

On-farm Anaerobic Digestion Systems 32

Design Types of Animal On-farm Anaerobic Digesters 32

Plug Flow Anaerobic Digester Sampling Site: 35

Sampling 35

Materials and Methods 38

Chapter Page

DNA Extraction, Quantification, and 16S rRNA Sequencing..... 38
Analysis and Visualization of Sequenced Data..... 40
Results: Microbial Composition and Diversity Analysis..... 44
Municipal Wastewater Treatment and Anerobic Co-digestion..... 55
Municipal Wastewater Co-digestion Sampling Site..... 56
Sampling..... 57
Materials and Methods..... 59
Results: Microbial Composition and Diversity Analysis..... 59
Experimental Benchtop Anaerobic Digesters: Value and Purpose..... 70
Experimental Benchtop Anaerobic Digester Sampling Site..... 70
Sampling..... 71
Materials and Methods..... 73
Results: Microbial Composition and Diversity Analysis..... 73
3. Microbial Composition Analysis of Core Microbiomes of All Sampled Digesters..... 82
Introduction..... 82
Microbial Composition..... 82
4. Discussion..... 95
References..... 102

List of Tables

Table Page

1.1. Overlap in microbial taxa between 136 digesters from 9 countries 24

2.1. Metadata table including operational information for four sampled and analyzed digestion sites used for QIIME2 analysis 41

4.1. The overlap in microbial taxa between four AD systems plus 136 digesters from 9 countries (marked with*) 97

4.2. The overlap in microbial taxa between four AD systems during anaerobic phase with 136 digesters from 9 countries (marked with*) 98

List of Figures

Figure Page

1.1. Historical trend of nominal coal, oil, and natural gas prices in the period from 1950 to 2008 8

1.2. The transformation of organic matter to methane.....17

2.1. Haubenschild farm schematic, anaerobic digester (AD) set-up, and main sampling points (*) 37

2.2. A summary of demultiplxed sequences (10000 randomly selected by QIIME2 v 2019.1) forward and reverse reads timed up to basepair 10 and truncated from basepair 280 and 240 accordingly 43

2.3. Haubenschild farm high abundance genera (>1%), low abundance genera grouped in one sector (<1%), and unassigned sequences in aerobic conditions.....46

2.4. Haubenschild farm high abundance genera (>1%), low abundance genera grouped in one sector (<1%), and unassigned sequences in anaerobic conditions.....47

2.5. Alpha diversity analysis of (A) observed OTUs, (B) Faith's phylogenetic diversity (Faith's PD) and (C) Evenness of Haubenschild farm sampling sites: pre- (green), during (yellow) and post- (purple) digestion 50

2.6. Beta diversity analysis of (A) Bray-Curtis, (C) Jaccard, and (E) un-weighted Unifrac distances during anaerobic and aerobic phases in the AD

Figure Page

system; (B) Bray-Curtis, (D) Jaccard, (F) un-weighted Unifrac distances between different sampling points at the Haubenschild farm sampling sites with the Prok primer points in a green frame and the Arc primer results outside the frame 53

2.7. St Cloud WWTF schematic, including AD set-up and main sampling points (indicated by the asterisks *) 58

2.8. St Cloud WWTF sampling site: high abundance genera (>1%), low abundance genera (<1%), and unassigned sequences in aerobic conditions.61

2.9. St Cloud WWTF sampling site: high abundance genera (>1%), low abundance genera (<1%), and unassigned sequences in anaerobic conditions.....62

2.10. Alpha diversity analysis of (A) observed OTUs, (B) Faith's phylogenetic diversity (Faith's PD) and (C) Evenness of St Cloud WWTF: pre- (yellow), during (green), and post- (purple) digestion.....65

2.11**.** Beta diversity analysis of (A) Bray-Curtis, (C) Jaccard, and (E) un-weighted Unifrac distances during anaerobic and aerobic phases in the AD system; (B) Bray-Curtis, (D) Jaccard, (F) un-weighted Unifrac distances between different sampling points at the WWTF sampling sites with the Prok primer points in a green frame and the Arc primer results outside the frame.....68

Figure Page

2.12. University of Minnesota, Dr Hu experimental benchtop AD set up; main sampling points (indicated by the asterisks *) 72

2.13. University of Minnesota benchtop AD sampling site: high abundance genera (>1%), low abundance genera (<1%), and unassigned sequences in aerobic conditions 75

2.14. University of Minnesota benchtop AD sampling site: high abundance genera (>1%), low abundance genera (<1%), and unassigned sequences in anaerobic conditions 76

2.15. Alpha diversity analysis of (A) observed OTUs, (B) Faith's phylogenetic diversity (Faith's PD) and (C) Evenness of University of Minnesota benchtop digester: pre- (white) and during (green) digestion with Digester 1 being an aerobic chamber and Digester 2 being an anaerobic chamber 78

2.16**.** Beta diversity analysis of (A) Bray-Curtis, (C) Jaccard, and (E) un-weighted Unifrac distances during anaerobic and aerobic phases in the AD system; (B) Bray-Curtis, (D) Jaccard, (F) un-weighted Unifrac distances between different sampling points at the University of Minnesota benchtop digester sampling sites with the Prok primer points in a green frame and the Arc primer results outside the frame 81

3.1. Haubenschild farm, WWTF, University of Minnesota benchtop AD, and SCSU food fermenter sampling sites: high abundance genera (>1%) and low abundance genera (<1%) 84

Figure Page

3.2. Haubenschild farm, WWTF, University of Minnesota benchtop AD, and SCSU food fermenter ADs sampling results: high abundance genera (>1%) and low abundance genera (<1%) 86

3.3. Alpha diversity analysis of Faith's phylogenetic diversity (Faith's PD) of Haubenschild farm (Farm), SCSU food fermenter (SCSU ISELF), University of Minnesota benchtop digester (UoM bechto AD), and St Cloud WWTF (WW treatment).....88

3.4. Alpha diversity analysis of species evenness (Pielou's diversity) of Haubenschild farm (Farm), SCSU food fermenter (SCSU ISELF), University of Minnesota benchtop digester (UoM bechto AD), and St Cloud WWTF (WW treatment).....89

3.5. Alpha diversity analysis of species evenness (Pielou's diversity) of Haubenschild farm, SCSU food fermenter, University of Minnesota benchtop digester, and St Cloud WWTF.....90

3.6 Alpha diversity analysis of species observed OTUs of Haubenschild farm, SCSU food fermenter, University of Minnesota benchtop digester, and St Cloud WWTF.....90

3.7. Beta diversity analysis of (A) Bray-Curtis, (C) Jaccard, and (E) un-weighted Unifrac distances during anaerobic and aerobic phases in the AD system; (B) Bray-Curtis, (D) Jaccard, (F) un-weighted Unifrac distances between all sampling points at the Haubenschild farm, SCSU food fermenter, University

of Minnesota benchtop digester, and St Cloud WWTF sampling sites with the Prok primer points in a green frame and the Arc primer results outside the frame 94

Chapter 1: Introduction

Traditional Energy Sources: Economic Impacts

The development of reliable, renewable, and cost-effective alternative energy sources is a defining priority of contemporary societies, both in developed and developing countries. This need stems from the now consolidated awareness that global reserves and deposits of fossil fuels (e.g. coal, oil, and natural gas) are finite and depleting at an ever-faster pace. Furthermore, the uneven distribution of these resources and the world overdependence on fossil fuels results in inequitable dependency of most countries on a handful of producers and erratic energy price fluctuations.

Currently, these non-renewable energy sources dominate the 1.5 trillion USD worldwide energy market, accounting for 86% of the global energy output (1). However, with the current rate of production and consumption, it is estimated that coal, crude oil, and natural gas will be potentially exhausted within 150, 60, and 40 years, respectively. Some models' estimates predict even shorter timelines (2). At this time, there is still conflicting evidence about the efficiency and cost-effectiveness of alternative energy sources compared to fossil fuels. As a result, a lot of economic and research resources are being deployed to develop equivalent substitutes (3).

It is important also to consider the impact of the global reliance on few non-renewable sources of energy. This is particularly evident when examining global historical market trends of crude oil, natural gas, and coal. In fact, it is rather evident that price fluctuations can swing dramatically within a single year. For instance, crude oil prices increased significantly from 1998 to 2008 (Figure 1.1) (4), with the prices going from 70

USD to 145 USD between July 2007 and November 2008 (5), then falling to 39 USD in February 2009 (6).

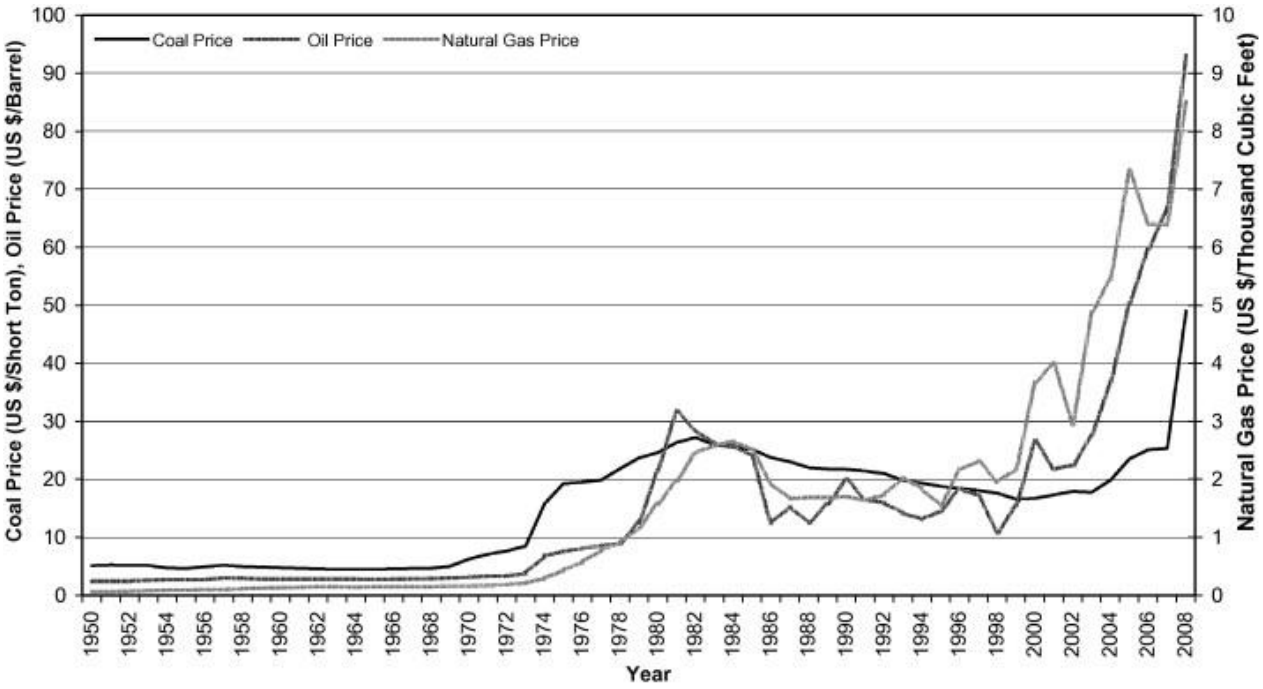


*Figure 1.1***.** Historical trend of nominal coal, oil, and natural gas prices in the period from 1950 to 2008 (adapted from (4)).

This issue is further compounded by the unequal geographic distribution of these sources in a handful of major producing countries (e.g., Russia, Australia, Canada, Saudi Arabia, etc.) which leads to a high polarization of the energy market. In total, only 15 countries are responsible for the production of 80% of crude oil and 38% of its processing is controlled by Russia, Saudi Arabia, and the USA. As for reserves of fossil fuels, it was estimated in 2017 that 47.6% of crude oil and 40.9% of natural gas are concentrated in the Middle East, compared to only 15.5% and 5.6% in North America, respectively (7). Obviously, these drawbacks are a major concern, but they cannot be compared to the magnitude of environmental impacts of fossil fuels use even remotely.

Traditional Energy Sources: Environmental Impacts

The environmental impact of the use of fossil fuels is difficult to quantify, but it is arguably very large (8). Since the 18th century, following the industrial revolution, fossil fuels have been the primary energy source for all aspects of production, transport, and manufacturing. There are multiple environmental threats associated with the use of fossil fuels as the main energy source: emission of greenhouse gases, spills during transportation, physical pollution of air with microscopic particles (smog), etc. Indeed, by the end of the 20th century, scientists and world leaders understood and acknowledged that the use of these energy sources was being harmful to the environment and human health, but, more importantly, was leading to unprecedented changes in the global climate (8, 9).

Global climate change is a consequence of increased percentage of greenhouse gases in the atmosphere due in large part to combustion of fossil fuels (10). Greenhouse gases are defined as such because (11, 12), very much like a greenhouse and the outside environment, they absorb and trap heat preventing its dissipation into outer space. The three main climate-affecting greenhouse gasses are carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O) (13). A certain percentage of greenhouse gases has always been present in the environment, however the increase in concentration observed in the past four decades has caused a significant rise in worldwide temperature in the atmosphere as well as on or below ground and in water bodies such as oceans and lakes (14).

The United States Environmental Protection Agency (EPA) determined that total emissions of greenhouse gases in 2010 in the USA were an estimated total of 6,457 million metric tons of CO_2 equivalent and originating from industry (22.2%), electricity production (27.5%), agriculture (9%), transportation (28.9%), residential and commercial sector (11.6%), and land use and forestry (11.1%). The EPA also reported that global carbon emissions went from 500 million metric tons in 1900 to almost 10,000 million in 2017 (15). It has also been shown that overall approximately 70% of anthropogenic emissions of greenhouse gases are coming from the global energy sector (16, 17).

Another negative effect on the environment comes from the production and transportation of fossil fuels as spills and leakage of crude oil in the ocean. It is estimated that the total amount of oil spilled into marine waters globally from various sources (e.g., oil tankers, drilling platforms, pipelines, storage, etc.) was approximately 1.5 billion gallons in the 1970s, 845 million gallons in the 1980s, and 943 million gallons in the 1990s. In 2000s the number of oil spilled is much lower than the previous decades, however, only one spill in 2018 was around 36 million gallons from one tanker and is considered the worst spill disaster in decades (18, 19). Crude oil discharged in the ocean negatively impacts marine ecosystems, human health, and the overall quality of potable water (20). Crude oil is toxic for fish and birds, it also can form a highly adhesive layer on birds' feathers impairing their ability to fly. It does not dissolve in water, forming a thick film that suffocates fish and stops the sunlight needed for aquatic plants. Additionally, when reaching shores of water bodies crude oil toxicity kills organisms residing there (21).

Lastly, smog created by the combustion of fossil fuels causes air pollution, with coal being the most damaging culprit. Smog consists of fine particulate, nitrogen oxides, sulfur oxides, ozone, smoke, and other trace chemicals. It is proven that there is a direct correlation between death rates and the severity of smog. It affects the respiratory, cardiovascular, and immune systems of people, animals, insects, etc. For example, Scientists proved that around 4000 deaths were caused by smog in London in the middle of the twentieth century. The smog in 1952 became the worst pollution case in the history of Great Britain involving thousands of fatalities and forced authorities to concentrate on finding alternative energy sources and decrease the concentration of smog in the atmosphere (22). However, high smog levels continue to be a deadly problem in many countries. For example, in Ulaanbaatar City, Mongolia with a population of 1.3 million people annually each family burns five tons of coal and three cubic meters of wood for heating annually and, in addition, there are three coal-fueled power plants, four hundred heat-only boilers, and around 300,000 vehicles (23). This mix of pollutants makes the air quality in Ulaanbaatar one of the worst in the world, according to the 2009 World Bank report (24). In fact, the daily average of fine particles in the air during winter is around 750 mg/m^3 , when the World Health Organization guideline is 10 mg/m^3 (25).

All the above mentioned issues caused by traditional energy sources have led to a lot of financial and scientific resources being devoted to the development of clean, affordable, and renewable alternative energy sources to achieve the following goals: protect and improve the environment, reduce chemical and thermal pollution, and decrease the overall impact. Furthermore, new energy sources will provide affordable

energy for people's essential needs in their everyday life: supplying potable running water, cooking food, heating houses, and much more (26–28).

Available Alternative Energy Sources

As summarized by Yinghao Chu from the Global Energy Network Institute, different alternative-energy (energy source that does not use fossil fuels(8)) technologies have been adopted more broadly than others. Those that utilize the power of wind, water, solar radiation, and biogas to generate electricity are the most common. Indeed, some of them have been known and used for many centuries. Wind energy, for example, has been known to people as a source of energy since ancient times. According to the Office of Energy Efficiency and Renewable Energy website (energy.gov), archeological and historic records show that wind has been used by ancient Egyptians to propel boats and mill grains since 5000 BC, and by Chinese to lift water from wells since 200 BC. The first industrial wind turbine producing electricity was built in the USA at the end of the 19th century. However, only in the last two decades wind turbines in the USA have become an important player in the generation of electricity. One fifth of all the electricity from wind turbines globally is produced in the USA and this capacity keeps growing across the country (energy.gov). In the European Union (EU) in 2018 wind power covered 14% of electricity demand with 362 TWh generated. In the same year, in Denmark, the share of wind power went up to 41% of energy generating sources and became the highest in the EU. Overall, in 2018, 63% of the fundings into renewable energy in the EU was invested in wind-based energy production (29).

The other most popular alternative energy technology is based on the use of solar energy to generate heat and electricity. This technology is clean, renewable, and has become increasingly cost effective. Indeed, the use of solar energy, if efficiently harvested, would have the potential of satisfying the energy needs of the entire planet. For example, the energy emitted from the sun in one day, as measured on Earth, could satisfy the current demand for electricity for more than 20 years.

As for the current state, according to 2015 data, the USA alone have 27,500 megawatts of solar power capacity. Furthermore, only in 2015 the solar power capacity to generate 7,260 megawatts was added in the USA following the global interest in this energy source. Overall, the total global investment into solar energy sector exceeded 25 billion USD in 2018. Also, the National Renewable Energy Laboratory estimated that the solar energy potential within the United States can be up to 400 zettawatt hours comparing to the current overall generation capacity of 22,813 terawatt-hours (3, 30). However, there is still a long way to go for these technologies to substitute traditional sources of energy. Still, there is a constant interest in research and development of different alternative ways to produce energy (31).

Biogas production from anaerobic digestion

Albeit not as commonly deployed as solar or wind-based energy producing technologies, anaerobic digestion is the only technology that addresses both the generation of clean heat and electricity and the reduction waste via the production of high-quality biogas from various organic feedstocks (e.g. manure, kitchen food waste, waste from food manufacturing, wastewater sludges, etc.) (32, 33).

Currently, the accumulation of solid organic waste produced by agriculture, food industry (including processing and retail), city sewage systems, and households ends up in landfills or other types of open-air storage facilities. Open-air decomposition of these materials generates additional greenhouse gases. Globally, farm manure alone contributes to 5% of anthropogenic greenhouse gas emissions. Global livestock contributes 18% of global methane emissions primarily through aerobic decomposition (34). At the same time, methane is considered to be the biggest contributor to global warming (35).

The World Bank estimates that by 2025, cities around the world will be producing 2.2 billion tons of solid waste annually. Even now, the amount that is disposed into landfills is very high. For instance, Hong Kong disposed of around 5.5 million tons of solid waste in 2016, 60%of which was municipal solid waste that could have been treated via anaerobic digestion to generate energy and fertilizers for agriculture (36). However, presently, the US is planning to “harvest” close to 36 billion gallons of biogas by 2022 and there is a considerable interest in the anaerobic digestion technology, its improvement, and understanding of its multiple stages is constantly growing (37). The local benefits

associated with the use of anaerobic digestion technology are the production of inexpensive biogas and reduction of green gas emissions. Globally recognized benefits include improving air quality by reducing odor emissions and in some instances even receiving carbon credit payments. Also, anaerobic digestion contributes to improving and protecting the quality of water by eliminating the quantity of pathogens from wastewater and manure entering water bodies and ground water and eliminates weed seeds in manure that reduces the use of herbicides on fields. Furthermore, biogas generates an excess of energy for extra revenue for digester operators and the outcome of anaerobic reactors (digested fiber can be used as a bedding on farms (32, 38).

Metabolic process of anaerobic digestion

Anaerobic digestion is a complex process of decomposition of organic matter by microorganisms (bacteria and archaea) in the absence of oxygen. This process occurs in controlled conditions and leads to the production of biogas (methane, carbon dioxide, and traces of other gases) and compost (decayed organic matter). Biogas then can be transformed by a generator to heat and electricity, and compost used as a high-value fertilizer for agriculture (27, 39–42).

The process of anaerobic digestion consists of four distinct stages. In the first stage, called hydrolysis, complex components of organic matter like fats, proteins, and carbohydrates are decomposed into monomers. In fact, enzymes released by fermentative bacteria hydrolyze macromolecules into sugars, amino acids, and long-chain fatty acids. During the second stage of this process, called acidogenesis, facultative and obligate anaerobic bacteria uptake and decompose these organic compounds into short-

chain fatty acids including volatile fatty acids (lactic acid, acetic acid, formic acid). During this stage, short-chain volatile fatty acids combine with alcohol and are converted into acetic acid, carbon dioxide and hydrogen (27). In the third stage, acetogenesis that is considered a secondary fermentation process, where acetogenic bacteria digests lactic acid and pyruvic acid into acetic acid and hydrogen. As well as in the second stage, during the acetogenesis some carbon dioxide is also produced. The fourth stage, methanogenesis, is critical in the anaerobic digestion process, because close to 70% of methane is produced during this step. Methanogenesis is a much slower stage comparing to other processes of AD and is considered to be rate-limiting step. During this stage, methanogenic precursors produced in the previous stages are utilized by methanogens that are obligate anaerobes. Two types of methanogens are considered to be responsible for the final step and represent two main biomethane production pathways: acetoclastic methanogens (major methane producers, up to 70% of methane produced) that utilize acetate for methane production; and hydrogen-utilizing methanogens that use hydrogen for the reduction of the process intermediates such as carbon dioxide, methyl animate, and methanol to produce methane (CH_4), the final product of anaerobic digestion (Figure 1.2) (27, 37, 43). However, there is also a third pathway of methane production during the AD process's final stage: methylotrophic pathway, where methanol can be transferred to a methyl carrier and reduced to methane.

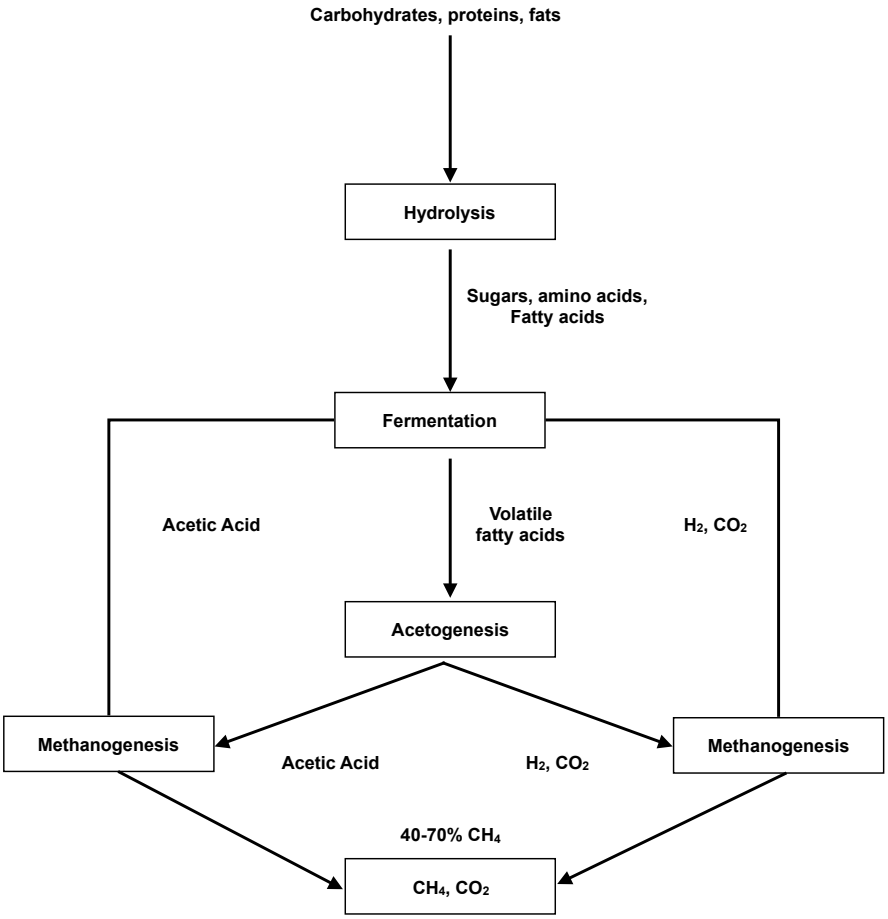


Figure 1.2 The transformation of organic matter to methane (Adopted from Cotana, Cavalaglio G, Petrozzi, Coccia (44))

The steps of methanogenesis mentioned above and illustrated in Figure 1.2 occur simultaneously. It is important to note that, methanogenic microorganisms have been found to be more sensitive to external conditions than acid-forming bacteria (45). There are, in fact, a number of parameters that influence methanogenesis; its speed, quality, and quantity of biogas produced. First of all, methanogenesis can only occur in an anaerobic environment. This means no oxygen can enter the anaerobic digester system, due to the main microbial groups taking part in the process being obligate anaerobes, as described above. Also, methanogens are mesophilic prokaryotes and, therefore, their

optimal temperature is 35-37°C and their growth stops at temperatures below 18°C. Acidity plays a crucial role in methanogenesis as well as these microorganisms are usually neutrophils and the environmental pH should be between 6.5 and 8.5. Finally, the ratio of carbon to nitrogen is extremely important with an optimal ratio of 16:19 Carbon to Nitrogen (46). Although all the chemical and physical parameters necessary for an efficient anaerobic digestion have been well characterized are known, the main challenge with this technology is the stability of the anaerobic digester (a.k.a., anaerobic reactor). This can be due to the lack of understanding of this complex process from the microbiological standpoint.

So far, scientists conventionally agree that a constant balance between main microbial consortia and their metabolic activity is required to maintain peak performance of the anaerobic digester (41). However, at the same time a detailed description of the main “players” of the AD process is missing. In addition, the level of dissimilarity between microbial communities among different types of anaerobic digesters is still not understood. Furthermore, in previously published papers, we find two completely opposite approaches. Some researchers say that ADs deal with completely different microbiomes depending on the design, organic infeed of the AD (physical and chemical), and temperature regulations (47, 48). On the other hand, there are publications that present evidence of a “core microbial community” – an essential microbiota that is similar in all of the ADs and is in charge of the main stages of methanogenesis (49–51), but base their conclusions comparing anaerobic digesters of similar designs, using similar infeed material and similar inoculum (52–56), or when performing analysis use incomprehensive

methods that make data from different sources lack consistency. This leads to the fact that preventing, handling, and resolving issues within the microbial community of digesters is still difficult and readily-applicable protocols from technology suppliers are not always available so far (50, 52, 57). And as a result, the potential inhibition of biogas production due to the fragile stability of the microbial community that moderates the process (58), can lead to the unexpected collapse of large-scale anaerobic digesters thus limiting the broad adoption of this technology (4, 3). Also, because of their unstable performance, anaerobic digesters can be complex and time consuming for operators to run without a fulltime assistance from the outside (assistance from companies that usually advertise and sell the technology as a self-running process) which can be costly. That is why the implications of clarifying a true ‘core’ microbial community inherent to bioreactors/anaerobic digesters holds promise to assist AD operators as well as technology producers in managing their reactors in a more systematic way

Microbial Composition: a Cross-sectional Meta-analysis

Recently, our research team conducted a meta-analysis of microbial communities reported in literature and characterized from a variety of anaerobic digesters (Melendrez-Vallard et al., in preparation). Our group considered 18 studies from 2011 to 2018 covering 136 anaerobic digesters from 9 countries. The designs of the digesters were very diverse as they fitted the needs for agriculture, wastewater treatment, food waste or household related processes. This implied also a broad variety of chemical and physical parameters and requirements. Their volumes spanned from 1 liter to 37 million liters, their pH range from 6.5 to 9, hydraulic retention times (HRTs) from 9.3 hours to 84 days,

operational temperatures from 19°C to 57°C. The sequencing data from these studies were characterized using two platforms Quantitative Insights into Microbial Environment (QIIME) or Metagenomic Rapid Annotations using Subsystems Technology (MG-RAST). Although the differences between the digesting systems (different starting inoculum, temperature controls, etc.) might bring into question the feasibility of a comparison, our group goal was to leverage those differences to highlight the commonalities between communities. The reasoning behind this was that in as much as anaerobic digestion is the core process common to all these systems, microorganisms present in all the systems would be the 'core' microbial community sustaining it. Therefore, the taxa overlap between serviced industries, sites, and reactor designs was identified, and it is represented in Table 1.1.

Table 1.1.

Overlap in microbial taxa between 136 digesters from 9 countries

Genus level taxa	Family level taxa	Class level taxa	Phylum level taxa
Methanoculleus, Methanosaeta, Bacillus, Clostridium, Ruminococcus, Methanospirillum, Methanosarcina, Bacteroides, Parabacteroides, Treponema, Anaerolinea, Fibrobacter, Methanobacterium, Flavobacterium, Syntrophomonas	Erysipelotrichaceae, Porphyromonadaceae, Verrucomicrobiaceae	Thermoplasmata	Chloroflexi, Aminicenantes, Candidatus Cloacamonas, Tenericutes, Proteobacteria, Euryarchaeota, Crenarchaeota, Firmicutes, Synergistetes and Spirochaetes

However, our team realized that in all these studies, batch fermentors were not represented. This was one of the major gaps in our understanding of the technology. In a

batch fermentor there is no introduction of fresh nutrients or new components to the microbial community. As a result, the inoculated microbial community changes as a direct consequence of the starting chemical and physical parameters without external interference. To address this and add more data to the analysis, our team set up two batch fermentors.

Benchtop Fermentor Study at SCSU: Set up

Two conical anaerobic fermentors were set up in Dr. Fink's laboratory, in the Integrated Science and Laboratory Facility (ISELF), on SCSU campus. Two fermentors were operational in the fall of 2018 and were sampled using a logarithmic method. Fermentor #1 and Fermentor #2 were biological replicates: same infeed, inoculated with the same inoculum, and the processes in both fermentors started at the same time.

The feedstock in these fermentors was food waste from SCSU campus cafeteria inoculated with cow manure from University of Minnesota Rosemount Research and Outreach Center in Rosemount, MN with a manure to food waste ratio of 1:5. First, however, food waste was sifted through and blended with 3.5L of sterile DI water to obtain a liquid consistency. Cow manure was also sifted to remove large particulates and blended with 2.5 L sterile DI water. The manure and food waste then stored at -80°C until the fermenters were ready for loading. Prior to loading into the batch fermenter system, the food and manure were removed from 80°C freezer and thawed at 4°C 48 hours.

After food waste mix and manure mixes were combined an additional 5.5 L of sterile DI water was added to make a mix consistency liquid enough to pour into the fermenters. Fermenters then were sealed. The anaerobic condition in both fermenters

was monitored by anaerobic test strips. Sampling started after sealing of fermenters and the following logarithmic schedule was followed: 0, 0.5, 1, 2, 4, 8, 16, 24, 41.5, 42, 42.5, 43.5, 45.5, 49.5, 57.5, 65.5, 89.5, 101.5, 128.5, 155.5 and 182.5 hours. The anaerobic conditions were detected at 41,5 hours and the logarithmic sampling model was restarted. Sampling stopped after no biogas production was detected any longer. After each sample was collected it was stored at -80°C until ready to send for microbiome characterization by 16S rRNA survey sequencing.

Benchtop Fermentor Study at SCSU: 16S Profiling

For sequencing of the 16S rRNA all samples were sent to the University of Minnesota Genomics Center (UMGC), St. Paul, MN USA. Three primer sets: two universal primers - Proc and V4 and Arc primer targeting Archaea were used for 16S rRNA sequencing (Melendrez-Vallard, et al., in preparation). A detailed description of sequencing and analysis for this study is presented in the Materials and Methods section in Chapter 2. After sequencing at UMGC raw demultiplexed data was analyzed using the QIIME2 software (v 2018.10).

As a result this study detected high abundant groups of microorganisms including *Clostridium, Fibrobacter, Bacteroides, Pseudomonas, Alistipes, Flavobacterium, Prevotella, Atopostipes, Bifidobacterium, Carnobacterium, Coxiella, Erysipelothrix, Myroides, Psychrobacter, Succinispira, 5-7N15, Akkermansia, Arthrobacter, Barnesiella, CF231, Dorea, Fluviicola, Isobaculum, Lactococcus, Lactonifactor, Leuconostoc, Lishizhenia, Pediococcus, Phascolarctobacterium, rc4-4, Sporobacter, Succinivibrio, vadinCA11, and Weisella.* It appears that many of these genera are highly specific to the food waste digestate used in the fermenters on SCSU campus. At the same time there were other genera identified in the batch fermenter experiment that were also present in other studies where digesters have a completely different design and infeed. For instance, *Bifidobacterium* was highly represented due to the food waste being very rich in fats and oils. At the same time *Clostridium* and *Fibrobacter* genera were seen in all the digesters represented in the comparative study. Also, at the family level Pseudomonadaceae, at the order level Desulfuromonadales was never mentioned in

fermentation studies before; however, were found in the SCSU batch fermentation study conducted by Melendrez-Vallard and colleagues. In previous studies, this family and order were associated with wastewater treatment and agriculture waste anaerobic digester systems only. It is also important to underline that members of the phylum Euryarchaeota (ie. *Methanobrevibacter, Methanobacterium, Methanoculleus, Methanosarcina, Methanosaeta*) that are directly in control of methane production which are consistently detected in all kinds of digester and fermenter designs and infeeds, were also confirmed to be present in the batch fermentation system on SCSU campus.

Analyzing the results presented above as well as review articles (56, 59, 60) we can conclude that further more experimentation and metagenomic sequencing of digester systems from a variety of use-cases using universal primers, as well as a primer set targeting Archaea alone should be performed to answer a question about existence of a core microbiome responsible for biogas production.

Chapter 2: Comparison of Microbial Representation in Dissimilar Anaerobic Digesters

Introduction

There is a great variety of designs for anaerobic digestors (ADs), however, they all generally perform the same set of biochemical reactions leading to biogas production. In our hypothesis, we proposed that there is a common (core) group of microorganisms that, albeit unrelated taxonomically, are interconnected in a metabolic network ultimately leading to methanogenesis. Characterization, and analysis of microbial composition within anaerobic digesters will help to make the process more stable by monitoring a core taxon in the community, the operator will be able to determine if a system is collapsing. Also, profiling the core microbial community across dissimilar anaerobic digesters to create a synthetic inoculum will enhance the productivity of an inoculated anaerobic digester, regardless of design of the reactor or of infeed consistency and source. Furthermore, if the synthetic microbiota composed of dominant microorganisms present in dissimilar anaerobic digesters it will be able to restart an anaerobic digester if it crashes completely.

To identify the microorganisms involved in the shared metabolic work, we needed to include diverse ADs to identify the few common members of the microbial community that are essential for the generation of biogas. For this reason, we identified sampling sites based on differences in AD designs, organic matter infeed, and scale of a reactors. All chosen sampling locations are in Minnesota and represent both industrial and laboratory settings. For the infeed of the digesters, we have food waste fermentor inoculated with cow manure, food waste inoculated with dry digester inoculum,

wastewater mixed with high strength food waste, and a digester running on cow manure only. As for the reactor sizes they range from one gallon to a hundred thousand gallons. Locations sampled included: two conical anaerobic fermentors operating under mesophilic conditions (description in Chapter 1); a plug flow system operating under thermophilic conditions with a 420,000 gallon capacity at Haubenschild Dairy Farms, Princeton, MN; two continuously-stirred tank reactors (CSTR) operating under thermophilic conditions at St. Cloud Wastewater Treatment Facility, St Cloud, MN; and a mini reactor which operates as a two-stage benchtop anaerobic digester under mesophilic conditions at Dr. Bo Hu's laboratory, Department of Bioproducts and Biosystems and Engineering, University of Minnesota.

To achieve the goals of describing microbial communities the following three objectives were targeted:

1. **Anaerobic digester infeed, digestate, and byproducts were sampled.** Three types of anaerobic digesters: industrial, and laboratory experimental digesters were selected. Samples were taken at all process stages to determine changes in abundance and structure of the microbiota.
2. **Metagenomic sequencing of collected anaerobic digester samples.** Metagenomic sequencing of the samples were performed at the University of Minnesota Genomics Center. This allowed us to examine the microbial communities and detect even low-abundance species. We targeted two variable regions of the full-length 16S rDNA – V3-V4 for bacteria and V4-V5 for archaea (61–63).

3. **Standardized common analysis of sequenced data.** After obtaining sequencing data the Quantitative Insights into Microbial Ecology (QIIME2)(64) software pipeline was used to analyze sequences and identify microorganisms involved in anaerobic digestion (62). Based on the microbial taxa detected and shared by each system, we will then determine the composition of a synthetic inoculum. Once the composition and relative abundance of the taxa in this inoculum are defined, tested and optimized, it should work with any reactor design and help overcome critical periods of failing biogas production due to disturbance of the activity of the microbiota.

Hypothesis

H_1: Although the microbial composition of anaerobic digesters of different designs and digesting organic infeed from different sources will show significantly different and unique characteristics, the taxa responsible for the production of methane, and their ancillary groups will form a core microbiota that can be identified.

H_0: The microbial composition of anaerobic digesters of different design and digesting organic infeed from different sources is different among digesters and impacted by different environmental factors such as infeed composition. As a result, there is no core microbiota specifically responsible for the production of methane.

On-farm Anaerobic Digestion Systems

Design Types of Animal On-farm Anaerobic Digesters

Livestock manure is one of the most common substrates used to feed anaerobic digesters. As described in Chapter 1, all anaerobic digesters perform the same basic function (store organic waste in the absence of oxygen, stable temperature, and other conditions for methane producing microbial consortia to grow and deliver high yields of biogas); however, there are a number of designs only for on-farm anaerobic digesters. There are three big classes of on-farm anaerobic digesters: passive systems where biogas yields are added to an existing treatment component; low rate systems where manure that streams through the system is the main contributor of methane-forming consortia; and high rate systems where the methane forming microbial community is trapped inside the digester to increase the efficiency even though the digestate is constantly changing.

Covered lagoon systems belong to passive systems. They comprise two cells where the first cell (lagoon) is air-tight and the second cell is open. The manure level of the first cell is constant to promote breakdown of organic matter and the second cell is a storage cell with a variable level of processed manure. The main advantage of these systems is the low maintenance. However, these systems are not heated and as a result they are more productive in warm climates. In climates where temperatures drop below 20°C a covered lagoon biogas production lowers to negligible levels. In addition, methane forming sludge stays in the first cell for many years (in some cases for 20 years) trapping fertilizer nutrients there for the same amount of time (39, 65).

As for low rate systems digesters there are two types of them: complete mix digesters and plug flow digesters. Complete mix digesters are heated (can be mesophilic or thermophilic), constantly or intermittently mixed tanks where incoming infeed displaces volume in the digester and the same amount of liquid overflows the tank. These digesters are more stable due to the constant inside temperature. Their hydraulic retention time (HRT) is usually from 20 to 30 days and recommended solids concentration is from 3 to 6 percent. Plug flow digesters have a similar design with the main difference in solids concentration, which is higher, from 10 to 15 percent, and in some cases up to 20 percent (38). This high rate of solids does not allow particles to settle at the bottom and manure simply travels through as a plug, no mechanical mixing is required. The length of these digesters is usually five times their section diameter. The HRT in this case is 15 - 20 days. This type of digesters work best for bigger farms with a high volume of manure production, such as dairy farms (66).

Finally, high rate systems can be divided in four groups: contact stabilization digesters, fixed film digester, suspended media digesters, and sequencing batch digesters. In contact stabilization digesters effluent leaving the digester moves to an outside clarifier where microbe-rich slurry settles down and is recycled back to the constantly mixed digester. This way a constantly high microbial concentration is achieved, which helps reduce the HRT time. Fixed film digesters are tall tanks (columns) filled with predefined of media that have a biofilm formation on it. This media keeps microbes from being washed off. Plastic bits, rings, or wood chips are mostly used as a filler. The extra high concentration of microbes in this type of digesters reduces the HRT to five days. The

only drawback is that this type of digesters is only efficient with infeed solid concentrations up to 5%, any higher and the solids would plug the system. With the loss of solids, the biogas yields also decreases. In suspended media digesters microbes are suspended in a constant upward flow of liquid. The flow is regulated this way so biogas-forming bacteria form biofilms on larger particles that stay inside, and smaller ones are washed out. The maximum solid concentration for type of reactors is 12%. Finally, anaerobic sequencing batch reactors are a variation of intermittently mixed reactors where biogas-forming microbial communities stay inside the reactor with settled solids. These reactors have four stages: digester feeding stage (filling a reactor with manure), react phase (mixing stage), settling stage (a time when solids settle down), decant stage (removal of decomposed liquid part). These types of digesters are very efficient (liquid retention time can be only five days) however they only work with manure solid concentrations of less than 1%(39).

According to the EPA, in 2015, 7 AD projects were located and operating in Minnesota on livestock farms. At the time of sampling for the current study, only one plug flow digester was operational in Princeton, MN. Overall, 42% of the anaerobic digester designs that are operating on livestock farms in the US are plug flow digesters (102 ADs) making it the most popular design for livestock farms including dairy farms.

Plug Flow Anaerobic Digester Sampling Site:

Haubenschild Farms is a 1000-acre family farm in Princeton, Minnesota. The farm operates a stable plug flow anaerobic digester (AD) that has been functioning since August 1999 with high biogas yields. The farm AD is a plug flow system originally designed to decompose manure from ~800 dairy cows. At the time of the sample collection there were ~1600 cows at the facility. The AD system's main component is a concrete covered tank installed in the ground which is heated with suspended water pipes. The total capacity of the AD tank is ~1.3 million liters. It is fed with manure infeed twice daily. Manure is scraped from the floors continuously by an automated scraper into troughs that carry the manure into the digester system (Figure 2.1). The digester functions constantly on the higher end of mesophilic conditions (~35-40°C) and the HRT is approximately 17 days. The percentage of solids in the infeed does not exceed 11% (standard solids percentage for plug flow digesters (32)), due to manure being mixed with water prior to entering the reactor.

Sampling

Haubenschild farm has four isolated barns that house animals of different age groups, fed different diets, and have different health conditions. Also, each barn has differences in bedding and temperature cycles. Barn #1 hosts mature animals that are milked three times daily. Barn #2 hosts heifers, replacement stock, and pregnant cows. These two barns use dried processed fibers - AD digestate (the digester outcome) as a bedding. Barn #3 hosts sick, older cows, and stock ready to be sent to a slaughterhouse. The animals in this barn receive more feed to help them recover and gain weight faster

and straw is used as a bedding. Barn #4 is completely isolated from the rest of the farm barn and hosts calves that are on a milk diet. Barn #3 and #4 do not contribute to the digester but were sampled as control points to compare sequencing results. The following sites were chosen to be sampled: each of the four barn floors (manure), the mixer (manure and H_2O mix), the digester exit point (digestate), the separator (solids and water mix), and the bedding storage (dried digestate) (Figure 2.1). Samples were collected into 50 ml falcon tubes, refrigerated on site and transferred to SCSU where they were stored at -80°C. These sampling sites covered the entire AD process from waste generation (barn) to the final exit of the AD (bedding storage site) which offered an opportunity to look at the change in diversity, richness, and evenness of the microbial community through the DNA sequencing and analysis techniques used.

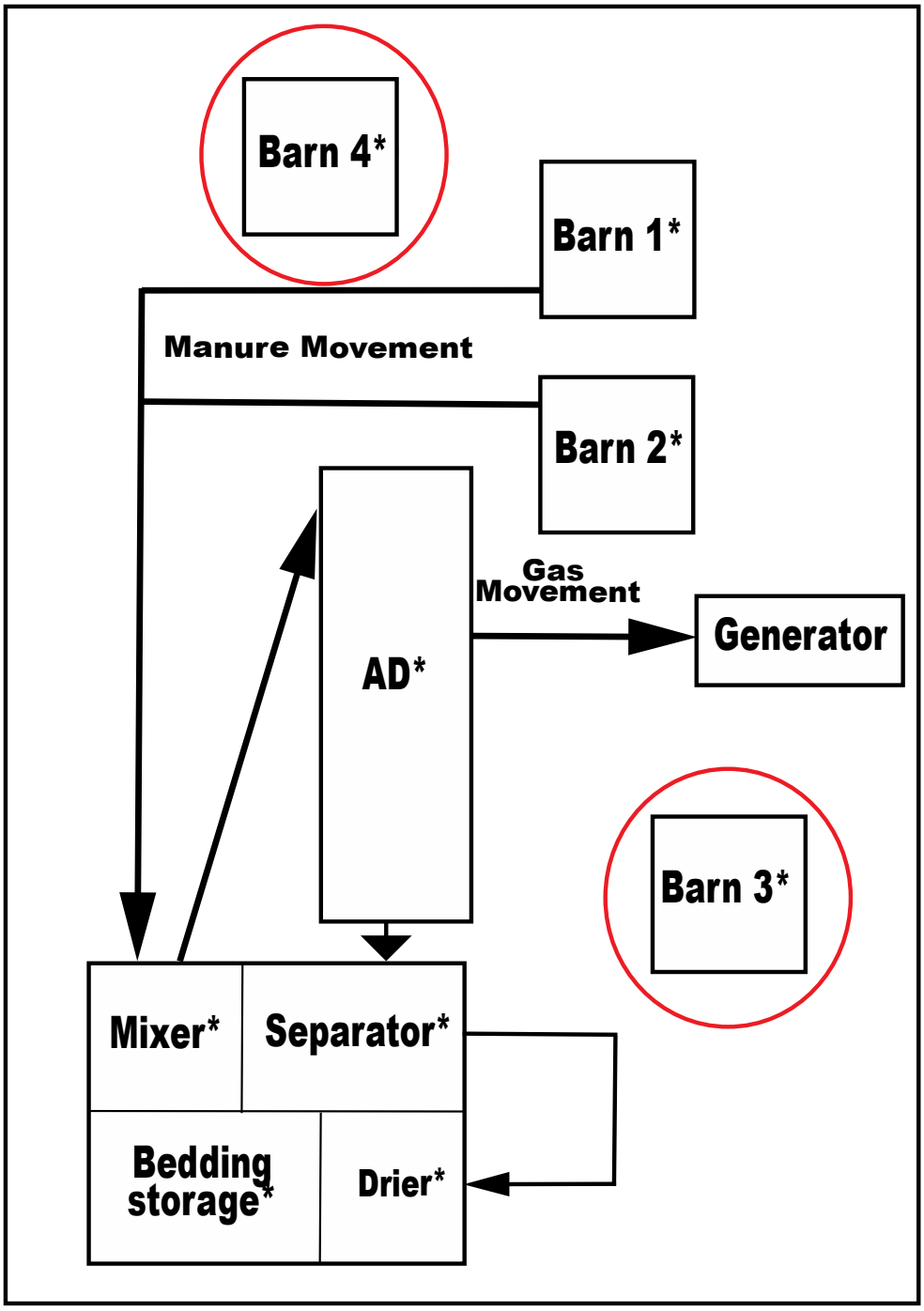


Figure 2.1. Haubenschild farm schematic, anaerobic digester (AD) set-up, and main sampling points (*). Arrows indicate manure movement from barns to the digester through the mixer and the return of the digestate through the separator, drier and ending up at the bedding storage section. Barns 3 and 4 (circled) are isolated and don't contribute to the anaerobic digester

Materials and Methods

DNA Extraction, Quantification, and 16S rRNA Sequencing

To prepare samples for DNA extraction, about 250 mg of each sample in duplicates were transferred to a 1.5 ml centrifuge tubes and filled up with deionized (DI) water following the University of Minnesota Genomics Center (UMGC) guidelines and stored at a -80°C freezer. Sequencing of the samples was performed at UMGC, St. Paul, MN. For DNA extraction, the PowerSoil DNA Isolation kit (MoBio Laboratories Inc., Carlsbad, California) was used according to the manufacture's protocol (67). DNA quality and quantity were determined using two different methods, spectrophotometric and fluorometric. For the spectrophotometric quantification of DNA, NanoDrop 2000 Spectrophotometer (ThermoFisher Scientific, Waltham, MA) was used (68). PicoGreen assay was used for DNA quantification by fluorescence tagging (69). Both quantification methods confirmed the availability of enough high-quality DNA from all the samples for Illumina library preparation and 16S rRNA sequencing.

Amplification of 16S rRNA was performed using 2 primer sets; one which targeted the domain Archaea, the primary microorganism responsible for methane production, and the other domain Bacteria, with some ability to detect Archaea. Previous studies have utilized primer sets that do not target Archaea specifically but rather attempt to pick up both Bacteria and Archaea. While these primer sets are capable of picking up both domains as a result Archaea are often in lower representation in sequence diversity surveys. As a result, we selected a primer set shown to target and characterize Archaeal diversity to ensure all potential diversity contributing to the process of methanogenesis

was represented in our analysis. The primer sets used included: a universal primer set - Pro341F_Nextera (16S targeting sequence CCTACGGGNBGCASCAG) and Pro805R_Nextera (16S targeting sequence GACTACNVGGGTATCTAATCC), and a primer set targeting Archaea - Arc515F_Nextera (16S targeting sequence TGYCAGCCGCCGCGGTAAHACCVGC) and Arc915R_Nextera (16S targeting sequence GTGCTCCCCCGCCAATTCCT) (70). The MiSeq Illumina benchtop system (Illumina Inc., San Diego, Ca) was used for sequencing. MiSeq 600 cycle v3 kit (Illumina) was used according to the manufacturer's protocol. PCR reactions were carried out using the KAPA HiFidelity Hot Start Polymerase (KAPA Biosystems, Wilmington, MA). PCR rounds start with a 5-minute at 95°C, then switching to 98°C 25 cycles for 20 seconds, 55°C for 15 seconds, 72°C for 1 minute, and finishes at 4°C hold. For the second round PCR products were diluted. 1:100. The second round of PCR starts with a 5-minute 95°C, then 10 cycles of 98°C for 20 seconds, 55°C for 15 seconds, 72°C for I minute, and ends with 4°C hold. Next, samples were denatured with NaOH, diluted to 8 pM in HT1 buffer (Illumina) and spiked with 15% PhiX. Last, when cluster generation wass completed sequencing took place. Sequencing reagents from Illumina kit were added (i.e. fluorescently labeled nucleotides) to the DNA and when the flow cell (a glass plate with Nano cells storing amplified DNA) is imaged each of the bases radiates a unique wavelength that makes the base identifiable (71). During the library preparation, samples were multiplexed using sequence barcoding. Following sequencing the raw sequencing data was demultiplexed by the UMGC and uploaded to a secure website. The v3 16S rDNA sequences were then available for downloading in a Cassava 1.8 FASTQ format,

which is a sequence file containing the sequences bases and quality scoring for each nucleotide (72).

Analysis and Visualization of Sequenced Data

First, a sample metadata table (Table 2.1) was prepared based on the information available about each sample: ID, collection date, physical state, sampling collection point, and primer used for sequencing. The Metadata table (Table 2.1) was then imported into a software platform used for analysis of a sequence set, QIIME2 (v2019.1). The metadata table uses the following sample-id legend: sample number at the sampling site in a duplicate: A, B; sampling point at the sampling site: B1-4 – Barn 1, Barn 2, Barn 3, Barn-4; D 1-4 – Digester 1, Digester 2, Digester 3, Digester 4, Mixer, HSIW – high strength waste, Lystek, Digestate and Sep – digester and separator at Haubenschild farm, Table – solid waste sampling point at WWTF; sampling site: HF – Haubenschild Farm, WW – St. Cloud Wastewater Treatment Facility, and FW - Dr. Bo Hu's laboratory, Department of Bioproducts and Biosystems and Engineering, University of Minnesota; Arc – primer targeting Archaea and Prok – universal primer used for sequencing. For data from fermentors in Dr. Ryan C. Fink's laboratory, at IESELF, St Cloud State University the following legend was used: F1-2 - Fermentor 1, Fermentor 2; hours count followed by day count when fermentor was sampled, and Arc, Prok – primers used for sequencing.

Table 2.1.

Metadata table including operational information for four sampled and analyzed digestion sites used for QIIME2 analysis

#	Sample ID	Site	Phase	Subject	Sample-type	Sampling-Date
1	1A-B1-HF	Farm	Aerobic	Barn-1	Manure	18-Oct-2018
2	1B-B1-HF	Farm	Aerobic	Barn-1	Manure	18-Oct-2018
3	2A-B2-HF	Farm	Aerobic	Barn-2	Manure	18-Oct-2018
4	2B-B2-HF	Farm	Aerobic	Barn-2	Manure	18-Oct-2018
5	3A-B3-HF	Farm	Aerobic	Barn-3	Manure	18-Oct-2018
6	3B-B3-HF	Farm	Aerobic	Barn-3	Manure	18-Oct-2018
7	4A-B4-HF	Farm	Aerobic	Barn-4	Manure	18-Oct-2018
8	4B-B4-HF	Farm	Aerobic	Barn-4	Manure	18-Oct-2018
9	5A-Mixer-HF	Farm	Aerobic	Mixer	Manure	18-Oct-2018
10	5B-Mixer-HF	Farm	Aerobic	Mixer	Manure	18-Oct-2018
11	6A-Sep-HF	Farm	Aerobic	Separator	Manure	18-Oct-2018
12	6B-Sep-HF	Farm	Aerobic	Separator	Manure	18-Oct-2018
13	8A-Digestate-HF	Farm	Anaerobic	AD	Manure	18-Oct-2018
14	8B-Digestate-HF	Farm	Anaerobic	AD	Manure	18-Oct-2018
15	9A-Bed-HF	Farm	Aerobic	Bedding	Manure	18-Oct-2018
16	9B-Bed-HF	Farm	Aerobic	Bedding	Manure	18-Oct-2018
17	1A-D1-WW	WW treatment	Anaerobic	Digester-1	Waste water/Brew waste	20-Feb-2019
18	1B-D1-WW	WW treatment	Anaerobic	Digester-1	Waste water/Brew waste	20-Feb-2019
19	2A-D2-WW	WW treatment	Anaerobic	Digester-2	Waste water/Brew waste	20-Feb-2019
20	2B-D2-WW	WW treatment	Anaerobic	Digester-2	Waste water/Brew waste	20-Feb-2019
21	3A-D3-WW	WW treatment	Anaerobic	Digester-3	Waste water/Brew waste	20-Feb-2019
22	3B-D3-WW	WW treatment	Anaerobic	Digester-3	Waste water/Brew waste	20-Feb-2019
23	4A-D4-WW	WW treatment	Anaerobic	Digester-4	Waste water/Brew waste	20-Feb-2019
24	4B-D4-WW	WW treatment	Anaerobic	Digester-4	Waste water/Brew waste	20-Feb-2019
25	5A-HSIW-WW	WW treatment	Aerobic	High strength waste	Brew waste	20-Feb-2019
26	5B-HSIW-WW	WW treatment	Aerobic	High strength waste	Brew waste	20-Feb-2019
27	6A-Table-WW	WW treatment	Aerobic	Gravity belt table	Waste water	20-Feb-2019
28	6B-Table-WW	WW treatment	Aerobic	Gravity belt table	Waste water	20-Feb-2019
29	7A-Lystek-WW	WW treatment	Anaerobic	Lystek	Waste water-Brew waste	20-Feb-2019
30	7B-Lystek-WW	WW treatment	Anaerobic	Lystek	Waste water-Brew waste	20-Feb-2019
31	8A-Mixer-WW	WW treatment	Aerobic	Mixer	Waste water-Brew waste	20-Feb-2019
32	8B-Mixer-WW	WW treatment	Aerobic	Mixer	Waste water-Brew waste	20-Feb-2019
33	1A-1R-FW	UoM AD	Aerobic	Digester-1	Food waste	3-Mar-2019
34	1B-1R-FW	UoM AD	Aerobic	Digester-1	Food waste	3-Mar-2019
35	2A-FW-FW	UoM AD	Aerobic	Infeed mix	Food waste	3-Mar-2019
36	2B-FW-FW	UoM AD	Aerobic	Infeed mix	Food waste	3-Mar-2019
37	3A-HR-FW	UoM AD	Anaerobic	Digester-2	Food waste	3-Mar-2019
38	3B-HR-FW	UoM AD	Anaerobic	Digester-2	Food waste	3-Mar-2019
39	F1-89-5-FF	SCSU ISELF	Anaerobic	Digester-1	Food waste/Manure	25-Feb-2018
40	F1-89-5-FF	SCSU ISELF	Anaerobic	Digester-1	Food waste/Manure	25-Feb-2018
41	F2-101-5-FF	SCSU ISELF	Anaerobic	Digester-2	Food waste/Manure	26-Feb-2018
42	F2-101-5-FF	SCSU ISELF	Anaerobic	Digester-2	Food waste/Manure	26-Feb-2018
43	FD1-FF	SCSU ISELF	Aerobic	Food waste	Food waste	21-Feb-2018
44	FD1-FF	SCSU ISELF	Aerobic	Food waste	Food waste	21-Feb-2018
45	M1-FF	SCSU ISELF	Aerobic	Infeed mix	Food waste/Manure	21-Feb-2018
46	M1-FF	SCSU ISELF	Aerobic	Infeed mix	Food waste/Manure	21-Feb-2018

To analyze the demultiplexed sequences, we used a protocol available in the open source plugin-based system package QIIME2 (v2019.1) (64). Overall there were 4,326132 sequences generated and the average number of sequences per sample was ~47,023 prior to the quality control step. To decide on the parameters used for the quality control step a summary of the demultiplexing results was generated (*demux.qzv* file) (Figure 2.4). Figure 2.4 is a visualization of Illumina sequencing by synthesis technology (SBS) where each read is assigned a quality score by a phred-like algorithm for a user to estimate the length of a high-quality sequence reads. Sequences with a quality score over 20 (“y” axes) have an accuracy rate of 99% and are reliable for further analysis. Next, low quality reads at the beginning of the forward and reverse reads were trimmed up to base 10. Also, forward and reverse reads were truncated at 280 and 240 bases respectively during the denoising step due to the appearance of lower quality reads after indicated basepair counts. The R-based Divisive Amplicon Denoising Algorithm (DADA2) was employed to denoise the samples which consists in correcting or removing noisy sequence reads (chimeras, PCR single-base errors, and errors in sequence readings) (73). DADA2 implements a novel algorithm “that models the errors introduced during amplicon sequencing, and uses that error model to infer the true sample composition” (73). The DADA2 protocol filters and trims the sequences (removes short reads and truncates longer reads to make all reads have a constant length), de-replicates them (collapses reads that encode the same sequence), determines error rates (parametric model), infers sample composition, merges forward and reverse reads, removes potential chimeras (sequences formed during PCR amplification that do not derive from the same

initial sequence), and builds a uniform sequence table. This sequence table is then linked to a sample metadata table to visualize the dataset (prescribes each sequence characteristics of a sample – location, time of sampling, primer used, etc.) and further analysis. Next, the taxonomic classification of sequences was performed (defining species present in a dataset) by comparing the dataset to a reference database – Greengenes (v13.8) (74, 75).

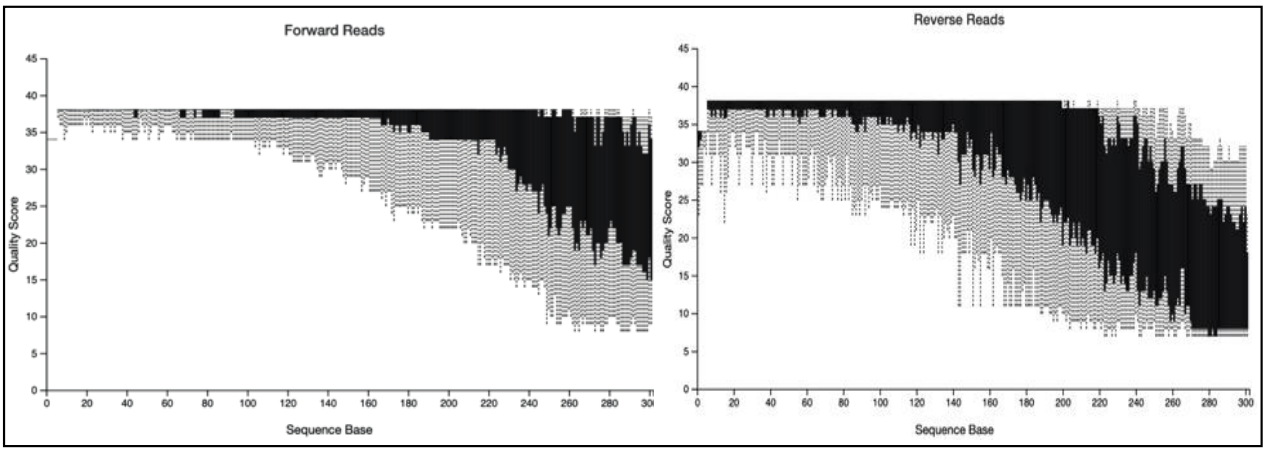


Figure 2.2. A summary of demultiplexed sequences (10000 randomly selected by QIIME v2019.1) forward and reverse reads trimmed up to basepair 10 and truncated from basepair 280 and 240 accordingly

Phylogenetic tree generation. To produce a phylogenetic tree a pipeline (align-to-tree-mafft-fasttree) that is a part of the q2-phylogeny plugin is used. First, a multiple sequence alignment (MSA) was produced using *MAFFT* (Multiple Alignment using Fast Fourier Transform) (76, 77) program from a q2-phylogeny plugin. *MAAFT* uses FFT-NS-2 method (Fourier transform, progressive method). This method has several steps: 1. Comparison of all sequences; 2. Building a distance matrix; 3. Building a tree-1 ; 4. Group alignment takes place and a new distance matrix is composed; 5. A tree-2 replaces tree-1 if more accurate (77). FastTree (78) program initially builds an unrooted tree, however, in

the final step it constructs a rooted tree based on mid-points of the two tips farthest apart. In fact, FastTree performs neighbor joining of aligned sequences, then reduces the size of the tree using nearest-neighbor interchanges (NNIs) and subtree-prune-regraft moves (SPRs). Furthermore, FastTree improves the tree by using maximum-likelihood rearrangements (Juke-Cantor model (79)). Finally, FastTree used Bayesian approach (80) with gamma prior to assign each site a category to prevent overfitting of small alignments.

As a result using QIIME2 q2-diversity plugin the following interactive visualizations were generated: genus level identification of the archaeal and bacterial sequences, alpha diversity indeces - Shannon's Diversity Index (81), observed OUTs index, Faith Phylogenetic Diversity index (82) Pielou's Evenness (83, 84); beta diversity - Jaccard distance (85), Bray-Curtis (86), and Unweighted Unifrac (87).

Results: Microbial Composition and Diversity Analysis

Microbial composition. For Haubenschild farm, overall 771,999 nonchimeric, good-quality sequences were recovered, representing 463 operational taxonomic units (OTUs) (24 archaeal OTUs, 439 bacterial OTUs). However, only 439,232 (~56%) sequences were identified at the genus level (279,195 by the Prok primer and 160,037 by the Arc primer). Also, 1,849 sequences were left unassigned (1,810 unassigned sequences recovered by the Arc primer and 39 sequences by the Prok primer).

Out of 160,037 sequences detected by the Arc primer, 106,445 archaeal sequences were identified at the genus level, representing 66,5% of the total recovered Arc sequences. Conversely, only 24% of the sequences detected by both primers were

characterized at the genus level. These archaeal sequences (n = 106,445) represented 12 OTUs. Also, the Arc primer recovered 51,782 bacterial sequences, they represented ~44% of overall Arc recovered sequences and ~12% of overall sequences recovered by both primers characterized to a genus level. The bacterial sequences (n = 51,782) recovered by Arc primer represented 132 OTUs.

As for the Prok primer, out of 279,195 sequences it recovered, only 33,663 were archaeal sequences. These sequences (n = 33,663) represent ~12% of overall Prok recovered sequences characterized to a genus level and ~7,7% of overall sequences recovered by both primers characterized to a genus level. These sequences (n = 33,663) represented 10 OTUs. Furthermore, the Prok primer recovered overall 245,493 bacterial sequences, they represented ~88% of overall Prok recovered sequences and ~32% of overall sequences recovered by both primers characterized to a genus level. The Prok recovered bacterial sequences represented 235 OTUs.

Generally, both primers together detected 13 high abundance OTUs (>1%)(60) at the genus level: *Methanobrevibacter* (~16%), *Bifidobacterium* (~3%), *Corynebacterium* (~3%), *Atopococcus* (~2.5%), *Lactobacillus* (~2%), *Atopostipes* (~2%), *Collinsella* (~1%), *Methanosphaera* (~1%), *Psychrobacter* (~1%), *Clostridium associated with Clostridiaceae family* (~1%), *Clostridium* associated with Ruminococcaceae family(~1%), *Clostridium* associated with Peptostreptococcaceae family (1%) (88). Only two of 13 high abundance genera represent Archaea (*Methanobrevibacter* and *Methanosphaera)* with *Methanobrevibacter* genus representatives being a dominating genus at Haubenschild farm sampling site. Furthermore, when looking at samples coming from the Haubenschild

farm site but were under the influence of different site conditions (anaerobic and aerobic) there is a noticeable change in richness, evenness and microbial abundance detected. There are overall 412.930 sequences associated with aerobic conditions that were identified to a genus level representing 257 OTUs (12 archaeal OTUs, 245 bacterial OTUs). Overall 21 of OTUs is considered high abundance (>1%) with two OTUs representing Archaea and 19 OTUs representing Bacteria. Archaeal genus *Methanobrevibacter* was determined to be a dominant genus at ~28% (Figure 2.2).

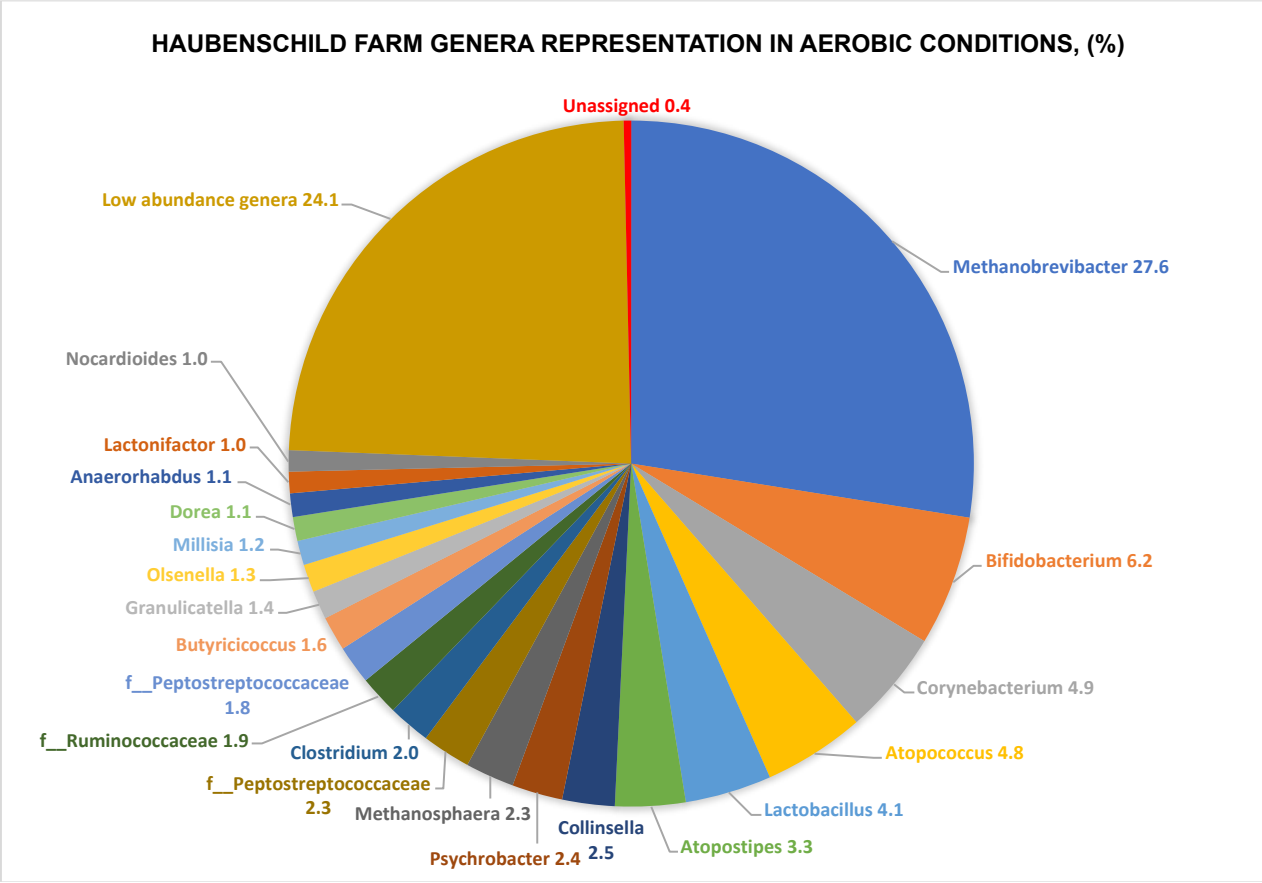


Figure 2.3. Haubenschild farm high abundance genera (>1%), low abundance genera grouped in one sector (<1%), and unassigned sequences in aerobic conditions.

Next, when looking at samples derived from the anaerobic sampling point (digester), the total number of OTUs characterized to a genus level was 109 OTUs which was 148 OTUs less than in aerobic conditions, confirming lower species richness in the anaerobic region of the AD. At the same time, the number of high abundant OTUs (>1%) stayed the same at 21 OTUs. Archaeal communities showed low diversity with only five OTUs and *Methanobrevibacter* dominating with ~36% over the whole microbial community of the AD (Figure 2.3).

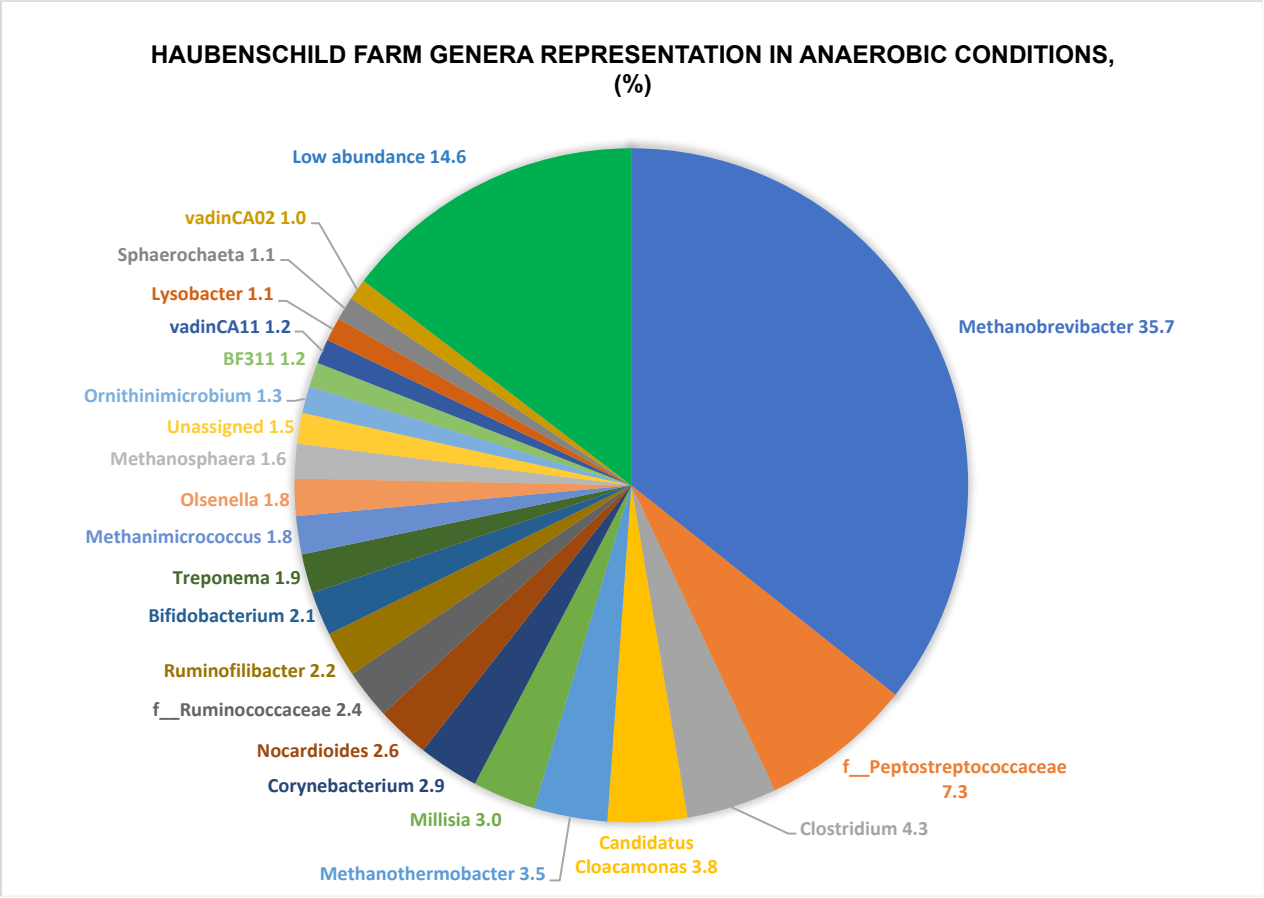


Figure 2.4. Haubenschild farm high abundance genera (>1%), low abundance genera grouped in one sector (<1%), and unassigned sequences in anaerobic conditions.

Alpha diversity. For the Haubenschild farm site, analysis alpha diversity (Figure 2.4) is a crucial instrument to understand how the community structure changes along with surrounding conditions and how these changes may contribute to the high, stable biogas generation. First, Barn 4 at the Haubenschild farm sampling site shows a very low Observed OTU count with an equal abundance for each defined OTU (Figure 2.4.a). At the same time Faith's PD (Figure 2.4.c) shows that this community has a closely related population with an evenness value of 0.68 (Figure 2.4.b) suggesting that it is spread out evenly with most OTUs having low abundance according to the Evenness index. The closely related, low OTU counts found within Barn 4 can be attributed to the isolation of calves from the grown-up population that are also on a very specific milk-based diet. Additionally, a clean uncontaminated saw dust is used as a bedding instead of a dried digestate from the AD. Both isolation, diet, and clean bedding can contribute to a unique microbiome make-up as compared to the adult cows in Barns 1, 2 and 3. This sampling point is considered a control point for this study to prove that sequencing and analysis was done correctly, because a low microbial richness with closely related species was expected at this location. Conversely, the main contributors to the AD Barns 1 and 2, as well as Barn 3 and the Mixer had a much higher Observed OTU count (Figure 2.4.a) with an equal distribution of abundance. As for the Faith's PD (Figure 2.4.c) it shows much more diverse population as compared to Barn 4, but it is also evenly distributed according to the Evenness index; which ranges from 0.78 to 0.84 for all three barns and the Mixer (Figure 2.4.b). The higher outlier measurement for these three barns can be attributed to the presence of high abundance Archaeal populations typically found in the gut of cows

such as *Methanobrevibacter*. Finally, when looking at the AD and post-AD (Separator and Bedding) sampling points they show lower Observed OTUs (Figure 2.4.a) for all three sampled points (AD, Separator, and Bedding) with the median being around 300 OTUs for the AD and the Bedding and less than 200 for the Separator with an even distribution of abundance that is confirmed by the Evenness index for all three sites (Figure 2.4.b); which ranges from ~0.76 to ~0.80. At the same time Faith's PD index (Figure 2.4.c) also shows close relatedness between the evenly distributed OTUs with outliers that can be pointing towards the less diverse Archaea representatives that are distant from bacterial species.

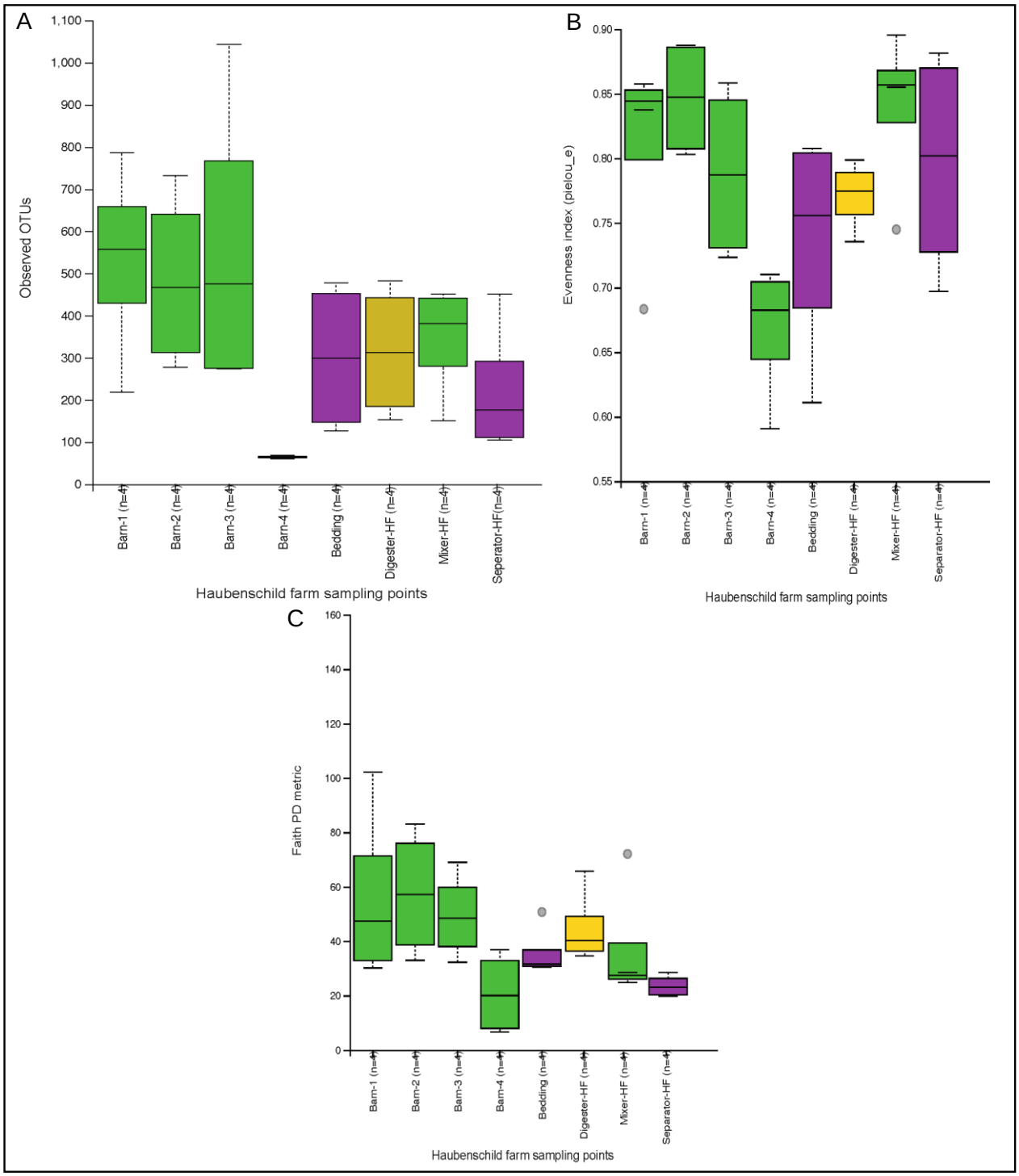


Figure 2.5. Alpha diversity analysis of (A) observed OTUs, (B) Faith's phylogenetic diversity (Faith's PD) and (C) Evenness of Haubenschild farm sampling sites: pre- (green), during (yellow) and post- (purple) digestion.

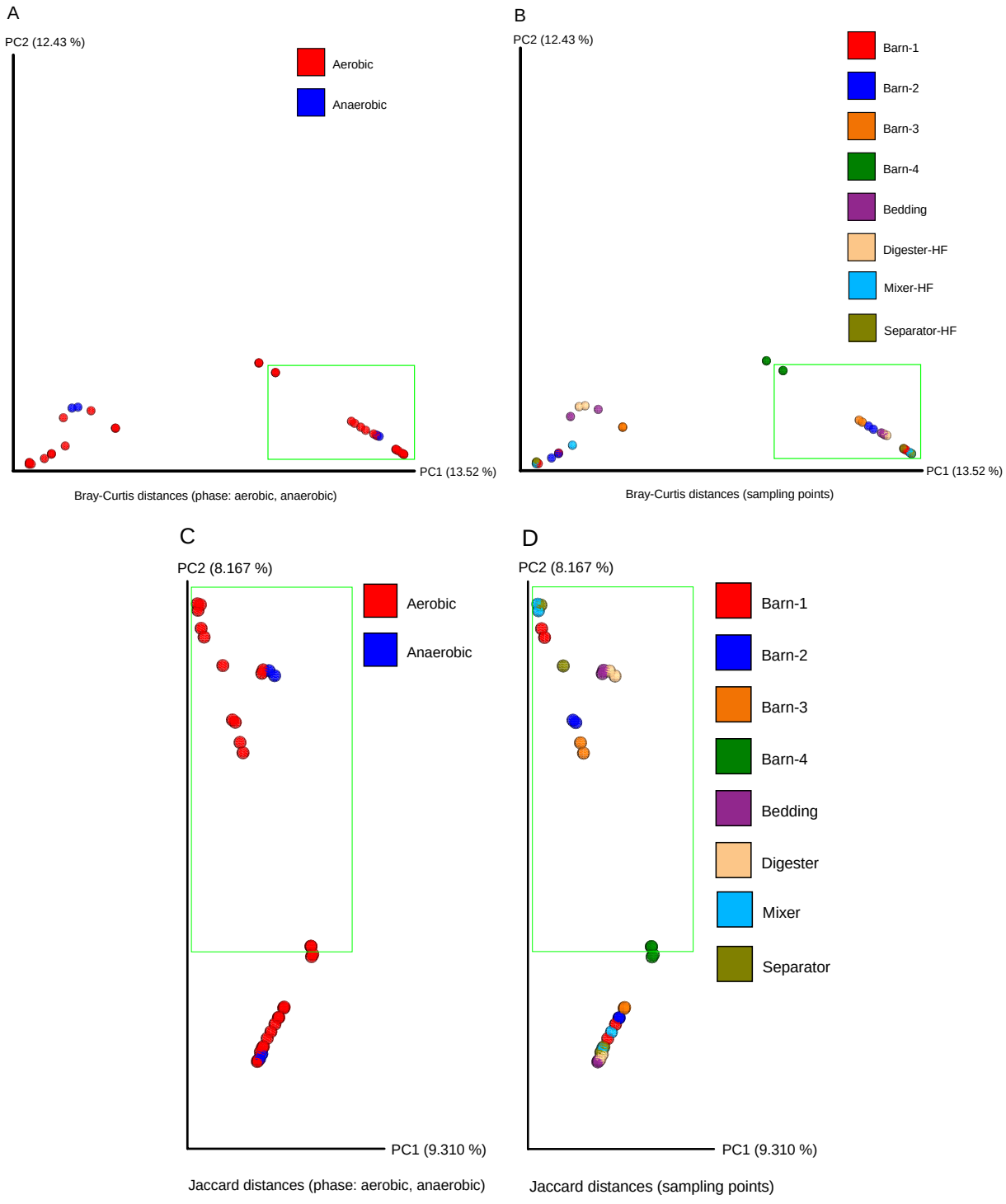
Beta diversity. Beta diversity was used to test for community differences between different phases (aerobic, anaerobic); between communities detected by the Arc and Prok primer sets; and between different sampling points at the Haubenschild farm AD system. This analysis was performed by examining taxonomic distinctiveness between microorganisms in a given taxon and as a result determine an existence of a core community shared between all sampling points on the farm. Figures for beta diversity were visualized on QIIME2 (v 2019.1) view (Figure 2.5) (64).

All the results of the Principal Coordinates Analysis (PCoA) for the Haubenschild farm sampling points were produced to demonstrate grouping of microbial communities based on similarities of communities via Bray-Curtis dissimilarity distances (Figure 2.5 A, B), Jaccard dissimilarity distances (Figure 2.5 C, D), and unweighted UniFrac phylogenetic distance metrics (Figure 2.5 E, F). The PCoA plots were visualized on EMPeror (89).First, PCoA plots (Bray-Curtis dissimilarity distances, Jaccard dissimilarity distances, unweighted UniFrac phylogenetic distance) are concordant. Also, PCoA align with Alpha diversity analysis and microbial composition analysis. All three analysis determined a significant taxonomic distinction between the microbiomes in manure of Barn 4 that hosts calves and the rest of the AD system (p-value=0.001) by both primer sets. There was no significant taxonomic differences detected according to Bray-Curtis (Figure 2.5 A, B) and Jaccard (Figure 2.5 C, D) PCoA plots between the rest of sampled points detected by the Prok and the Arc primer sets including sampling points representing aerobic and anaerobic phases.

The Unweighted Unifrac PCoA (Figure 2.5 E, F) plot for the Prok primer set showed a distinctively group including Barn 1-3, mixer and separator communities, as well as Bedding and Digester communities with the Barn-4 being clearly separated from the rest of sampling points; this observation is concordant with the Bray-Curtis (Figure 2.5 A, B) and Jaccard (Figure 2.5 C,D) PCoA plots. However, the Unweighted Unifrac PCoA plot for the Arc primer set placed Barn 4 close to the Digester and Bedding communities. This might be due to the very specialized nature of the primer set that targets Archaea representatives and was not able to detect multiple bacterial species of the samples and focused only on archaeal species placing this sampling point closer to points dominated by Archaea.

Overall, the trend for Haubenschild farm sampling site is that the microbial composition significantly changed when comparing aerobic and anerobic phases of the system. The first three most abundant genera in the aerobic phase were: *Methanobrevibacter* (~27.6%), *Bifidobacterium* (~6.2%) – both anaerobic microorganisms, and *Corynebacterium* (~4.9%) – aerobic bacteria. At the same time, three most abundant genera in the anaerobic phase were: *Methanobrevibacter* with the increased abundance to ~35.7% (~8% increase comparing to the aerobic phase), *Clostridium* associated with Clostridiaceae family (~4.3%), and *Clostridium* associated with Peptostreptococcaceae family (~7.3%) – all three representing obligate anaerobes. Although, manure is the only infeed component of the Haubenschild farm AD system that is itself coming from the anaerobic environment (cattle gut) the microbial composition

changed and became completely dominated by anaerobic microorganisms - proven participants of the AD process.



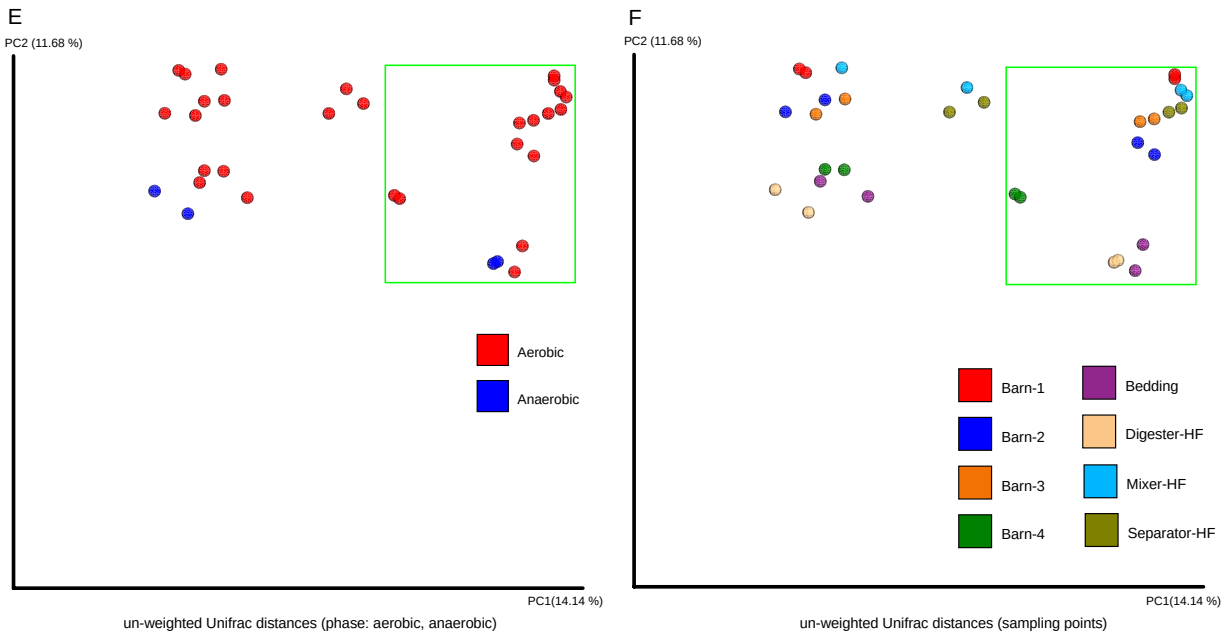


Figure 2.6. Beta diversity analysis of (A) Bray-Curtis, (C) Jaccard, and (E) un-weighted Unifrac distances during anaerobic and aerobic phases in the AD system; (B) Bray-Curtis, (D) Jaccard, (F) un-weighted Unifrac distances between different sampling points at the Haubenschild farm sampling sites with the Prok primer points in a green frame and the Arc primer results outside the frame.

Municipal Wastewater Treatment and Anerobic Co-digestion

The most conventional way of treating wastewater in the USA as well as in EU is aerobic treatment by applying suspended growth process also known as activated sludge (90) . These types of systems typically consist of four main elements: a constantly aerated bioreactor, a settling tank, a solids recycling route, and a solids removal route. Inside the first element of the process – bioreactor, aerobic bacteria process organic waste into three main components: carbon dioxide, water, and biomass. After that, wastewater moves to the second element where microbial flocs or granules are removed. The biomass that is left either goes back to the aerated bioreactor or gets removed for further processing (91). This process is efficient and reliable. The process removes a high percentage of nitrogen and phosphorus and yields water that meets regulatory standards. However, this process is costly (requires a lot of energy to aerate the bioreactor and circle sludge back to the tank). In addition, the cost of an excess sludge treatment that is left after the process can be accounting for up to 60% of the overall wastewater treatment. One of the ways of dealing with excess sludge is adding an anaerobic step to the activated sludge process that drastically reduces the amount of sludge and makes it usable as a fertilizer (92). This step takes place in an anaerobic continuous stirred tank reactor (CSTR) under mesophilic or thermophilic conditions. CSTRs are highly customizable depending on the requirements and size of the operation. Also, to make the process more energy productive (increase volume of the biogas produced) it is often designed as a co-digestion of municipal wastewater and high strength waste streams: agricultural waste, slaughterhouse waste, fruit and vegetable waste, manure and fats,

oils, and grease (FOG) (93, 94). This approach not only increases productivity but also resolves issues that arise with mono-digestion. These problems may include imbalanced components in the digestate or, accumulation of toxic components like ammonia. Within the last two decades research of co-digestion has significantly increased in order to find optimal proportions to the co-digestion recipe, as well as the most appropriate ingredients to co-digest (95).

Municipal Wastewater Co-digestion Sampling Site

St Cloud Wastewater Treatment Facility (WWTF) (Figure 2.5) has been operating since 1976 servicing the communities of St. Cloud, St. Augusta, St. Joseph, Sartell, Sauk Rapids, and Waite Park in Minnesota. The facility is capable of processing ~50 million liters of sewage per day from a population of over 120 thousand people. This facility operates reactors that substantially reduce the amount of electricity that it purchases from traditional sources for internal use. WWTF produce stable and high yields of biogas, that is capable of the generating 65% of electricity used by the facility according to 2018 data (retrieved from the facility webpage). In 2013 WWTF purchased 6.6 million kWh for its operational needs and in 2018 ~4.3 million kWh were generated internally. According to the WWTF documentation (https://www.ci.stcloud.mn.us/336/Wastewater-Treatment-Process) they use a common aerobic wastewater treatment system as a first step in the sewage processing treatment. There are four stages that wastewater goes through before biosolids are processed in ADs at the St Cloud WWTF. These stages are preliminary treatment, primary treatment, secondary treatment, and final clarification. The main goal of these steps is separation of solids to constantly feed two primary reactors (Digester 1

and 2). The primary reactors operate on a mix of sewage and high strength food waste with an HRT time of 15 days. The high strength food waste is delivered from local breweries and other food processing plants. The ratio of food waste to sewage is 1:6. ADs 1 and 2 work under anaerobic thermophilic conditions and are constantly stirred. At the same time Digester 3 and 4 are secondary reactors that work on a digestate from primary reactors without being heated or stirred. The final stage of the whole treatment process is liquefaction of the outcome of ADs by the biosolids thermo-alkaline hydrolysis treatment Lystek. This process is added to produce ready and easy to use liquid fertilizer to satisfy the needs of local farmers (96).

Sampling

For the St Cloud WWTF we followed the same protocol for sampling as at Haubenschild farm: we sampled of all the sites that feed ADs, digestate from digesters themselves, and finally the outcome of the AD process. This approach gave us a full picture of how microbial community richness, evenness, and overall composition changes on all stages of the process. At the WWTF the following sites were sampled (Figure 2.6): each of the four digesters (digestate), food waste from breweries, municipal thickened sludge (at the gravity belt table), mix of food waste and sludge, and the thermal hydrolysis reactor Lystek (digester outcome).

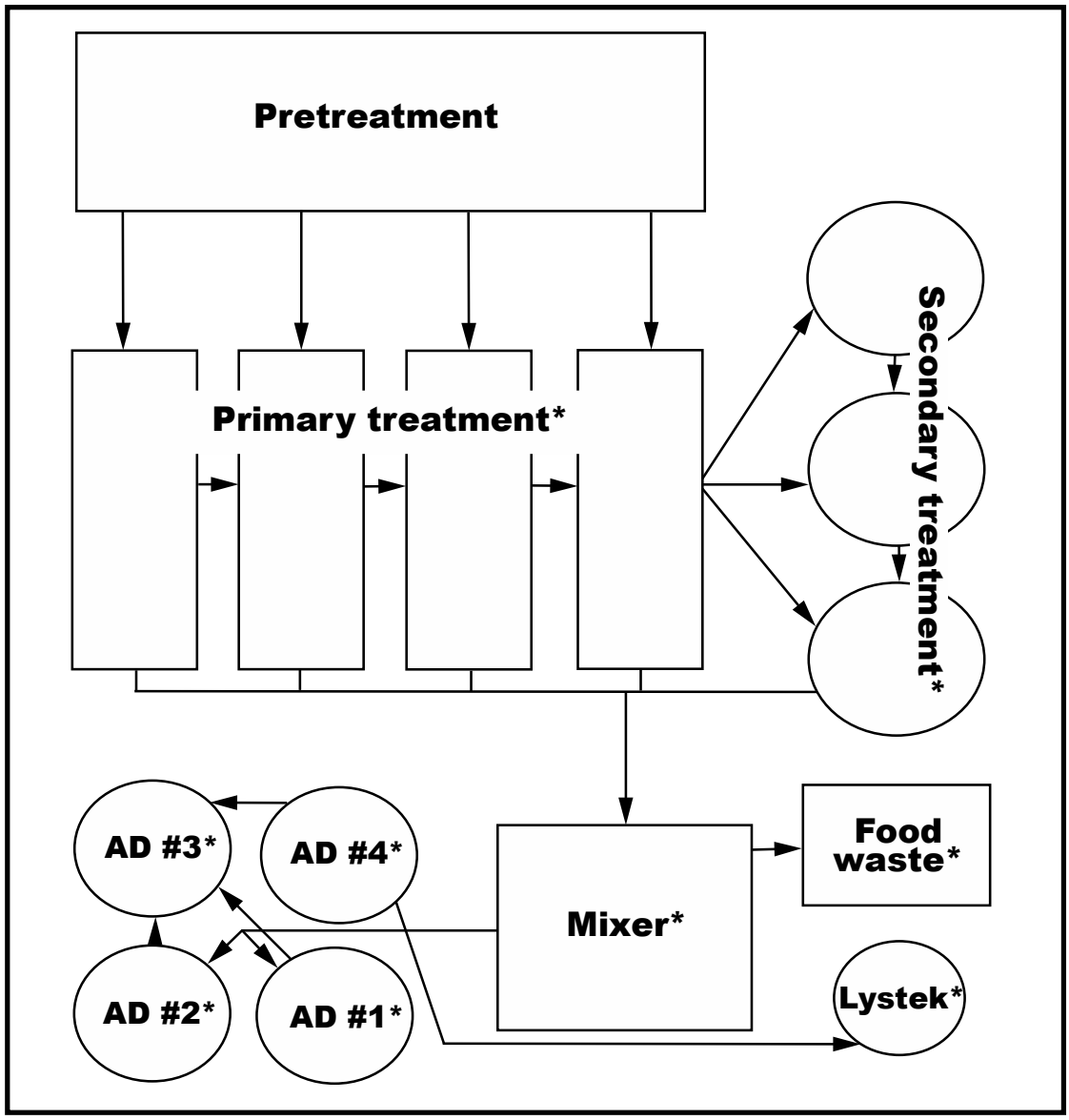


Figure 2.7. ST Cloud WWTF schematic, including AD set-up and main sampling points (indicated by the asterisks *).

Materials and Methods

Materials and methods used for St Cloud WWTF are identical to Haubenschild farm site (Chapter 2, Section: Material and methods: DNA extraction, quantification, and 16S rRNA sequencing).

Results: Microbial Composition and Diversity Analysis

Microbial composition. For St. Cloud WWTF overall 477,900 nonchimeric, good-quality sequences were recovered. However, only 184,418 of all recovered good-quality sequences (~39%) sequences were characterized to a genus level. 103,865 sequences were recovered by a Prok primer, 80,553 sequences were recovered by an Arc primer, and 4,410 sequences were left unassigned: 3,854 unassigned sequences recovered by an Arc primer and 556 sequences by a Prok primer. A total of 424 OTUs were assigned for WWTF (33 archaeal OTUs, 401 bacterial OTUs).

Overall, all Arc primer recovered sequences that were characterized to the genus level (n=80,553sequences) represent ~42% of overall Arc recovered sequences and ~17% of overall sequences recovered by both primers. 12 archaeal OTUs recovered by Arc primer (n = 52,116 sequences) and 115 bacterial OTUs (n = 28,437 sequences) were characterized to the genus level. The number of archaeal OTUs recovered by the Arc primer is higher compared to the Prok primer yields, which is expected.

As for the Prok primer, it was able to recover 103,865 sequences that were characterized to the genus level; ~37% of overall Prok recovered sequences; ~22% of overall sequences recovered by both primers. Overall, 11,415 sequences recovered by the Prok primer represent 5 archaeal OTUs and 92,450 sequences represent 181

bacterial OTUs. The universal Prok primer recovered 66 more bacterial OTUs compared to the Arc primer.

Together both primers detected 23 high abundance OTUs (>1%)(60) on a genus level: *Methanosaeta* (~26%), *Lactobacillus* (~6%), *Granulicatella* (~5%), *Candidatus Cloacamonas* (~4%), *Methanobrevibacter* (~4%), *Candidatus Microthrix* (~3%), *BHB21* (~3%), *Methanomassiliicoccus* (~3%), *Petrimonas* (~2%), *Syntrophus* (~2%), *W22* (~2%), *Streptobacillus* (~2%), *Levilinea* (~2%), *Trachelomonas* (~1,5%), *Leptolinea* (~1,5%), *Treponema* (~1%), *Paracoccas* (~1%), *Tetrasphaera* (~1%), *Helicobacter* (~1%), *Litorilinea* (~1%), *Papillibacter* (~1%), *Clostridium* associated with Ruminococcaceae family(~1%), (88). Only three of 23 high abundance genera represent Archaea (*Methanosaeta, Methanobrevibacter,* and *Methanomassiliicoccus)* with *Methanosaeta* genus representatives being a dominating genus at St Cloud WWTF sampling site.

Furthermore, when looking at samples coming from the St Cloud WWTF site but were under the influence of different site conditions (anaerobic and aerobic) there is a noticeable change in richness, evenness and microbial abundance detected (Figure 2.7 and Figure 2.8). There are overall 126.465 sequences associated with aerobic conditions that were identified to a genus level representing 201 OTUs: 12 archaeal OTUs and 189 bacterial OTUs. However, only 23 OTUs are found to be highly abundant (>1%) in aerobic conditions with three OTUs representing Archaea and 20 OTUs representing Bacteria. Bacterial genus *Lactobacillas* was determined to be a dominant genus at ~11% in the aerobic stage of the AD process (Figure 2.7).

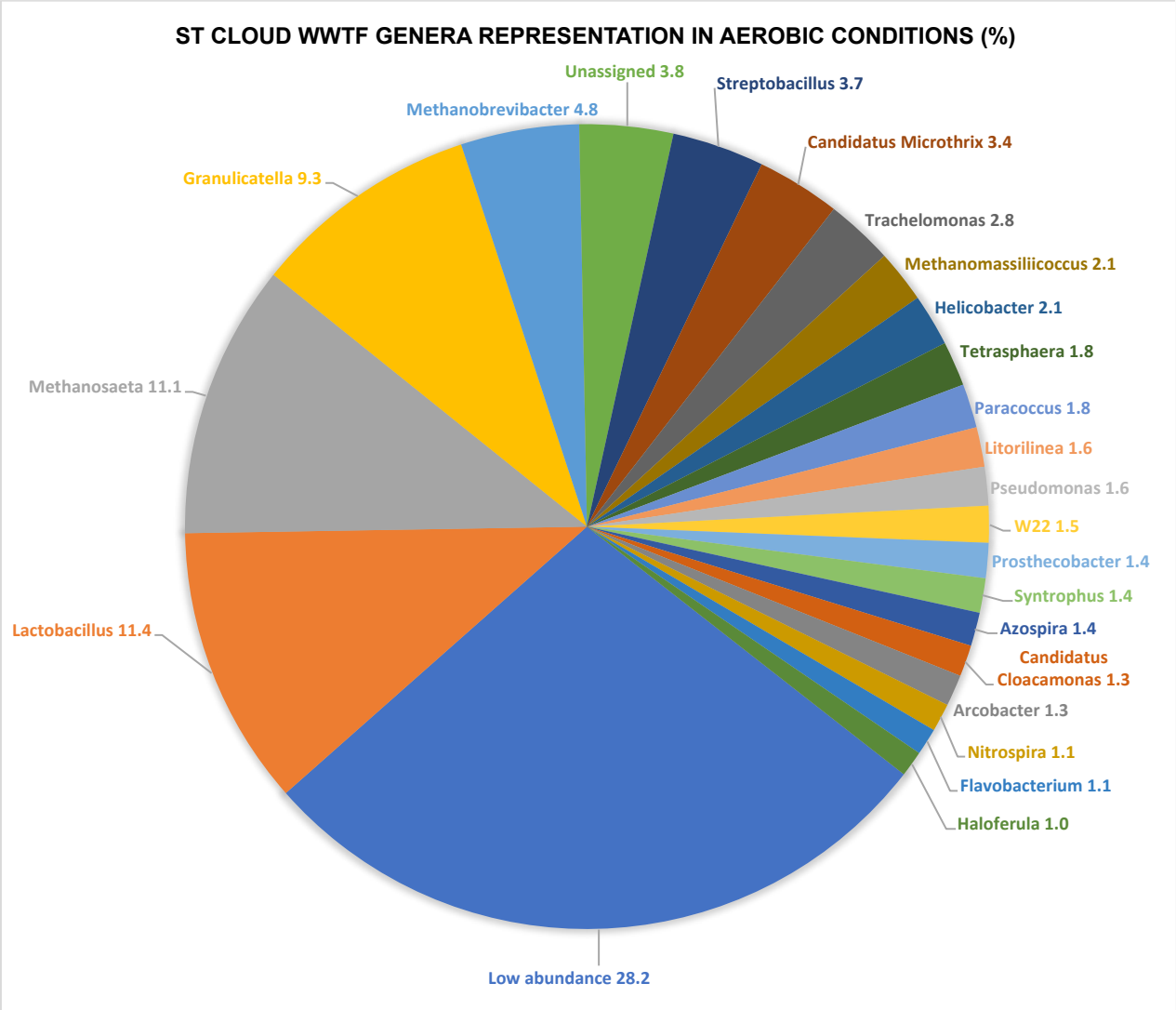


Figure 2.8. St Cloud WWTF sampling site: high abundance genera (>1%), low abundance genera (<1%), and unassigned sequences in aerobic conditions.

Next, when looking at samples derived from anaerobic sampling points (digesters) the total number of OTUs characterized to a genus level is 111 OTUs which is 90 OTUs less than in aerobic conditions, confirming lower species richness in the ADs. At the same time, the number of high abundant OTUs (>1%) also decreased to 20 OTUs. Archaea communities showed low diversity with three OTUs present. At the time archaeal genus

Methanosaeta was a dominant genus with ~41% over the whole microbial community of the AD (Figure 2.8).

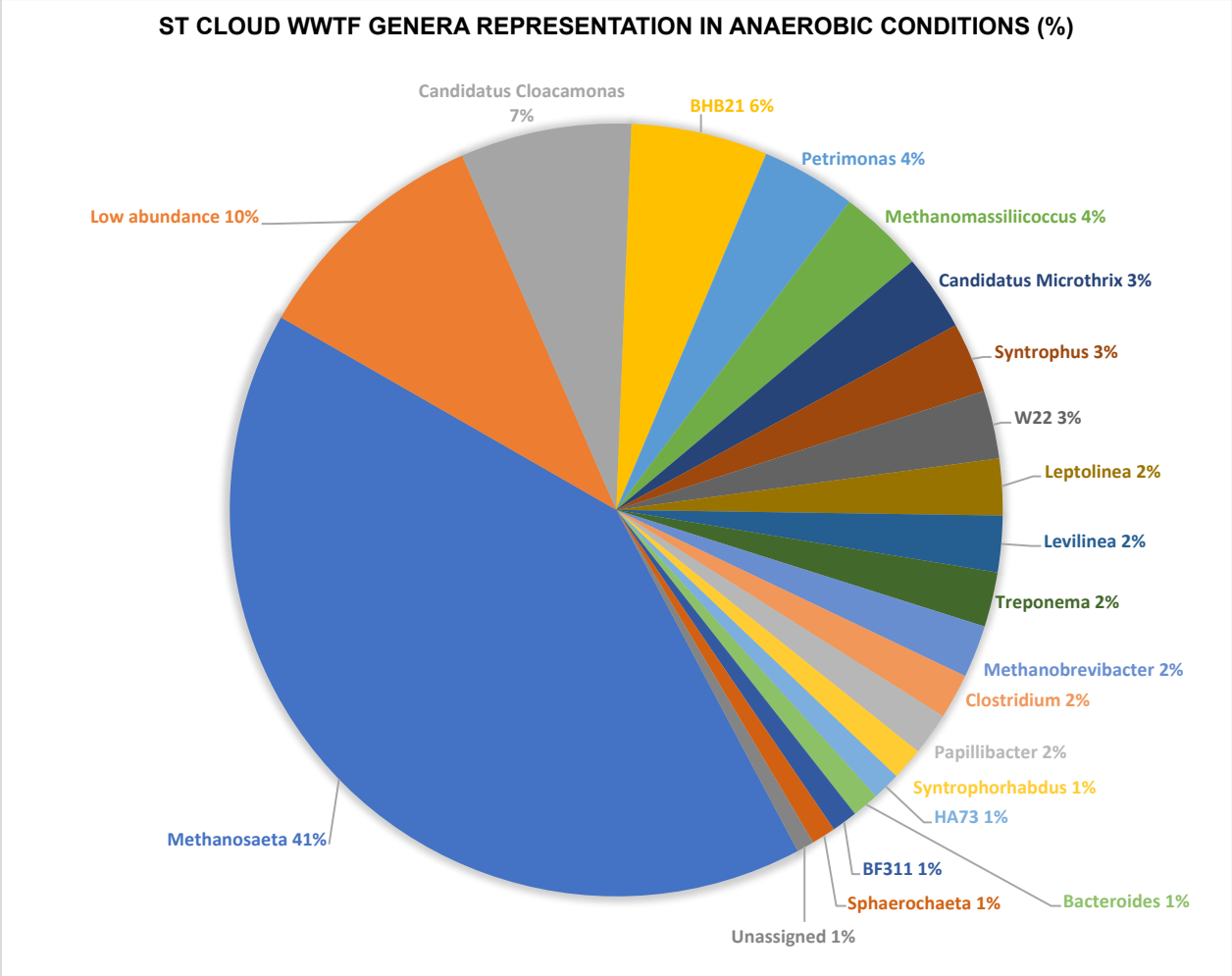


Figure 2.9. St Cloud WWTF sampling site: high abundance genera (>1%), low abundance genera (<1%), and unassigned sequences in anaerobic conditions.

Alpha diversity. For St Cloud WWTF site analysis, alpha diversity measurement (Figure 2.9) is an instrument that helps understand how the community structure changes due to temperature changes, addition of a co-digestant, oxygen availability or absence, and most importantly how these changes contribute to a constantly high and stable biogas generation at this facility. First, two main pre digestion sampling points at WWTF

site showed a very different count of observed OTUs. The Gravity belt table sampling point that carries wastewater solids showed fewer OTUs being highly abundant with a median being at ~350 OTUs. At the same time, a co-digestant (High strength waste sampling point) had median at ~250 OTUs with an equal abundance for each defined OTU. Finally, at a mixing point of both waste types a median went up to ~410 OTUs with most of the OTUs were present at low abundance. As for the Faith's PD, all three pre-digestion sampling points showed diverse populations with medians in the range from 58 (Gravity belt table) to 80 (High strength waste) nucleotide substitutions. However, after mixing of two co-digestants Faith's PD index went down, with the median being at 58 nucleotide substitutions (the level of Faith's PD index for the Gravity belt sampling point), this can be because a diverse community of waste coming from breweries was outcompeted by wastewater microbial community. Finally, all three sites showed high evenness values in the range from 0.775 to 0.875, with most OTUs being equally abundant in the High strength waste and the Mixer sampling points. Furthermore, when looking at both primary and secondary digesters they showed equal abundance of observed OTUs for all four sites. At the same time, primary digesters Observed OTUs index (300 for Digester 1 and 310 for Digester 2) is higher than secondary digesters' index. The absence of mixing and temperature decrease can contribute to this change within secondary digesters. Also, Faith's PD being on the same level at ~45 nucleotide substitutions for at least three digesters suggests that digesters not only have similar Observed OTU levels but are also operated by very similar communities. Finally, Evenness index (>0.8) continued proving that all four digesters have very similar evenly

distributed microbial communities. The outliers for the Evenness index suggested dominance of certain Archaea species, specifically species that belong to *Methanosaeta* genus. Finally, the post- digestion sampling point (Lystek) showed a decrease in Observed OTUs as a result of a much higher operational temperature. At the same time PD diversity decreased as well showing close relatedness between microorganisms at the post- digestion site. However, the Evenness index increased slightly to 0.825 when comparing to the communities inside of digesters.

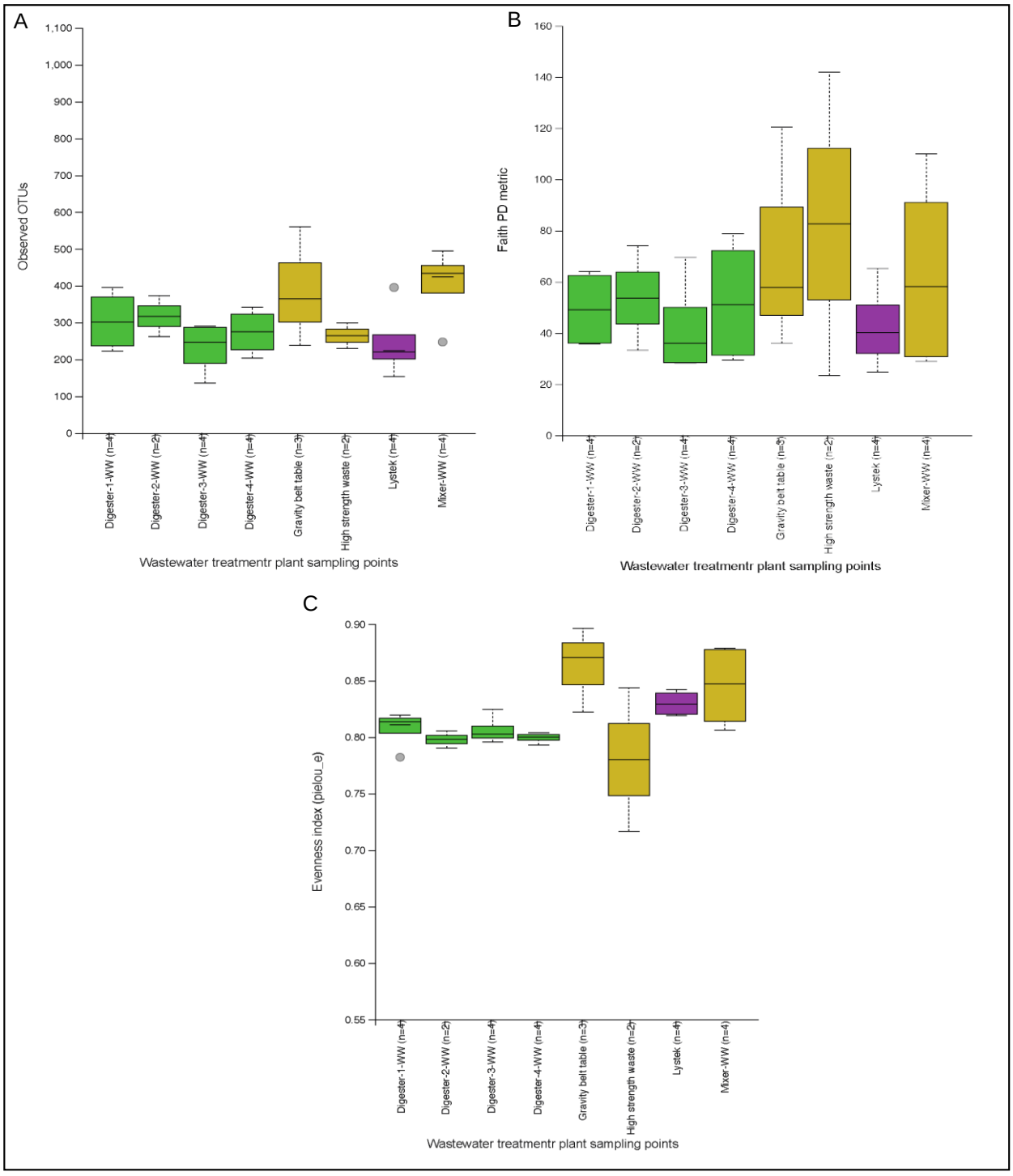


Figure 2.10. Alpha diversity analysis of (A) observed OTUs, (B) Faith's phylogenetic diversity (Faith's PD) and (C) Evenness of St Cloud WWTF: pre- (yellow), during (green), and post- (purple) digestion.

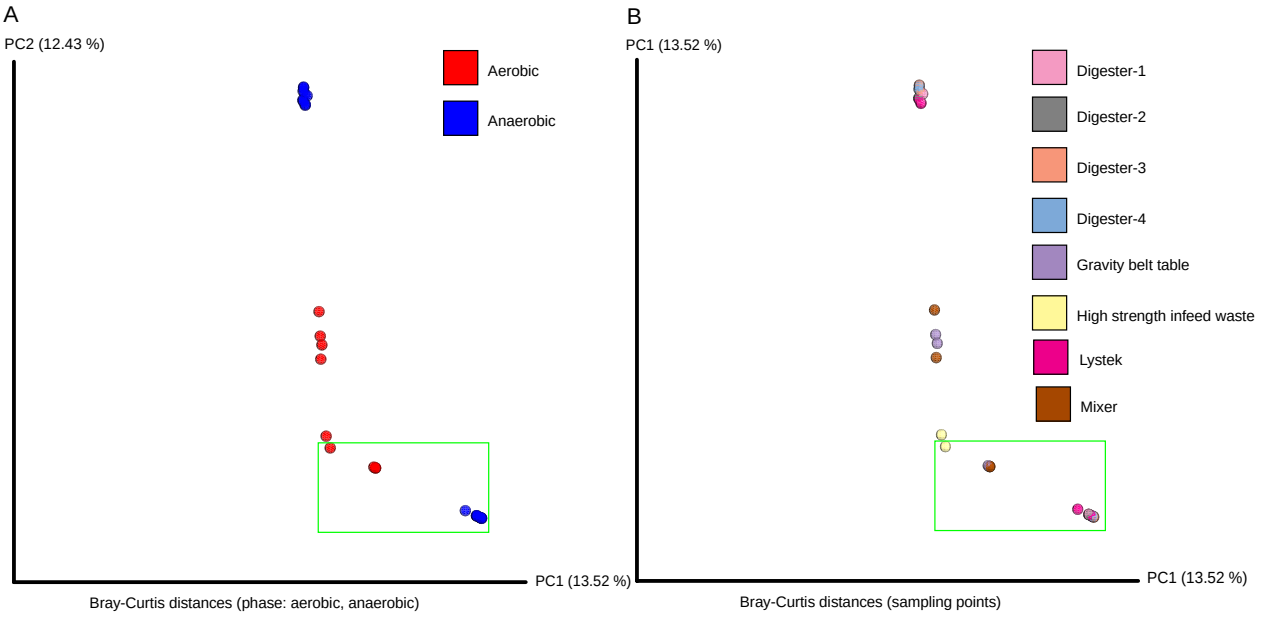
Beta diversity. Beta diversity was used to test for community differences between different phases (aerobic, anaerobic); between communities detected by the Arc and Prok primer sets; and between different sampling points at the St Cloud WWTF AD system. This analysis was performed by examining taxonomic distinctiveness between microorganisms in a given taxon and as a result determine an existence of a core community shared between all sampling points at the WWTF AD system. Figures for beta diversity were visualized on QIIME2 (v 2019.1) view (Figure 2.10) (64).

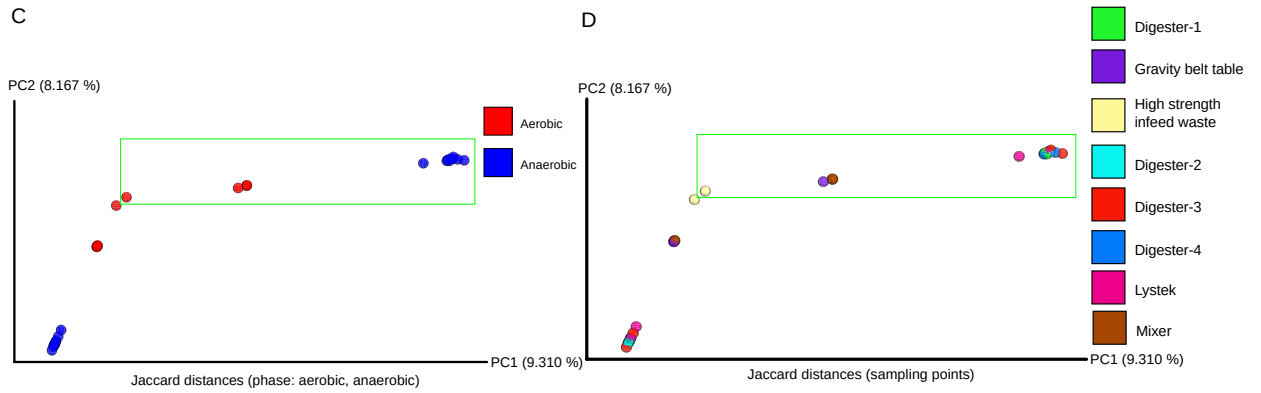
All the results of the Principal Coordinates Analysis (PCoA) for the WWTF sampling points were produced to demonstrate grouping of microbial communities based on similarities of communities via Bray-Curtis dissimilarity distances (Figure 2.10 A, B), Jaccard dissimilarity distances (Figure 2.10 C, D), and unweighted UniFrac phylogenetic distance metrics (Figure 2.10 E, F). The PCoA plots were visualized on EMPeror (89).

First, PCoA plots (Bray-Curtis dissimilarity distances, Jaccard dissimilarity distances, unweighted UniFrac phylogenetic distance) are concordant. All three approaches determined a significant taxonomic distinction between the microbiomes in High strength infeed waste sampling point (p-value = 0.001), sampling points in anaerobic phase and the rest aerobic phase sampling points by both primer sets. Basically, there are three distinct groups: samples coming from aerobic sampling points, samples coming from anerobic sampling points and samples coming from the High strength infeed waste with the last one being the most distanced from the rest according to Bray-Curtis (Figure 2.10 A, B), Jaccard (Figure 2.10 C,D), and Unweighted Unifrac (Figure 2.10 E, F) PCoA

plots. Furthermore, both Arc and Prok primer sets grouped samples according to the above trend.

Overall, the trend for St Cloud WWTF site is there is a significant variation in microbial communities depending on the phase they are in (aerobic or anaerobic). This aligns with the taxonomic analysis (Figure 2.7 and 2.8) that displays very distinct domination of *Methanosaeta* genus with ~41% in anaerobic sampling points, when the same genus had only ~11% percent at aerobic sampling points, an increase of ~30%. Also, the other two most abundant genera that follow *Methanosaeta* genus in the aerobic phase of the system, *Lactobacillus* genus and *Granulicatella* genus that are both aerotolerant were substituted with anaerobic genera from Cloacamonaceae family (*Candidatus Cloacamonas* and BHB21). These two genera are both syntrophic fattyacid fermenters and previously were detected in swine manure treating reactors(97). This overlap is an evidence that despite chemically different infeed, as well as very different designs of on-farm manure treating reactors and WWTF reactors still share microbial communities responsible for methane productions process.


A
PC2 (12.43 %)
Aerobic
Anaerobic
PC1 (13.52 %)
Bray-Curtis distances (phase: aerobic, anaerobic)
B
PC1 (13.52 %)
Digester-1
Digester-2
Digester-3
Digester-4
Gravity belt table
High strength infeed waste
Lystek
Mixer
PC1 (13.52 %)
Bray-Curtis distances (sampling points)


C
PC2 (8.167 %)
Aerobic
Anaerobic
PC1 (9.310 %)
Jaccard distances (phase: aerobic, anaerobic)
D
PC2 (8.167 %)
Digester-1
Gravity belt table
High strength infeed waste
Digester-2
Digester-3
Digester-4
Lystek
Mixer
PC1 (9.310 %)
Jaccard distances (sampling points)

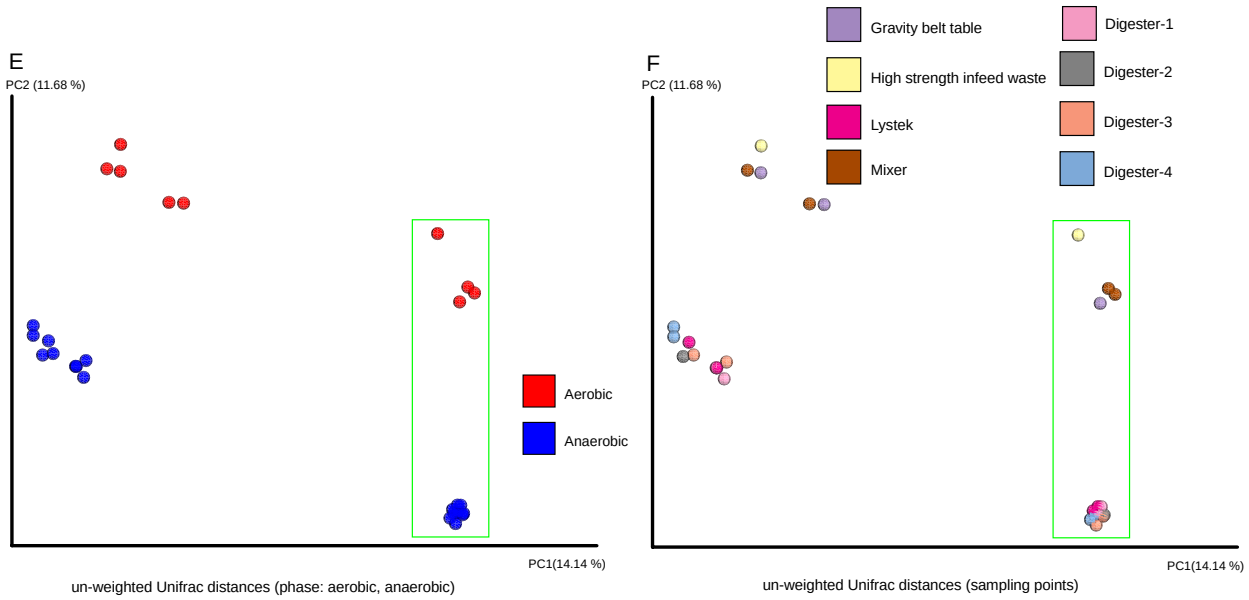


Figure 2.11. Beta diversity analysis of (A) Bray-Curtis, (C) Jaccard, and (E) un-weighted Unifrac distances during anaerobic and aerobic phases in the AD system; (B) Bray-Curtis, (D) Jaccard, (F) un-weighted Unifrac distances between different sampling points at the WWTF sampling sites with the Prok primer points in a green frame and the Arc primer results outside the frame.

Experimental Benchtop Anaerobic Digesters: Value and Purpose

In order to improve already established approaches to AD, test new formulas of co-digestion, measure potential benefits by changing operational temperatures, rates or volumes of infeed, and observe dynamics of microbial community diversity commercial and scientific institutions often use laboratory benchtop size ADs. The use of benchtop digesters is especially valuable when experimenting with microbial community responses to changing AD standard operational parameters (temperature, pH, alkalinity, etc.) (98). This type of digester can be designed to completely resemble an industrial size AD, however, in case of the benchtop digester microbial community collapse it is much easier to restart the process with the minimum loss of time and no revenue losses. The size of the benchtop digesters varies from 100 ml to 10-15 liters. Normally, a new experiment can start with multiple 100 ml digesters to test as many scenarios as possible. After rejecting a number of hypotheses, the size of a benchtop experiment can be scaled up to test the reminder of hypotheses in a larger AD to increase the accuracy and in many cases mimic the industrial AD setting as close as possible.

Experimental Benchtop Anaerobic Digester Site

A mini AD that operates at the Department of Bioproducts, Biosystems, and Engineering at the University of Minnesota laboratory as a two-stage benchtop reactor under mesophilic conditions is a stable biogas producer and a test reactor for a future 100-liter experimental AD. The substrate for the AD consisted of fruits and vegetables that were inoculated with an inoculum from a dry industrial digester operating on food waste and yard trimmings at the University of Wisconsin, Oshkosh. The feedstock for the benchtop reactor

consisted precisely of 10% apple, 1% banana, 7% bread, 8% beef, 15% cabbage, 7% cereals, 2% cheese, 16% milk, 3% onion, and 31% potato. This feedstock was prepared by mixing all the ingredients in a blender and then frozen. The amount of the feedstock prepared was enough to satisfy the need for the whole length of the experiment conducted at Dr Hu's laboratory and guarantee a consistency in the infeed composition.

As for the design, the benchtop solid-state AD consisted of two separated chambers. The first chamber held the mix of food waste and inoculum and operated in aerobic conditions for 38 days. After that the digestate was transferred to a second chamber that operated under anaerobic mesophilic conditions. The second chamber would continue digestion for another 18 days and resulted in stable biogas yields.

This digester was chosen because of the several major differences from the ADs described above: the digestate composition being solely food waste and not inoculated with microorganisms from cow manure, this AD is a dry digester (>20% of total solids), its digestate chemical composition is constant, it operates under mesophilic conditions (no heating elements used for reactors) (41, 99).

Sampling

For the benchtop AD we followed the same sampling protocol as for the Haubenschild farm site and the St Cloud WWTF: we sampled all the sites that feed the AD, and the digestate of both chambers of the benchtop set up. However, an outfeed of the AD was not available.

At this location samples from the following points were collected: feedstock from a storage container kept in a freezer, digestate from the aerobic chamber of the reactor, and

digestate from the anaerobic stage of the reactor (Figure 2.11).

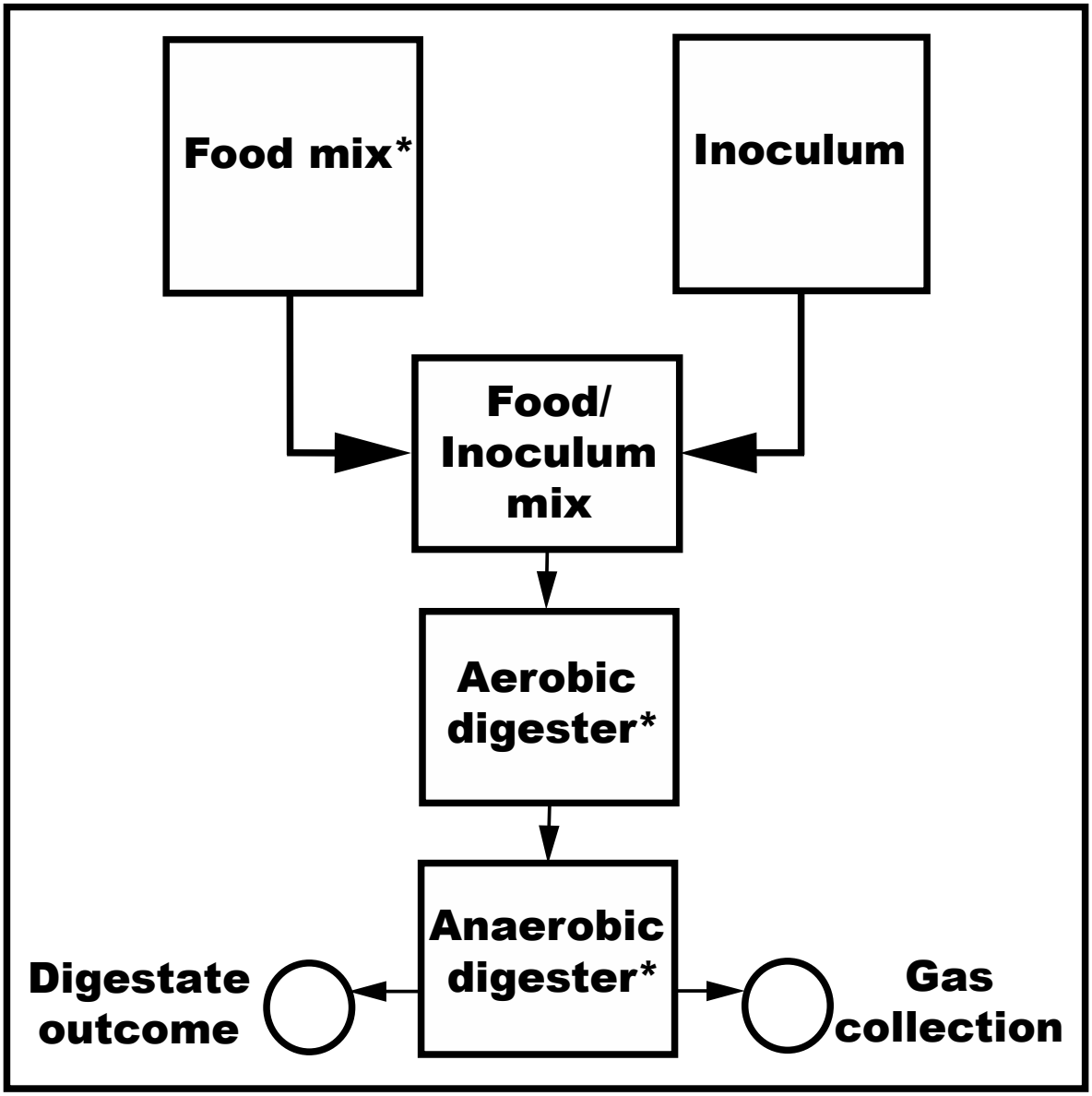


Figure 2.12. University of Minnesota, experimental benchtop AD set up; main sampling points (indicated by the asterisks *).

Materials and Methods

Materials and methods used for the University of Minnesota benchtop digester are identical to Haubenschild farm site (Chapter 2, Section: Material and methods: DNA extraction, quantification, and 16S rRNA sequencing).

Results: Microbial Composition and Diversity Analysis

Microbial composition. For dry benchtop anerobic digester at the University of Minnesota site overall 218,356 nonchimeric, good-quality sequences were recovered. These sequences (n = 218,356 sequences) represent 142 OTUs: 8 archaeal OTUs and 134 bacterial OTUs. However, only ~44% of sequences (n = 96,232) were characterized to a genus level: 13,209 sequences were recovered by an Arc primer and 83,496 sequences were recovered by a Prok primer. Overall, only 473 sequences were left unassigned. All unassigned sequences were detected by the Arc primer, which can be because of still limited knowledge about Archaea species.

Overall, the Arc primer recovered 13,209 sequences that were characterized to the genus level and represent ~27.5% of overall Arc recovered sequences (n = 48,021) and ~22% of overall sequences recovered by both primers. These sequences (n = 13,209) represent 8 archaeal OTUs (n = 3,616 sequences) and 39 bacterial OTUs (n = 9,593 sequences).

As for the Prok primer, it was able to recover 83,496 sequences that were characterized to the genus level and represent ~49% of overall Prok recovered sequences and ~38% of overall sequences recovered by both primers. These Prok recovered sequences (n = 83,496) represent 5 archaeal OTUs (n = 316 sequences) and

129 bacterial OTUs (n = 83180 sequences). The number of detected bacterial OTUs by the universal Prok primer is higher by 90 OTUs compared to the results of the Arc primer.

Overall, together both primers detected 16 high abundance OTUs (>1%)(60) on a genus level: *Prevotella* (~20%), *Petrimonas* (~13%), *Lactobacillus* (~13%), *Clostridium* (~5%), *Sporanoaerobacter* (~5.5%), *Acholeplasma* (~4%), *Clostridium* associated with Ruminococcaceae family (~4%)(88), *Pseudomonas* (~4%), *Bifidobacterium* (~3%), Eryspelotrichaceae family, genus [gut] (~3%), *Raphanus* (~2%), Aminobacterium (~2%), *Methanosaeta* (~2%), *Streptococcus* (~2%), *Bacteroides* (~2%), and *Paenibacillus* (~1%). Only one genus out of 16 high abundance genera was an archaeal genus (*Methanosaeta*) and the rest were bacterial genera. Furthermore, when looking at samples from the above site but were under the influence of different site conditions (anaerobic and aerobic) there is a noticeable change in richness, evenness and microbial abundance detected. There are overall 58,875 sequences associated with aerobic conditions that were identified to a genus level representing 124 OTUs: 8 archaeal OTUs, 116 bacterial OTUs. However, only 16 OTUs are found to have high abundance (>1%) in aerobic conditions with two OTUs representing Archaea and 14 OTUs representing Bacteria. Bacterial genus *Lactobacillus* was determined to be a dominant genus with 11,821 sequences that represent ~22.5% of all sequences assigned a genus (Figure 2.12).

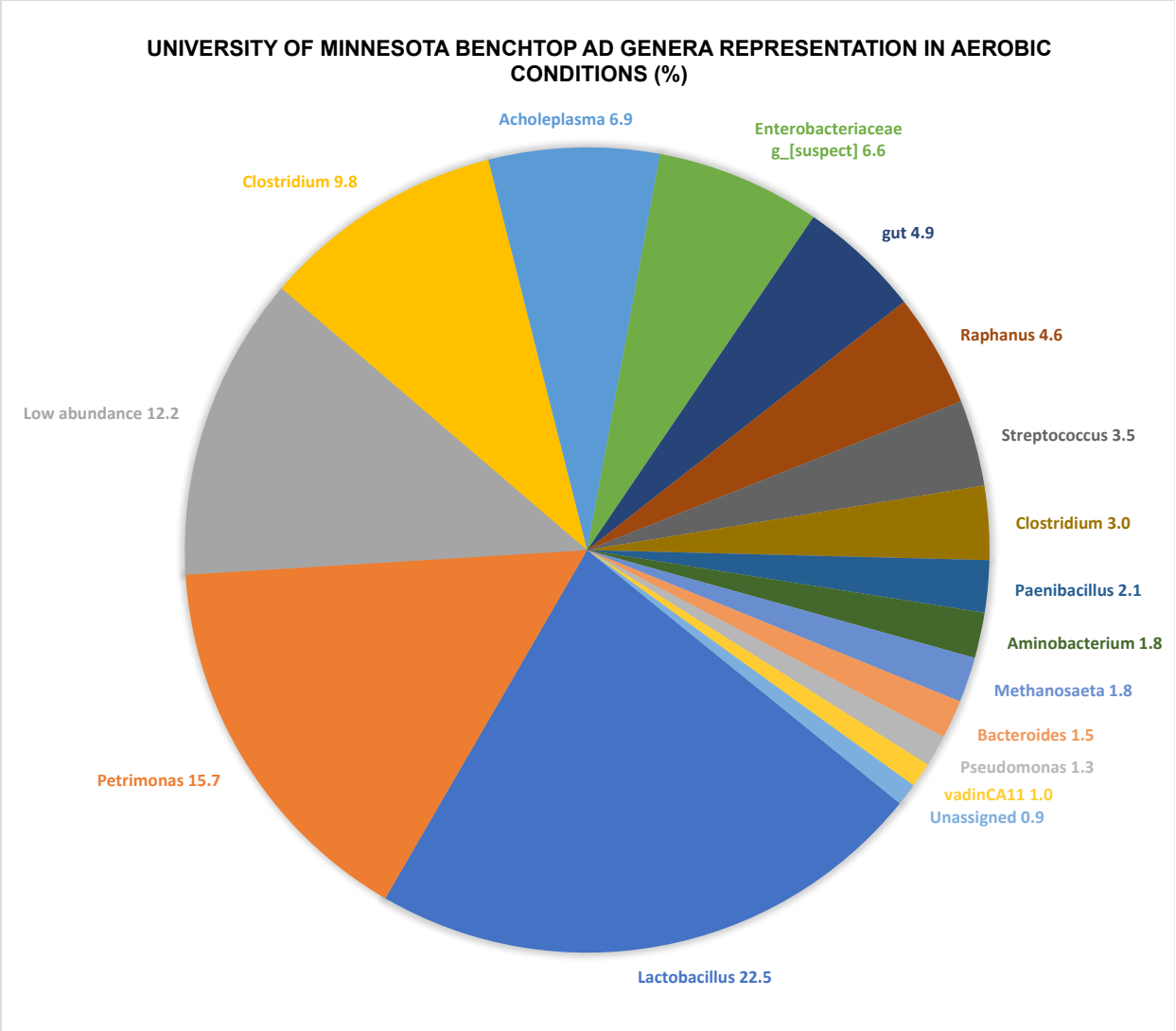


Figure 2.13. University of Minnesota benchtop AD sampling site: high abundance genera (>1%), low abundance genera (<1%), and unassigned sequences in aerobic conditions.

Next, when looking at samples derived from anaerobic sampling points (second chamber of the digester) the total number of OTUs characterized to a genus level is 87 OTUs which is 37 OTUs less than in aerobic conditions, confirming lower species richness in the ADs. At the same time, the number of high abundant OTUs (>1%) also decreased to 11 OTUs. Also, Archaea communities showed low diversity again with only

one OTUs. It is the first digester in this study that was dominated with a bacterial genus (*Prevotella* ~44%, gut microbiota representative) in anaerobic conditions (Figure 2.13).

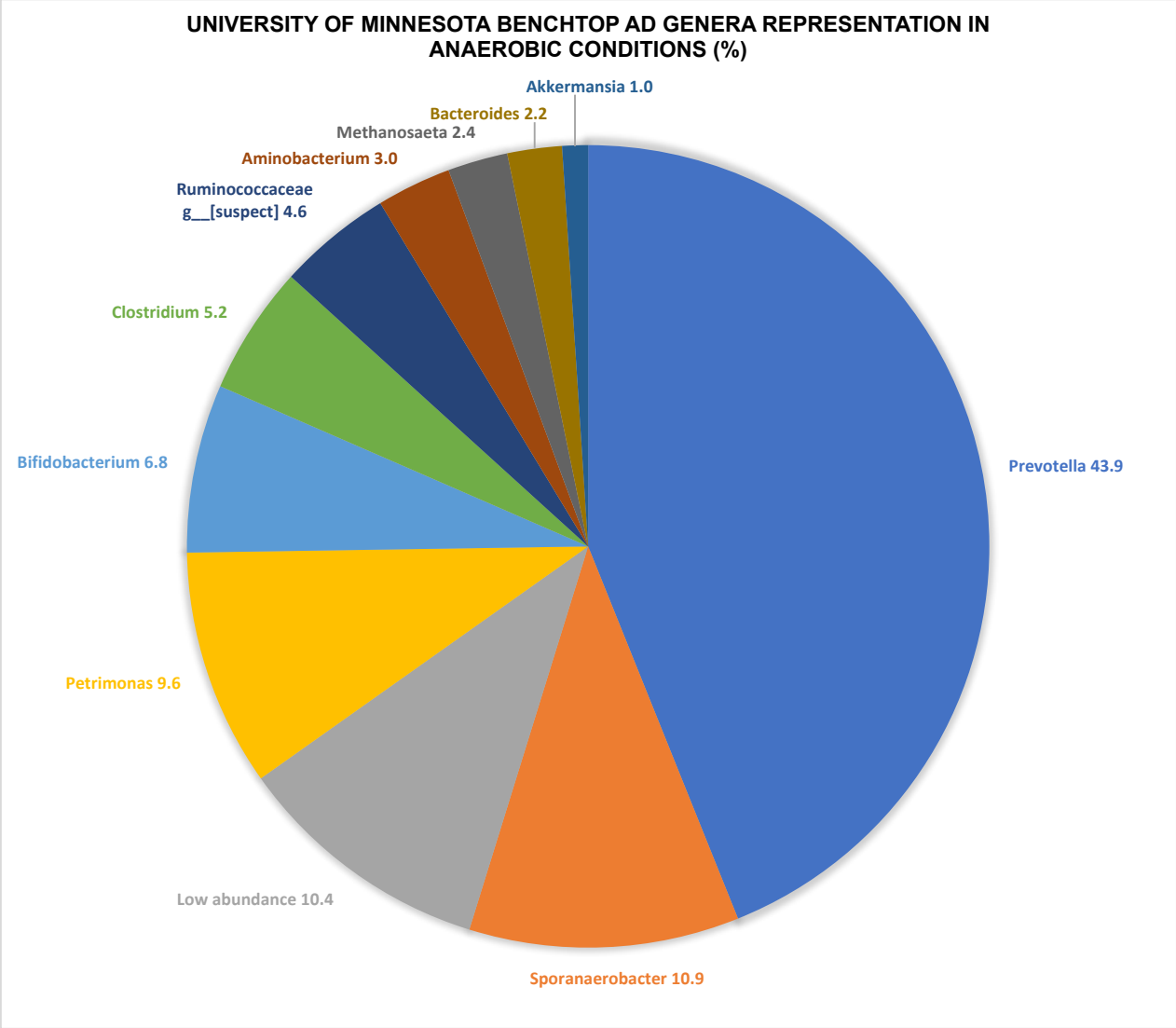


Figure 2.14. University of Minnesota benchtop AD sampling site: high abundance genera (>1%), low abundance genera (<1%), and unassigned sequences in anaerobic conditions.

Alpha diversity. For the University of Minnesota benchtop AD site analysis, alpha diversity measurement (Figure 2.14) is a tool that might help to understand how the community structure of the infeed mix changes inside the digester that was inoculated with another dry feed digester inoculative material only at the start of the AD operation but continues to be stable.

First, the infeed mix sample (pre digestion sampling point) at the University of Minnesota benchtop AD demonstrated a lower count of observed OTUs comparing to the both chambers of the AD reactor. Also, the OTUs demonstrated a lower evenness between genera at this spot with the median being at ~0.625; outliers can be attributed to *Lactobacillus* being a dominating genus (22.5% sequences). Finally, Faith PD demonstrated low diversity (no Archaea at this point, only closely related Bacteria genera). As for two samples coming from the aerobic and anerobic chambers, they demonstrated a much higher counts (expected result due to chambers being inoculated with a digestate from another dry AD). However, aerobic chamber (Digester 1) had the observed OTUs equally abundant with the median in the middle of the boxplot. At the same time, anaerobic chamber (Digester 2) matched the OTU count of the Digester 1 with the median being at ~280, but the abundance shifted towards most of OTUs being in low abundance. At that point anaerobes became dominant and aerobic organisms were still present but in much smaller quantities. This observation is also supported by the Evenness index: aerobic chamber has a higher result for evenness.

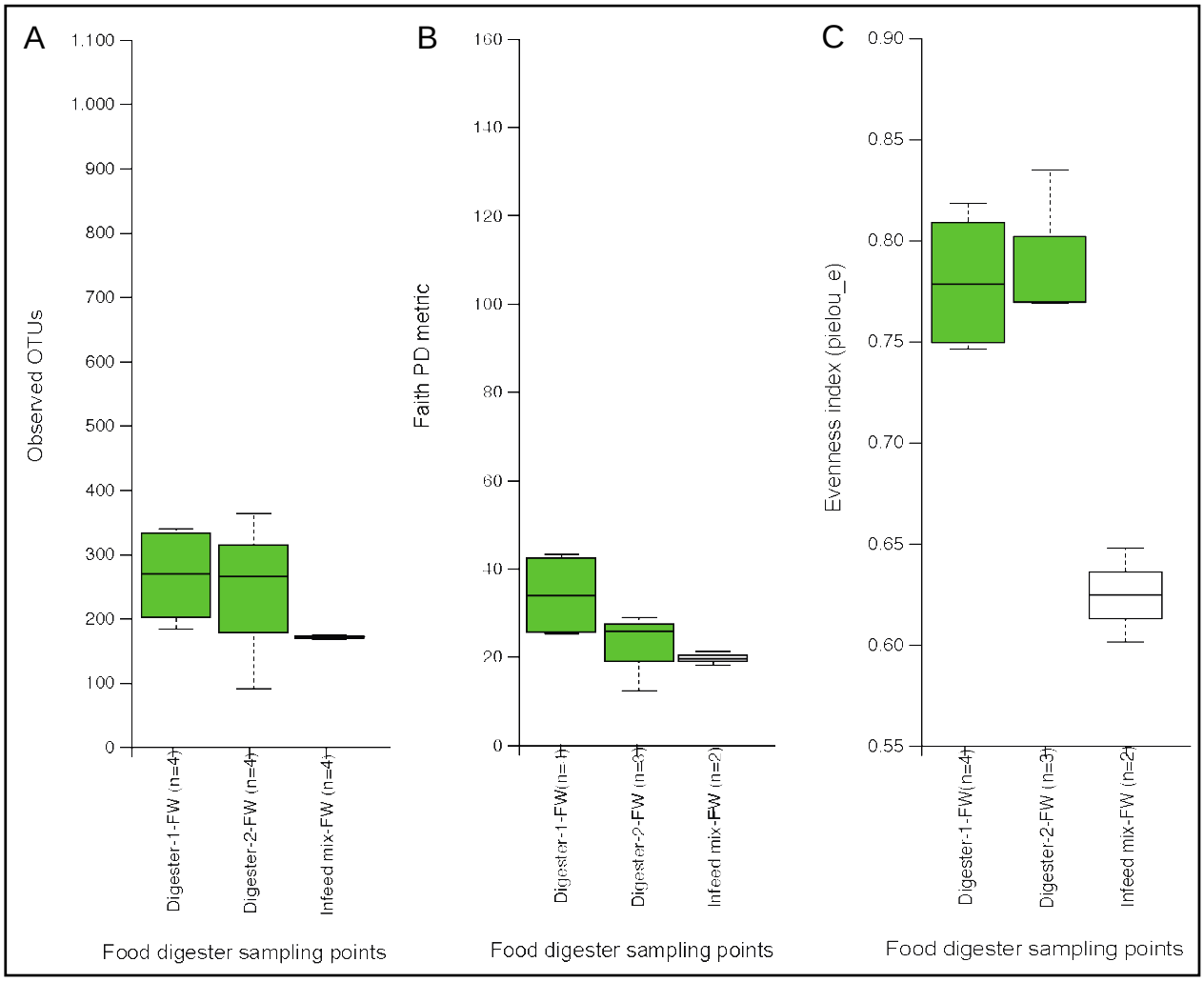


Figure 2.15. Alpha diversity analysis of (A) observed OTUs, (B) Faith's phylogenetic diversity (Faith's PD) and (C) Evenness of University of Minnesota benchtop digester: pre- (white) and during (green) digestion with Digester 1 being an aerobic chamber and Digester 2 being an anaerobic chamber.

Beta diversity. Beta diversity was used to test for community differences between different phases (aerobic, anaerobic); between communities detected by the Arc and Prok primer sets; and between different sampling points at the University of Minnesota benchtop AD system. This analysis was performed by examining taxonomic distinctiveness between microorganisms in a given taxon and as a result determine an existence of a core community shared between all sampling points at the benchtop AD system. Figures for beta diversity were visualized on QIIME2 (v 2019.1) view (Figure 2.15) (64).

All the results of the Principal Coordinates Analysis (PCoA) for the University of Minnesota benchtop AD sampling points were produced to identify grouping of microbial communities based on similarities of communities via Bray-Curtis dissimilarity distances (Figure 2.15 A, B), Jaccard dissimilarity distances (Figure 2.15 C, D), and unweighted UniFrac phylogenetic distance metrics (Figure 2.15 E, F). The PCoA plots were visualized on EMPeror (89).

First, PCoA plots (Bray-Curtis dissimilarity distances, Jaccard dissimilarity distances, unweighted UniFrac phylogenetic distance) of the Arc primer set are concordant. All three approaches determined a significant taxonomic distinction between the microbiomes in anaerobic and aerobic sampling points (p-value = 0.001). However, Infeed mix reads samples were disqualified at a quality step of the pipeline for this sampling site and did not take part in the Beta diversity analysis. As for the Prok primer set, Bray-Curtis dissimilarity and unweighted UniFrac phylogenetic distance do look alike by putting all the samples relatively close to each other. However, Jaccard dissimilarity for

the Prok primer distanced Digester-1 (aerobic reactor of the AD system) from the rest of the samples.

The general trend for the University of Minnesota benchtop digester site, there is a variation in microbial communities of the benchtop digester sampling points depending on the phase they are in (aerobic or anaerobic). However, unweighted UniFrac distances suggest that this difference is not significant, as well as PERMANOVA results suggesting p-value = 0.091 between anerobic Digester-2 samples and the infeed mix of the digester samples. These communities were still relatively close because phylogenetically speaking there was not much change occurring. This was the first site where methanogens (Archaea representatives) did not become dominant and bacterial representatives (aka *Prevotella* ~44%,) were dominating both phases (aerobic and anaerobic). Bacteria species of *Prevotella* genus are gram-negative anaerobic bacteria associated with human microbiota. Also, it was previously detected in anaerobic digesters processing food waste and responsible for degradation of cellobiose, glucose, and mannose (100).

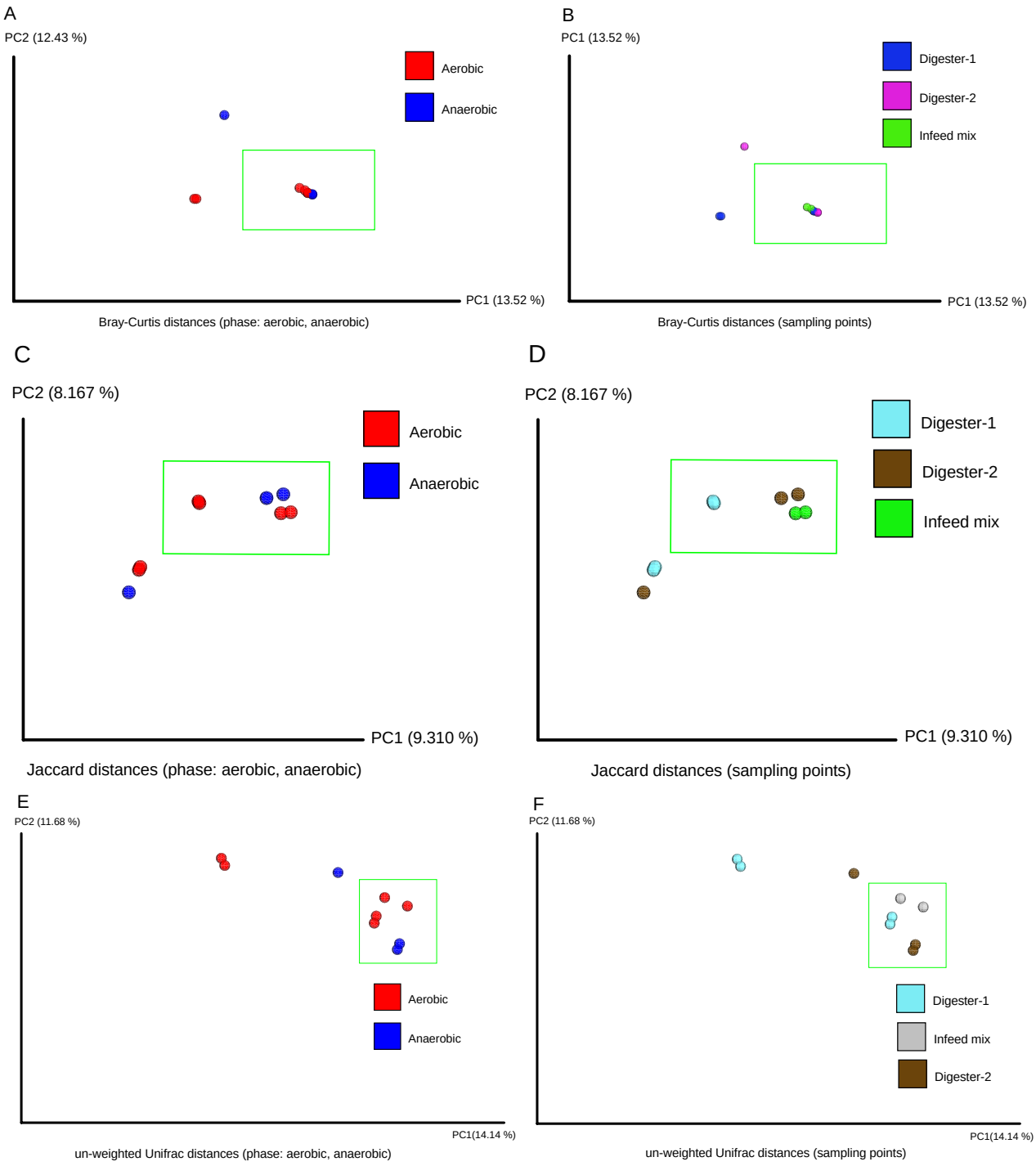


Figure 2.16. Beta diversity analysis of (A) Bray-Curtis, (C) Jaccard, and (E) un-weighted Unifrac distances during anaerobic and aerobic phases in the AD system; (B) Bray-Curtis, (D) Jaccard, (F) un-weighted Unifrac distances between different sampling points at the University of Minnesota benchtop digester sampling sites with the Prok primer points in a green frame and the Arc primer results outside the frame.

Chapter 3: Microbial Composition Analysis of Core Microbiomes of All Sampled Digesters

Introduction

In this chapter, the results of comparison of amplicon sequencing of all collected DNA samples including a SCSU food fermenter site are reported. To assess predominant OTUs, sequencing data were combined for all samples from each anaerobic reactor, all pre-, and post digester points.

Microbial Composition

Overall, there were 4,326,132 sequences recovered by MiSeq Illumina by for all sampled sites. After QIIME2 (v 2019.1) quality control step 1,787,804 sequences were reported to be nonchimeric and good quality. These sequences (n = 1,787,804 sequences) were assigned to 33 archaeal OTUs, 690 bacterial OTUs, and 7,259 sequences, were left unassigned. Overall, there were 365,615 archaeal sequences and 1,422,189 bacterial sequences. Sequences that belong to genus *Methanobrevibacter* was found to be the most abundant with 160,621 sequences total. Bacteroidales oreder appeared to be the most abundant bacterial OTU that was characterized to the order level with 117,203 sequences total. Number three in the most abundant OTUs list is the OTU with 96,980 sequences, however it was only assigned a kingdom – Bacteria. The number of sequences that were not characterized deeper may be explained by the limit of information in the database used for this study (Greengenes v 13.8 database is used by QIIME2 2019.1 as a default). For further taxonomic analysis and comparison, only OTUs

characterized to the genus level were used. Genus level is the highest taxonomic rank that is recommended to go to with the methods used for this study. There were no attempts taken to assign species at this time. Overall, there were 397 OTUs (894,403 sequences, ~50% of overall detected sequences) characterized to the genus level: 15 archaeal OTUs and 382 bacterial OTUs. Among these OTUs only 16 are highly abundant OTUs (>1%) (Figure 3.1): *Methanobrevibacter, Lactobacillus, Methanosaeta, Bifidobacterium, Prevotella, Clostridium, Corynebacterium, Atopococcus, Petrimonas, Psychrobacter, Weisella, Treponema, Methanospaera, Collinsella, Alistipes, Pseudomonas.*

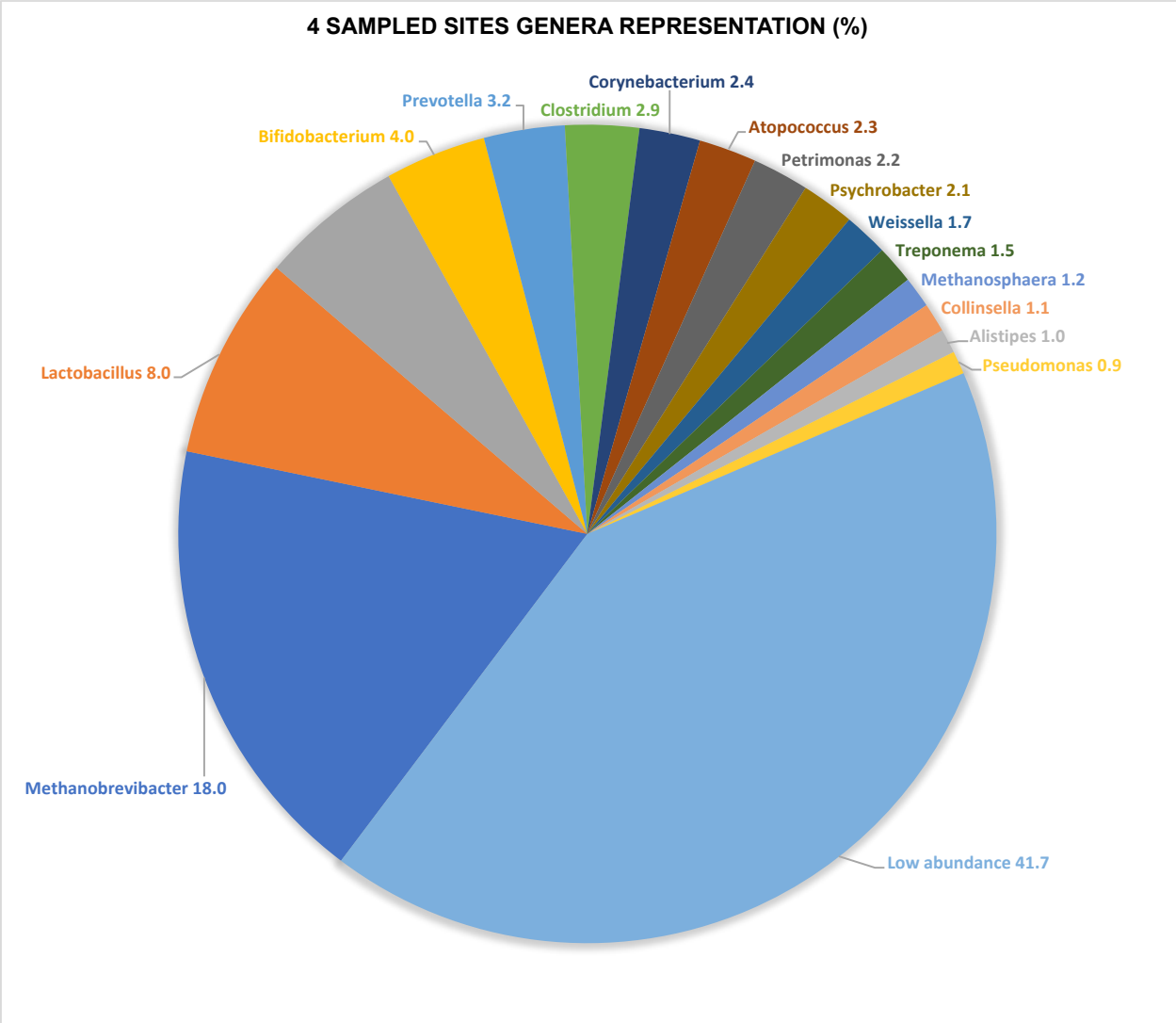


Figure 3.1. Haubenschild farm, WWTF, University of Minnesota benchtop AD, and SCSU food fermenter sampling sites: high abundance genera (>1%) and low abundance genera (<1%).

As for samples that were coming from the anaerobic phase of sampled sites, they show different abundance ratios. First, there were 648,255 sequences reported to be nonchimeric and good quality by QIIME (v 2019.1) for ADs of all sampled sites after the quality control step. At anaerobic locations the following number of OTUs were recovered: 23 archaeal OTUs that were represented by 148,770 sequences; 457 bacterial OTUs that were represented by 497,800 sequences; and 1,685 sequences, were left unassigned.

However, a much smaller number of sequences were assigned a genus. A total of 291,261 sequences that represent ~45% of all sequences reported from ADs sites were assigned a genus. Overall, QIIME2 (v 2019.1) assigned 94,303 archaeal sequences and 196,958 bacterial sequences to a genus (this data was used for further taxonomic analysis and comparison). The above sequences represent 13 archaeal and 247 bacterial OTUs characterized to a genus level. There were 21 OTUs that were considered high abundance OTUs (>1%), which suggested a higher diversity among high abundant OTUs in ADs. It was determined that the most abundant (>1%) OTUs in anaerobic phase of sample sites are (Figure 3.2): *Methanosaeta, Methanobrevibacter, Prevotella, Lactobacillus, Clostridium, Petrimonas, Candidatus Cloacomonas, Treponema, Bifidobacterium, Psychrobacter, BHB21, Mathanomassliicoccus, Sporanaerobacter, W22, Syntrophus, Alistipes, Akkermansia, Candidatus Microthrix, Levilinea, Clostridiodes,* and *Leptolinea.*

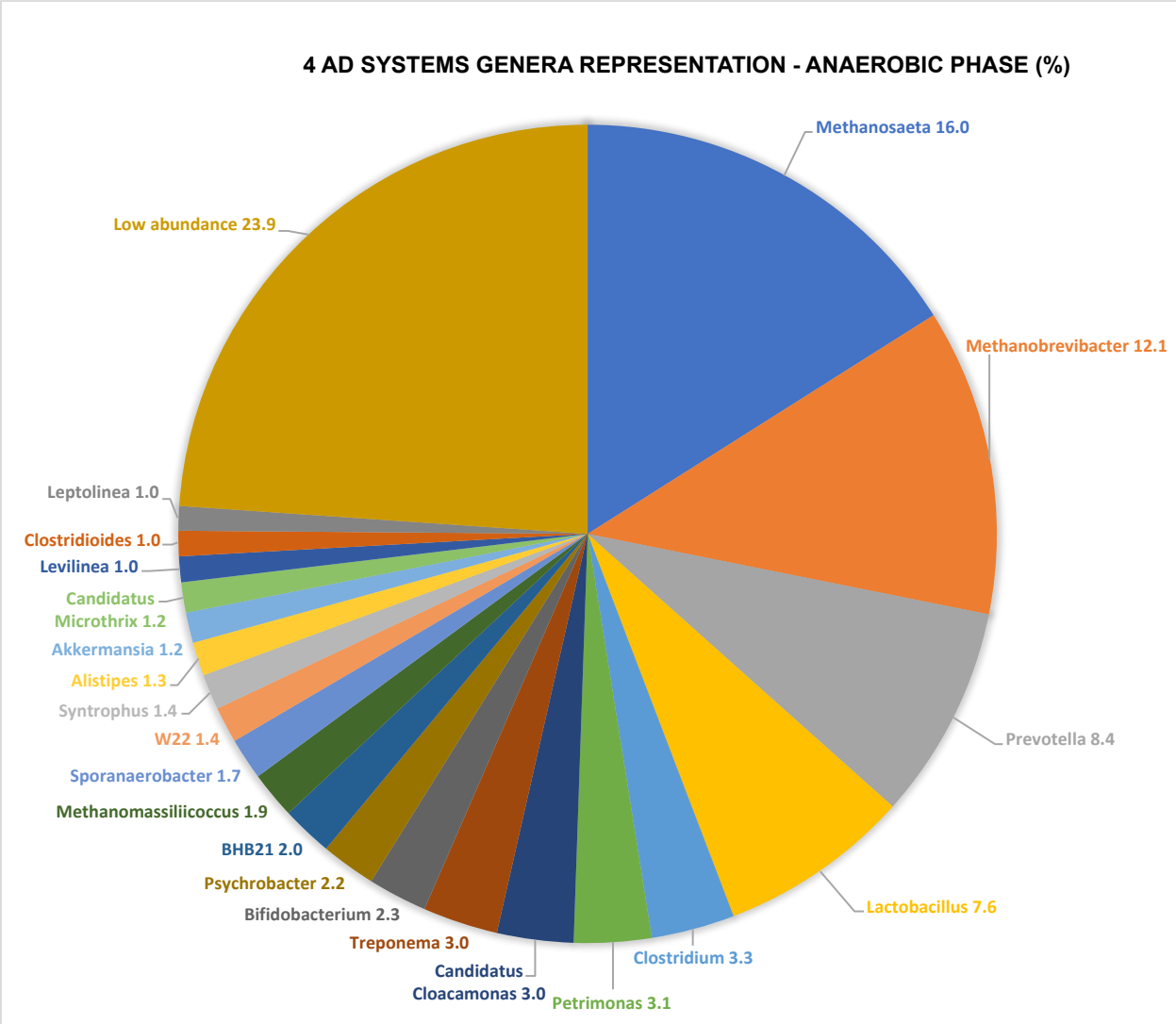


Figure 3.2. Haubenschild farm, WWTF, University of Minnesota benchtop AD, and SCSU food fermenter ADs sampling results: high abundance genera (>1%) and low abundance genera (<1%).

Combined Microbial Diversity of All Sampled Sites

Alpha diversity. Alpha diversity is a crucial instrument to understand how the community structures of different sampling sites compare to one another. Haubenschild farm, SCSU food fermenter, St Cloud WWTF, and UoM benchtop AD alpha diversity indices (Faith PD, Evenness, and Observed OTUs) were assessed for each sampling site and for each individual sampling point within a site. First, phylogenetic diversity index was measured using Faith's PD index. Faith's PD is based on the length summary of all the branches of the phylogenetic tree that covers a particular species and comparing this sum to the sum of branch length of other species present in the same sample (82). Faith's PD plot (Figure 3.3) suggests that the University of Minnesota has the lowest phylogenic diversity among all four sites with (p-value = 0.01). This observation might be due to the ingredient's composition used for the benchtop digester (solely fresh food inoculated with the dry digester inoculum at the start of the project).

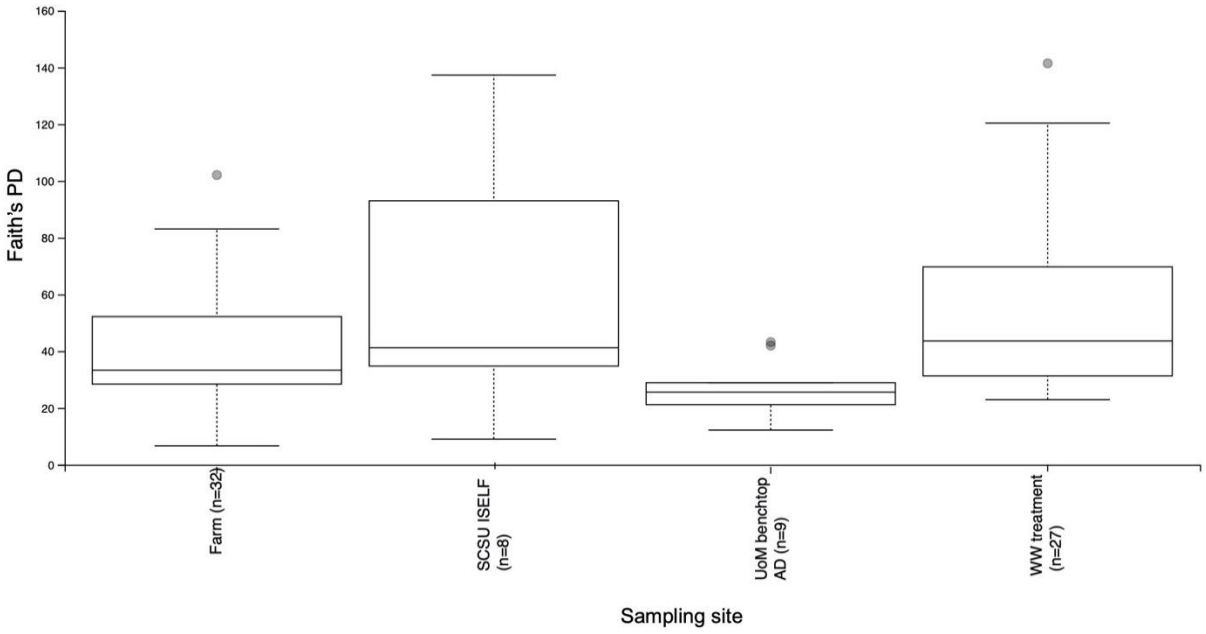


Figure 3.3. Alpha diversity analysis of Faith's phylogenetic diversity (Faith's PD) of Haubenschild farm (Farm), SCSU food fermenter (SCSU ISELF), University of Minnesota benchtop digester (UoM bechto AD), and St Cloud WWTF (WW treatment).

As for the evenness index (Pielou's evenness: 0 to 1, where 1 is complete evenness) that measures diversity along with species richness (richness - the number of different species in one area vs evenness is the number of organisms of each species in the same area) it shows high measurements for all for sites with the p-value being insignificant (p-value = 0.15) (Figure 3.4). However when looking at the sampling points at all the sites, p-value is significant (p-value = 0.007) and three sampling points (Barn 4 at the Haubenschild farm, food waste for the food fermenter at the SCSU, and infeed mix for the University of Minnesota benchtop digester) were observed to show lower measurements than the rest 20 sampling points (< 0.7) (Figure 3.5). Finally, observed OTUs index showed insignificant differences between four sampling sites (p-value = 0.18) but according to the graph benchtop digester at the University of Minnesota had the

lowest count of OTUs (Figure 3.6). At the same time, the outliers that were present at all four sites suggested that there were dominating species at all sampling sites, which was confirmed when looking at each site individually in Chapter 2.

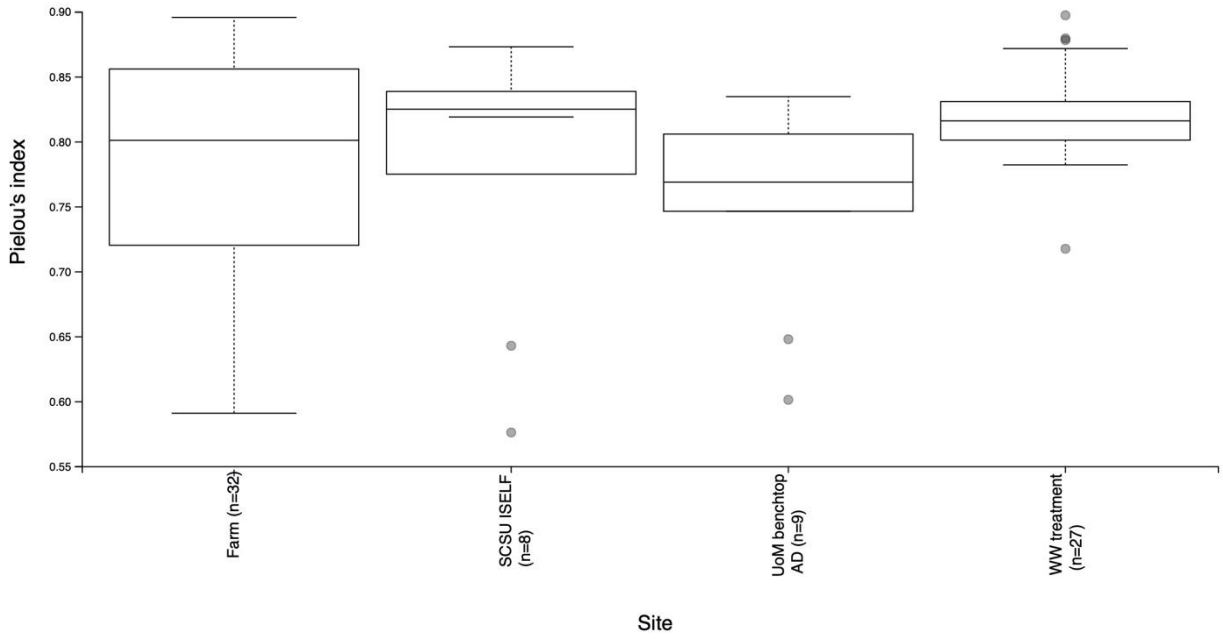


Figure 3.4. Alpha diversity analysis of species evenness (Pielou's diversity) of Haubenschild farm (Farm), SCSU food fermenter (SCSU ISELF), University of Minnesota benchtop digester (UoM bechto AD), and St Cloud WWTF (WW treatment).

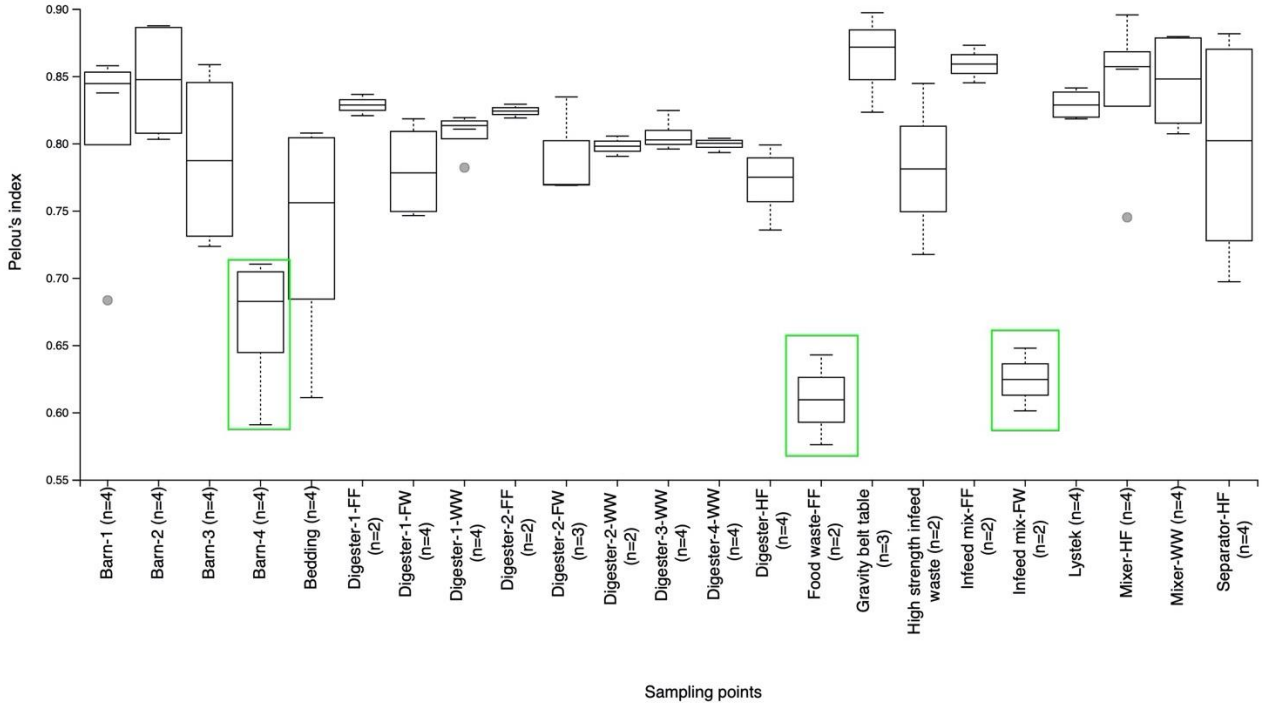


Figure 3.5. Alpha diversity analysis of species evenness (Pielou's diversity) of Haubenschild farm, SCSU food fermenter, University of Minnesota benchtop digester, and St Cloud WWTF 23 sampling points.

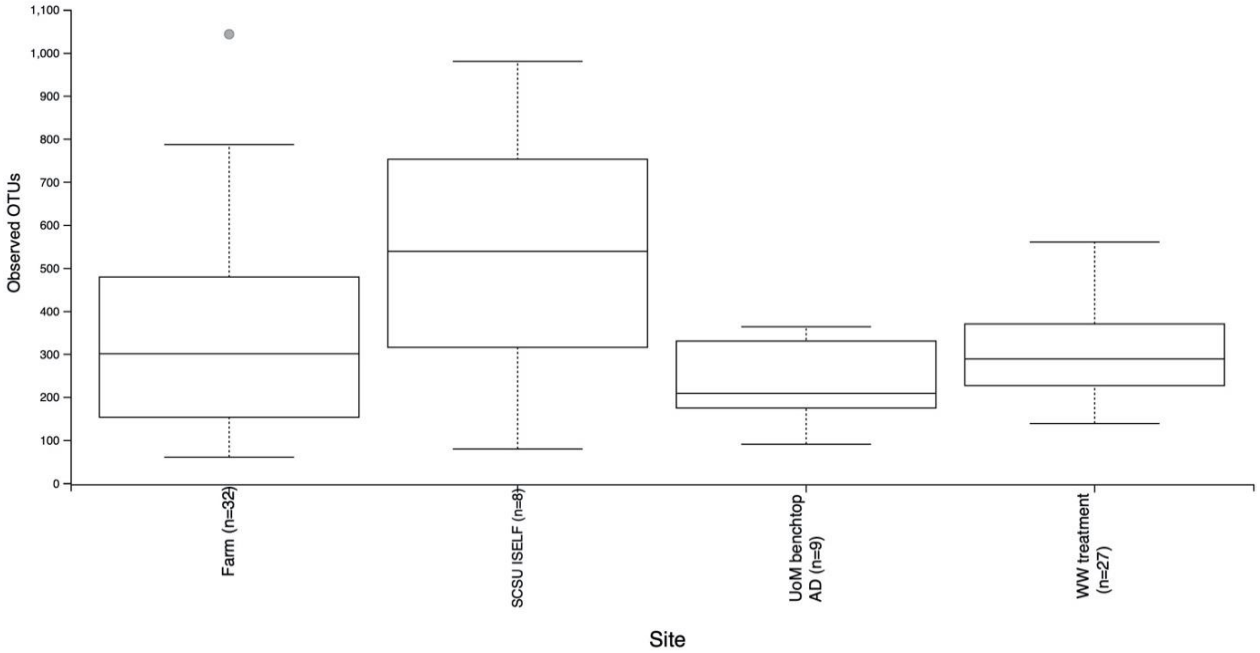
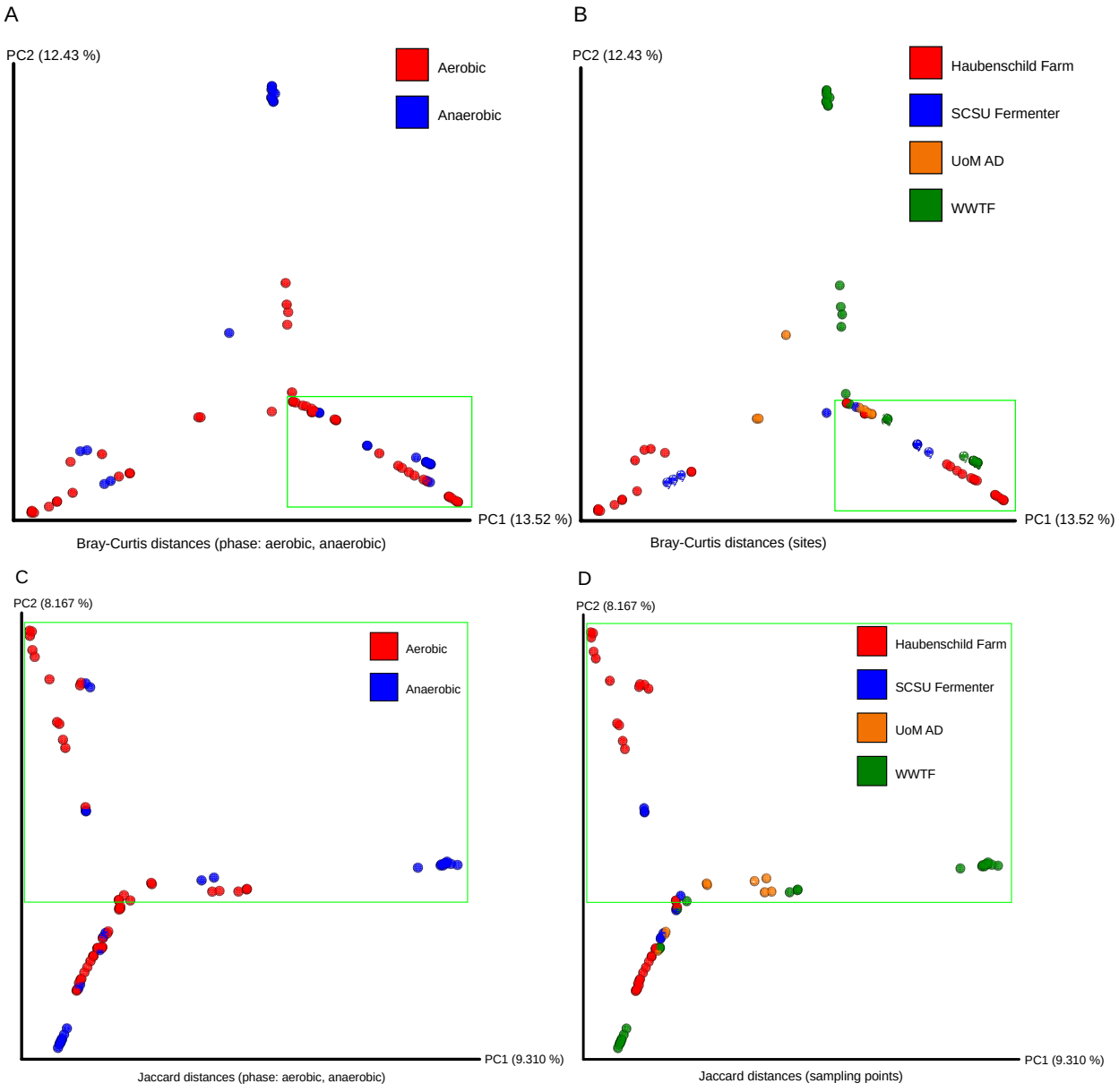


Figure 3.6 Alpha diversity analysis of species observed OTUs of Haubenschild farm, SCSU food fermenter, University of Minnesota benchtop digester, and St Cloud WWTF.

Beta Diversity (83). Beta diversity was used to test for community differences between communities detected by the Arc and Prok primer sets; and between different phases in AD systems (aerobic and anaerobic). This analysis was performed by examining taxonomic distinctiveness between microorganisms in a given taxon and as a result determine an existence of a core community shared between all sampling sites. Figures for beta diversity were visualized on QIIME2 (v 2019.1) view (Figure 3.7) (64). First, Bray-Curtis index (a quantitative measure of dissimilarity between sampling points) was estimated. Bray-Curtis index measured the dissimilarity between samples by summarizing the number of in common species and dividing the sum by the total number of species present in the samples. First, the Arc primer detected that species associated with samples from anaerobic stage at the WWTF were least common at the rest of the sampling sites. Also, the same primer grouped Haubenschild farm samples and SCSU food fermenter very close to each other due to both sites using cow manure at their AD systems. However, the Prok primer located all the sampling sites relatively close to each other. This might be due to the poor ability of the Prok primer to determine Archaea representatives. On the other hand, when calculating Jaccard index (a qualitative measure of dissimilarity between sampling points, that uses presence-absence data) based on the Prok primer results it placed WWTF and Haubenschild farm site furthest from each other and two other sampling sites relatively close. These results are justified by the similarity or dissimilarity of infeed nature used for each individual site. Finally, unweighted Unifrac index that “measures the difference between two collections of sequences as the amount of evolutionary history that is unique to either of the two, which

is measured as a fraction of branch length in phylogenetic tree that leads to descendants of one sample or the other but not both” (87), according to the Prok and Arc primer results WWTF anaerobic samples were located furthest from the rest of sampling sites proving the uniqueness of their core communites. However, unweighted Unifrac index located the rest of sampled communities identified by both primers relatively close to each other including aerobic and anaerobic samples that gives a ground to believe that the core microbial community does exist for these sampling sites.


A
PC2 (12.43 %)
Aerobic
Anaerobic
PC1 (13.52 %)
Bray-Curtis distances (phase: aerobic, anaerobic)
B
PC2 (12.43 %)
Haubenschild Farm
SCSU Fermenter
UoM AD
WWTF
PC1 (13.52 %)
Bray-Curtis distances (sites)
C
PC2 (8.167 %)
Aerobic
Anaerobic
PC1 (9.310 %)
Jaccard distances (phase: aerobic, anaerobic)
D
PC2 (8.167 %)
Haubenschild Farm
SCSU Fermenter
UoM AD
WWTF
PC1 (9.310 %)
Jaccard distances (sampling points)

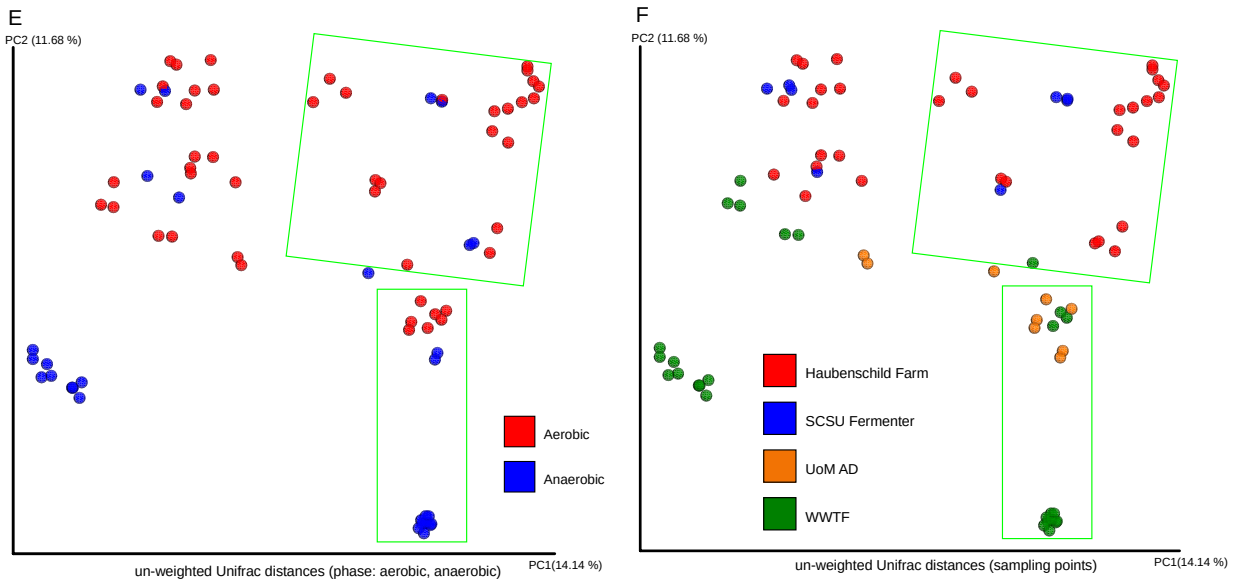


Figure 3.7. Beta diversity analysis of (A) Bray-Curtis, (C) Jaccard, and (E) un-weighted Unifrac distances during anaerobic and aerobic phases in the AD system; (B) Bray-Curtis, (D) Jaccard, (F) un-weighted Unifrac distances between different sampling points at the Haubenschild farm, SCSU food fermenter, University of Minnesota benchtop digester, and St Cloud WWTF sampling sites with the Prok primer points in a green frame and the Arc primer results outside the frame.

Chapter 4: Discussion

This study characterized the microbial community of four AD systems that had different design, operational scale, digesting infeed of different nature, and operating on inoculums of different origin. This study revealed that sampled AD systems have rich diverse microbial communities and also that this richness varies markedly between phases of the AD process (aerobic vs anaerobic) with anaerobic phase had been mainly dominated by representatives associated with Archaea. At the same time, it was revealed a much modest richness of microbial communities during anaerobic phase of the AD process. Furthermore, this study was able to identify highly abundant genera in common for all four sampling sites and are participants in two stages of AD (acidogenesis, acetogenesis).

First, we detected that eight genera were overlapping between SCSU fermenter study and the University of Minnesota benchtop digester study (*Bifobacterium, Prevotella, Lactobacillus, Streptococcus, Clostridium – Clostridiaceae* family, *Clostridium – Ruminococcaceae* family, *Raphanus,* and *Pseudomonas*) (Table 4.1.). This result is most likely due to the similar source of infeed that was used (food waste) (101). However, at the same time, the main methane producing genus - *Methanobrevibacter*, that was shared between three of four sampling sites and is often associated with mammal intestinal tract (102), was not detected in high abundance at the University of Minnesota benchtop digester experiment. This digester was never inoculated with microorganisms associated with mammalian intestinal tract (manure, wastewater, etc.). Alternatively, a *Methanosaeta* genus was a dominant methanogenic genus at the above site, as well as

at the WWTF. The presence of *Methanosaeta* genus both at WWTF and at a benchtop digester site aligns with the previously conducted studies focused on the description of its role in the AD process. Besides, in the previous studies this genus was detected as a dominant genus in ADs with mesophilic conditions (103) and also became a dominant genus in studies characterizing microbiomes of ADs operating on activated sludge (104).

Furthermore, when comparing the most abundant genera (>1%) of all four sampled sites, it was determined that only two genera (*Lactobacillus, Clostridium* from the Ruminococcaceae family) were overlapping in all four AD systems; three genera (*Methanobrevibacter, Bifidobactyerium*, *Clostridium* from the *Clostridiaceae* family) were in common in three AD systems, and nine genera (*Methanosaeta, Petrimonas, Prevotella, Streptococcus, Clostridium from Peptostreptococcacea family, Raphanus, Psychrobacter, Pseudomonas, Treponema*) were shared among two AD systems. Although both genera (*Lactobacillus* and *Clostridium*) present in all four analyzed AD systems, their involvement only in two stages of the AD process (acidogenesis and acetogenesis) was confirmed (105, 106). At the same time, it is fair to hypothesize that representatives of a core microbiome would be involved in all the stages of the AD process. Thus, because the above genera are not involved in hydrolysis, as well as the most crucial stage of the biogas formation - methanogenesis, these 2 genera solely cannot be considered a core microbiome. With that, our research displayed heterogeneity in AD microbial communities, and did not support the possibility to determine a core microbiome for all digesters that differed in a starting inoculum at this time. However, the knowledge of the above genera function and presence in the variety of AD systems can prompt how it can

be used by the operator to improve the stability and productivity of the system's particular phases, namely: acidogenesis and acetogenesis.

Table 4.1

*The overlap in microbial taxa between four AD systems plus 136 digesters from 9 countries (marked with *)*

High abundance genera	Haubenschild farm (# of sequences)	SCSU fermenter study (# of sequences)	St Cloud WWTF (# of sequences)	UoM benchtop AD (# of sequences)
Methanobrevibacter	123,424	30,040	6,721	N/A
Methanosaeta*	N/A	N/A	48,031	2,027
Bifidobacterium	26,004	6633	N/A	3,069
Petrimonas	N/A	N/A	4,104	12,443
Prevotella	N/A	7,628	N/A	19,478
Lactobacillus	16894	31,541	11,231	12,197
Streptococcus	N/A	841	N/A	1,867
Clostridium*	9,406	7,521	N/A	7,431
Peptostreptococcaceae;genus [Clostridium]	8,232	2,164	N/A	N/A
Ruminococcaceae;genus Clostridium	8,407	2,494	1,871	3,592
Raphanus	N/A	3,743	N/A	2,387
Psychrobacter	9921	7,875	N/A	N/A
Pseudomonas	N/A	4,173	N/A	3,464
Treponema*	N/A	8,862	2,363	N/A

We also determined that the overlapping microbial composition was changing in the anaerobic phase of the AD process (Table 4.2). In fact, one genus (*Clostridium* from the *Clostridiaceae* family) that overlaps in all four sampled digesters was confirmed to be a core genus involved in acidogenesis and acetogenesis (107, 108). In addition, there were four genera that overlapped at least in three sampled AD systems (*Methanobrevibacter, Bifidobacterium*, *Treponema*, and genus *Clostridium* from the Ruminicocccaceae family). It is worth mentioning that *Treponema* genus associated with hydrolysis and acetogenesis stages of AD process (109) was also found to be negatively correlated with the volume of methane emissions by Cuncha et al. (110). The same research though confirmed a positive correlation of a dry matter intake associated with

Treponema that makes this genus effective in increasing the volume of matter processed but not the volume of CH_4 produced.

Table 4.2

*The overlap in microbial taxa between four AD systems during anaerobic phase with 136 digesters (marked with *)*

High abundance genera	Haubenschild farm (# of sequences)	SCSU fermenter study (# of sequences)	St Cloud WWTF (# of sequences)	UoM benchtop AD (# of sequences)
Methanobrevibacter	9,380	23,649	1,980	N/A
Peptostreptococcaceae; genus Clostridium	1,930	1,422	N/A	N/A
Clostridium*	1,138	5,690	1,704	2,297
Ruminococcaceae;genus Clostridium	624	1,947	N/A	2,037
Bifidobacterium	553	3,265	N/A	3,002
Treponema*	494	5,755	2,063	N/A
BF311	310	N/A	974	N/A
Sphaerochaeta	295	N/A	912	N/A
Prevotella	N/A	5,175	N/A	19,425
Methanosaeta*	N/A	N/A	37,043	1,073
Petrimonas	N/A	N/A	3,572	4,231
Bacteroides	N/A	N/A	1,016	972

When comparing our results received during the anaerobic phase (Table 4.2.) to the studies that overall involve 136 AD systems, and with some of them supporting a core microbiome hypothesis (49–51, 104, 111, 112), and some that do not find the evidence of the core microbiome existence (47, 48, 113), we detected only three genera to be in common: *Methanosaeta* (49, 51, 104, 112) – methanogenesis (St Cloud WWTF, University of Minnesota benchtop digester)*; Clostridium* associated with *Clostridiaceae* family (49) – acidogenesis, acetogenesis (Haubeschild farm, SCSU fermenter, University of Minnesota benchtop digester)*; and Treponema* (49) – hydrolysis, acetogenesis (SCSU fermenter, St Cloud WWTF) (Table 4.1). However, we again were not able to detect genera that overlaps in all four digesters in our study, as well as studies mentioned above and at the same time involved in all four stages of AD process.

As for *Methanosaeta* genus that was confirmed to be a dominant genus in mesophilic reactors processing food waste or a combination of food waste and manure in the review by Rabii et al. (95), it was not found to be present in the benchtop reactor at SCSU. Although, SCSU benchtop fermenter operated under mesophilic conditions and the infeed consisted of manure and food waste, it was dominated by *Methanobrevibacter* genus, that was not detected as an active methanogen in the above review. At the same time, in the review by Wang et al. (114) *Methanobrevibacter* genus was only associated with thermophilic conditions but not mesophilic. Nevertheless, in our research *Methanobrevibacter* was found both in mesophilic (Haubenschild farm, SCSU fermenter) and thermophilic (St WWTF) conditions. These observations confirm that there is still no complete description of genera involved in AD under different environmental conditions and processing the variety of infeed formulas and might be suggesting uniqueness of microbial compositions of AD systems.

Furthermore, in this study, microbial communities were analyzed in a random time point and the stability of gas production that directly depends on the stability of the microbial community (115), was not considered. Although, both Haubenschild farm and St WWTF were operating for decades and according to the operators the yields of biogas received were stable. At the same time, the University of Minnesota benchtop AD system was operational for less than six months, as well as the SCSU fermenter and for these reactors there were no evidences collected about the stability of the process and the biogas yields. According to the research conducted by Wu et al. (116), time is the critical factor influencing the microbial composition inside the digester, hence the volume and the

quality of the biogas produced. The above research suggests a drastic change in microbial composition when comparing it at the start of the operation and after the first year of operation. A relative abundance of the crucial for biogas production genus *Methanosaeta*, for example, between 289 and 501 days was significantly higher than that between 45 and 167 days. The above example suggests more sampling has to be performed for digesters that were operational for less than a year. In our case, to test the primary hypothesis again, detect the stability of the microbial composition and perform an analysis that would include strictly mature, stable microbial communities (> 1 year) the University of Minnesota benchtop reactor has to resampled at least once in the summer of 2020.

In summary, at this time the results of the research cannot support the primary hypothesis of a universal core microbial community existence in AD systems of different design digesting infeed from different sources. There are, in fact, two major reasons that do not allow us to support the primary hypothesis. Firstly, although we determined the genera that is present in all sampled digesters (*Lactobacillus* and *Clostridium*), these genera are not involved in all four stages of the AD process. Thus these 2 genera cannot be viewed as a universal core microbial community completely responsible for the stability of the AD system. Rather, these two genera can be considered a steady component of the unique core microbial communities of various AD systems (116).

Secondly, our research involved an AD system (University of Minnesota benchtop digester) that was operational for less than a year and an evidence of the stability of its microbial community was not available. This particular AD system needs to be resampled

after a year of operation to perform analysis of sequences that involve only mature, stable digesters, due to the evidence presented above proving the occurrence of significant compositional changes in newly established microbial communities with time.

References

1. Abas N, Kalair A, Khan N. 2015. Review of fossil fuels and future energy technologies. Futures 69:31–49.
2. Shafiee S, Topal E. 2009. When will fossil fuel reserves be diminished? Energy Policy 37:181–189.
3. Center for Sustainable Systems. 2016. U.S. Renewable Energy Factsheet.
4. Shafiee S, Topal E. 2010. A long-term view of worldwide fossil fuel prices. Appl Energy 87:988–1000.
5. Ghosh S. 2011. Examining crude oil price - Exchange rate nexus for India during the period of extreme oil price volatility. Appl Energy 88:1886–1889.
6. Baumeister C, Kilian L. 2016. Forty years of oil price fluctuations: Why the price of oil may still surprise us. J Econ Perspect 30:139–160.
7. Dudley B. 2018. Statistical Review of World Energy 1–56.
8. Owusu PA, Asumadu-Sarkodie S. 2016. A review of renewable energy sources, sustainability issues and climate change mitigation. Cogent Eng 3:1–14.
9. Johnsson F, Kjärstad J, Rootzén J. 2019. The threat to climate change mitigation posed by the abundance of fossil fuels. Clim Policy 19:258–274.
10. Proposers GFOR. 1999. Energy, Environment and Sustainable Development Sub-programme Environment and Sustainable Development. Sustain Dev 64:0–41.
11. Cassia R, Nocioni M, Correa-Aragunde N, Lamattina L. 2018. Climate change and the impact of greenhouse gasses: CO2 and NO, friends and foes of plant oxidative stress. Front Plant Sci 9:1–11.
12. Yang L, Lu F, Zhou X, Wang X, Duan X, Sun B. 2014. Progress in the studies on the greenhouse gas emissions from reservoirs. Acta Ecol Sin.
13. Environmental Protection Agency. 2013. Inventory of U.S. Greenhouse Gas Emissions and Sinks 1990-2017 53.
14. Anderson TR, Hawkins E, Jones PD. 2016. CO2, the greenhouse effect and global warming: from the pioneering work of Arrhenius and Callendar to today's Earth System Models. Endeavour.
15. Metz B, Meyer L, Bosch P. 2007. Climate change 2007 mitigation of climate changeClimate Change 2007 Mitigation of Climate Change.
16. Cellura M, Cusenza MA, Longo S. 2018. Energy-related GHG emissions balances: IPCC versus LCA. Sci Total Environ 628–629:1328–1339.
17. Raupach MR, Marland G, Ciais P, Le Quéré C, Canadell JG, Klepper G, Field CB. 2007. Global and regional drivers of accelerating CO2 emissions. Proc Natl Acad Sci U S A 104:10288–10293.
18. Etkin DS. 2005. Analysis of oil spill trends in the united states and worldwide. 2005 Int Oil Spill Conf IOSC 2005 293–302.
19. Shifei Kang, Mengya Chen, Yuting Wang, Fan Tang LC and, Dong M. 2019. Easy Synthesis of Light-weight Hierarchical Mesoporous Carbon Sponges for Efficient Oil Spill Recovery 0–20.
20. O'Rourke D, Connolly S. 2003. Just Oil? The Distribution of Environmental and Social Impacts of Oil Production and Consumption. Annu Rev Environ Resour

28:587–617.
21. Vikas M, Dwarakish GS. 2015. Coastal Pollution: A Review. Aquat Procedia 4:381–388.
22. Whittaker A, BéruBé K, Jones T, Maynard R, Richards R. 2004. Killer smog of London, 50 years on: Particle properties and oxidative capacity. Sci Total Environ 334–335:435–445.
23. Amarsaikhan D, Battsengel V, Nergui B, Ganzorig M, Bolor G. 2014. A Study on Air Pollution in Ulaanbaatar City, Mongolia. J Geosci Environ Prot 02:123–128.
24. World Bank, Institute for Health Metrics and Evaluation. 2016. The Cost of Air Pollution: Strengthening the Economic Case for Action. World Bank Inst Heal Metrics Eval Univ Washington, Seattle 122.
25. Guttikunda SK, Lodoysamba S, Bulgansaikhan B, Dashdondog B. 2013. Particulate pollution in Ulaanbaatar, Mongolia. Air Qual Atmos Heal 6:589–601.
26. Castellano-Hinojosa A, Armato C, Pozo C, González-Martínez A, González-López J. 2018. New concepts in anaerobic digestion processes: recent advances and biological aspects. Appl Microbiol Biotechnol 102:5065–5076.
27. Ren Y, Yu M, Wu C, Wang Q, Gao M, Huang Q, Liu Y. 2018. A comprehensive review on food waste anaerobic digestion: Research updates and tendencies. Bioresour Technol 247:1069–1076.
28. Moya D, Aldás C, López G, Kaparaju P. 2017. Municipal solid waste as a valuable renewable energy resource: A worldwide opportunity of energy recovery by using Waste-To-Energy Technologies. Energy Procedia 134:286–295.
29. Komusanac I, Fraile D, Brindley G. 2019. Wind energy in Europe in 2018. Trends and Statistics.
30. Kabir E, Kumar P, Kumar S, Adelodun AA, Kim KH. 2018. Solar energy: Potential and future prospects. Renew Sustain Energy Rev 82:894–900.
31. Chu Y. 2011. Review and Comparison of Different Solar Technologies.
32. Debruyn J, Hilborn D. 2007. Anaerobic Digestion Basics. Small 2007:1–6.
33. Cecchi F, Cavinato C. 2015. Anaerobic digestion of bio-waste: A mini-review focusing on territorial and environmental aspects. Waste Manag Res 33:429–438.
34. Steinfeld H. GPWTDCVRMHCD et al. 2006. Livestock's long shadow : environmental issues and options / Henning Steinfeld ... [et al.].
35. Liu Y, Ni Z, Kong X, Liu J. 2017. Greenhouse gas emissions from municipal solid waste with a high organic fraction under different management scenarios. J Clean Prod 147:451–457.
36. Dong YH, An AK, Yan YS, Yi S. 2017. Hong Kong's greenhouse gas emissions from the waste sector and its projected changes by integrated waste management facilities. J Clean Prod 149:690–700.
37. Molino A, Larocca V, Chianese S, Musmarra D. 2018. Biofuels production by biomass gasification: A review. Energies 11.
38. Carl N, Lamb J. 2002. Final Report: Haubenschild Farms Anaerobic Digester 35.
39. Liebetrau J, Sträuber H, Kretzschmar J, Denysenko V, Nelles M. 2019. Anaerobic digestionAdvances in Biochemical Engineering/Biotechnology.
40. O'Callaghan K. 2016. Technologies for the utilisation of biogenic waste in the

bioeconomy. Food Chem 198:2–11.
41. Town JR, Links MG, Fonstad TA, Dumonceaux TJ. 2014. Molecular characterization of anaerobic digester microbial communities identifies microorganisms that correlate to reactor performance. Bioresour Technol 151:249–257.
42. Clarke WP. 2018. The uptake of anaerobic digestion for the organic fraction of municipal solid waste – Push versus pull factors. Bioresour Technol 249:1040–1043.
43. Nelson MC, Morrison M, Yu Z. 2011. A meta-analysis of the microbial diversity observed in anaerobic digesters. Bioresour Technol 102:3730–3739.
44. Cotana F, Cavalaglio G, Petrozzi A, Coccia V. 2015. Lignocellulosic biomass feeding in biogas pathway: State of the art and plant layouts. Energy Procedia 81:1231–1237.
45. Sousa DZ, Salvador AF, Ramos J, Guedes AP, Barbosa S, Stams AJM, Alves MM, Pereira MA. 2013. Activity and viability of methanogens in anaerobic digestion of unsaturated and saturated long-chain fatty acids. Appl Environ Microbiol 79:4239–4245.
46. Ward AJ, Hobbs PJ, Holliman PJ, Jones DL. 2008. Optimisation of the anaerobic digestion of agricultural resources. Bioresour Technol.
47. Calusinska M, Goux X, Fossépré M, Muller EEL, Wilmes P, Delfosse P. 2018. 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. Biotechnol Biofuels 11:1–19.
48. Mei R, Nobu MK, Narihiro T, Kuroda K, Muñoz Sierra J, Wu Z, Ye L, Lee PKH, Lee PH, van Lier JB, McInerney MJ, Kamagata Y, Liu WT. 2017. Operation-driven heterogeneity and overlooked feed-associated populations in global anaerobic digester microbiome. Water Res 124:77–84.
49. Peces M, Astals S, Jensen PD, Clarke WP. 2018. Deterministic mechanisms define the long-term anaerobic digestion microbiome and its functionality regardless of the initial microbial community. Water Res 141:366–376.
50. Treu L, Kougias PG, Campanaro S, Bassani I, Angelidaki I. 2016. Deeper insight into the structure of the anaerobic digestion microbial community; The biogas microbiome database is expanded with 157 new genomes. Bioresour Technol.
51. Lucas R, Kuchenbuch A, Fetzer I, Harms H, Kleinsteuber S. 2015. Long-term monitoring reveals stable and remarkably similar microbial communities in parallel full-scale biogas reactors digesting energy crops. FEMS Microbiol Ecol 91:1–11.
52. Tang YQ, Shigematsu T, Morimura S, Kida K. 2015. Dynamics of the microbial community during continuous methane fermentation in continuously stirred tank reactors. J Biosci Bioeng 119:375–383.
53. Goux X, Calusinska M, Lemaigre S, Marynowska M, Klocke M, Udelhoven T, Benizri E, Delfosse P. 2015. Microbial community dynamics in replicate anaerobic digesters exposed sequentially to increasing organic loading rate, acidosis, and process recovery. Biotechnol Biofuels 8:1–18.
54. de Jonge N, Moset V, Møller HB, Nielsen JL. 2017. Microbial population dynamics

in continuous anaerobic digester systems during start up, stable conditions and recovery after starvation. Bioresour Technol.

55. Lee J, Shin SG, Han G, Koo T, Hwang S. 2017. Bacteria and archaea communities in full-scale thermophilic and mesophilic anaerobic digesters treating food wastewater: Key process parameters and microbial indicators of process instability. Bioresour Technol 245:689–697.
56. Li YF, Nelson MC, Chen PH, Graf J, Li Y, Yu Z. 2015. Comparison of the microbial communities in solid-state anaerobic digestion (SS-AD) reactors operated at mesophilic and thermophilic temperatures. Appl Microbiol Biotechnol 99:969–980.
57. Amha YM, Anwar MZ, Brower A, Jacobsen CS, Stadler LB, Webster TM, Smith AL. 2018. Inhibition of anaerobic digestion processes: Applications of molecular tools. Bioresour Technol.
58. Ward AJ, Hobbs PJ, Holliman PJ, Jones DL. 2008. Optimisation of the anaerobic digestion of agricultural resources. Bioresour Technol 99:7928–7940.
59. Wirth R, Böjti T, Lakatos G, Maróti G, Bagi Z, Rákhely G, Kovács KL. 2019. Characterization of Core Microbiomes and Functional Profiles of Mesophilic Anaerobic Digesters Fed With Chlorella vulgaris Green Microalgae and Maize Silage. Front Energy Res 7:1–18.
60. Luo G, Fotidis IA, 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 9:1–12.
61. Case RJ, Boucher Y, Dahllöf I, Holmström C, Doolittle WF, Kjelleberg S. 2007. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. Appl Environ Microbiol 73:278–288.
62. Plummer E, Twin J. 2015. A Comparison of Three Bioinformatics Pipelines for the Analysis of Preterm Gut Microbiota using 16S rRNA Gene Sequencing Data. J Proteomics Bioinform 8:283–291.
63. Chachkhiani M, Dabert P, Abzianidze T, Partskhaladze G, Tsiklauri L, Dudauri T, Godon JJ. 2004. 16S rDNA characterisation of bacterial and archaeal communities during start-up of anaerobic thermophilic digestion of cattle manure. Bioresour Technol 93:227–232.
64. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang K Bin, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik A V, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N,

Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857.

65. Wright P. 2001. Overview of Anaerobic Digestion Systems for Dairy Farms. Nat Resour Agric Eng Serv 15.
66. Lazarus WF, Rudstrom M. 2007. The economics of anaerobic digester operation on a Minnesota dairy farm. Rev Agric Econ 29:349–364.
67. Mahmoudi N, Slater GF, Fulthorpe RR. 2011. Comparison of commercial DNA extraction kits for isolation and purification of bacterial and eukaryotic DNA from PAH-contaminated soils. Can J Microbiol 57:623–628.
68. ThermoScientific. 2011. NanoDrop: Assessment of Nucleic Acid Purity. Protoc Prod Manuals 1–2.
69. Ahn SJ, Costa J, Emanuel JR. 1996. PicoGreen quantitation of DNA: Effective evaluation of samples pre- or post-PCR. Nucleic Acids Res 24:2623–2625.
70. Raymann K, Moeller AH, Goodman AL, Ochman H. 2017. Unexplored Archaeal Diversity in the Great Ape Gut Microbiome. mSphere 2:1–12.
71. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. Appl Environ Microbiol 79:5112–5120.
72. Cock PJA, Fields CJ, Goto N, Heuer ML, Rice PM. 2009. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. Nucleic Acids Res 38:1767–1771.
73. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583.
74. McDonald D, Price MN, Goodrich J, Nawrocki EP, Desantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 6:610–618.
75. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72:5069–5072.
76. Katoh K, Rozewicki J, Yamada KD. 2018. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160–1166.
77. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 30:772–780.
78. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 - Approximately maximum-likelihood trees for large alignments. PLoS One 5.

79. Phylogenetic T, Methods R, Shallow R, Relationships E, Well E. 2000. Traditional Phylogenetic Reconstruction Methods Reconstruct Shallow and Deep Evolutionary Relationships Equally Well 1823–1827.
80. Rasmussen MD, Kellis M. 2011. A Bayesian approach for fast and accurate gene tree reconstruction. Mol Biol Evol 28:273–290.
81. Dickman M. 2019. Some Indices of Diversity 49:1191–1193.
82. Faith DP, Baker AM. 2006. Phylogenetic Diversity (PD) and Biodiversity Conservation: Some Bioinformatics Challenges. Evol Bioinforma 2:117693430600200.
83. Ricotta C. 2017. Of beta diversity, variance, evenness, and dissimilarity. Ecol Evol 7:4835–4843.
84. Smith B, Wilson JB, A JB. 2016. Nordic Society Oikos A Consumer ' s Guide to Evenness Indices Author (s): Benjamin Smith and J . Bastow Wilson Published by : Wiley on behalf of Nordic Society Oikos Stable URL : http://www.jstor.org/stable/3545749 Accessed : 31-03-2016 18 : 44 UTC Your 76:70–82.
85. Baselga A, Leprieur F. 2015. Comparing methods to separate components of beta diversity. Methods Ecol Evol 6:1069–1079.
86. Society E. 2019. A Quantitative Evaluation of the Bray-Curtis Ordination Author (s): Hugh G . Gauch , Jr . Published by : Wiley on behalf of the Ecological Society of America Stable URL : https://www.jstor.org/stable/1935677 REFERENCES Linked references are available on 54:829–836.
87. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. 2011. UniFrac: An effective distance metric for microbial community comparison. ISME J 5:169–172.
88. Galperin MY, Brover V, Tolstoy I, Yutin N. 2016. Phylogenomic analysis of the family peptostreptococcaceae (Clostridium cluster xi) and proposal for reclassification of Clostridium litorale (Fendrich et al. 1991) and Eubacterium acidaminophilum (Zindel et al. 1989) as peptoclostridium litorale gen. nov. Int J Syst Evol Microbiol 66:5506–5513.
89. Vázquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. 2013. EMPeror: A tool for visualizing high-throughput microbial community data. Gigascience 2:2–5.
90. Kulkarni P, Olson ND, Paulson JN, Pop M, Maddox C, Claye E, Rosenberg Goldstein RE, Sharma M, Gibbs SG, Mongodin EF, Sapkota AR. 2018. Conventional wastewater treatment and reuse site practices modify bacterial community structure but do not eliminate some opportunistic pathogens in reclaimed water. Sci Total Environ.
91. Lozano S, Garrett MJ, Zhang Y, Mosoke EV, Steinfeld C, Del Porto D, Code S, Maksimova S, Kosaurova D, Pesheva A, Gabiana CP, Gomes K, Ebrary I, Qian Y, Sudasinghe MI, Galagedara LW, Gunawardena ERN, Junge-berberovic R, Graber A, Town FN, Bradbury MA, Lukoyaa NB. 2014. Primer for Municipal Wastewater Treatment Systems. Master thesis 22:55.
92. Wei Y, Van Houten RT, Borger AR, Eikelboom DH, Fan Y. 2003. Minimization of excess sludge production for biological wastewater treatment. Water Res.
93. Li C, Champagne P, Anderson BC. 2015. Enhanced biogas production from

anaerobic co-digestion of municipal wastewater treatment sludge and fat, oil and grease (FOG) by a modified two-stage thermophilic digester system with selected thermo-chemical pre-treatment. Renew Energy.
94. Murto M, Björnsson L, Mattiasson B. 2004. Impact of food industrial waste on anaerobic co-digestion of sewage sludge and pig manure. J Environ Manage.
95. Rabii A, Aldin S, Dahman Y, Elbeshbishy E. 2019. A review on anaerobic co-digestion with a focus on the microbial populations and the effect of multi-stage digester configuration. Energies 12.
96. Elbeshbishy E, Aldin S, Nakhla G, Singh A, Mullin B. 2014. Impact of alkaline-hydrolyzed biosolids (Lystek) addition on the anaerobic digestibility of TWAS in lab - And full-scale anaerobic digesters. Waste Manag.
97. Esquivel-Elizondo S, Parameswaran P, Delgado AG, Maldonado J, Rittmann BE, Krajmalnik-Brown R. 2016. Archaea and Bacteria Acclimate to High Total Ammonia in a Methanogenic Reactor Treating Swine Waste. Archaea 2016.
98. Niu Q, Kobayashi T, Takemura Y, Kubota K, Li YY. 2015. Evaluation of functional microbial community's difference in full-scale and lab-scale anaerobic digesters feeding with different organic solid waste: Effects of substrate and operation factors. Bioresour Technol.
99. Cho SK, Im WT, Kim DH, Kim MH, Shin HS, Oh SE. 2013. Dry anaerobic digestion of food waste under mesophilic conditions: Performance and methanogenic community analysis. Bioresour Technol.
100. Guo X, Wang C, Sun F, Zhu W, Wu W. 2014. A comparison of microbial characteristics between the thermophilic and mesophilic anaerobic digesters exposed to elevated food waste loadings. Bioresour Technol.
101. Nelson MC, Morrison M, Schanbacher F, Yu Z. 2012. Shifts in microbial community structure of granular and liquid biomass in response to changes to infeed and digester design in anaerobic digesters receiving food-processing wastes. Bioresour Technol 107:135–143.
102. Enzmann F, Mayer F, Rother M, Holtmann D. 2018. Methanogens: biochemical background and biotechnological applications. AMB Express 8:1–22.
103. Yu D, Kurola JM, Lähde K, Kymäläinen M, Sinkkonen A, Romantschuk M. 2014. Biogas production and methanogenic archaeal community in mesophilic and thermophilic anaerobic co-digestion processes. J Environ Manage 143:54.
104. De Vrieze J, Christiaens MER, Walraedt D, Devooght A, Ijaz UZ, Boon N. 2017. Microbial community redundancy in anaerobic digestion drives process recovery after salinity exposure. Water Res 111:109–117.
105. Karthikeyan OP, Selvam A, Wong JWC. 2016. Hydrolysis-acidogenesis of food waste in solid-liquid-separating continuous stirred tank reactor (SLS-CSTR) for volatile organic acid production. Bioresour Technol.
106. Tzun-Wen Shaw G, Liu AC, Weng CY, Chou CY, Wang D. 2017. Inferring microbial interactions in thermophilic and mesophilic anaerobic digestion of HOG waste. PLoS One 12:1–22.
107. Romanazzi V, Bonetta S, Fornasero S, De Ceglia M, Gilli G, Traversi D. 2016. Assessing Methanobrevibacter smithii and Clostridium difficile as not conventional

faecal indicators in effluents of a wastewater treatment plant integrated with sludge anaerobic digestion. J Environ Manage 184:170–177.
108. Park C, Lee C, Kim S, Chen Y, Chase HA. 2005. Upgrading of anaerobic digestion by incorporating two different hydrolysis processes. J Biosci Bioeng 100:164–167.
109. Guo J, Peng Y, Ni BJ, Han X, Fan L, Yuan Z. 2015. 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 14:1–11.
110. Cunha CS, Veloso CM, Marcondes MI, Mantovani HC, Tomich TR, Pereira LGR, Ferreira MFL, Dill-McFarland KA, Suen G. 2017. Assessing the impact of rumen microbial communities on methane emissions and production traits in Holstein cows in a tropical climate. Syst Appl Microbiol.
111. GuangLiang T, WuDi Z, MingHua D, Bin Y, Rui Z, Fang Y, XingLing Z, YongXia W, Wei X, Qiang W, XiaoLong C. 2017. Metabolic pathway analysis based on high-throughput sequencing in a batch biogas production process. Energy (Oxford) 139:571–579.
112. Rivière D, Desvignes V, Pelletier E, Chaussonnerie S, Guermazi S, Weissenbach J, Li T, Camacho P, Sghir A. 2009. Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. ISME J 3:700–714.
113. Theuerl S, Klang J, Heiermann M, De Vrieze J. 2018. Marker microbiome clusters are determined by operational parameters and specific key taxa combinations in anaerobic digestion. Bioresour Technol 263:128–135.
114. Wang P, Wang H, Qiu Y, Ren L, Jiang B. 2018. Microbial characteristics in anaerobic digestion process of food waste for methane production–A review. Bioresour Technol.
115. Shi X, Guo X, Zuo J, Wang Y, Zhang M. 2018. A comparative study of thermophilic and mesophilic anaerobic co-digestion of food waste and wheat straw: Process stability and microbial community structure shifts. Waste Manag 75:261–269.
116. Wu L, Yang Y, Chen S, Zhao M, Zhu Z, Yang S, Qu Y, Ma Q, He Z, Zhou J, He Q. 2016. Long-term successional dynamics of microbial association networks in anaerobic digestion processes. Water Res 104:1–10.