CHARACTERISATION OF ANAEROBIC DIGESTER MICROBIOME IN VARIOUS OPERATING CONDITIONS USING NEXT-GENERATION AND THIRD-GENERATION SEQUENCING TECHNOLOGIES

YU-CHUNG LAM

MPhil

The Hong Kong Polytechnic University

2020

The Hong Kong Polytechnic University

Department of Civil and Environmental Engineering

Characterisation of anaerobic digester microbiome in various operating conditions using next-generation and third-generation sequencing technologies

Yu-Chung Lam

A Thesis Submitted in Partial Fulfilment of the Requirement for the degree of Master of Philosophy

Sep 2019

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(Signed)

_____Lam Yu Chung_____(Name of Student)

ABSTRACT

Anaerobic digestion is a common wastewater treatment technology that alleviates pressure on sludge disposal and generates energy via combustion of biogas. Diverse groups of microorganisms are responsible for the biodegradation, To manipulate the process better, we study the anaerobic digestion process through the lens of genomic analysis.

In the past decade, the characterisation of the microbial community in anaerobic digestion was primarily achieved by using the next-generation sequencing technologies. However, the next-generation short-read approach has inherent primer bias and low phylogenetic resolution. A more accurate phylogenetic identification is desired. In this study, we employed the next generation sequencing offered by Illumina and the third-generation sequencing technology offered by Pacific Bioscience to characterise microbiome in anaerobic digesters operated under different conditions.

From sample preparation to data analysis, this work serves as a reference to utilise third-generation sequencing for 16S survey. The high-quality circular consensus sequences provided a superior resolution for diversity profiling. Unexpectedly, *Methanothrix concilii* and *Smithella* were found to have universal presence in different operating conditions. Unique methanogenic and syntrophic species (*Methanosarcina horonobensis, Methanosarcina flavescens, Thermanaerovibrio acidaminovorans, Ca.* Cloacimonas acidaminovorans, etc.) were observed in divergent operating conditions. The species-level information facilitated our understanding of their survival niches and how they become dominant in respect of the operating conditions. We also compared the two sequencing technologies in terms of taxonomic classification and diversity indices. The results reflected the use of different technology might lead to inconsistent interpretation of the microbiome. Last but not least, shotgun metagenomics sequencing and *mcrA* gene amplicon sequencing were performed on lab-scale microaerated fermenters. The facultative bacteria promote hydrolysis by switching to more thermodynamic favourable pathways under microaerated condition, ultimately facilitating methane production.

PUBLICATION ARISING FROM THE THESIS

Journal Papers:

Lam, T.Y.C., Mei, R., Wu, Z., Lee, P.K.H., Liu, W.T., Lee, P.H., 2020. Superior resolution characterisation of microbial diversity in anaerobic digesters using full-length 16S rRNA gene amplicon sequencing. Water Research.

ACKNOWLEDGEMENTS

First and foremost, I owe my most enormous gratitude to Dr Po-Heng (Henry) LEE for his continuous encouragement and enlightenment throughout the years and Dr Yi JIANG for his unconditional support in my time of need.

I have been working under the guidance of Dr LEE since my sophomore year. In that summer, my fellow students and I reached to Dr LEE, hoping to get a sense of direction to our future career. Dr LEE kindly offered an opportunity for me to work his project and experience science. A few years later, I graduated from an MPhil programme and have a profound understanding of what it means to be a researcher. From academic prowess to the philosophy of life, I learnt a lot from Dr LEE.

Thank you, Dr Zhuoying WU and Dr Giin Yu Amy TAN, for opening the window of microbiology to a naive engineering student. It was challenging to step into a new domain, and you made it easy.

Thank you, Prof Wen-Tso LIU and Ran MEI from the University of Illinois Urbana-Campaign, and Dr Duc NGUYENG from the University of Hawai‘i at Mānoa for the collaboration. Your expertise is invaluable.

I would like to thank Dr Hang LIU, Huichuan ZHUANG, Ling LENG, and Xiaoming JI for advising on bioinformatics issues; Mr WS LAM and Celine CHE for the technical (and non-technical) support in the lab; Jonathan Yuen-Him CHAN for the moral support and being an excellent lunch buddy. It has been a blessing to grow in such a nourishing environment.

A special thanks to Dr Jiachi (Amber) CHIOU from the department of ABCT, for introducing me to the world of the gut microbiome and the opportunity to visit the Microbiome Centre in the University of Chicago.

And, all my love to my friends and family for their support and patience.

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AD	anaerobic digestion
BMP	bio-methane potential
COD	chemical oxygen demand
CCS	circular consensus sequence
CRT	cyclic reversible termination
CSTR	continuous stirred-tank reactor
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA
dNTPs	deoxynucleotides
FL	full-length
HRT	hydraulic retention time
LDA	linear discriminant analysis
LEfSe	LDA effect size
mcrA	Methyl coenzyme M reductase
NCBI	National Centre for Biotechnology Information
NGS	next-generation sequencing
ORP	oxidation-reduction potential
OTU	operational taxonomic unit
PCR	polymerase chain reaction
rRNA	ribosomal ribonucleic acid
SAS	secondary activated sludge
SR	short-read
ssDNA	single-strand DNA
TGS	third-generation sequencing
TS	total solids
VFA	volatile fatty acid
VVS	volatile suspended solids

Chapter 1 Introduction

1.1 The roles of anaerobic digestion

Anaerobic digestion (AD) refers to the fermentation process that organic matters are degraded and converted into biogas. The occurrence of AD is ubiquitous, and it plays a quantitatively important role in the global carbon cycle. It can happen in the digestive system of ruminants, municipal landfills and sewers, also in natural anoxic environments such as freshwater sediments and swamps. AD is a good strategy in wastewater treatment as it significantly diminishes the organic content in sewage sludge, the most abundant wastewater treatment by-product. It subsequently reduces the pressure on sludge disposal and incineration and allows resource recovery with application like fertiliser. In comparison with energy-consuming aerobic treatment, the operating cost of anaerobic treatment is lower. Also, the biogas produced during the AD process can be used for energy generation as it mainly consists of combustible methane gas (Fig 1.1). Therefore, AD has been widely adopted as an economical option for wastewater treatment (Appels et al., 2011; Vasco-Correa et al., 2018).

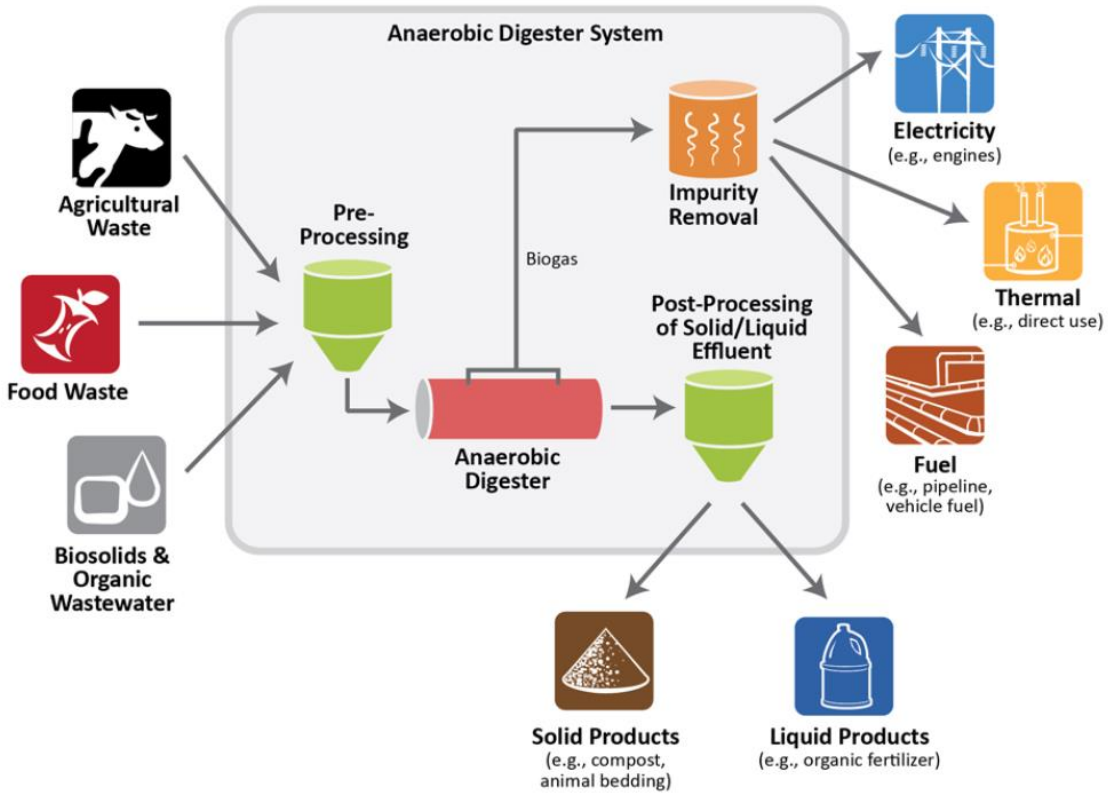


Figure 1.1 Applications of anaerobic digestion (GMI, 2013)

The growth of population and economic activities lead to immense stress on the energy demand and sanitation at a global scale. The United Nation predicts the level of urbanisation will continue to grow, and approximately 68% of the world population will live in congregated urban area by 2050 (DESA/UN-WUP, 2018). To combat the potential energy crisis and environmental issues, AD serves as a promising technology to alleviate the pressure and should be recognised as one of the long-term strategies.

1.2 Limitations and opportunities

Despite the advantages of AD, its slow reaction and instability under fluctuating condition are not desirable. It limits the application of AD as an option for electricity generation. Therefore, an improvement in biogas production performance is desired to increase its treatment capacity further. There are several research directions in improving the efficiency of AD process, including operation optimisation, reactor design, influent pre-treatment, and the addition of trace elements (Mao et al., 2015). The approaches mentioned above aim at providing a better environment or simpler substrate for the AD microbial community to thrive, where their effectiveness is often evaluated by the traditional monitoring indicators of the system performance, including chemical oxygen demand (COD) removal rate, volatile suspended solids removal rate, biogas composition, and methane yield. However, these indicators are the phenotypes of underlying biological functions. The success of AD process relies on the balance between the complicated microbial interaction of the microorganisms, while it is challenging to study their intricacy via classical microbiology methods.

In recent years, the improvement of the efficiency and the reducing cost of the sequencing technologies allows researchers in the field of environmental engineering to adopt a bioinformatics approach to decipher the genomic code of mixed culture samples. It helps to elucidate the complexity of the AD microbiome and improves our understanding of the AD metabolic pathways (Cai et al., 2016). In this study, we employed two sequencing technologies, the next-generation sequencing (NGS)

and the third-generation sequencing (TGS), to investigate the AD microbiome in various operating conditions. This study lays down the foundation of how to manipulate the AD process for more stable operation, more efficient organic solids removal and methane production, by providing genomic insight on AD communities.

1.3 Aims and objectives

This study aims to characterise the core AD microbiome under different operating conditions using bioinformatics approaches and evaluate the difference between NGS and TGS.

1. Develop a pipeline for full-length 16S rRNA amplicon sequencing from sample preparation to data analysis;
2. Evaluate the application potential of the full-length 16S rRNA amplicon sequencing for species-level identification of AD microbiome;
3. Evaluate the discrepancy between the NGS and TGS on taxonomic classification, diversity indices, and subsequent effect(s) on understanding AD process;
4. Deepen our understanding of the effect of microaeration on AD microbiome via shotgun metagenomic sequencing and *mcrA* gene amplicon sequencing.

1.4 Scope of the research

This study uses PacBio Sequel single-molecule real-time (SMRT) sequencing as a representation of TGS and Illumina MiSeq/HiSeq high-throughput sequencing by synthesis as a representation of NGS. The expected outcome of this study is to serve as a reference for utilising full-length 16S rRNA amplicon sequencing from sample preparation to bioinformatics analysis and providing a characterisation of AD microbiomes in coordination with their operating conditions. The AD microbiomes characterised in chapter 4 and 5 are from full-scale anaerobic digesters, and the microaerated microbiome is from lab-scale fermenters in University of Hawai‘i at Mānoa. Fig. 1.2 illustrates the project overview.

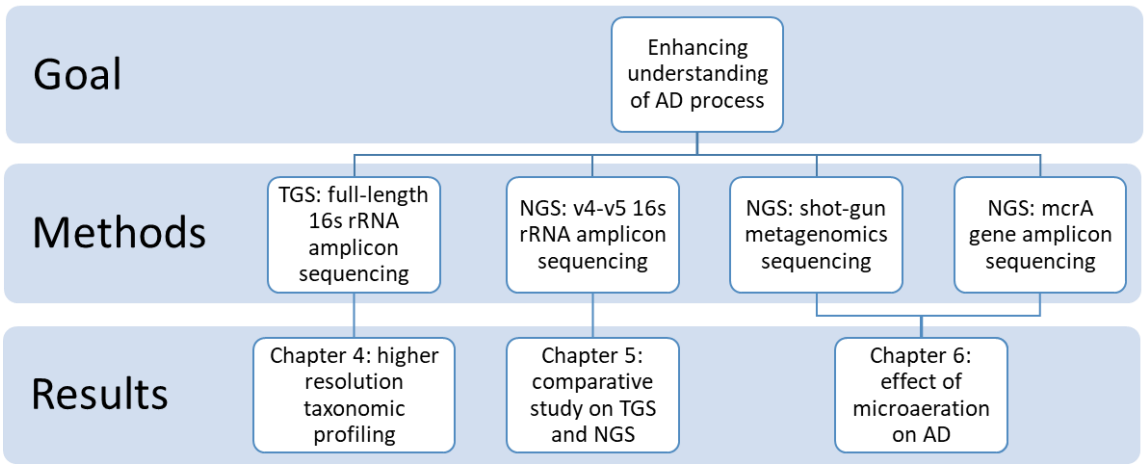


Figure 1.2 Project overview with methods corresponding to the research results

1.5 Organisation of the Thesis

This thesis is composed of seven chapters. Chapter one includes the background, motivation, objectives, scope, and the structure of this thesis.

Chapter two provides a comprehensive literature review on the important biological processes of anaerobic digestion and information regarding the sequencing technologies employed in this study.

Chapter three described the experiment flows and the analytical methods used in this study.

Chapter four provides the result and discussion regarding higher phylogenetic resolution profiling of the global anaerobic digestion microbiome using full-length 16S rRNA amplicon sequencing. Microbiome clustering, co-occurrence network analysis, phylogenetic tree construction, and species level identification were performed.

In chapter five, the discrepancies between the full-length approach (TGS) and the short-read approach (NGS) are evaluated by their difference in taxonomic classification and diversity indices. Their impact on performance prediction is illustrated.

Chapter six provides shotgun metagenomics and *mcrA* gene analysis on microaerated samples. Attention was paid to the facultative bacteria and how they facilitate the hydrolysis and fermentation process.

In the last chapter, we summarise the conclusions of each chapter and provide directions to future study.

Chapter 2 Literature Review

2.1 Anaerobic digestion

2.1.1 Major biochemical process

AD is a biodegradation process involving various groups of archaea and bacteria. The overall conversion of complex organic matter to methane in the AD process is most commonly described in four stages, namely: hydrolysis, acidogenesis/fermentation, acetogenesis and methanogenesis. Figure 2.1 illustrates a simplified schematic view of the major metabolic pathways in AD.

Different microbial groups have different capacity on the metabolism, microbes with larger genome size are more versatile in substrate utilisation. Conventionally, the microbial communities in AD are categorised into hydrolytic bacteria, fermentative bacteria (fermenters), syntrophic bacteria (syntrophs) and methanogenic archaea (methanogens). Hydrolytic bacteria hydrolyse complex organic compounds into monosaccharides, amino acids, higher fatty acids and alcohols. Fermenters further breakdown the substrates to volatile fatty acids (VFAs), also known as short-chain fatty acids. Methanogen species utilise a range of the intermediate electron shuttle (e.g. acetate, formate and hydrogen) produced by fermenters and syntrophs from VFAs. The syntrophic association between the methanogens and syntrophs is imperative in AD, as the intermediate electron shuttles need to be maintained at low substrate levels for a thermodynamically favourable environment (Thauer et al., 2008; Sieber et al., 2010). It should be noted that the grouping of the microbes is not definitive as microbes behave differently and adopt different metabolic pathways under different environment. For instance, the BCR-I group can reduce complex organic compounds to acetate and hydrogen (Kadnikov et al., 2019).

There is no universal definition for syntrophs. Herein, syntrophs are defined as bacteria having a syntrophic association with hydrogenotrophic and acetoclastic methanogens to maintain a

thermodynamically favourable substrate concentration (low concentration of hydrogen and acetate) for specific metabolism to occur (e.g. propionate oxidation) (Sieber et al., 2010).

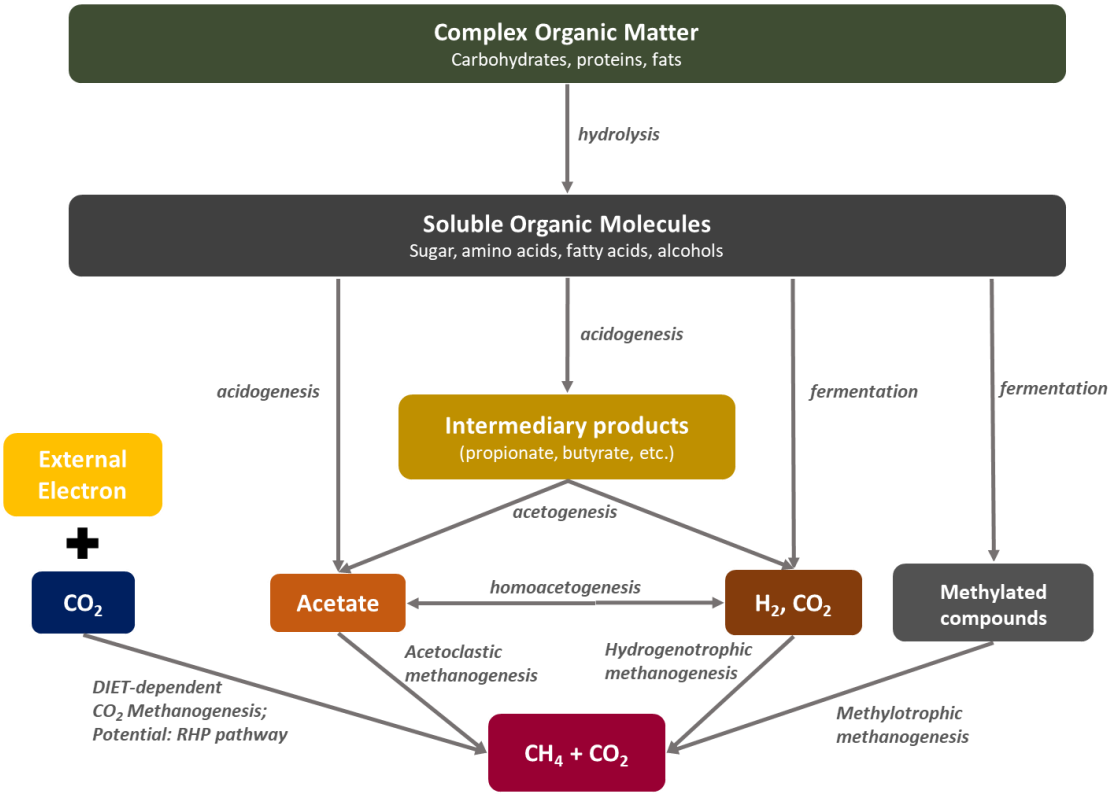


Figure 2.1 Simplified view of the major metabolic pathways in AD

2.1.2 Hydrolysis

AD process begins with groups of hydrolytic and fermentative bacteria hydrolysing complex, particulate matter such as carbohydrates, proteins and fat into dissolved compounds with a lower molecular weight. As bacteria cannot consume particulate organic matter directly, the hydrolysis requires the mediation of exoenzymes, enzymes that are secreted by the bacteria and function outside of the cell. Proteins are degraded via (poly) peptides to amino acids; carbohydrates are transformed into soluble sugars (mono and disaccharides); and lipids are converted to long-chain fatty acids and glycerine (Batstone and Jensen, 2010). Hydrolysis is often considered as one of the rate-limiting processes in AD (another one being methanogenesis) due to limited access to the particulate surface

(Chandler et al., 1980; Zeeman et al., 1997) or if the influent is rich in slow-degrading lignocellulosic materials (McCarty and Rittmann, 2001). Reactor designs like continuous stirred-tank reactor (CSTR) and upflow anaerobic sludge blanket (UASB) aims to promote the chance of surface contact; the introduction of microaeration and thermal pre-treatment also facilitate the degradation of lignocellulosic material.

2.1.3 Acidogenesis

The acidogenesis here is referring to the general fermentation process that occurs during AD, as some consider the term "acidogenesis" only the acid production process. Dissolved compounds generated in the hydrolysis are taken up in the cells of hydrolytic, fermentative bacteria. The products of acidogenesis consist of low molecular weight compounds. Volatile fatty acids (VFAs), also known as short-chain fatty acids, are the major acidification products, including acetate, propionate and butyrate (Table 2.1). The production of VFAs is highly dependent on the hydrogen partial pressure in the reactor. A sufficient amount of H_2 scavenging organisms (e.g., hydrogenotrophic methanogens) can maintain the hydrogen partial pressure at a low level so that acetate is the primary end product. Conversely, if the hydrogen partial pressure is maintained at a high level, propionate and butyrate will be produced at a higher rate, along with other more reduced compounds, such as alcohols and lactate.

The overall conversion rate of acidogenesis is estimated to be the highest among all the stages (13 gCOD/gVSS·d), making anaerobic digesters subject to a sudden pH drop with a vicious circle (fig. 2.2). The pH drop will lead to suppression on methanogens as they are sensitive to the pH level. Subsequently, it will escalate the accumulation of VFAs and leading to a continuous drop of pH. Therefore, a sufficient amount of buffer material is needed to provide alkalinity to withstand the potential pH drop.

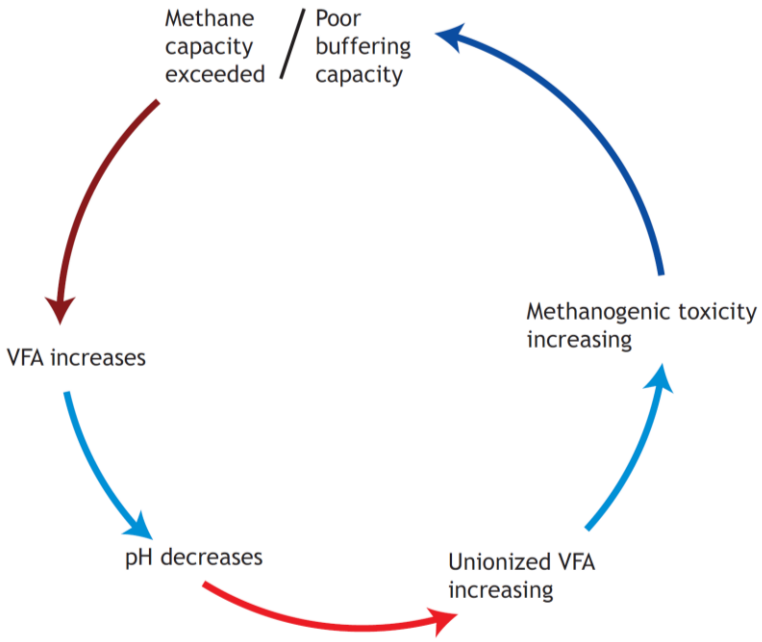


Figure 2.2 Continuous pH drop caused by VFAs accumulation and methanogenic overloading (van Lier et al., 2008)

The acidogenic process is carried out by a diverse group of bacteria, most of which are obligate anaerobe. However, some are facultative and can also metabolise organic matter via the oxidative pathway, that the organic matter is degraded along with oxygen, noted that the oxygenated zone exists in the large-scale fermenter.

2.1.4 Acetogenesis & syntrophic association

The syntrophic association between hydrogen-producing bacteria and hydrogen-utilising microbes is crucial for acetogenesis. In AD, syntrophs and hydrogenotrophic methanogens are key players to maintain such association. The biochemical process of AD is governed by the rule of thermodynamics, where the metabolic pathways that require less energy input or generate more energy are always preferred. Under the methanogenic condition, the association between syntrophs and methanogens help to maintain a thermodynamically favourable environment for both communities to thrive. Syntrophs like *Smithella propionica* degrade VFAs to simple substrates like acetate and hydrogen;

methanogens then consume these simple substrates for methanogenesis. Together, they maintained a low hydrogen partial pressure for the community to thrive. Therefore, syntrophs and methanogens are thermodynamically interdependent (McInerney et al., 2009).

Figure 2.3 and Table 2.1 illustrate the stoichiometry relationship between the change of free energy and hydrogen partial pressure. The result was summarised by van Lier, Mahmoud and Zeeman (2008). As shown in table 2.1, the change of free energy (ΔG) of ethanol, butyrate, and propionate are positive under standard condition. These reactions are unlikely to occur without the syntrophic association to maintain a low hydrogen partial pressure. Under a stabilised digester condition, the hydrogen partial pressure is often maintained below 10^{-4} atm. This interaction is also referred to as interspecies electron transfer, as hydrogen acts as an electron shuttle.

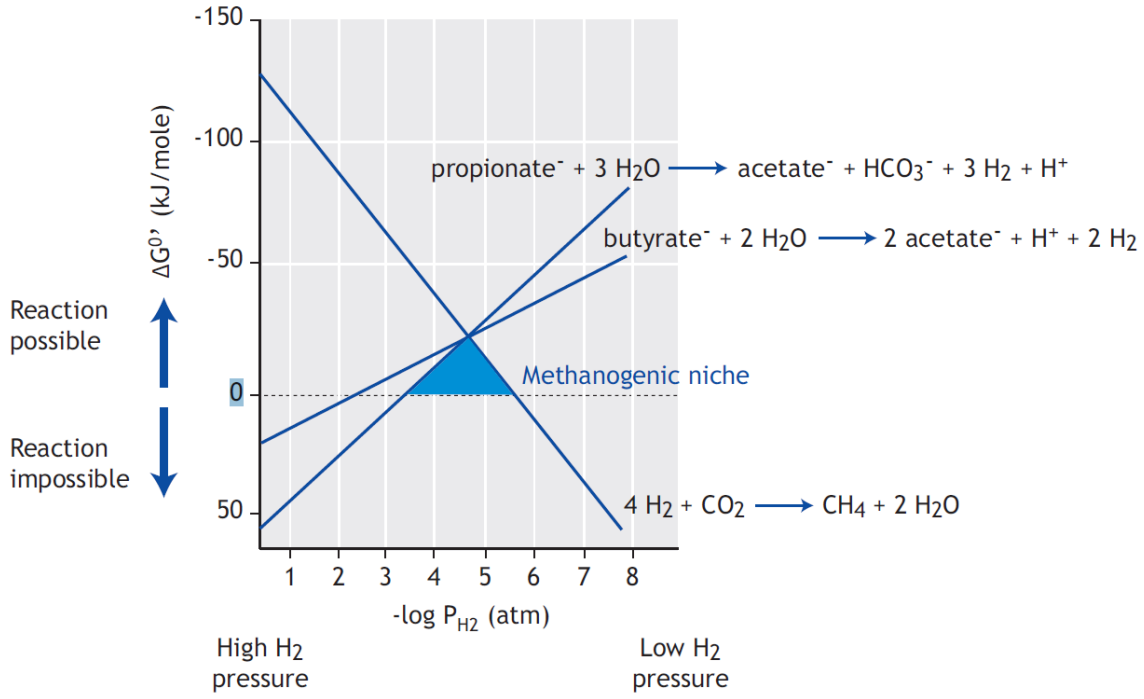


Figure 2.3 Stoichiometry relationship between free energy and H_2 partial pressure. Low H_2 partial pressure promotes the degradation of VFAs (van Lier et al., 2008).

Table 2.1 Stoichiometry of common simple substrates at standard conditions (temperature = 25°C, pressure = 1 atm)

Compound	Reaction	ΔG°' (kJ/mol)
Lactate	$CH_3CHOHCOO^- + 2H_2O \rightarrow CH_3COO^- + HCO^- + H^+ + 2H_2$	-4.2
Ethanol	$CH_3CH_2OH + H2O \rightarrow CH_3COO^- + H^+ + 2H_2$	9.6
Butyrate	$CH_3CH_2CH_2COO^- + 2H_2O \rightarrow 2CH_3COO^- + H+ + 2H_2$	48.1
Propionate	$CH_3CH_2COO- + 3\,H_2O \rightarrow CH_3COO^- + HCO_3^- + H+ + 3H_2$	76.1
Methanol	$4\,CH_3OH + 2\,CO_2 \rightarrow 3CH_3COOH + 2H_2O$	-2.9
Hydrogen-CO_2	$2\,HCO_3^- + 4\,H_2 + H^+ \rightarrow CH_3COO^- + 4\,H_2O$	-70.3
Palmitate	$CH_3^-(CH_2)_{14}\text{-}COO^- + 14H_2O \rightarrow 8CH_3COO- + 7H^+ + 14H_2$	345.6

2.1.5 Methanogenesis

Methanogenesis often is the rate-limiting step in the overall digestion process. Under the optimal thermodynamic condition, methanogens produce methane as the primary product of anaerobic biomass degradation instead of anaerobic respiration. Most of the methane in nature originates from acetate as it is a unique methanogenic substrate, which can serve as both electron donor and acceptor after converted to acetyl-CoA and cleaved into a carbonyl group and a methyl group. Other common electron donors include hydrogen, formate, several alcohols, such as ethanol, propan-2-ol, butan-2-ol and 3-methylbutanol. For electron acceptors, carbon dioxide, methanol, and several different methylamines and methylsulfides are the common acceptors (Deppenmeier, 2002; Ferry, 1993; Pritchett and Metcalf, 2005). Acetoclastic pathway, hydrogenotrophic pathway and methylotrophic pathway are the three major methanogenic pathways (Fig. 2.2). The biochemistry of methanogenesis involves multiple enzymes and coenzymes. To date, the formation of methane must occur by converting methyl-CoM with the help of coenzyme B, while the reactions prior to that would be dependent on the enzyme encoded with the methanogen.

2.1.5.1 Acetoclastic pathway

The acetoclastic pathway is the most utilised pathway in methanogenesis (Lv et al., 2015). Although the methanogens under the order *Methanosarcinales* can use acetate since they contain cytochromes, only two genera *Methanosarcina* and *Methanothrix*, are able to conduct the acetoclastic pathway. It is

estimated that aceticlastic methanogens are responsible for two-thirds of the methane production annually on earth (Fournier and Gogarten, 2008). Acetate is the product of the acetogenesis conducted by acetogenic bacteria (fermenters) and syntrophs. One clear feature to distinguish the methanogens that reduce CO_2 with H_2 or acetate to methane is whether the methanogens possess coenzyme A to enable the Acetyl-CoA conversion, and the genome information can be found on National Centre for Biotechnology Information (NCBI). Acetyl-CoA phosphorylates and activates acetate, which allows it to be processed by acetyl-CoA decarbonylase complex. The C-C and C-S bonds are cleaved by this large enzyme complex and oxidized to CO_2. The transformation of the methyl group to tetrahydrosarcinapterin is carried out after the Acetyl-CoA conversion, which is further converted into methyl-CoM. The coenzyme M and B then demethylate the complex and convert it into methane (Ferry, 1993, 1992; Thauer, 1998).

2.1.5.2 Hydrogenotrophic (CO_2 reduction) pathway

In the hydrogenotrophic (CO_2 reduction) pathway with H_2 as the electron donor, CO_2 is activated and transformed into formylmethanofuran by methanofuran. The formyl group is then transferred to tetrahydromethanopterin, forming the methyl group. The methyl group is then transferred by methyl-H4MPT: coenzyme M methyltransferase to coenzyme M, resulting in methyl-CoM, which is then disproportionate into methane and CO_2. The net reaction of the hydrogenotrophic pathway is the conversion of CO_2 to methane, as six electrons are required for the reduction of three methyl-CoM molecules to methane, only the remaining one methyl-CoM is oxidized to CO_2 (Vorholt and Thauer, 1997).

2.1.5.3 Methylotrophic pathway

Methane production from methylated C1 compound carries a similar reaction route. All methylotrophic pathwats transfer the methyl-group compound (e.g. methanol, methylamine, methanethiol and methylthiopropanoate) from the electron acceptor to coenzyme M. This process is

conducted by an enzyme system: two methyltransferases (MT1 and MT2) and a corrinoid prosthetic group (Sauer et al., 1997). MT1 catalyses the transfer of the methyl group from the substrates to the corrinoid group when the substrate binds to this complex. The methyl group are then transferred from the corrinoid to coenzyme M by MT2. The following reactions are slightly different for each substrate, but in general, coenzyme M generates methyl-CoM which is then converted into methane and CO_2 (Ferguson et al., 2000; Ferguson and Krzycki, 1997).

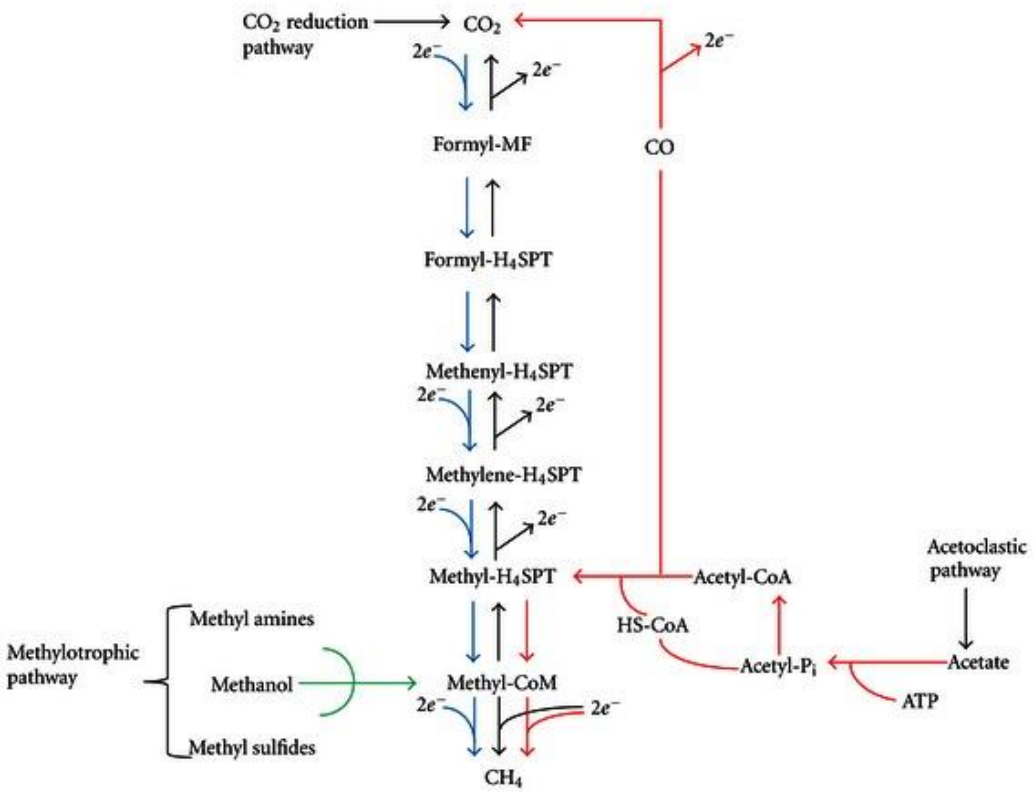


Figure 2.4 Schematic illustrating the major substrates (H2/CO2, methanol, methylamines, methyl sulfides, and acetate) and the respective pathways utilised for methanogenesis (Ellis et al., 2012).

2.1.5.4 DIET-dependent CO_2 reduction pathway

Direct interspecies electron transfer (DIET) is a syntrophic metabolism in which electrons flow from one cell to another directly without the reliance on reduced molecules such as molecular hydrogen or formate. The first report on microbial electricity generation was in 1911, while the attention in DIET started to gain in recent year as the academic community recognised its potential as a new source of

energy, especially in relation to AD. The occurrence of DIET has been demonstrated with various species of exoelectrogenic bacteria and methanogenic archaea in both co-culture and mixed culture environments, and substrate types varied from simple (e.g. ethanol, glucose) to complex ones (e.g. leachate, waste sludge) (Barua and Dhar, 2017; Park et al., 2018). The DIET pathway showed superior performance on methane production, as mentioned previously. One possible explanation is the DIET pathway promotes the cell growth of methanogen (Holmes et al., 2017). It was also demonstrated that the electric conductivity of the conductive materials was not the determining factor for DIET to function. Increase in methane yield was reported for all DIET studies in AD, while there was no significant correlation between the relative methane formation rate and the dosage and types of conductive material.

To date, three types of DIET mechanisms were reported. The mechanisms are microbial aggregate-based: the short distance between the exoelectrogenic bacteria and methanogenic archaea allows the electron transfer. Regardless of the DIET mechanisms, how the methanogens receive the electron has not been determined. The *Methanothrix* species are generally lack of the biological structure or secretion to facilitate the electron reception. Moreover, the presence of conductive material can facilitate either as a bedding material for aggregation or as an electron transfer medium.

(A) DIET through conductive pili. The first observation of DIET phenomenon via conductive pili was the co-culture of *Geobacter metallireducens* (electron-donating) and *G. Sulfurreducens* (electron-accepting) with ethanol and fumarate. Enhanced OmcS (cytochrome-c) was identified in the electrically-conductive aggregate via conductive pili (Summers et al., 2010). It facilitates DIET between the two bacteria and hydrogen was ruled out as an electron donor as both species could not metabolise hydrogen. The similar phenomenon was observed in a co-culture of *G. metallireducens* and *Methanothrix harundinacea*, and methane production was observed (A.-E. Rotaru et al., 2014; A.

E. Rotaru et al., 2014). This finding indicates that DIET via conductive pili is a possible pathway of methanogenesis.

(B) DIET through conductive materials. The first notion of conductive material could stimulate DIET was in 2012 after the observation of the increase in methane production in anaerobic serum bottles supplemented with conductive material (Kato et al., 2012). The enrichment of microbial population was observed. It was demonstrated later that in a co-culture experiment, *M. barkeri* and *G. metallireducens* were not able to produce methane with the absence of conductive material (Liu et al., 2012). The species selected in their study ruled out the possibility of conductive pili formation, and metabolism on hydrogen and formate, suggesting conductive material can act as an electron transfer medium during DIET. The addition of conductive material can facilitate the occurrence of DIET as a novel methanogenic pathway, which the electron from exoelectrogenic *Geobacter* species acted as electron donor, enabling acetoclastic *Methanothrix* species to use carbon dioxide as the electron acceptor, a group of methanogens was known to unable to reduce CO_2.

(C) DIET through membrane-bound electron transport proteins. A transmission electron microscopic images of a co-culture of *Prosthecochloris aestaurii* and *G. sulfurreducens* revealed intimate cell connection between two bacteria during DIET (Ha et al., 2017). No notable conductive pili were observed and the multiheme outer-surface cytochrome OmcZ was believed to be responsible for the electron transfer. The similar physical association was observed in a co-culture of a sulfate-reducing bacterium and a methanogenic archaeon (McGlynn et al., 2015), but no methane production was detected, suggesting this type of DIET may not facilitate methanogenesis.

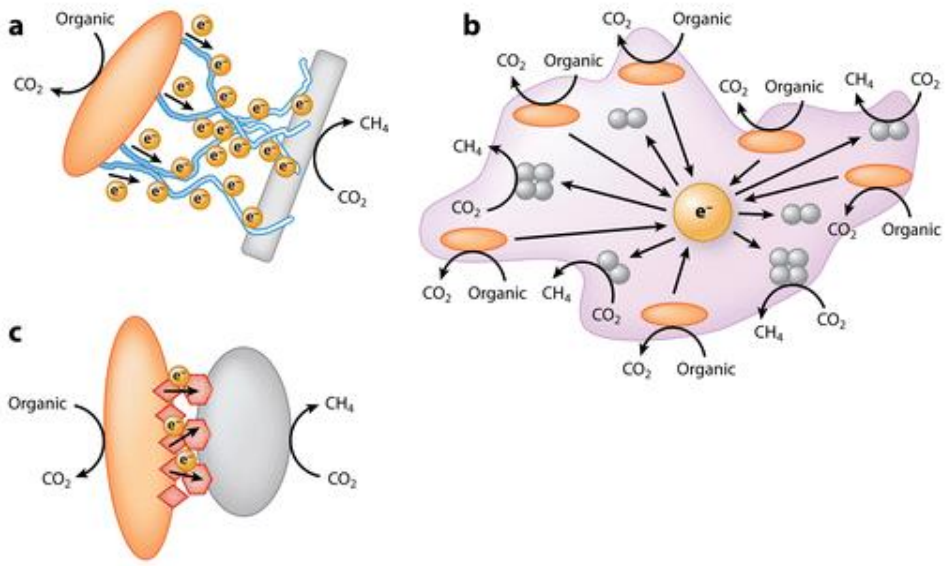


Figure 2.5 Illustration of the DIET mechanisms: (a) conductive pili, (b) electrically conductive material, (c) membrane transfer (Lovley, 2017)

A recent study also suggested that methanogens capable of conducting DIET-dependent CO_2 methanogenesis can carry out reductive hexulose-phosphate (RHP) pathways, a pathway similar to the Calvin-Benson cycle, a common carbon fixation pathway found in plant cells (Yang et al., 2019). The carbon fixation pathway in RHP pathway requires less energy investment for cell synthesis comparing to the traditional pathway. The RHP pathway facilitates the population growth of certain methanogen species in the community and promoting the overall activities. Furthermore, the 5,6,7,8-tetrahydromethanopterin hydrolase gene responsible for formaldehyde condensation was highly expressed in the recovered genome bin of *Methanothrix concilii*, suggesting that the RHP pathway potentially involves in methanogenesis. As RHP pathway consume electron during the process, the delivery of this pathway might rely on the DIET. Nevertheless, the question of how *Methanothrix* receives electron remains unsolved.

2.2 Bioinformatics: analysing microbiome

With the advancement of computer system and molecular biology in combination with genetics, the research practice of environmental engineering is increasingly switching to "take a few days to carry out in the laboratory, then several months in the office to process the resulting data" (Hodgman, 2017). The definition of bioinformatics is surprisingly broad, and its functional definition can be summarised as *"seeking of generate knowledge of the properties, population, and processes of biological entities"*. It is an application of computing and mathematics to the management, analysis, and understanding to solve biological questions. Since AD is essentially a series of biological process, the bioinformatics approach helps to elucidate the mechanism behind the complex microbial interaction.

Genome sequencing is one important approach to studying the AD process. The classical cultivation-based microbiological methods bring a little insight into the complex biological interactions. Firstly, most anaerobes are slow-growing and require strict cultivation conditions. Secondly, the syntrophic bacteria are dependent on hydrogen scavenging micrograms, so the classical pure culture method is ineffective to study their physiology. Third, it is ineffective in studying microbial interaction and microbial composition. Instead, omics-driven approaches allow researchers to delve deeper to uncover the biological system, and to decipher their functions without the hardship of cultivating (Narihiro and Kamagata, 2013). Marker gene (e.g. 16S rRNA) and whole metagenome are the common sequencing methods for surveying microbial communities (Knight et al., 2018). The 16S rRNA amplicon sequencing offers cheap and robust taxonomic profiling, and the untargeted ('shotgun') metagenomics sequencing allows the reconstruction of metabolic pathways given sufficient sequencing coverage.

The advancement of sequencing technology provides convenient platforms to study environmental microbiology. From the first generation Sanger sequencing to the high-throughput next generation to the third generation long-read sequencing, each technology has its merit and short-coming.

2.2.1 16S rRNA amplicon sequencing

In the 1970s, Carl R. Woese and George E. Fox discovered ribosomal RNA (rRNA) (Woese and Fox, 1977) that shaped the current evolutionary biology with the three domains of life, *Eukarya*, *Archaea* and *Bacteria* (Woese et al., 1990). The 16S rRNA (Fig. 2.6) is a component of the 30S small subunit of a ribosome that binds to the Shine-Dalgarno sequence. The term 16S denotes the size of the 1600-nt-long small subunit. The universality and extremely conserved structure of this gene in archaea and bacteria (Coenye and Vandamme, 2003) makes it an excellent candidate for phylogenetic analysis. Being a catalytic element for protein synthesis, the concentration of rRNA is proportional to the RNA and the number of cellular ribosomes (Bremer and Dennis, 1996; Kerkhof and Ward, 1993; Poulsen et al., 1993). As a result, the 16S rRNA reads from the sequencing result can be roughly treated as the relative abundance of the sequenced microbial community, and this application was actualised by the development of the next-generation high throughput sequencing technologies.

In the environmental samples, the amount of bacterial and archaeal DNA is in a trace amount. It is difficult to obtain sufficient DNA for sequencing under normal circumstances. The invention of polymerase chain reaction (PCR) solved this problem. Since the conserved regions of the 16S rRNA are stable, they serve as the primer binding sites for PCR to occur. With the help of PCR, we can even generate enough DNA from low biomass and highly host-contaminated samples. The copies of DNA are referred to as **amplicon**.

In comparison to the first-generation sequencing (e.g., Sanger sequencing), the next-generation high throughput sequencing can generate a significantly larger dataset, thus allowing the inference on the microbial composition. However, high-throughput sequencing traded off accuracy for more massive data sets. For instance, the Illumina platform, which is the most popular sequencing platform, with 60 to 90% of the market share over the years (Wood, 2016), produces short read length ranging from 50 to 500 bp. The primer set targets the different hypervariable regions, including V4, V3-V4 and V4-

V5 in 16S rRNA gene (Caporaso et al., 2012). This short-read approach has inherent primer bias. In regard to surveying AD microbiome, it was shown that the V4-V5 primer set was inefficient in amplifying *Euryarchaeota* due to low gene copy number (Campanaro et al., 2018), a phylum that contains the majority of the methanogenic archaea. Besides, the limited phylogenetic resolution of the short-read approach cannot classify microbes below the genus level (Kim et al., 2011; Yarza et al., 2014). Most importantly, the taxonomic classification via 16S gene is limited by the existing reference.

For the third generation sequencing technology, Pacific Bioscience (PacBio) single-molecule real-time (SMRT) Sequencer provides a long-read sequencing approach to achieve full-length 16S rRNA gene coverage and generates sufficient read copies to interpret the microbial composition. Full-length 16S that covers hypervariable region V1-V9 and conserved regions, avoid primer bias and provide higher phylogenetic resolution with the consideration of the non-homogeneous distribution of mutation (Earl et al., 2018; Narrowe et al., 2017; Shin et al., 2016; Singer et al., 2016). The extended read length provides higher confidence in the classification of the phylogeny (Wagner et al., 2016), subsequently provides additional OTU when the OTU-clustering strategy is used (Franzén et al., 2015). Also, the circular consensus sequencing (CCS) approach allows the generation of highly accurate CCS read by aligning multiple subreads (Frank et al., 2016; Travers et al., 2010), facilitating the use of singleton for diversity profiling and identification of novel phylotype.

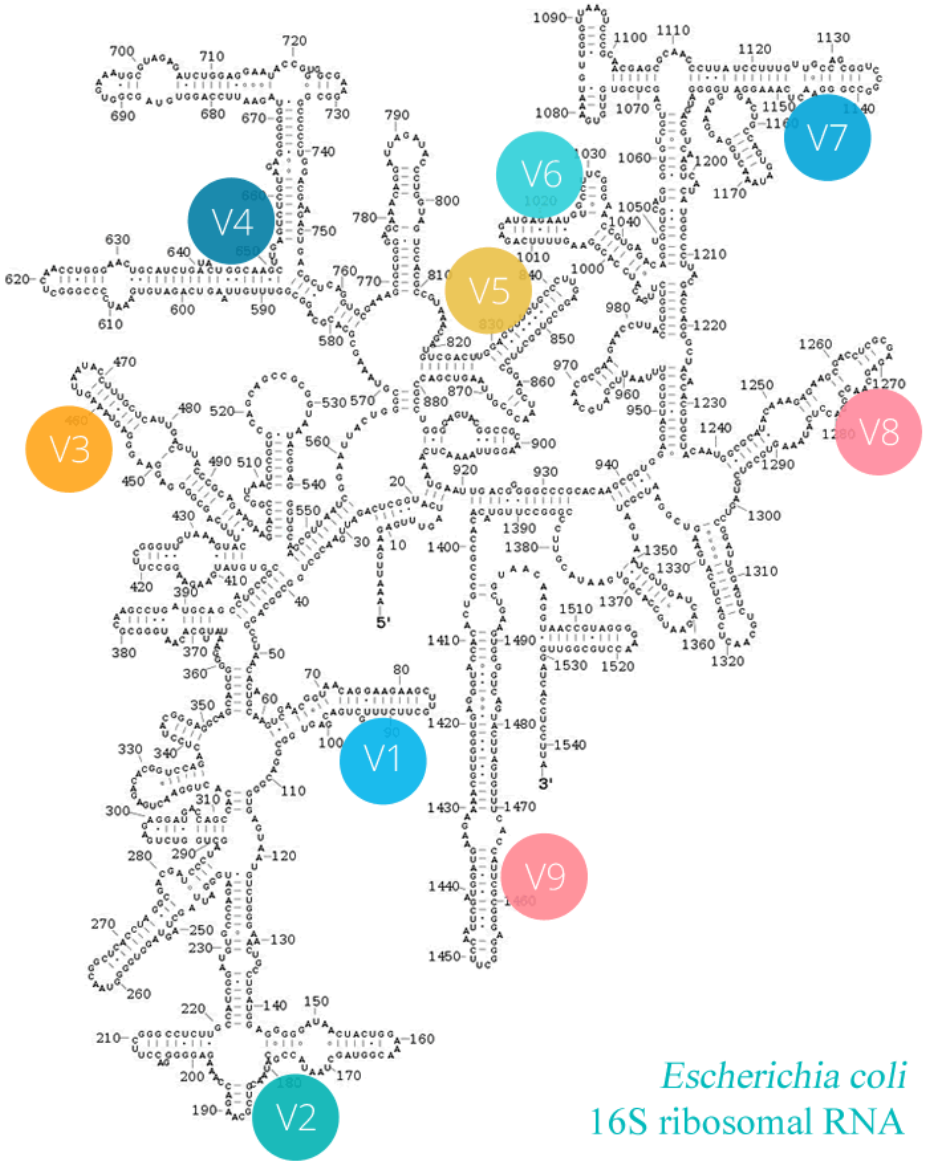


Figure 2.6 Illustration of *E. coli* 16S rRNA gene (Era7, 2016)

2.2.2 Shotgun metagenomics sequencing

Shotgun (untargeted) metagenomics is referring to the sequencing methods that sequence all microbial genome within a DNA sample, which also includes viral and eukaryotic DNA. The use of metagenomics sequencing can eliminate the primer bias in the 16S rRNA amplicon sequencing as it is a PCR-free sequencing method. In addition, the recovery of the whole genome information allows the reconstruction of the metabolic pathway. This is extremely valuable to a mixed-culture study (e.g.,

AD microbiome), as there is a large number of uncultivated micrograms present in the sludge. Metagenome allows the inference of their functions in the AD microbiome.

With that being said, there are a few challenges in using metagenomics sequencing. Firstly, it requires higher quality DNA than in the amplicon approach. Secondly, the sequencing cost of metagenome samples is significantly higher than the amplicon samples. Thirdly, it requires much more computational power to process the data. Fourthly, sufficient sequencing coverage is required for the genome bin recovery if metabolic pathway reconstruction is the purpose (Knight et al., 2018). Therefore, it is "more expensive" to perform and analyse a metagenome sample than an amplicon sample.

An economical and also common approach is to conduct a taxonomic survey via the 16S rRNA amplicon sequencing to evaluate the relative abundance of the target microbe in the sample, then we can calculate the on-plate concentration required to reach the sufficient coverage.

2.2.3 Mechanism of Illumina SBS: CRT technology (NGS)

NGS sequencing has become the standard of the industry in the field of genomic study, because the development over the decades led to a reduction in cost and improvement of data output, which enables the understanding of the complexity of genome (Goodwin et al., 2016). Illumina (Illumina, Inc., San Diego, California, US) is the predominant NGS platform (Timmerman, 2015). The broad definition of short-read sequencing approaches can be categorised into sequencing by ligation (SBL) and sequencing by synthesis (SBS). Illumina uses the cyclic reversible termination (CRT) as their SBS DNA-polymerase-dependent method (Fig. 2.7). In CRT, elongation is prevented by blocking the ribose 3'-OH group. A complementary primer is attached to an adapter region of a DNA template to begin the process, and the single-stranded DNA (ssDNA) is bound and becomes the double-stranded DNA (dsDNA) by the polymerase. The mixture of four 3'-blocked deoxynucleotides (dNTPs) are

added during each PCR cycle as template material for amplification. Every single dNTP incorporates to their complementary strand for elongating, then imaging of the surface identifies which dNTP is incorporated at each cluster and the unbound dNTPs are removed. A new cycle is then conducted after the removal of fluorophore blocking group.

The overall accuracy rate of the CRT approach is over 99.5% (Bentley et al., 2008) with less observable homopolymer errors compared to other platforms. However, it was criticised for its under-representation in AT-rich (Dohm et al., 2008; Harismendy et al., 2009) and GC-rich region (Harismendy et al., 2009; Nakamura et al., 2011) and substitution errors (Minoche et al., 2011).

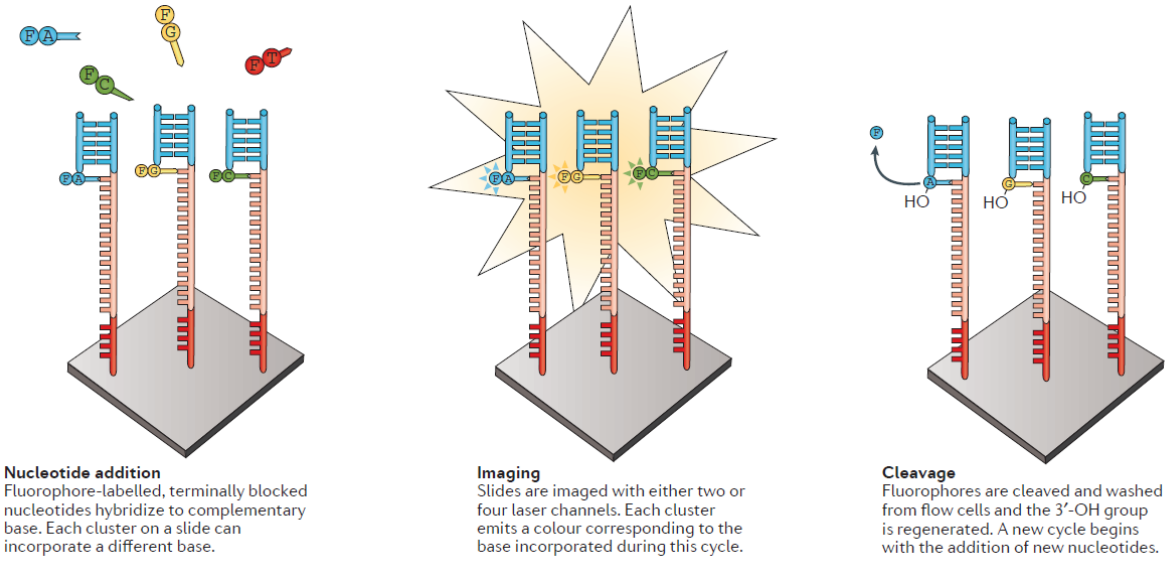


Figure 2.7 Illustration of sequencing by synthesis (SBS): cyclic reversible termination (CRT) approach (Goodwin et al., 2016)

It should be noted that although most of the Illumina sequencers employ the CRT approach, there are Illumina HiSeq X that uses the single-nucleotide addition (SNA) approach, and Illumina Synthetic Long-Read that uses single-molecule real-time (SMRT) long read approach. The above-mentioned sequencers were not used in this study.

2.2.4 Mechanism of Pacific Bioscience SMRT sequencing technology (TGS)

The praise of long-read approach comes from its ability to deliver observation of large structural variations that are relevant to evolution and adaptation (Pollard et al., 2018). The long read length with an average of 10 to 15 kb is ideal for *de novo* genome assembly application (Schatz et al., 2010). Pacific Bioscience (Pacbio) is the most widely used platform that employs single-molecule real-time (SMRT) sequencing technology as the long-read approach (Korlach et al., 2010). Instead of a DNA cluster, a single molecule of DNA is fixed with a single DNA polymerase enzyme at the bottom of a zero-mode waveguide (ZMW), which illuminated observation volume that is tiny enough to enable the observation of a single nucleotide of DNA being incorporated.

Since the raw read error rate of SMRT sequencing is as high as 15% with indel (insert and delete) error as the dominant factor (Carneiro et al., 2012), PacBio employed an approach called Circular Consensus Sequence (CCS) reads (Fig. 2.8) to improve the accuracy. Each CCS read was determined by multiple subreads aligned without reference sequence from a single ZMW. Higher the number of passes, more accurate the CCS read determined. With this approach, SMRT sequencing delivers the longest read lengths and highest consensus accuracy comparing to any other sequencing technology (Rhoads and Au, 2015). Yet, the high cost of reagents and sequencing service and the stringent requirement on DNA quality remain as the most significant obstacle in applying SMRT sequencing in environmental research.

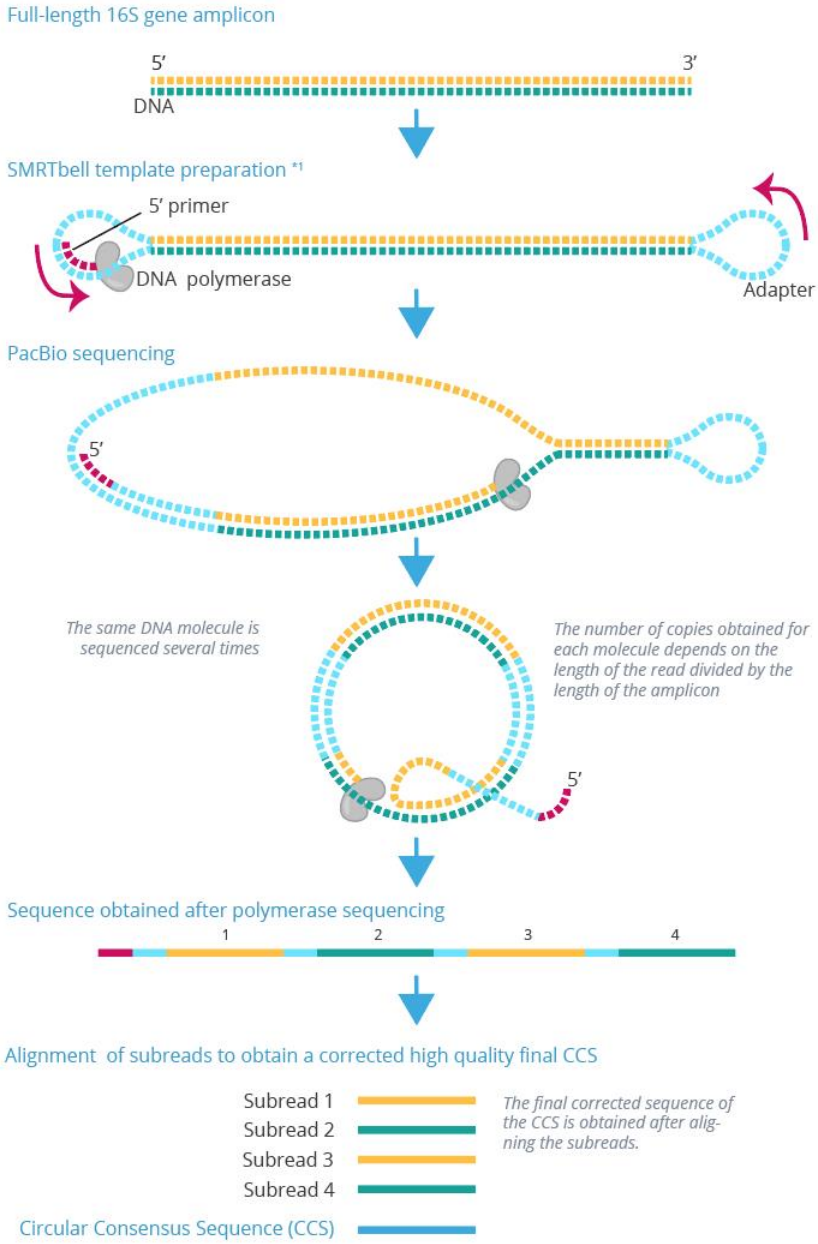


Figure 2.8 Illustration of CCS curation using full-length 16S rRNA amplicon sequencing as an example (Era7, 2016)

Chapter 3 Materials and methodology

We employed three strategies in this study:

1. full-length 16S rRNA amplicon sequencing to study the effect of full-length sequences on the phylogenetic resolution. The result is discussed in **chapter 4**
2. Short-read (v4-v5) 16S rRNA amplicon sequencing to compare the discrepancy between the full-length and short-read sequences on the microbial analysis. The result is discussed in **chapter 5**
3. Shotgun metagenomics and methyl coenzyme M reductase (*mcrA*) gene amplicon sequencing to study the microbial community and functional pathway on micro-aerated anaerobic digesters. The result is discussed in **chapter 6**

3.1 Full-length 16S rRNA amplicon sequencing preparation, library construction and analysis

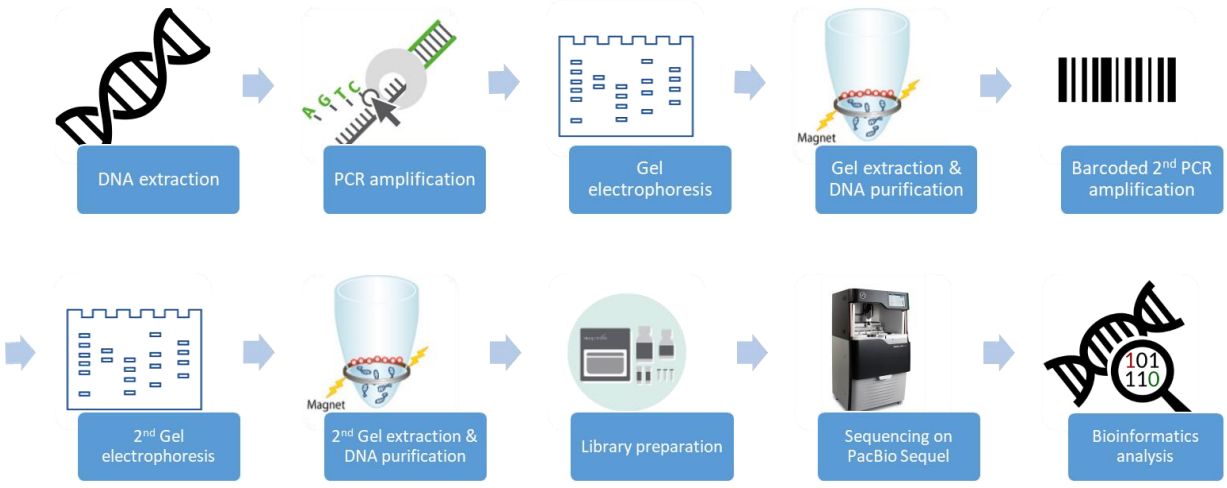


Figure 3.1 Developed pipeline for the full-length 16S gene amplicon sequencing, from DNA extraction to bioinformatics analysis

3.1.1 Sample collection

In total, 48 anaerobic digester sludge DNA samples were collected from 48 full-scale municipal wastewater treatment plants in five countries/region: Canada, Hong Kong, Japan, Netherlands and the United States. The samples represented a wide range of digester operating conditions. The DNA extraction was performed on the samples collected from Hong Kong with the PowerSoil DNA Isolation Kit (Qiagen, Venlo, NL). The input mass was increased from 0.25 g suggested by the protocol to 0.4 g for more yield on DNA material.

3.1.2 Universal primers, 16S rRNA gene amplification and purification

FL 16S amplification was conducted on the DNA samples with KAPA HiFi HotStart PCR Kit (KAPA Biosystems, MA, USA). The $MgCl_2$ concentration was set at 3 mM for the first and second polymerase chain reaction (PCR). The reaction volume for the PCR template was 50 μL, and 50 ng genomic DNA was added for the first PCR with a universal primer set 27F (AGRGTTYGATYMTGGCTCAG) & 1492R (RGYTACCTTGTTACGACTT) for the bacteria amplification, and A1F (GKTTGATCCYGSCRGAG) & 1490R (GGYTACCTTGTTACGACTT) for archaea amplification, resulted in 38 unique samples. The thermal cycle of the PCR was initiated with denaturation at 97 °C for 7 minutes, followed by 20 cycles of denaturation (97 °C, 30s), annealing (56 °C, 30s), extension (72 °C, 60s), and finished with extension at 72 °C for 10 minutes. Gel electrophoresis was performed for gel-based size selection to remove contaminants and undesired PCR products. The gels were then extracted with QIAquick Gel Extraction Kit (Qiagen, Venlo, NL) and purified with AMPure PB beads (PacBio, CA, USA). The purified DNA was then amplified in the second PCR with 10 ng as DNA template and barcoded primers. The thermal cycle for PCR amplification remained constant, proceeded with another round of gel electrophoresis size selection, gel extraction and AMPure PB beads purification. DNA qualification and quantification were performed using a Nanodrop 2000c Spectrophotometer and Qubit 3.0 Fluorometer (Thermo Fisher Scientific, MA, USA), respectively. The absorbance ratios were maintained >1.8 for 260/280 ratio

and 260/230 ratio to ensured the purity of the DNA samples. The purity of the DNA samples was ensured by having the absorbance ratio exceeded >1.8 for 260/280 ratio and >2.0 for 260/230 ratio.

After the purification step, 19 out of the 48 DNA samples were qualified to be sequenced on the PacBio Sequel. The detailed plant operating condition of the sequenced samples was listed in Table A1.

3.1.3 PacBio full-length 16S rRNA amplicon sequencing

The purified barcoded amplicons were pooled for multiplexing and proceeded to library construction with SMRTbell™ Template Prep Kit v1.0-SPv3. The sequencing was performed by the PacBio Sequel platform at the Hong Kong Polytechnic University using standard protocol "Procedure & Checklist - FL 16S Amplification, SMRTbell® Library Preparation and Sequencing" with extended DNA damage repair (37ºC for 2 hours) and blunt-ligation reaction (25ºC for 24 hours). Insert size was confirmed using Agilent bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA) for optimising the binding concentration and the on-plate loading concentration. Sequencing Primer v3 was used for primer annealing, Sequel™ Binding Kit 2.0 was used for polymerase binding, and Sequel DNA Internal Control 2.1 was added to the mixture. Diffusion loading was selected with specified on-plate concentration ranging from 8 pM to 16 pM. Sequel™ Sequencing Kit 2.1 was used for sequencing with 4 Sequel™ SMRT® Cell 1M v2. Immobilisation time, pre-extension time, and movie time were set at 120 mins, 44 mins, and 10 hours, respectively.

3.1.4 Microbial community analysis

The subreads were demultiplexed and then were aligned to the Circular Consensus Sequences (CCS) by the SMRT Analysis function. The CCS reads were screened using Mothur v.1.39.5 (Schloss et al., 2009), sequences outside of the range (<1300 bp, >1700 bp) were removed. As CCS reads could be in

mixed-orientation, a check orientation command was added (Schloss et al., 2016). Chimaeras were removed using VSEARCH (Rognes et al., 2016). Taxonomy alignment was performed against the SILVA 132 reference (Quast et al., 2013) with 80% cut-off value, Operational Taxonomic Unit (OTU) classification was performed with a 99% identity. Alpha diversity analysis included diversity indices (Inverse Simpson and Shannon diversity index), richness abundance estimators (Chao1 and ACE), evenness indices (Simpson evenness and Shannon evenness) and good's coverage. Principal Coordinate Analysis (PCoA) was used to evaluate the beta diversity with weighted and unweighted Unifrac distance metrics.

Linear discriminant analysis effect size (LEfSe) (Segata et al., 2011) was used in identifying the abundant taxa to characterise the AD microbiome among different operating conditions. LEfSe is an algorithm to discover high-dimension biomarker and explain the distinctions of genomic features between multiple biological conditions. Both biological relevance and statistical significance are emphasised to allow researchers to identify the abundant feature with biological meaning. It should be noted that the application of prior knowledge to the biological process is needed for a fair evaluation. Threshold on the logarithmic LDA score for discriminative features was set at 2 with alpha value set at 0.05. All-against-all strategy for multi-class analysis was selected.

Co-occurrence networks were constructed using the CoNet method (Faust and Raes, 2016). Edges were obtained from the calculation using four similarity measures: Pearson and Spearman rank correlations; Bray Curtis and Kullback-Leibler non-parametric dissimilarity indices. Randomisation was performed with 100 iterations of permutation and bootstrap with a *p*-value threshold of 0.05. The networks were visualised by Cytoscape v3.0 (Shannon et al., 2003). Positive-edges-only networks were clustered via CytoCluster (Li et al., 2017) using OH-PIN algorithm (Wang et al., 2012). Noted that the relationships visualised by the network does not necessarily suggest any causality of the underlying mechanism. Prior knowledge should always be applied for such analysis.

ARB (Ludwig et al., 2004) was used for the phylogenetic tree construction. The representative sequences from OTU classifications were aligned using SINA aligner v1.2.11 (Pruesse et al., 2012) and imported to ARB. The sequences were inserted to the SILVA 132 SSU Ref NR 99 reference tree using parsimony algorithm; phylogenetic trees were built using the neighbour-joining algorithm with Felsenstein correction and bootstrap value set at 5000. Species level identification was performed by BLASTn v2.8.1 (Zhang et al., 2000) against SILVA 132 reference database using Megablast algorithm.

3.2 Short-read 16S rRNA amplicon sequencing preparation and analysis

3.2.1 Sample collection, v4-v5 primers, 16S gene amplification and sequencing

Thirteen DNA samples were used in the comparison study. The DNA samples were selected from the collection mentioned in session 3.1.1. PCR amplification was done using a universal primer set 515F (5'- GTGCCAGCMGCCGCGGTAA-3') & 909R (5'-CCCCGYCAATTCMTTTRAGT-3') targeting the V4-V5 region of both bacterial and archaeal 16S rRNA gene. The thermal cycle of the PCR was initiated with denaturation of 97 °C for 3 minutes, followed by 20 cycles of denaturation (94 °C, 30s), annealing (55 °C, 45s), extension (72 °C, 60s), and finished with extension at 72 °C for 10 minutes. Sequencing was performed on Illumina Miseq Bulk 2 × 300 nt paired-end system.

3.2.2 Microbial community analysis

The short-read sequences with a quality score below Q20 were filtered using Mothur (Schloss et al., 2009). Taxonomy alignment was performed of the Silva reference files Release 132 (Quast et al., 2013) with chimaeras removed by the UCHIME algorithm (Edgar et al., 2011). OTU clustering was performed to obtain the relative abundance of the microbial communities and the phylogenetic view of community composition using 97% similarity. OTU-based alpha, beta diversity analysis and

rarefaction were performed to compare the diversity of the samples. Bray-Curtis dissimilarity based dendrograms were drawn using MEGA X (Kumar et al., 2018); the visualization of OTUs composition was done using iTOLv3 (Letunic and Bork, 2016). The discrepancy between the full-length and short-read data sets was evaluated using Spearman's rank correlation calculated by the vegan package in R (Dixon, 2003) with Bray-Curtis distance.

3.3 Shotgun metagenomics and *mcrA* gene amplicon sequencing preparation and analysis

3.3.1 Sample collection, shotgun metagenomics sequencing and mcrA gene amplicon sequencing

Five biomass samples were collected from lab-scale AD fermenters in the University of Hawai'i at Mānoa at anaerobic, microaerobic and anaerobic-after-microaeration conditions. DNA was extracted using the PowerSoil DNA Isolation Kit (Qiagen, Venlo, NL). Primer set mcrA3F & mcrA3R was used for mcrA gene amplicon sequencing to investigate the methanogenic communities. Both shotgun metagenomics sequencing and mcrA gene amplicon sequencing were performed on the Illumina HiSeq2000 platform.

3.3.2 Microbial community analysis

The shotgun metagenomics data was uploaded to a Linux server with 48 processors for data processing. Trimmomatic (Bolger et al., 2014) was used to remove adapters and base pair reads with an average quality score lower than 30 or below 36 bases long. The assembly of trimmed data was performed using SPAdes (Bankevich et al., 2012) with k-mer sizes set at 19, 33, 47, 61 and 75. The binning of the assembled contigs files was performed by MaxBin (Wu et al., 2014), and the completeness of the bins was assessed using CheckM (Parks et al., 2015). Strainseeker (Roosaare et al., 2017) was used to identify the key microbe in species level for bin selection. The selected bins

were further analysed through PhyloPhlAn (Segata et al., 2013) and BlastKOALA (Kanehisa et al., 2016) for metabolic pathways identification.

For *mcrA* gene analysis, primer removal, dereplication and chimaera check were performed using Mothur prior to the FunGene Pipeline Tools (Fish et al., 2013) were used on the *mcrA* data. Framebot was used to translate the nucleotides sequence to protein sequence using FunGene in-build *mcrA* database. The alignment was performed by HMMER3 and clustering was performed at a distance cutoff of 0.05. Representative sequences were selected and compared to the NCBI database by drawing a phylogenetic tree using MEGA7 (Kumar et al., 2016).

Chapter 4 High resolution in characterisation of global anaerobic digester microbiome using full-length 16S rRNA gene amplicon sequencing

4.1 Overview

Anaerobic digestion (AD) is capable of converting various organic compounds into methane mediated by specific microbial communities for a series of biological reactions. Conventionally, the microbial communities are categorised into fermentative bacteria (fermenters), syntrophic bacteria (syntrophs) and methanogenic archaea (methanogens). One of the critical pathways in AD is the syntrophic association between the methanogens and syntrophs, which interact by the change of key metabolites, e.g., acetate, formate and hydrogen. This mutualism maintains low levels of these key metabolites so that syntrophic metabolism is thermodynamically favourable (McInerney et al., 2009; Sieber et al., 2010). Determining the microbial communities involved in syntrophic metabolism is often carried out by 16S rRNA gene profiling using the Illumina platform (e.g., Liu et al., 2016; Müller et al., 2016; You et al., 2016). Our previous study (Mei et al., 2017) revealed the microbial community structures present in 90 different full-scale anaerobic digesters using the Illumina platform and suggested that the AD microbiomes composition were operation-driven. However, two major shortcomings of the Illumina sequencing approach limit its resolution in taxonomic identification. Firstly, the Illumina platform can only produce short read lengths ranging from 50 to 250 bp, targeting the different hypervariable regions of the 16S rRNA gene (typically 16–33% of the total length). Because many closely related bacteria share similar sequences in the amplified region (Fuks et al., 2018), the taxonomic classification is often limited to genus-level identifications. Also, the non-homogeneous distribution of mutation, which occurs across the entire hypervariable regions (Singer et al., 2016), prevents confident identification of some short-read sequences. Secondly, the V4-V5 primer set is inefficient in amplifying *Euryarchaeota* due to low 16S rRNA gene copy number (Campanaro et al., 2018), which contains most of the methanogenic species. Consequentially, the limited phylogenetic resolution obtained with the Illumina platform does not allow, in most cases, the classification of

microbes below the genus level (Kim et al., 2011; Yarza et al., 2014). Hence, more accurate taxonomic identification is needed to understand the complexities of AD microbial communities.

To obtain a better resolution of the microbial composition of AD digesters, Pacific Bioscience (PacBio) single-molecule real-time (SMRT) sequencing approach was used for full-length 16S rRNA gene sequences, achieving at the species-level identification. The full-length 16S rRNA gene sequence covers both hypervariable regions V1-V9 and conserved regions, avoids primer bias, and provides more considerable phylogenetic information (Angle et al., 2017; Earl et al., 2018; Shin et al., 2016). Moreover, such extended read length enables higher confidence in phylogenetic classifications (Wagner et al., 2016) and offers additional OTUs for clustering improvement (Franzén et al., 2015). The circular consensus sequencing (CCS) approach generates highly accurate sequences via multiple subreads alignments (Frank et al., 2016; Travers et al., 2010), allowing the use of singletons for diversity profiling and identification of novel phylotype. Although near full-length 16S rRNA gene sequences can be obtained through Sanger sequencing, it is time-consuming and unsuitable for processing large numbers of samples and overwhelming with a large number of sequences per sample (Youssef et al., 2009). Additionally, SMRT sequencing can provide more coverage than Sanger sequencing for microbial profiling at a lower cost (Hebert et al., 2018). To this end, the full-length 16S rRNA amplicon sequencing using PacBio Sequel was employed to re-sequence selected AD samples from our previous study (Mei et al., 2017). The sample preparation and data analysis of the full-length 16S rRNA amplicon PacBio Sequel sequencing can be used as a reference for other future studies. The unique diversity in fermenters, syntrophs, and methanogens of anaerobic digesters associated with operating conditions was addressed. Overlooked methanogenic archaea, including the methanogens outside of *Euryarchaeota*, and the comparison to our previous Illumina MiSeq result (Mei et al., 2017) were discussed. In this regard, results gleamed identify overlooked microbiomes and cross-link with digestion operations.

4.2 Results and discussion

4.2.1 Sequence quality, diversity indices

Table 4.1 summarises the number of sequences obtained from samples extracted from seven operating conditions, including high salinity, pre-treatment, temperature (thermophilic, <30°C, ~40°C and mesophilic) and food waste co-digestion. High-quality CCS reads with an average read score of 0.995 ±0.0004 (>99% accuracy) were achieved by setting the minimum passes at ten, suggesting that such subread alignment can significantly improve the accuracy of PacBio sequence. In total, 113,970 CCS reads were obtained after quality-filtering and chimaera removal from 19 bacterial and 17 archaeal amplicons. Among them, two archaeal amplicons from USSF and HKSW were removed due to exceptionally low CCS reads (455 and 131), despite their high DNA quality after the two-step purification and confirmed gel imaging. The average CCS length for the archaeal and bacterial sequence was 1493.43 ± 81.31 and 1509.32 ± 43.63 bp, respectively. Further, we selected the SILVA rRNA gene database for being the most comprehensive 16S reference database with a vast collection of the environmental genome (Segota and Long, 2019). We retained singletons for diversity profiling, considering the advantage of full-length 16S gene sequencing. Thus, 10,490 archaeal and 45,775 bacterial OTUs were classified using 99% sequence similarity. Overall, the bacterial communities had the Shannon diversity (*H'*) and evenness (*J'*) of 5.27±0.88 and 0.78±0.10, respectively, while archaeal communities had values of 2.95±0.79 and 0.60±0.05, respectively for these indices. Similarly, Chao 1 richness, ACE richness, Inverse Simpson and Inverse Simpson evenness were higher in bacterial communities than archaeal ones (Table 4.2), indicating a higher alpha-diversity in bacterial communities over its counterpart. Also, the PCoA results showed that the archaeal communities mostly clustered according to the operating conditions, while the distribution of bacterial communities was more random (Fig. 4.1). For the identification, the results were varied among different microbial groups. Nearly half of the methanogen-related OTUs (abundance >1%) had percent identities higher than 98.7% (Yarza et al., 2014), achieving the identification at the species-level. Even lower percent identities were observed in most syntroph-related OTUs. The results still generate somewhat insights at species-level. On the other hand, the remaining fermenter-related OTUs could be identified at the

genus-level. In regard to profile the AD community, it is more practical to focus on identifying methanogens from archaeal amplicon at this stage.

Table 4.1 Summary of PacBio full-length 16S sequences conversion

		Bacteria			Archaea		
Operating Condition	Sample	Subreads	CCS reads	No. of Pass	Subreads	CCS reads	No. of Passes
High Salinity	HKTP	174984	11491	10	192684	9210	15
Pre-treatment	USSF	674129	25122	17	12728	455*	16
Thermophilic, >50°C	CALG	1146804	52792	14	297191	14262	15
	JPMR	182696	12080	10	570190	27525	15
	NEDR	82983	5265	10	200446	9961	14
	USOA	796904	36708	14	285454	13374	15
<30°C	JPHW	500927	19908	16	660552	27952	15
	USRA	210465	9429	14	217939	11424	14
Non-specific	HKSW	444212	16355	17	4364	131*	17
~40°C	JPHG	132816	8769	10	458315	22829	15
	JPSS	717825	26391	17	630942	23667	17
	JPTO	390437	14566	17	702570	26640	17
	JPTR	841400	31544	17	451245	17263	17
Mesophilic	CAII	276389	12684	14	307518	14942	15
	JPMU	213965	14041	10	863903	41665	15
	NEBR	284607	12894	14	243295	11904	15
	USTU	418929	19126	14	333704	16077	15
	USWA	355638	16189	14	654763	32721	14
Food Waste	NEAV	136603	6158	14	415828	20150	15

*** The two archaeal samples had been removed from analysis due to abnormal sequence count.**

Table 4.2 Summary of alpha diversity indices for bacterial and archaeal communities

Operating Condition	Inverse Simpson(1/D)		Shannon Diversity(H')		Inverse Simpson evenness(E')		Shannon evenness(J')		Good's coverage (%)		Chao 1 richness		ACE richness	
	Archaea	Bacteria	Archaea	Bacteria	Archaea	Bacteria	Archaea	Bacteria	Archaea	Bacteria	Archaea	Bacteria	Archaea	Bacteria
High Salinity	1.90	43.49	1.77	5.05	0.02	0.06	0.39	0.77	0.79	0.73	1746.36	4738.23	4237.19	11414.55
Pre-treatment	N/A	61.04	N/A	5.78	N/A	0.06	N/A	0.83	N/A	0.53	N/A	21877.91	N/A	66585.54
Thermophilic	6.77	17.27	2.93	4.54	0.05	0.03	0.61	0.70	0.71	0.72	2965.34	6984.22	11834.70	18305.75
<30°C	52.60	227.96	4.21	6.47	0.18	0.18	0.78	0.91	0.46	0.50	3771.02	10124.07	20236.82	27508.29
Non-specific	N/A	72.85	N/A	5.83	N/A	0.07	N/A	0.84	N/A	0.55	N/A	12413.44	N/A	37142.03
~40°C	8.85	71.89	3.07	5.44	0.07	0.08	0.64	0.81	0.72	0.65	2298.51	8093.35	8667.40	22235.02
Mesophilic	4.50	88.36	2.65	5.44	0.04	0.10	0.55	0.81	0.72	0.66	2345.77	8470.57	10071.50	21619.44
Food Waste	6.83	7.24	3.08	3.60	0.05	0.01	0.63	0.57	0.69	0.74	2210.09	19063.55	9372.80	60200.81

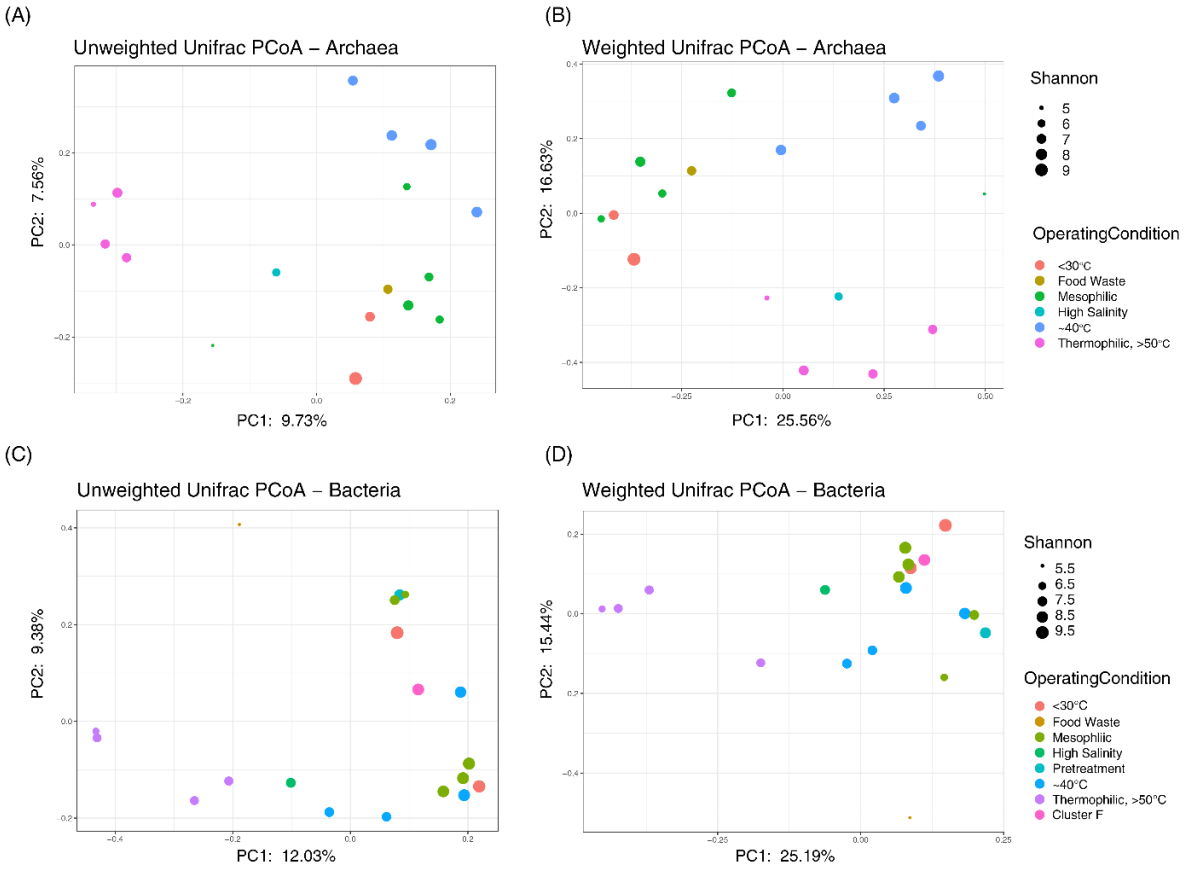


Figure 4.1 Principal coordinate analysis using Unifrac metric to measure the phylogenetic community distance and visualise the diversity of archaeal and bacterial communities. (A) & (C) are the unweighted Unifrac based on the abundance of the OTUs; (B) & (D) are the weighted Unifrac based on the presence and absence of the OTUs.

4.2.2 Characterisation of bacterial communities in various operating conditions

Covering 23 established and 30 candidate phyla in the List of Prokaryotic names with Standing in Nomenclature (Parte, 2018), fifty-three bacterial phyla (Fig. 4.2, *Appendix II* Table A2) were identified. The taxa recovered in the studies using Sanger sequencing (Chouari et al., 2005; Rivière et al., 2009; Zhang et al., 2014) were observed in our result. *Bacteroidetes*, *Firmicutes* and *Proteobacteria* with broad metabolic potentials were ubiquitous and abundant. Fortunately, due to the expansion of the constructed draft genome done, herein, somewhat metabolic interpretation of candidate phyla is possible. Plant-polysaccharide-degrading *Ca.* Patescibacteria (Taubert et al., 2019) was particularly abundant in pre-treatment digesters (19.71%), digesters operated at <30°C (21.49%) and HKSW (30.56%). The Candidate phyla, Acetothermia, Atribacteria, BRC1 (proposed name

Sumerlaeota), Cloacimonetes, Hydrogenedentes, Margulisbacteria, Modulibacteria and Zixibacteria, are likely involved in fermentative reaction and acetate and/or hydrogen formation (Balashov et al., 2018; Castelle et al., 2013; Eijsink et al., 2016; Hao et al., 2018; Kadnikov et al., 2019; Nobu et al., 2015; Pelletier et al., 2008a; Sekiguchi et al., 2015; Utami et al., 2019). Moreover, eleven candidate phyla with metabolically reconstructed genomic information are not associated with methanogenesis, while the genomic information of the remaining eleven is not available. Other than these, further information is unavailable.

Alternatively, to illustrate their taxonomic profiling further, we used the LEfSe analysis to characterise the AD microbiome specific to particular operating conditions. As shown in Fig. 4.3A, taxa highlighted by small circles and shadings in the cladogram were discriminately abundant in the respective conditions. To supplement it, we performed a co-occurrence network analysis to explore the potential interspecies interactions (Fig. 4.4). This generates fifty-eight subnetworks clustering into 433 nodes and 972 edges (*Appendix V* Figure A1). Using the all-against-all strategy, the LEfSe analysis revealed that a diverse group of the class *Clostridia*, including *Lanchnospiracea*, *Peptococcaceae*, *Ruminococcaceae*, and uncultured D8A-2 groups, had a more significant role in the thermophilic digesters (Fig. 4.3A). Representative species of these genera can survive under extreme conditions such as high-temperature environment (Yutin and Galperin, 2013). Notably, in one of the sub-networks, these *Clostridia*-related OTUs were co-present with a group of *Coprothermobacteraeota*-related OTUs (Fig. 4.4A). Together, we hypothesise that they are the potential volatile fatty acids (VFAs) producers (mainly propionate and acetate) in the thermophilic digesters. Another group of fermentative *Anaerolineae*-related OTUs was abundant in the high-salinity digester (Fig 4.3A). This lineage was also observed in a study with high abundance in high-salinity environments (Yang et al., 2016). In addition, Fig. 4.4B shows the close connection between the OTUs related to *Anaerolineae* and *Propionibacteriaceae* via *Rikenellaceae*. Such synergetic association could be due to one's end products can be used as the substrates for others (e.g., acetate, propionate, and alcohol), prior to methanogenesis (Graf, 2014; McIlroy et al., 2017; Stackebrandt, 2014). Other than the aforementioned, Fig. 4.4C showed the syntroph-related OTUs (e.g., *Smithella*,

Ca. Cloaimonas, and uncultured *Syntrophaceae*) were connected with the OTUs related to *Kiritimatiellaeota*, *Spirochaetes*, and *Acidobacteria* via BRC1-related OTUs in the mesophilic condition. They are also directly associated with the OTUs related to vadinHA17 group and other *Bacteroidetes* (Fig. 4.4D). Although the function of vadinHA17 group remains unknown, its wide appearance with its connection with syntrophs suggested that it involves in VFAs production.

In order to characterise the bacterial composition at species-level, taxonomic identification using BLAST was performed on the OTUs with a relative abundance >1%. Six out of sixty-one OTUs were identified with an average percent identity of 86.05 ± 7.65 %. Interestingly, three of the identified bacteria, *Coprothermobacter proteolyticus*, *Acetomicrobium mobile*, and *Defluviitoga tunisiensis,* were isolated from anaerobic digesters (Alexiev et al., 2014; Ben Hania et al., 2016; Hania et al., 2012). To fill the gap, a customised reference dataset derived from the specific environment is necessary, because the classical cultivation-based microbiological methods for digestion systems are challenging. In compliance with the full-length 16S rRNA sequences, constructing draft genome from cultivation-free whole metagenomic sequencing is one of the practical options. Therefore, such reference dataset for species level identification cross-mapping with operating conditions and performance would facilitate unprecedented digestion optimisation.

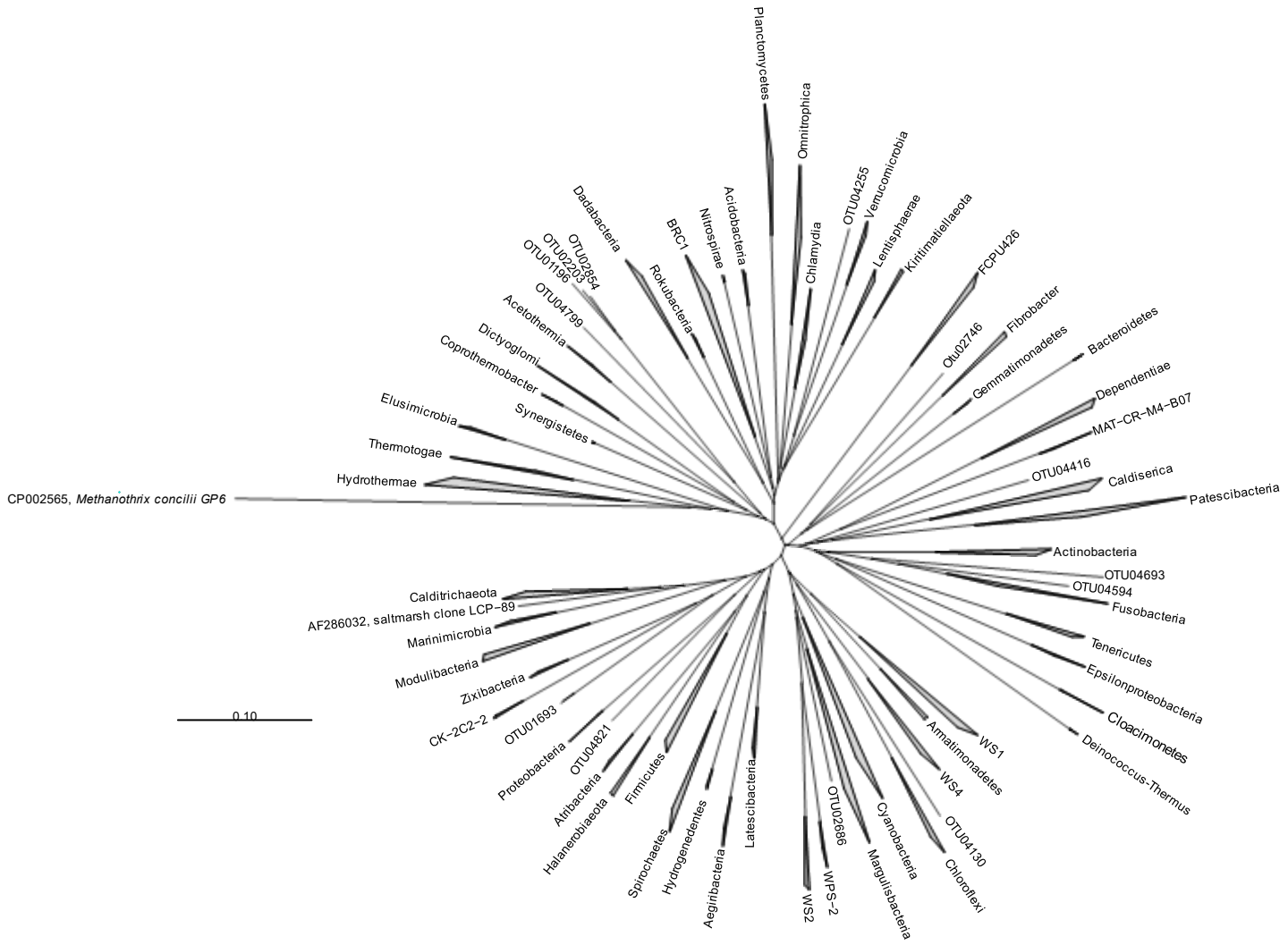


Figure 4.2 Full-length 16S rRNA-based phylogeny of 53 identified bacterial phyla. *Methanothrix concilii* was the outgroup for root.

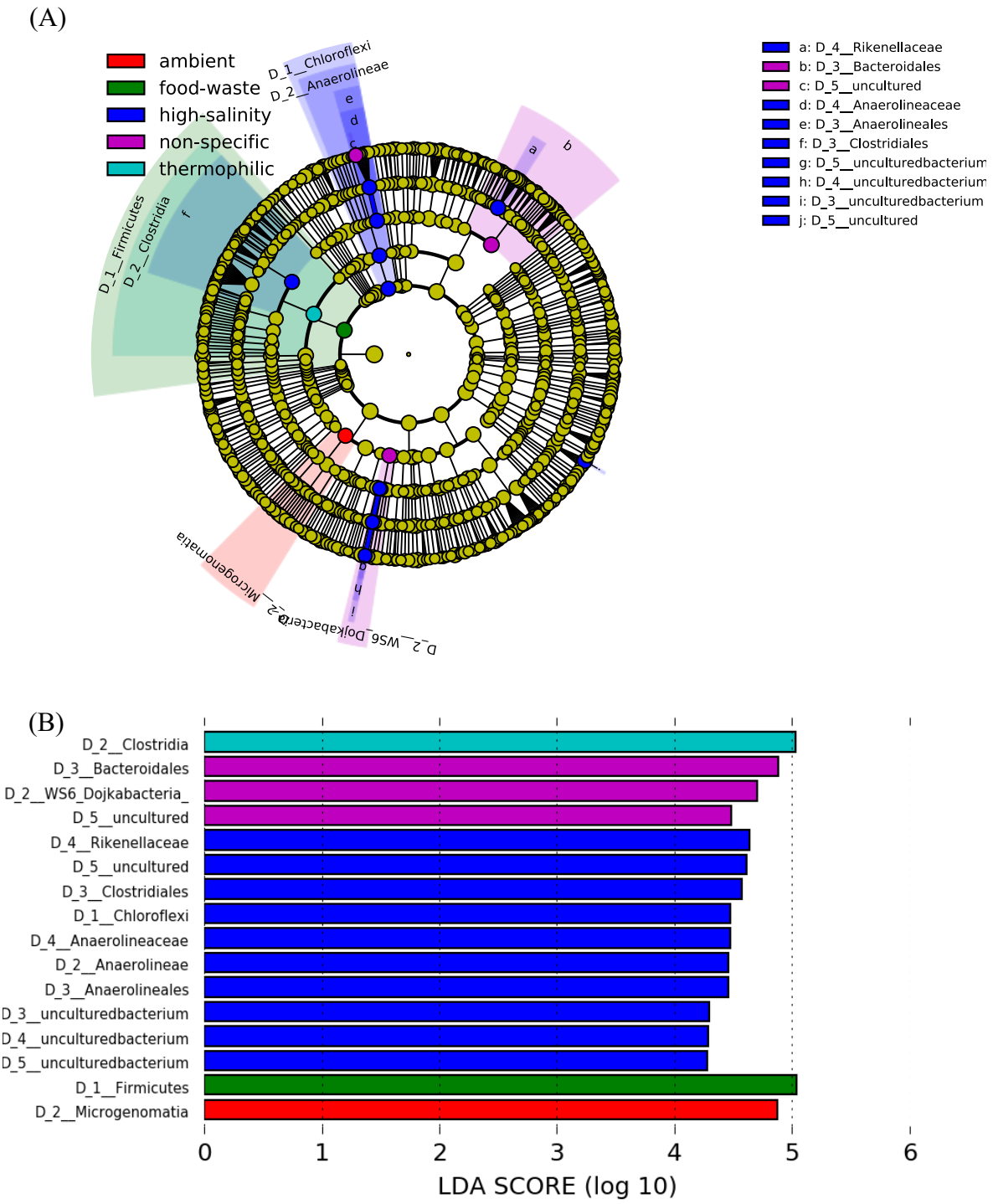


Figure 4.3 The taxonomic difference among different operating conditions evaluated using LEfSe. Panel A shows a cladogram illustrating the statistically significant taxa in different operating conditions. Panel B shows the LDA scores of the differentially abundant taxa.

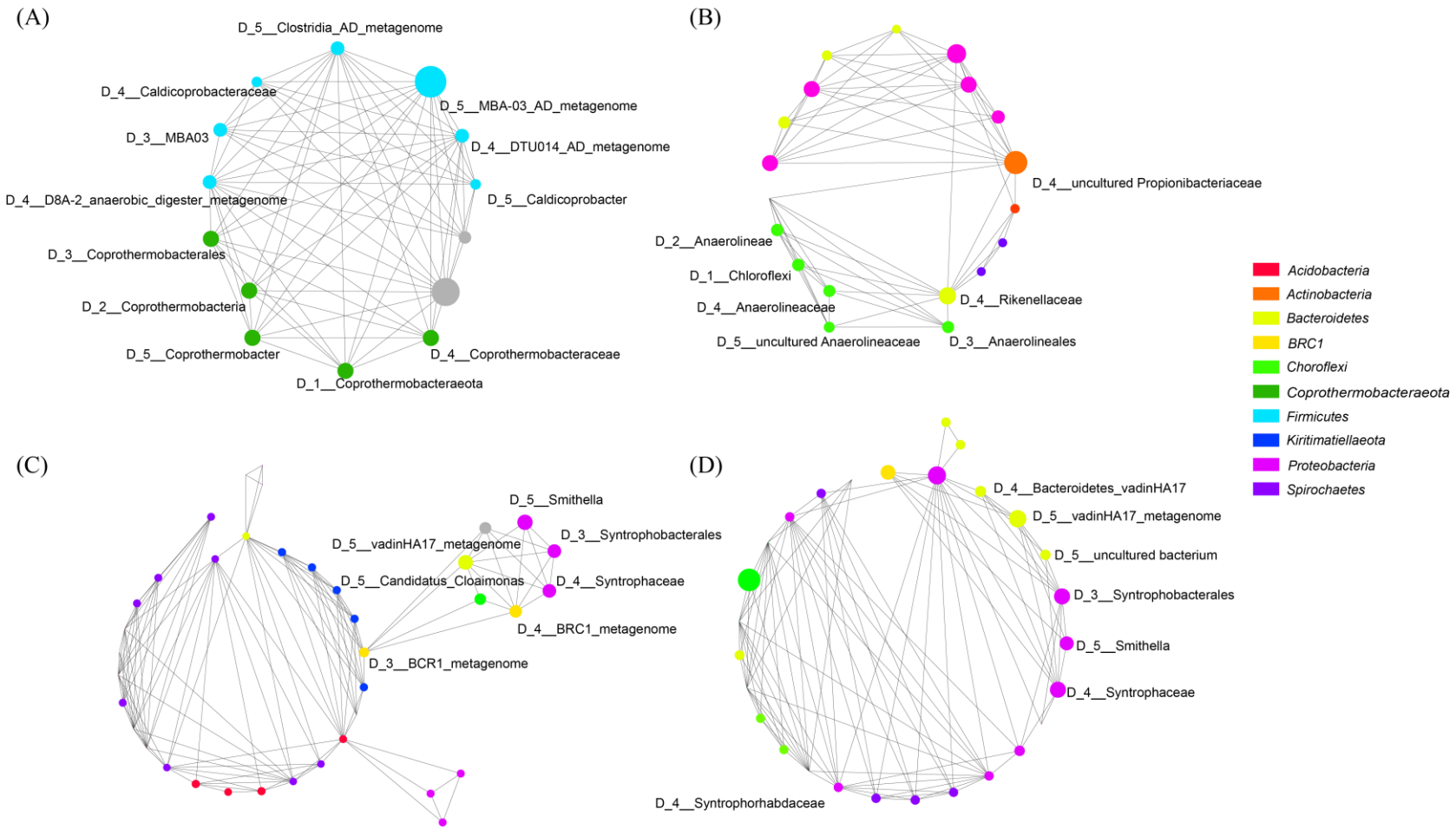


Figure 4.4 Four subnetworks were extracted from the network constructed using CoNet, clustered by CytoCluster, and visualised via Cytoscape. Each node denotes an OTU at 99% sequence similarity. The size of the nodes is proportional to their relative abundance. Colour legend is set at phylum level. Discussed OTUs are labelled with their identified taxonomy. The labels "D_1" to "D_5" designate taxonomic level from phylum to genus levels. Panel A contains 14 nodes and 71 edges. Panel B contains 19 nodes and 63 edges. Panel C contains 40 nodes and 135 edges. Panel D contains 28 nodes and 101 edges.

4.2.3 Characterisation of syntrophs population

Among the 1601 syntroph-related OTUs detected, they were associated with the following phyla: *Proteobacteria* (1339 OTUs), *Firmicutes* (229 OTUs), *Synergistetes* (9 OTUs) and *Ca.* Cloacimonetes (24 OTUs). Syntroph-related OTUs represented 4.69 ± 3.88 % of the bacterial population with 10.17% in mesophilic digesters and 10.04% in digesters operated at ~40°C (Table S6). Noticeably, we also observed that their abundance in high-salinity (2.70%) and thermophilic (2.94%) conditions were low. In an attempt to obtain taxonomic identification at species-level, we selected syntroph-related OTUs clustered in the same lineage with cultivated species in the phylogenetic tree for BLAST. As a result, seven out of forty-six of those syntroph-related OTUs were identified at species-level (Table S4B). The average percent identity for taxonomic identification was 96.15 ± 2.09 %. Overall, most syntroph-related OTUs were classified under fourteen genera; the remaining OTUs formed uncultured lineages within the families *Syntrophomonadaceae*, *Syntrophaceae*, and *Syntrophobacteraceae* (Fig. 4.5A).

Within *Proteobacteria*, mesophilic *Smithella* was the most abundant genus, accounting for 38.3 to 47.5% of the syntrophic population in the pre-treatment, digesters operated at <30 °C, ~40 °C, and mesophilic condition. Six uncultured lineages within the genus *Smithella* were detected under various conditions, indicating its universal presence (Fig. 4.5B). Its phylogenetic diversity could contribute to its universality. Specifically, over 90% of the *Smithella*-related OTUs were closely related to *Smithella* strains D17 and M82, and only one is closely related to *Smithella propionica.* Notably, *S. propionica* possesses a unique propionate degradation pathway (3 mol of hydrogen produced per mol of propionate degraded), distinguished from the methylmalonyl-CoA pathway (1 mol of hydrogen per mol of propionate consumed) (Dolfing J, 1987; Koch et al., 1983). Even so, the lack of key genes like succinate dehydrogenase in strain M82 indicates its inability to degrade propionate (Qin et al., 2017), as was *Smithella* strain D17 (Tan et al., 2014). Instead, strains D17 and M82 may be involved in alkanes degradation based on their draft genomes (Embree et al., 2014; Qin et al., 2017; Wawrik et al., 2016). This interesting observation is divergent from our previous understanding of the genus

Smithella metabolising in methanogenic propionate degradation (Bergey et al., 2005). With that being, both strains still help to maintain a thermodynamic favourable environment for methanogenesis to occur (Qin et al., 2017; Tan et al., 2014). As for the OTUs formed five uncultured lineages within the family *Syntrophaceae*, they are phylogenetically closely related to *Syntrophus* (Fig. 4.5C), known for benzoate and fatty acids degradation (Jackson et al., 1999; Liu et al., 1999). In addition, the OTUs related to the aromatic compound-degrading *Syntrophorhabdus* (Qiu et al., 2008) were also in all digesters except for the co-digestion of food waste. On the opposite, the OTUs related to acetate-oxidising *Syntrophaceticus* (Westerholm et al., 2010) and fatty-acid-oxidising *Syntrophothermus* (Sekiguchi et al., 2000) were uniquely detected in thermophilic digesters at low abundance (<0.1%). Other low abundance syntroph-related OTUs were identified under the *Proteobacteria* genera, including *Desulfovibrio*, *Geobacter*, *Pelobacter*, *Syntrophus*, and *Syntrophobacter*, were in low relative abundance (Table 4.3). These results show the diverse survival niches of syntrophs in the various engineered systems sequenced.

The abundance of *Ca.* Cloacimonas-related OTUs was seconded to *Smithella*-related OTUs in the syntrophic population. Within this genus, OTU 12 identified as *Ca.* Cloacimonas acidaminovorans (99.74% identity) was abundant at ~40°C (2.14% out of 4.27%) and mesophilic conditions (1.96% out of 3.79%). The implication of its pivotal role in propionate-degradation and amino acid fermentation (Pelletier et al., 2008b) were supported by a few of studies (Calusinska et al., 2018; Solli et al., 2014; Xia et al., 2018; Ziels et al., 2018). Moreover, three OTUs within *Synergistets* were identified as *Thermanaerovibrio acidaminovorans* (Chovatia et al., 2009) with a percent identity of 99.93%, 99.10% and 97.15%. *T. acidaminovorans* ferments a range of amino acids to acetate and propionate with hydrogen as the only reduced end-product (Guangsheng et al., 1992). They were present in high salinity, thermophilic, and ~40°C-operated digesters at low abundance (<0.5%). Only the fatty acid degrading *Syntrophomonas*-related OTUs were detected crossing all the operating conditions. *Syntrophomonas sapovorans* and *Syntrophomonas palmitatica* were identified in the mesophilic digesters with a percentage identity of 98.26% and 100%, respectively, potentially utilise fatty acids

with 4 to 18 carbon atoms (Hatamoto et al., 2007). Inclusively, our results show the importance of syntrophy in amino acids degradation in digestion.

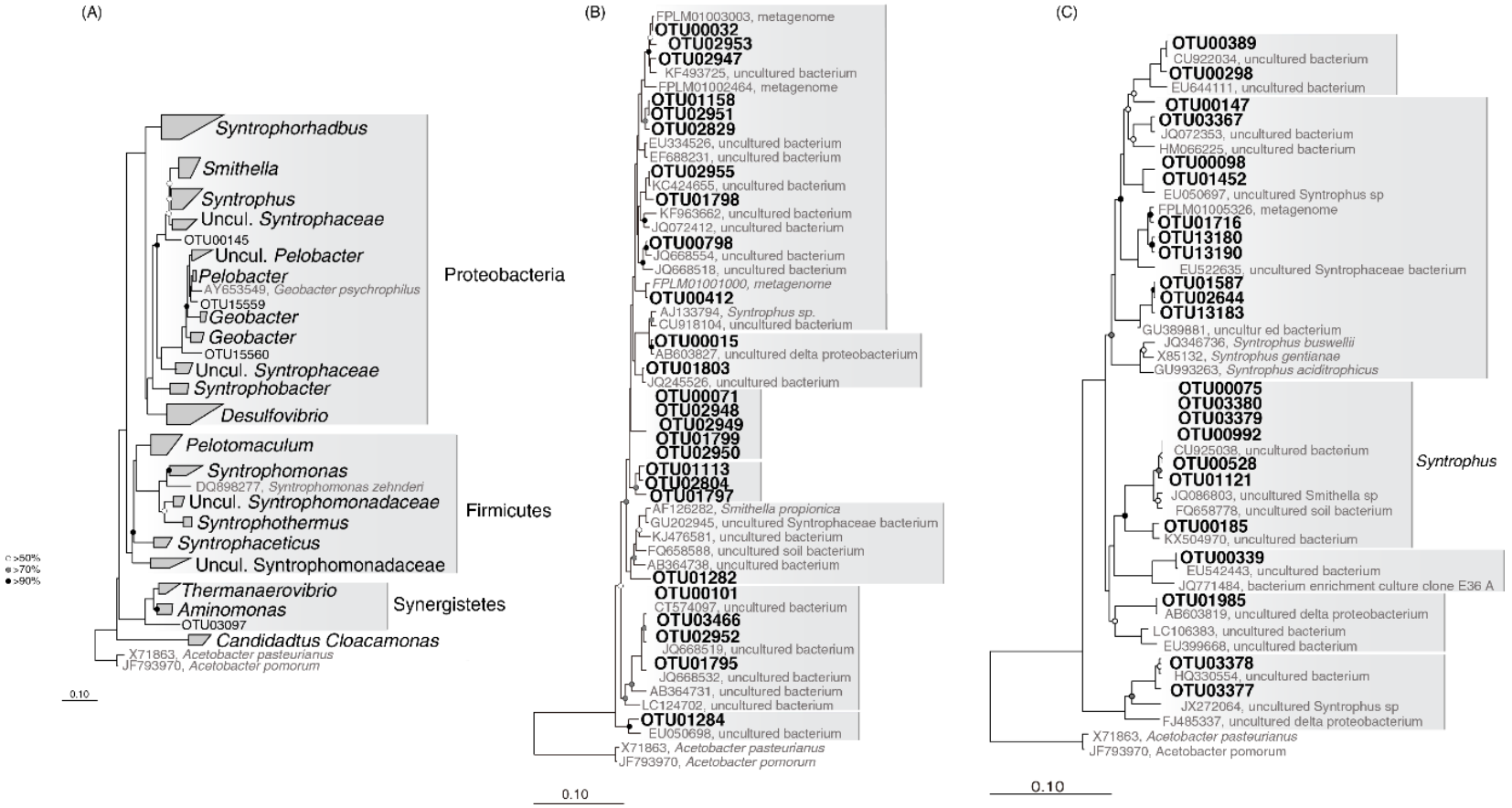


Figure 4.5 Full-length 16S rRNA-based phylogeny of (a) identified OTUs related to syntrophs in genus-level, (b) 7 lineages of *Smithella* related OTUs within the genus-level, (c) 5 lineages of OTUs related to uncultured *Syntrophaceae* within the family-level. *Acetobacter pasteurianus* and *Acetobacter pomorum* are the outgroups for root. Bootstrap values >90% (black), 75% (gray) and 50% (white) are indicated

Table 4.3 Relative abundance (%) of identified syntrophs related genera by Mothur

Genus	High Salinity	Pre-treatment	Thermophilic	<30°C	HKSW	~40°C	Mesophilic	Food Waste
Candidatus Cloacimonas	1.7043	1.5181	0.0012	1.0084	0.7528	4.2718	3.7916	0
Pelotomaculum	0	0	0.235	0	0	0.0315	0.0144	0
Syntrophomonas	0.4427	0.0123	0.1113	0.3812	0.0525	0.067	0.0814	0.0614
Syntrophothermus	0	0	0.0048	0	0	0	0	0
Uncul. *Syntrophomonadaceae*	0	0	0.949	0.0668	0	0.0123	0	0
Syntrophaceticu	0	0	0.011	0	0	0	0	0
Syntrophorhabdus	0.1992	0.3826	0.356	0.1476	0.035	0.7728	0.3418	0
Desulfovibrio	0.0221	0	0.0091	0.083	0	0.0127	0.0213	0
Geobacter	0	0.0247	0	0.0794	0	0.0084	0.0611	0
Smithella	0	2.666	0	2.0369	0	3.891	4.3673	0
Syntrophus	0	0	0	0.232	0	0.1535	0.1029	0
Uncul. *Syntrophaceae*	0	1.8761	1.2431	0.2568	0	0.9301	1.1866	0
Syntrophobacter	0.0885	0	0	0	0	0.0108	0.0653	0
Uncul. *Syntrophobacteraceae*	0	0	0	0	0	0	0.0075	0
Aminomonas	0.0221	0	0	0	0	0	0	0
Thermanaerovibrio	0.2213	0	0.0219	0	0	0.0059	0	0
Total	**2.7003**	**6.4799**	**2.9423**	**4.2921**	**0.8403**	**10.1678**	**10.0413**	**0.0614**

4.2.4 Diversity of archaeal population

With archaea-specific primers used, the majority of the OTUs (10,446 out of 10,490 OTUs) could represent 99.76% of total archaeal sequences. Most were associated with the phylum *Euryarchaeota*. So far, all the cultivated methanogens are classified to seven orders within *Euryarchaeota*. Among these seven orders, three orders were not observed in the present dataset, i.e., *Methanocelles*, *Methanopyrales*, and *Methanococcales* (Fig. 4.6A). *Methanocelles* species have been obtained from rice paddy fields (Lu and Conrad, 2005, Lü and Lu, 2012; Sakai et al., 2008, 2010), whereas isolates of *Methanopyrales* and *Methanococcales* were all obtained from marine environments, especially geothermally heated ecosystems, e.g., deep-sea hydrothermal vents (Huber et al., 1982; Kurr et al., 1991; Stewart et al., 2015; Takai et al., 2004). It is, thus, expected the absence of these three orders in anaerobic digesters conditions. The remaining four orders (i.e., *Methanomicrobiales*, *Methanosarcinales*, *Methanobacteriales*, and *Methanomassiliicoccales*) contained OTUs were detected in this study (Fig. 3B). Most OTUs were assigned to 14 different genera with cultivated representatives. Among them, *Methanothrix* (4450 OTUs), *Methanolinea* (1673 OTUs), and *Methanosarcina* (1086 OTUs) were the most abundant, accounting for 38.14%, 17.92%, and 12.33% of total archaeal sequences, respectively. Distinctly, there are sequences related to two candidate genera found in highly enriched cultures have not been isolated yet, i.e., *Ca.* Methanogranum and *Ca.* Methanoplasma. Both belong to the order *Methanomassiliicoccales* and include two lineages lacking cultivated representatives within the family *Methanomethylophilaceae* and *Methanomassiliicoccaceae*. The OTUs associated with *Methanomassiliicoccales* were commonly observed in our study. They are abundant in NEDR (>50°C) and ESRA (<30°C) with 11.91% and 11.27% of the archaeal population, respectively. *Methanomassiliicoccales* require hydrogen for methylated compounds metabolisation, such as methanol, methylamines and methylated thiols (Adam et al., 2017), suggesting its linking carbon, nitrogen and sulfur cycles in methanogenesis. Besides the aforementioned orders being cultivated, a class-level clade, *Ca.* Methanofastidiosales (previously known as WSA2 or ArcI) (Chouari et al., 2005; Hugenholtz, 2002) is capable of performing methanogenesis via methylated thiol reduction (Nobu et al., 2016). Many OTUs (1,048 OTUs) related to this clade were detected with

relative abundances ranging from 7.26% to 40.23% among the digesters sampled (Fig. S2A). The OTUs in this clade were phylogenetically diverse and spanned six genus or family level lineages (shadowed in Fig. 4.6B). Our data reveal that *Ca.* Methanofastidiosales is a significant player in anaerobic digestion, albeit being overlooked previously.

In addition to *Euryarchaeota*, the sequences related to other archaeal phyla (Fig. 4.6C) were observed. Twenty-two observed OTUs forming five genus and family level lineages are affiliated with *Ca.* Bathyarchaeota, containing members capable of methanogenesis based on a metagenomics study (Evans et al., 2015). While the members of *Ca.* Bathyarchaeota are mostly associated with natural environments, e.g., marine and freshwater ecosystems (Zhou et al., 2018), their presence in engineered systems warrants further investigations to understand the metabolism fully. Likewise, the phylum *Ca.* Verstraetearchaeota, not yet been detected in digesters previously, was observed in our mesophilic digesters. It is believed to be one of the few examples of non-Euryarchaeota methanogens (Vanwonterghem et al., 2016). Seven OTUs were closely related to the proposed methylotrophic *Ca.* Methanomethylicus in this phylum. In addition, fourteen OTUs associated with *Ca.* Lokiarchaeota and one OTU associated with *Ca.* Odinarchaeota were detected, but no evidence is available about their methanogenic capacity or ecological functions in digesters. Collectively, with the high resolution of the full-length sequences, a few of the overlooked methanogens were identified. Further attention is required to understand their roles in AD.

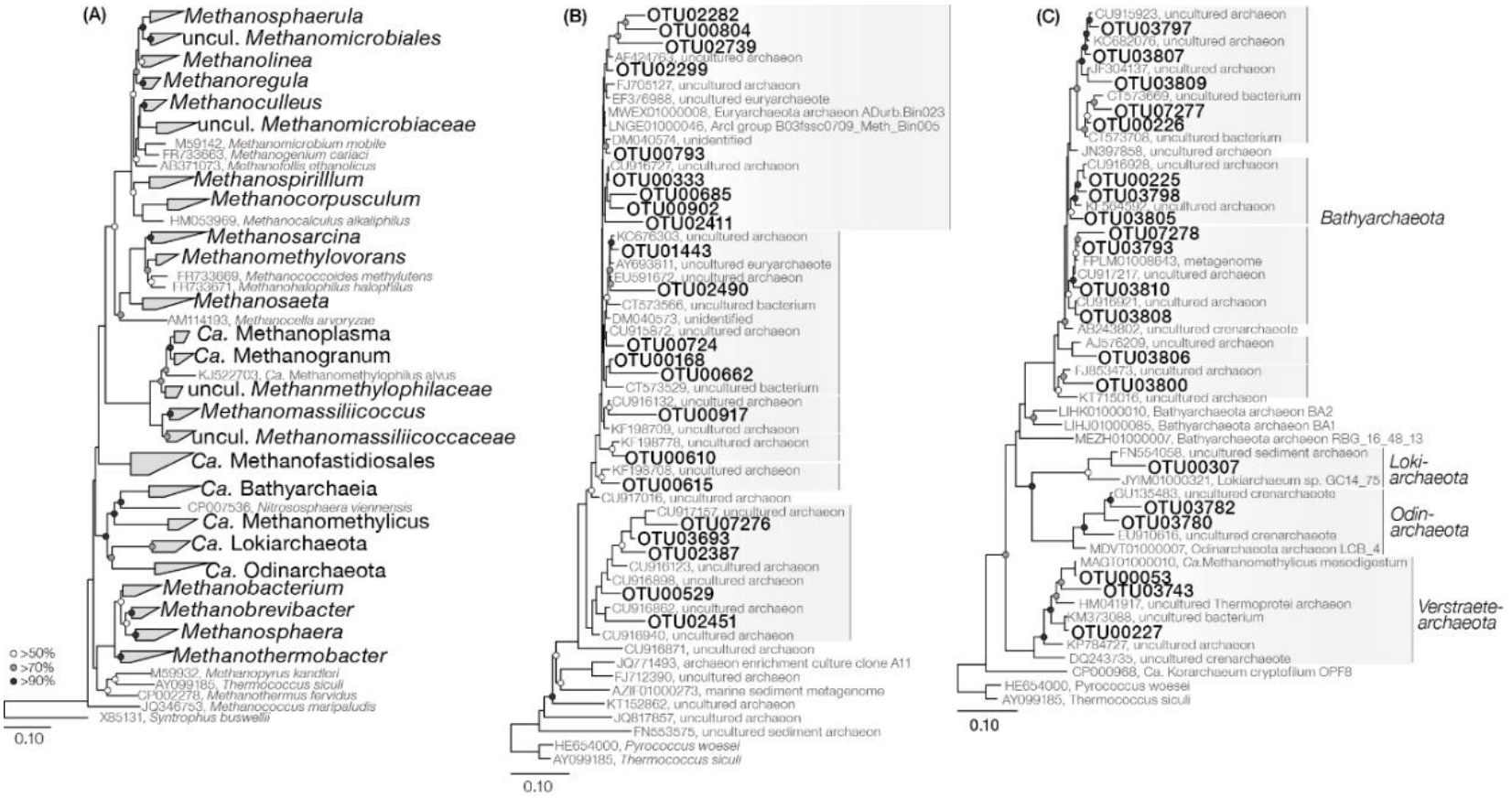


Figure 4.6 Full-length 16S rRNA-based phylogeny of (a) OTUs related to methanogens at phylum or genus level, (b) *Ca.* Methanofastidiosales related OTUs and (c) OTUs relatd to non-Euryarchaeota. *Syntrophus buswellii, Pyrococcus woesei* and *Thermococcus siculi* are the the outgroups for root. Bootstrap values >90% (black), 75% (gray) and 50% (white) are indicated.

4.2.5 Operation-driven methanogenic species

Considering species-level identification performed on 35 abundant archaeal OTUs with relative abundance >1% (Fig. 4.7, *Appendix III* Table A3b), all had high alignment score (2451.17 ± 96.45), query cover percentage (99.86 ± 0.69 %) and percent identity (98.35 ± 1.26 %) to known archaeal species. This allows further exploration of their potential functions such as methanogenic pathways, optimal growth conditions and substrate utilisation (*Appendix IV* Table A4). Below, we highlight three species by cross-referencing the dominance of certain methanogens with the digester operating conditions.

4.2.5.1 *Methanosarcina horonobensis* in high salinity digester

An OTU related to *Methanosarcina horonobensis* with 98.67% identity percentage was identified and represented 72.51% of the archaeal population in the high salinity digester (HKTP). As a member of *Methanosarcina*, it is generally assumed that it can conduct both hydrogenotrophic and acetoclastic methanogenesis. On the contrary, a physiology study reported that no methane was produced with H_2 and CO_2 by *M. horonobensis* (Shimizu et al., 2011). Although, according to the genome mapping in Kyoto Encyclopedia of Genes and Genomes, *M. horonobensis* contains the key enzymes of utilising H_2 and CO_2 for hydrogenotrophic methanogenesis, these genes were likely not transcribed. Another interesting finding reported by Shimizu et al. (2011) is that *M. horonobensis* grows optimally at 1.05% NaCl and can tolerate up to 1.75% NaCl, while our general understanding is that high salinity has an inhibitory effect on methanogenesis (Liu and Boone, 1991; Patell and Roth, 1977; Pattnaik et al., 2000). It can also grow on dimethylsulfide (Shimizu et al., 2011), which is plentiful in HKTP. In any event, it is rational to link that HKTP at a salinity of 0.8% and the high content of dimethylsulfide might serve as the selection measures for the dominance of *M. horonobensis* detected in this study. This observation of such a methanogen with the ability to thrive in a particular environment can serve as a guide for inoculum selection when treating specific substrate.

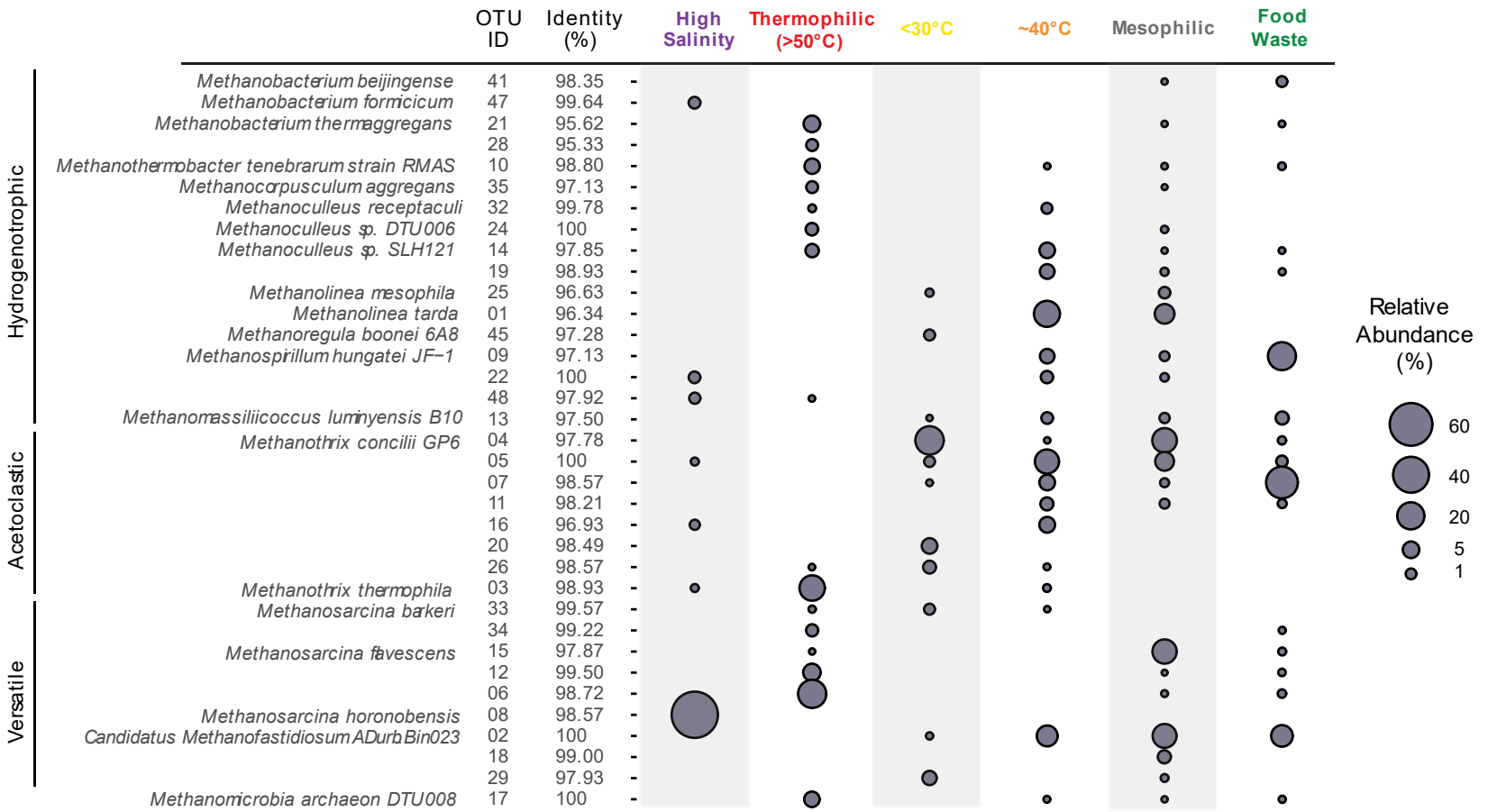


Figure 4.7 Relative abundance of abundant methanogenic OTUs (relative abundance >1%) in species level. The sequence identity was determined using BLAST against SILVA 132 Reference database.

4.2.5.2 *Methanosarcina flavescens* in the thermophilic digesters

Methanosarcina flavescens was the most abundant species (27.96%) in the thermophilic digesters, along with other thermophilic species including *Methanobacterium thermaggregans*, *Methanothermobacter tenebrarum*, *Methanoculleus receptaculi,* and *Methanothrix thermophila* (Cheng et al., 2008; Kamagata et al., 2009; Schupp et al., 2018; Takahashi et al., 2012). *Methanosarcina* contains cytochrome *c*, a high intermembrane electron transfer medium, serving as a distinguished character compared with other methanogens. It is, thus, postulated that cytochrome *c*-containing *Methanosarcina* could withstand high hydrogen partial pressures, compared to methanogens without cytochrome *c* (Thauer et al., 2008). As thermophilic conditions facilitate the production of hydrogen (Mir et al., 2016), it provides a favourable environment for *M. flavescens* for its abundant presence in thermophilic environments, even over its optimal growth temperature at 45°C (Kern et al., 2016).

4.2.5.3 Universal Methanothrix concilii

Aceticlastic *Methanothrix concilii* sequences (Olson et al., 2011; Patell and Sprott, 1990) were present in digesters under all operating conditions with an average abundance of 18.63%, accounting nearly half of the *Methanothrix* population, except for that in thermophilic digesters. Its universal presence leads us to the genome plasticity of *Methanothrix*. By elongating the genome, *Methanothrix* enriches the favourable phenotypes to adapt to the changing environment (Maeder et al., 2006). *Methanothrix* contains cytochrome c as does *Methanosarina* and both genera are often reported to involve direct interspecies electron transfer (DIET) with *Geobacter* (Holmes et al., 2017; Yang et al., 2019). Moreover, a recent metatranscriptomics analysis showed that the DIET pathway is associated with ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) mediated CO_2 reduction and cell synthesis in *Methanothrix concilii,* resulting in a high and efficient methane yield (Olson et al., 2011; Patell and Sprott, 1990). This pathway was found to be active in a highly efficient pilot-scale anaerobic fluidised-bed membrane bioreactor using real sewage as feed, demonstrating that it is possible to generate energy from such dilute stream (Yang et al., 2019). Whether the enzyme systems

involved in DIET are unclear. However, CO_2 reduction to methane via DIET requires coupling with exoelectrogenic *Geobacter* species. In the samples of this study, *Geobacter* was low in abundance, suggesting its methane yield was mainly via the acetoclastic pathway.

The low abundance of *Methanothrix* and high abundance of *Methanosarcina* is an indication of high acetate concentration environment (< 70 mg/L) in the thermophilic digesters, as *Methanothrix* has a higher substrate affinity but a lower specific utilization rate (Speece, 1996).

4.3 Chapter Summary

This study demonstrated the methodology in dealing with low-quality environmental DNA samples for the full-length 16S rRNA gene amplicon sequencing using PacBio Sequel. Results provide a higher resolution on the identification, facilitating the description of microbial composition with species-level specificity without bias from different primer sets in the short-read platforms. The merit of using the full-length 16s rRNA gene amplicon sequencing as a survey tool on AD diversity profile discloses various operation-driven microbiomes, allows an in-depth phylogenetic tree construction, and the inference of optimal growth conditions on the microbial composition. Through the expansion of the genome database and improvement of the sequencing chemicals, the application of the full-length sequencing on environmental samples can be even more promising.

Chapter 5 Comparison between the NGS and TGS 16S amplicon sequencing results

5.1 Overview

Although the advantage of using PacBio SMRT sequencing over NGS platform in phylogenetic resolution was discussed in chapter 4, being able to reach to species level identification might not always be the priority depending on the research interest. The stringent DNA input requirement (especially for environmental sample), expensive sequencing reagents and services of PacBio SMRT sequencing would often come into considerations, as these criteria hinder the application of SMRT sequencing as a high-throughput option for statistic evaluation. For large scale microbiome survey, for example, our previous study surveyed sludge samples from 148 digesters in order to obtain statistically significant data for concluding that the AD microbiome is operation-driven (Mei et al., 2017). If the same sample size was sequenced using the SMRT sequencing, it would cost 2 to 10 times more than using Illumina, depending on the sample quality, multiplex strategy and the required coverage.

In this chapter, the discrepancies of 16S rRNA gene amplicon sequencing between the NGS and TGS platforms were evaluated. NGS data from the previous study were reanalysed and compared to the TGS data. Beside the phylogenetic resolution discussed in chapter 4, the difference in taxonomic classification, coverage, alpha and beta diversity were discussed. Twelve DNA samples from chapter 4 were sequenced by the NGS Illumina MiSeq platform using v4-v5 primers, covering six operating conditions. As both sets of universal primers, 515F & 909R and 27F & 1492R were not effective in hybridising *Euryarchaeota*, only bacteria classification was discussed in this chapter. The discrepancy in taxonomic classification at phylum level was observed with uniform trend according to the thermophilic and mesophilic samples. Additional OTUs were obtained from the TGS data set. CCS

reads were not saturated in all the samples, indicating that the coverage of TGS data set was not optimised. The difference between alpha and beta diversity was insignificant; digester samples were clustered similarly for both data sets.

5.2 Results and discussion

5.2.1 Sequence quality

The total sequence reads output should be one of the considerations because the high-throughput feature of NGS allows the determination of microbial composition. For TGS to be relevant in the current industry, it has to provide an equivalent or higher capacity in determining microbial composition while providing higher resolution. While the high phylogenetic resolution of full-length sequences had been demonstrated by several studies and in chapter 4, it is still in question that how fit will the full-length sequences be in determining the microbial composition in a specific environment. In a glimpse of the sequence reads output (table 5.1), the PacBio SMRT sequencing subreads were comparable to the Illumina raw reads in terms of copy number. However, we cannot directly compare the sequence reads output between the two platforms as PacBio employs the CCS approach to improve the accuracy of the subreads, while the number of CCS reads generated from the subreads depends on the number of passes. Currently, there is no universal rule in comparing the sequence reads output between Illumina and PacBio platforms.

Surprisingly, the percentage difference of sequence classified between full-length (FL) sequences and short-read (SR) sequences was marginal (Table 5.2), which is contradictive to the result of a previous study (Singer et al., 2016). Noted that sequences could not be classified at kingdom level were discarded. The possible explanation could be the specificity of the AD microbiome which limited the classification of both data sets, and the fact that species level identification was not performed on the SR data set.

Table 5.1 Summary of sequence reads

Operating Conditions	Sample	Illumina (SR)	PacBio (FL)		
		Raw Reads	Subreads	CCS reads	No. of Pass
Food Waste	NEAV	73148	136603	6158	14
High Salinity	HKTP	123541	174984	11491	10
Thermophilic (>50°C)	JPMR	105995	182696	12080	10
	USOA	116170	796904	36708	14
	NEDR	221040	82983	5265	10
	CALG	87557	1146804	52792	14
<30°C	USRA	96148	210465	9429	14
~40°C	JPHG	85680	132816	8769	10
Mesophilic	JPMU	70806	213965	14041	10
	USWA	130933	355638	16189	14
	USTU	157812	418929	19126	14
	NEBR	78382	284607	12894	14
	CAII	118379	276389	12684	14

Table 5.2 Percentage of SR and FL sequences classified at various taxonomic levels

Taxonomic Level	% SR sequences classifed	% FL sequences classified
Phylum	99.2	99.5
Class	99.0	99.4
Order	98.6	98.6
Family	97.9	98.0
Genus	94.6	97.1

5.2.2 Discrepancy in bacterial taxonomic classification

The overall community composition between the FL and SR data sets was moderately similar (Spearman's rank correlation, $R^2 = 0.78$, $p = 0.0004$) The taxonomic classification of Illumina MiSeq and PacBio Sequel data sets were compared at the phylum level. The relative abundance of the ten most abundant phyla was illustrated (Fig. 5.1) and a heat map (Table 5.3) was used to compare the discrepancy.

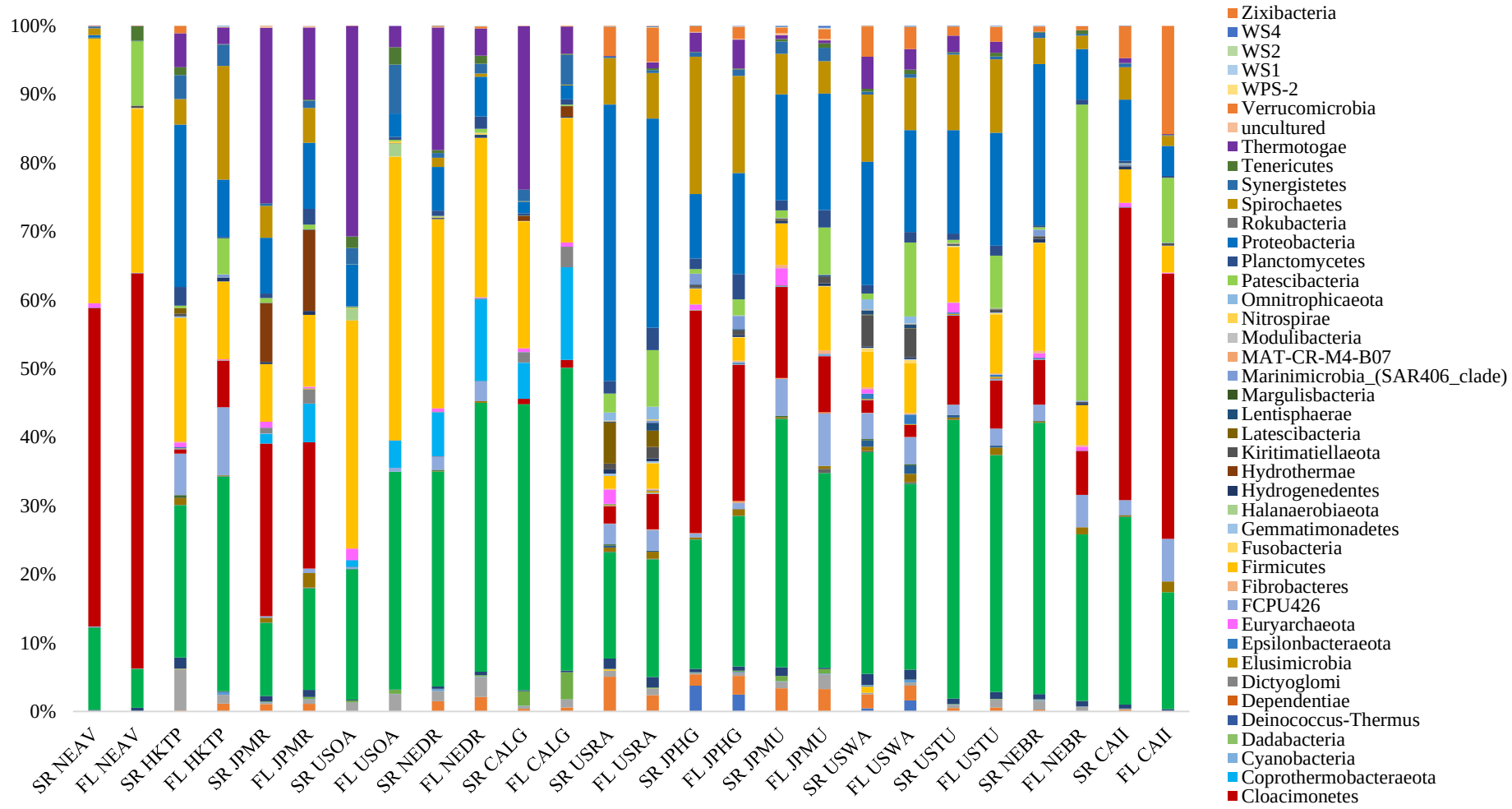


Figure 5.1 Taxonomic classification of the FL and SR results at the phylum level

Table 5.3 Discrepancy heat map of the top ten phyla between FL and SR data sets

	Food-waste	High Salinity	Thermophilic				<30°C	~40°C	Mesophilic				
Taxon	NEAV	HKTP	JPMR	USOA	NEDR	CALG	USRA	JPHG	JPMU	USWA	USTU	NEBR	CAII
Bacterioidetes	-6.35	9.11	4.15	12.76	7.85	2.51	1.70	3.15	-7.80	-5.08	-6.06	-15.23	-10.35
Ca. Cloacimonetes	11.19	6.13	-6.80	0.00	0.02	0.33	2.62	-12.57	-5.09	-0.11	-5.91	-0.12	-4.05
Firmicutes	-14.70	-6.67	2.03	8.10	-4.36	-0.38	1.86	1.13	3.28	2.09	0.56	-9.93	-0.94
Ca. Patescibacters	9.35	4.92	0.01	0.03	0.50	0.18	5.50	1.70	5.78	9.94	7.17	42.87	9.42
Proteobacteria	-0.30	-15.30	1.54	-2.63	-0.56	0.33	-9.74	5.34	1.54	-3.05	1.38	-16.13	-4.58
Spirochaetes	-0.96	12.86	0.37	-0.02	-0.89	0.04	-0.06	-5.86	-1.22	-2.11	-0.30	-1.83	-3.28
Thermotogae	-0.04	-2.54	-15.03	-27.55	-13.95	-19.85	0.79	1.39	-0.18	-1.64	-0.82	-0.02	-0.64
Verrucomicrobia	0.01	-0.97	0.05	-0.02	0.23	0.03	0.76	0.89	0.61	-1.15	0.80	-0.18	11.11
Ca. Coprothermobacteraeote	0.04	0.00	4.21	2.98	5.45	8.24	0.00	0.00	0.00	0.00	0.00	0.00	-0.06
Chloroflexi	-0.08	3.86	0.46	0.24	0.97	0.01	0.08	0.35	2.22	0.15	0.99	2.33	4.05

*The percentages were calculated by subtracting the relative abundance of FL result with the SR result

There was a noticeable discrepancy in thermophilic phyla *Thermotogae* and *Coprothermobacteraeota* in the thermophilic digesters. The *Thermotogae* had a higher representation in the SR result from 15.03% to 27.55%; it was the opposite for *Ca*. Coprothermobacteraeota from 2.98% to 8.24%. *Bacterioidetes* was less represented in the mesophilic samples. These discrepancies were observed uniformly according to the operating conditions.

Figure 5.2 and 5.3 illustrated the compositions of the abundant OTUs from the FL and SR results with relative abundance >2%. Within the abundant OTUs, 15 additional OTUs were observed in the FL data set (FL: 73; SR: 58). This indicated the FL 16S gene amplicon sequencing provided a higher specificity in microbial identification, despite the use of 97% similarity. The classification results between FL and SR data sets was partially similar. For instance, group W27, belonged to phylum *Ca.* Cloacimonetes, was assigned similarly in digester 38 NEAV (FL: 51.46%; SR: 46.04%). *Ca*. Cloacimonetes is a candidate phylum, the metabolic pathway reconstruction from metagenomics analysis indicated it is a propionate-degrading consortium (Ahlert et al., 2016; Pelletier et al., 2008a). Fermentative *Lentimicrobiaceae* (Sun et al., 2016) was also distributed similarly in 5 digesters between two data sets.

On the other hand, *Defluviitoga* was not assigned, and *Fervidobacterium* was assigned with less relative abundance in the FL result. The two genera are sulfur-reducing bacteria belonging to the phylum *Thermotogae*. There was also a specific genus, *Ca.* Coprothermobacter, only assigned in the FL result(s) with higher abundance. This genus was previously misclassified under the phylum *Firmicutes*; it was reclassified under a new phylum *Ca*. Coprothermobacterota proposed in 2018 (Pavan et al., 2018). *Coprothermobacter* has a strong association in the anaerobic digestion process by producing intracellular and extracellular protease responsible for protein and peptide degradation, and produce hydrogen (Gagliano et al., 2015). *Ca.* Coprothermobacter was assigned as the second most abundant OTU in the FL result, and it was the same for *Fervidobacterium* in the SR result. It is worth

noting that the two genera were described as “sister” in phylogenetic placement, explaining the significant discrepancy between *Thermotogae* and *Ca.* Coprothermobacterota. Both genera can survive under anaerobic and thermophilic environments, and since the samples are from anaerobic digesters without control, it is difficult to conclude which classification results are more accurate.

Regarding the performance prediction, *Ca.* Coprothermobacter is more beneficial to the methanogenesis as the process of sulfur reduction consumes intermediate substrates like acetate and therefore, creates a competition between methanogenic archaea and sulfur-reducing bacteria. In theory, one could observe the performance of the digester (i.e., methane production and hydrogen sulfur production) to infer the result of the classification. This observation provided an example of how to utilise the genomic information to optimise the digester operation.

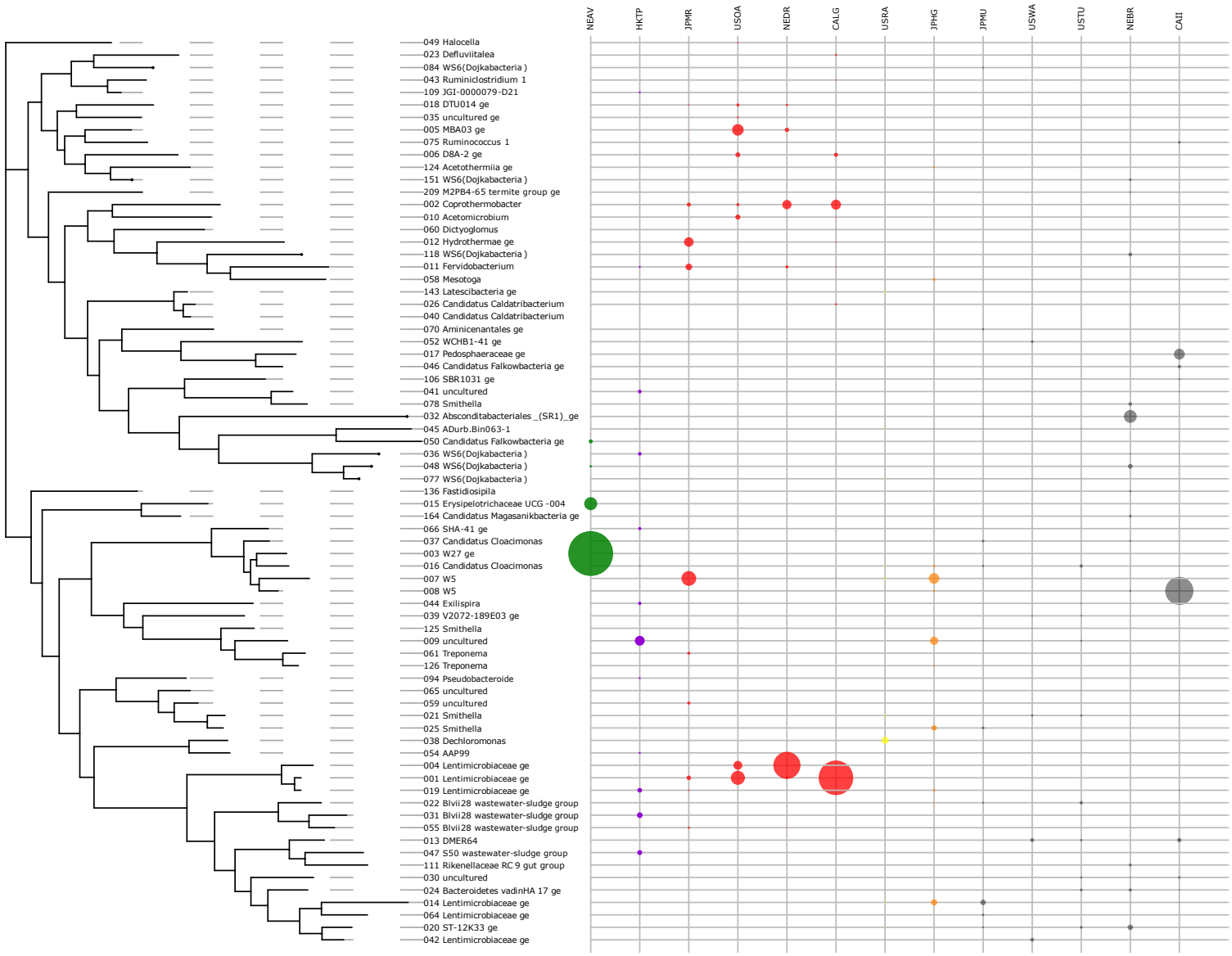


Figure 5.2 FL bacterial microbial composition of abundant OTUs with relative abundance >2%

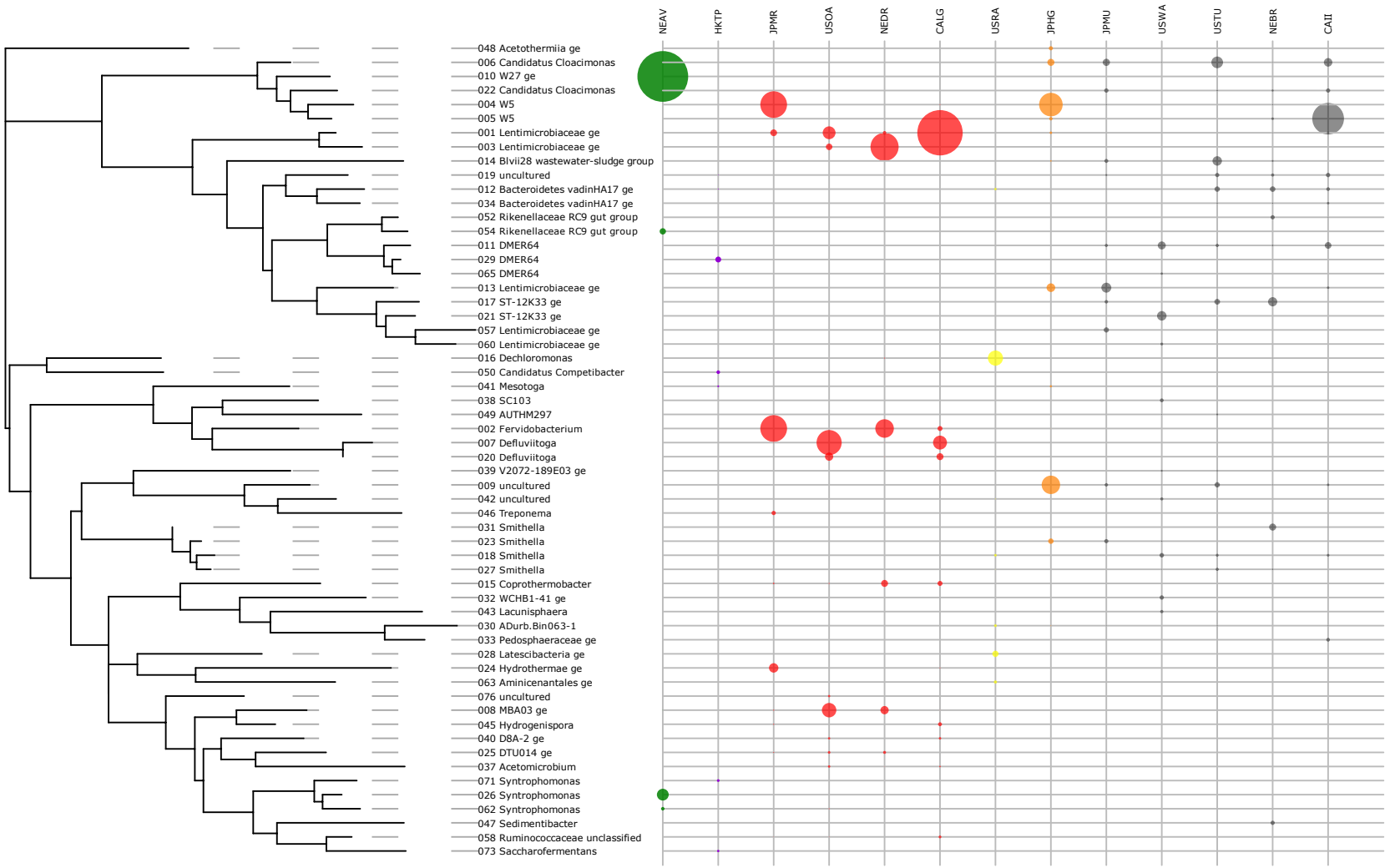


Figure 5.3 SR bacterial composition of abundant OTUs with relative abundance >2%

5.2.3. Diversity indices

Phylogenetic diversity analysis was performed to determine the difference between two sets of data using biodiversity as an indicator.

5.2.3.1 Alpha-diversity

Rarefaction curves were drawn using three different parameters to assess the communities of the two data sets (Fig. 5.4 & 5.5). The maximum sequence depth was set at 5,000 as determined by the sample with the fewest sequences for the FL results. For the observed OTUs curve, we can observe that the curve for SR results was convergent, meaning that the sequencing depth was saturated and most OTUs were observed (Fig. 5.4a). On the other hands, the trend of the observed OTUs curve for FL results is still inclining, meaning the coverage was insufficient to cover the scarce species in the samples. Considering the CCS reads generated from the PacBio Sequel is higher than the value suggested by the PacBio technician and the sufficient number of pass indicated the high sequence quality, it might require updates on the library preparation procedures and sequencing reagents to improve the coverage of the FL sequences.

However, the FL results had higher scores in both the Shannon index (SR: 4.28; FL:8.56) and Faith's Phylogenetic Diversity (SR:14.81; FL:156.61), indicating that the length of the sequence indeed affects the evaluation of the community structure. The number of observed OTUs largely contributes to the discrepancies, and because the additional OTUs observed were distributed among the genera, the FL results showed a higher communities evenness. Although the lack of reference genome hindered the taxonomic classification at the species level, the phylogenetic resolution of FL sequences was still exceptionally higher than that of the SR sequences. Hence, the Faith's PD of the FL results was significantly higher than the SR result, as it represented the species richness of a community.

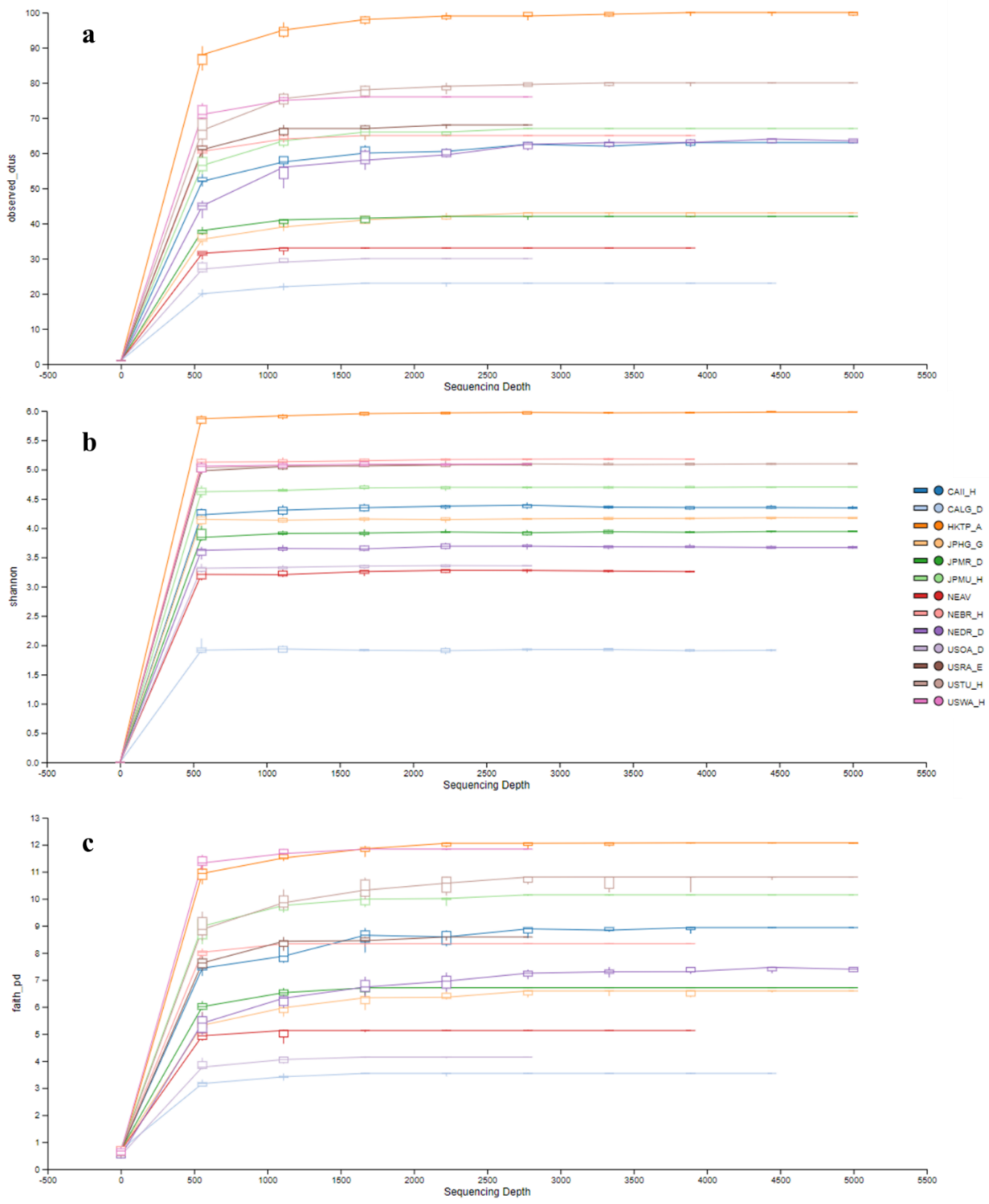


Figure 5.4 Rarefaction curve for SR result. Panel (a) shows the observed OTUs to evaluate the sequencing coverage. Panel (b) shows the Shannon index to evaluate the community evenness. Panel (c) shows the Faith's Phylogenetic Diversity to evaluate the species richness.

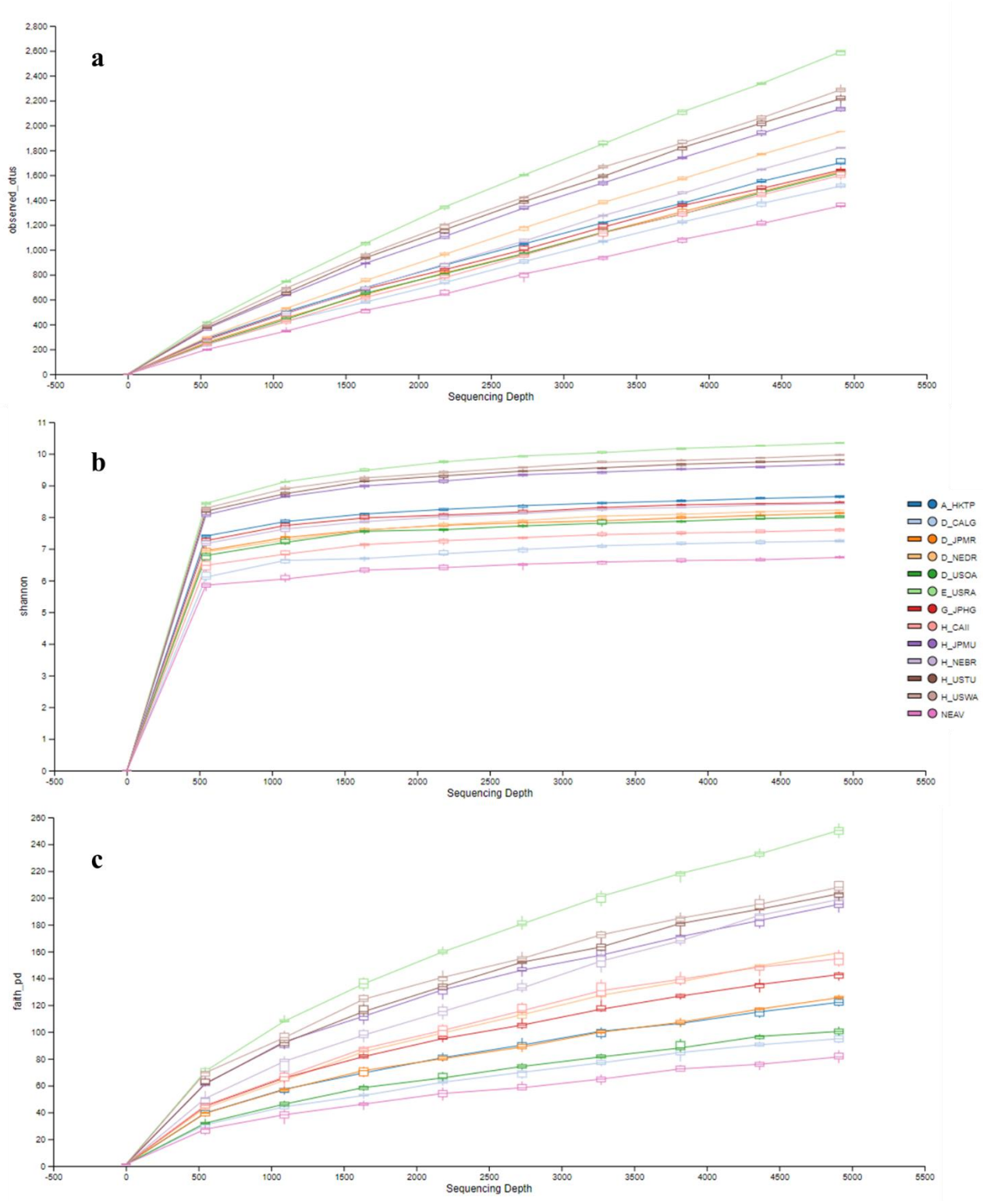


Figure 5.5 Rarefaction curve for FL result. Panel (a) shows the observed OTUs to evaluate the sequencing coverage. Panel (b) shows the Shannon index to evaluate the community evenness. Panel (c) shows the Faith's Phylogenetic Diversity to evaluate the species richness

5.2.3.2 Beta-diversity

In our previous study (Mei et al., 2017), the operation-driven heterogeneity of AD microbiome was suggested by the beta-diversity clustering. Here, we employed the same method to examine the effect of phylogenetic resolution on beta-diversity clustering. Dendrograms were drawn using the Bray-Curtis dissimilarity coefficient with a distance cutoff of 0.03 (Fig. 5.6 & 5.7). Samples were clustered following the characteristic of operation-driven heterogeneity for both data sets with slight distinctions on the distance. For instance, the distance between NEAV and HKTP was 0.060 in the FL result, and it was 0.014 in the SR result. The result indicated the effect of sequencing platforms on biodiversity clustering was minimal.

The Bray-Curtis distance calculates the dissimilarity among microbial population using abundance-based approach and does not consider the phylogenetic distance. As illustrated in the previous section, the discrepancy of the taxonomic classification was compositional. As the discrepancy mainly affected the phylogenetic identification instead of the microbial composition, the choice of distance coefficient has a large impact on the analysis result. The choice of Bray-Curtis distance was to maintain consistency as the same measure was employed in the previous study.

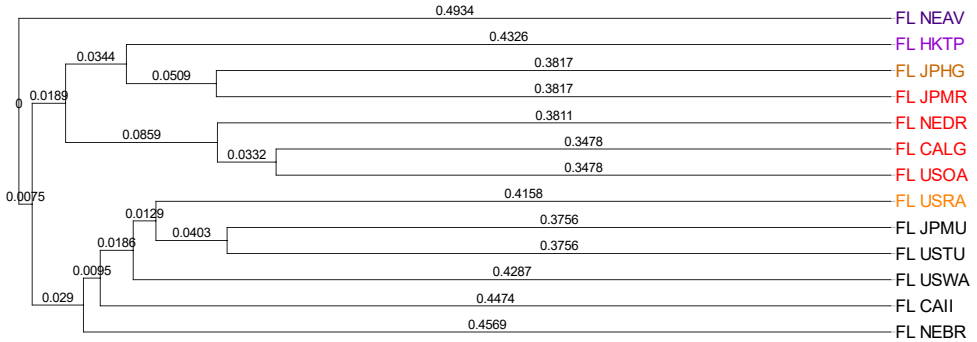


Figure 5.6 Bray-Curtis dissimilarity dendrogram for FL result, the colour of the labels denoted the operating conditions

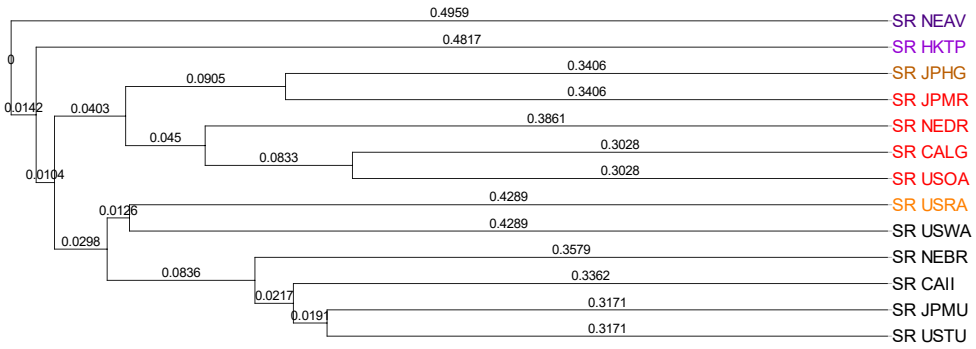


Figure 5.7 Bray-Curtis dissimilarity dendrogram for SR result, the colour of the labels denoted the operating conditions

5.3 Chapter summary

In the comparison between the NGS and TGS data sets, it was demonstrated that the FL sequences could provide more observed OTUs with relative abundance >2%, a feature contributed by its higher phylogenetic resolution. There was a discrepancy in the taxonomic classification that could impact the prediction and evaluation of the digester performance.

Considering the cost and effort required for SMRT sequencing, one should carefully determine the purpose of performing FL 16S gene amplicon sequencing for environmental samples. NGS platforms can provide economical and robust results for microbial composition characterisation and survey, and it also would be more cost-effective if one desires to perform statistical evaluation that requires large sample size (n > 20). The advantage of the full-length sequence is to allow higher phylogenetic resolution identification, meaning identifying microbes at species or even at the strain level. Therefore, if one is interested in species level phenomenon, for example, the DIET between *Geobacter* and *Methanosarcina*, it would be suitable to apply the SMRT sequencing as not all the species in the above-mentioned genera are capable of conducting DIET.

Chapter 6 Shotgun metagenomic and *mcrA* amplicon gene analysis on microaerated AD microbial communities

6.1 Overview

Two intermittent oxidation-reduction potential (ORP) controlled microaeration-anaerobic digesters were set up to evaluate the effectiveness of microaeration as pre-treatment method to promote cellulose degradation. The operating condition of the two digesters was identical: 12 days cycle of the intermittent ORP-controlled microaeration operation was implemented following a 24-hours-on and 24-hours-off pattern after the digesters reached the steady state on day 117. The selection of the ORP level for microaeration condition (-468 ± 9 mV) and the recovery for the anaerobic condition (-522 ± 12 mV) was analogous to previous studies (Jenicek et al., 2011; Khanal and Huang, 2006). The organic loading rate (OLR) was increased gradually to reach the target ORL at 5 g VS/L/day before the microaeration operation. The details for the operating condition and performance are listed in Table 6.1. The decrease in methane yield at OLR 5.0 g VS/L/day indicated that the system was overloaded, pH value was dropped gradually from 6.97 to 6.68 and the volatile fatty acids (VFAs) accumulated drastically by 10 times. The effect of microaeration operation on digester stability is discussed in the following session.

Two anaerobic (AD) effluent samples were collected on day 97 and 242, and 2 microaerated (mA) effluent samples were collected on day 147 and 192. Genomic DNA was extracted, and the DNA samples were shipped to the Novogene company for metagenomics sequencing on the Illumina HiSeq Platform. The raw data was retrieved and processed by metagenomics pipeline, and totally 178 high-quality genome bins were recovered. The quality of the genome bins was accessed by CheckM (Parks et al., 2015).The genome bins with completeness higher than 80% and contamination lower than 20% were selected for taxonomic identification and metabolic pathway reconstruction.

The AD samples contained 1182 operational taxonomic units (OTUs) on average, and 30 of them were abundant OTUs with relative abundance higher than 1%, the mA samples contained 1129 OTUs on average, and 32 of them were abundant OUTs. The biodiversity score was evaluated by the Shannon index. The anaerobic samples scored 4.3, and microaerated samples scored 3.4, which suggested the microbial community in the anaerobic samples were more diverse than microaerated samples.

Table 6.1 Digesters performance at stable operating period without microaeration at various organic loading rates (Nguyen, 2018)

	Organic loading rate (g VS/L/day)*		
	1.5	3.0	5.0
pH	6.97 ± 0.02a	6.89 ± 0.02b	6.68 ± 0.02c
ORP (mV-E_{Ag})	-472 ± 7a	-498 ± 7b	-504 ± 7b
Methane yield (NmL/g VS added)	100.1 ± 8.9a	98.9 ± 8.9a	62.8 ± 8.9b
CH_4/CO_2	1.26 ± 0.04a	1.25 ± 0.04a	1.03 ± 0.04b
Total VFAs (mg/L as HAc)	384 ± 253a	838 ± 248b	3,485 ± 259c
Total alkalinity (mg/L as $CaCO_3$)	1,996 ± 250a	2,017 ± 235a	2,787 ± 268b
VFA/ALK	0.19 ± 0.07a	0.42 ± 0.07b	1.29 ± 0.08c
Acetic acid (mg/L)	85.6 ± 32.1a	163.8 ± 32.1b	570.9 ± 10.7c
Propionic acid (mg/L)	178.8 ± 167.0a	478.2 ± 153.9a	2,066.5 ± 151.1b
Digestate TS (%)	2.6 ± 0.7a	5.4 ± 0.5b	10.4 ± 0.6c
VS removal (%)	62.3 ± 6.2a	51.3 ± 4.7ab	34.5 ± 5.5b
Digestate hemicellulose (% TS)	22.77 ± 0.47a	24.33 ± 0.47a	23.19 ± 0.47a
Digestate cellulose (% TS)	51.14 ± 0.85a	53.26 ± 0.85a	55.44 ± 0.85a

*Data are shown as mean ± standard error of duplicate digesters with 5 replications for each parameter. Data in each row with different letters indicate a significant difference (Tukey's HSD test with $p < 0.05$).

6.2 Results and Discussion

6.2.1 Biomass degradation with microaeration

Stoichiometric mass balance calculation was performed to evaluate the conversion efficiency from biomass to methane between anaerobic condition and microaerated condition. The methane yield was increased from 40 to 114 NmL/gVS with microaeration. The VFAs degradation was improved as the total VFAs level dropped from 9 to 3 gHAc/L and the acetate concentration dropped from 4 to 0.7 g/L after the digester reaching to the steady state with microaeration. The efficiency of VFAs degradation determines the methane yield of an anaerobic digester with the acetoclastic methanogenesis as the dominant pathway. In addition, while accumulation of VFAs induces pH drop and destabilises the microbial community (Kim et al., 2002), microaeration prevented the further decrease in pH value and re-established the pH balance at high OLR. It should be noted that a further increase in ORP level (further oxygenation) could cause system failure due to inhibition of the anaerobic methanogens.

Microaeration facilitates the breakdown of the large polymer chain (Sawatdeenarunat et al., 2017), and the mass balance calculation revealed the enhanced cellulose degradation in the digesters. The residual carbon content in digesters was reduced from 3.69 to 2.63 g/L, and the hemicellulose composition of digested fibre was 23% lower in the mA samples. The improved hydrolysis of the fibre content was the major contributor to improved biomass degradation.

The cycle was repeated with a replicate experiment, and a similar trend of methane yield and VFA degradation was observed (CH_4: 25.3 to 77.8 NmL/gVS, VFAs: 10.5 to 5 g HAc/L). This validated the effectiveness of microaeration on VFAs degradation, a similar change in microbial community composition was observed as described below.

6.2.2 Taxonomic summary

The taxonomic identification was done using Mothur (Schloss et al., 2009) and classified to the genus level. The abundant OTUs with relative abundance higher than 1% were included in the taxonomic identification and are showed in Figure 6.1. For bacteria composition, phylum Bacterioidetes was the most dominant hydrolytic and fermentative bacteria group across all five samples. Phyla *Proteobacteria, Firmicutes* and *Spirochaetae* were also detected. *Clostridia,* a bacteria class belongs to *Firmicutes,* was found and it was reported by several studies for its cellulose degradation property (Jia et al., 2016; Strang et al., 2017; Xia et al., 2014). Both sample types shared a common core community with similar OTUs numbers, and the majority of the abundant OTUs were involved in hydrolysis, acidogenesis and acetogenesis in anaerobic digestion.

Two significant difference between the AD samples and the mA samples were observed as the genus *Proteiniphlium* was doubled in microaerated samples (AD: 21.2%, mA: 44.3%). In addition, the genus *Bacteroides* in anaerobic samples was six times of microaerated samples (AD: 8.0%, mA: 1.4%). As being facultative anaerobes, *Proteiniphlium* and *Bacteroides* are capable of adjusting their metabolic pathways according to the ORP levels; they were more advantageous to replicate under microaeration condition in comparison to obligate anaerobes. It was noted that the syntrophic bacteria responsible for the acetogenesis was halved in mA samples (AD: 0.43%, mA: 0.24%). Typically, syntrophic bacteria held a crucial role in anaerobic digestion as they convert VFAs such as valerate, butyrate and propionate into acetate and hydrogen for methanogens to consume. The lack of syntrophic bacteria population explained the VFAs accumulation under high OLR condition. However, the improvement of VFAs degradation under the microaeration condition was observed, while there was a further decrease in syntrophic bacteria population. This result was contradictory to the traditional model and suggested there was an alternative mechanism involved.

The gene mcrA analysis, for a specific gene only found in methanogens, was carried out for identifying the archaea composition. The genus *Methanosarcina* was dominant in both samples with relative abundance higher than 70% (AD: 73.2%, mA: 73.7%). *Methanosarcina* is acetotrophic methanogen, which suggested that the dominant pathway in the digesters was the acetoclastic methanogenesis from acetate reduction, which relied on VFAs degradation by hydrolytic and fermentative bacteria to provide acetate for this methanogenesis to occur. Genera *Methanobacterium* and *Methanospirillum* were also detected (AD: 19.6%, mA: 20.5%), and unlike *Methnosarcina*, they are hydrogenotrophic methanogens relying on hydrogen and carbon dioxide for methanogenesis. The three methanogens constituted over 90% of the archaea population (AD: 92.8%, mA: 94.2%) in the digesters.

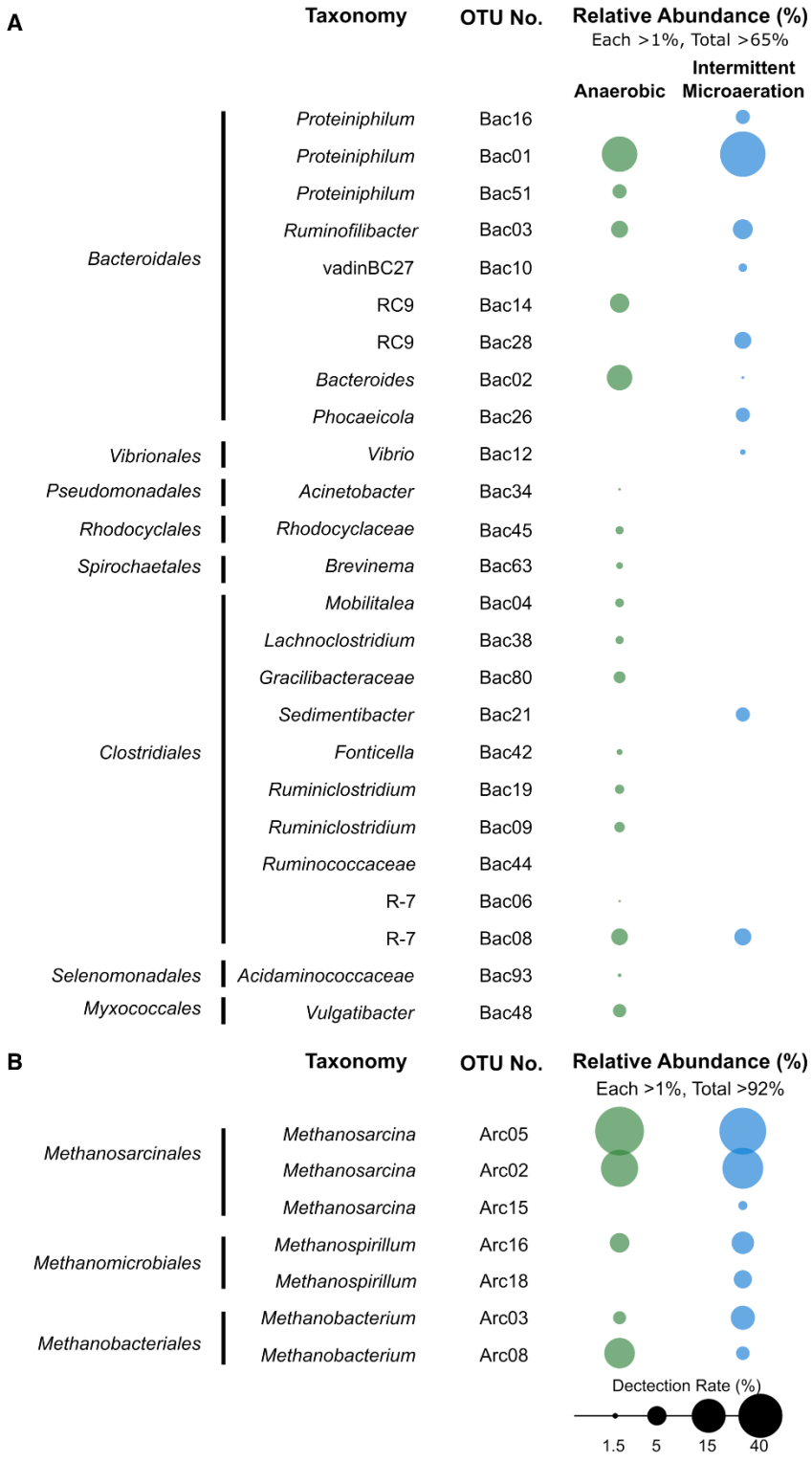


Figure 6.1 Microbial community structure (A) and archaeal community structure (B). Biomass samples collected from digesters of without and with intermittent ORP-controlled microaeration, respectively (Nguyen, 2018).

6.2.3 Reconstruction of the microaeration metabolic pathway

The metabolic pathway was determined by the methanogens presented in the digesters. The methanogenesis adopted were mostly acetoclastic methanogenesis and hydrogenotrophic methanogenesis.

The experimental data showed a decrease in acetate concentration under microaeration condition; it could be consumed in acetoclastic methanogenesis. The question remained as in the traditional digester, the synthesis of acetate relies on the syntrophic bacteria. However, the taxonomic result showed lack of syntrophic bacteria population under microaeration condition. To determine the metabolic pathway of the acetate synthesis, the distinctive difference between AD samples and mA sample in microbial community composition on genus *Proteiniphlium* was further investigated. The SPAdes assembled contigs files from all four samples were annotated by Strainseeker for the strain level identification; the result was then utilised as the whole genome reference. The high genome bins were crosschecked with the whole genome reference by drawing a phylogenetic tree with PhyloPhyLan. The association between 16 draft genome bins and species *Fermentimonas caenicola* ING2-E5B was identified with genetic change less than 0.01, which was a facultative bacterium that genetically neighboured with *Proteiniphilum saccharofermentans* M3/6 (Hahnke et al., 2016). It indicated the *Proteiniphlium* doubled in the AD and mA samples were genetically closest to *F. caenicola*. The specified genome bins were then submitted to BlastKOALA for metabolic pathway reconstruction. Coenzyme A (CoA) was identified in all the submitted genome bins. CoA is a key enzyme for acetoclastic methanogenesis to occur, because the process requires acetate as the feed substrate and CoA is participated in the degradation of long-chain VFAs molecule to acetate (Daugherty et al., 2002). The annotation result corresponded with the experimental result. The improved efficiency of VFAs degradation under microaeration condition could be explained by the change in microbial community composition. The microaerated environment provided a thermodynamically favourable condition for hydrolysis and acidogenesis to occur as the facultative bacteria could utilise oxygen as an electron donor. The role of syntrophic bacteria could be substituted

by the *F. caenicola* as it can produce CoA for the VFAs degrade to acetate. This could explain the population decline of the syntrophic bacteria in mA sample; there was a chance that the syntrophic bacteria were likely out-competed by *F. caenicola* under the microaeration condition.

The VFAs degradation also led to increasing hydrogen and carbon dioxide content in the digesters. As facultative bacteria could utilise oxygen as the electron donor, the aerobic respiration could take place and degrade the accumulated VFAs by converting oxygen to carbon dioxide for the exothermic reaction. In a traditional anaerobic digester, carbon dioxide is produced as biogas through hydrolysis conducted by the H_2-oxidising bacteria. The Gibbs free energy of converting oxygen to water ($\Delta G'_R$ = -92.7 kJ/e^-) is higher than reducing hydrogen to methane ($\Delta G'_R$ = -4.0 kJ/e^-), the hydrogenotrophic methanogens are constantly out-completed by H_2-oxidising bacteria in the traditional digester. However, the H_2-oxidising bacteria were not observed in any abundant OTUs in mA samples according to the taxonomic result, and it was probably due to the fast-growing facultative bacteria (μ_m = 13.2 day^{-1}) out-completed the H_2-oxidising bacteria (μ_m = 10.1 day^{-1}) under microaeration condition. Furthermore, gas chromatography analysis was performed on the biogas produced during the microaeration period, and no significant change in carbon dioxide percentage was found. Combined with the fact that H_2-oxidising bacteria was low in population, the carbon dioxide in the digesters was likely to be reduced to methane using hydrogen as an electron donor. Another possibility is that the carbon dioxide was still in the aqueous form, which was somewhat challenging to determine.

A proposed metabolic pathway is illustrated in Figure 6.2, where the facultative bacteria as the dominant bacteria in the digesters, they were responsible for the hydrolysis, acidogenesis and acetogenesis. A syngeneic relationship was formed between the facultative bacteria and methanogens, both acetoclastic and hydrogenotrophic methanogenesis could be exploited in the intermittent ORP controlled microaeration-anaerobic digesters.

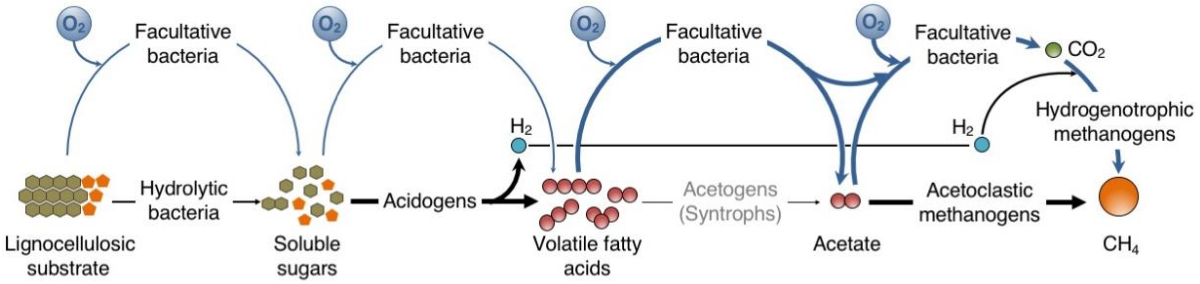


Figure 6.2 Proposed methane-producing pathway using lignocellulosic biomass via intermittent ORP-based micro-aeration (Nguyen, 2018).

6.3 Chapter summary

The impact of microaeration in high organic loading anaerobic digesters was investigated in this study through bioinformatics and stoichiometric mass balance calculation. Family *Porphyromonadaceae* was dominant in the microaeration samples and potentially shaped an alternative metabolic pathway in the cellulose-based anaerobic system.

- The anaerobic system under high VFAs stress was stabilised by the introduction of microaeration. The microaeration promoted hydrolysis, acidogenesis, and acetogenesis, which consequently improved the methane yield.
- The composition of the microbial community in the anaerobic digesters adopted the increase in oxygen level and shifted from strict anaerobes to facultative bacteria, as it was more thermodynamically favourable for facultative bacteria to grow.
- The presence of facultative bacteria replaced the traditional commensalism between syntrophs and methanogens with a more efficient VFAs degradation mechanism and favouring the acetoclastic methanogenesis.

Chapter 7 Conclusions and Recommendations

7.1 Conclusions

Conclusion on **full-length 16S rRNA amplicon sequencing pipeline development**

- In this study, a complete pipeline from sample preparation to bioinformatics analysis for full-length 16S rRNA amplicon sequencing was developed. SMRT sequencing has a way more stringent requirement on DNA quality comparing to the sequencing by synthesis approach. The method developed effectively provided high quality and highly concentrated DNA templates. The alteration on the library preparation helped to maintain the integrity of the DNA fragments, resulting in full-length sequences with high accuracy.
- Since there was no existing bioinformatics pipeline for analysing full-length 16S sequences, we integrate the strong suites of two popular 16S analysis platforms, Mothur and Qiime 2, for taxonomic classification, diversity evaluation, and data visualisation.

Conclusion on **the application of full-length in species level identification**

- Most abundant methanogenic archaea were identified at the species level, owing to the comprehensive genome reference. Unique methanogenic species like *Methanosarcina horonobensis* and *Methanosarcina flavescens* were identified in high salinity digester and thermophilic digester, respectively. The two species expressed different survival niches and substrate utilisation compared to other species within the genus *Methanosarcina*. The result demonstrated the value of species level identification as it provided a more accurate description of the engineered “black box”.
- The species level identification of the bacterial communities was hindered by the lack of reference genome. The development of environmental draft genome via metagenomics could facilitate the species level identification in the future.

Conclusion on **the discrepancy of NGS and TGS in taxonomy profiling in environmental samples**

- The full-length sequences generated by TGS were able to provide more observed OTUs than the short-read sequences generated by NGS. it indicated that the full-length allowed the taxonomic classification with higher phylogenetic resolution.
- Discrepancies between NGS and TGS on taxonomic classification was observed. Several major AD phyla, including *Bacterioidetes*, *Thermotogae*, and *Ca.* Coprothermobacteraeote. We observed that *Ca.* Coprothermobacter was identified in the full-length sequences, while its phylogenetic neighbour *Fervidobacterium* was identified in the short-read sequences. Such discrepancy severely affects the performance prediction of AD as the former is potential acetate-producing bacteria that facilitate methanogenesis, and the latter is sulfur-reducing bacteria that complete with methanogens for substrates.
- The effect on diversity dices was marginal. Similar clustering was observed using Bray-Curtis distance for beta diversity. The richness and evenness of the microbial communities observed in the full-length sequence were higher than the short-read sequences.

Conclusion on **the effect of microaeration on AD microbiome**

- The introduction of microaeration promoted the hydrolysis of cellulose compounds and stabilised the fermenter by alleviating the VFAs stress from accumulation. Oxygen supplied to the system can be utilised by the facultative bacteria as an electron acceptor, facilitating the degradation of organic matters.
- *Fermentimonas caenicola*, a facultative bacterium, was identified as the significant acetate producer under the microaerated environment.
- Under the microaerated environment, the syntrophic association between syntrophs and methanogens might not be imperative, as the facultative bacteria could also efficiently produce simple substrates (i.e., acetate, hydrogen and CO_2) for methanogenesis to occur.

7.2 Recommendations for future study

- From this study, we acknowledged the limitations and the potential of various sequencing technologies. Aside from being resource-intensive and lack of mature bioinformatics pipeline, the full-length 16S rRNA amplicon sequencing has advantages in phylogenetic resolution over the short-read approach. Species level identification provides more information on the metabolic pathway (e.g., *Methanosacrina horonobensis*). However, we also observed that the genome database for hydrolytic and fermentative bacteria are far less comprehensive compared to that of the methanogens. Since hydrolysis is considered as a rate-limiting stage in AD process, effort should be made to supplement this knowledge gap. Large-scale multi-omic analysis, including metagenomics, metatranscriptomics, metaproteomics, metabolomics, and metaepigenome, should be carried to investigate their roles in AD and evaluate their activities, including response to external stimulation (e.g., microaeration), in situ growth rate, and structural variant. We should recognise the importance of understanding the underlying mechanisms behind the microbial physiology. The completion of AD microbiome characterisation will allow us to manipulate the digester stability and methane production. Ultimately, we can shape the AD microbiome for optimising the AD process with utility under different circumstance by controlling the environmental stress.

- The experimental design should be meticulous in order to obtain accurate and meaning results. This study is considered as a survey research -- the main contribution is to describe the observation and explain the observed phenomena. In this case, the sample size is crucial for manifesting a result with statical significance. The smaple size in this study (i.e. 19 in chapter 4, 13 in chapter 5, and 4 in chapter 6) limited the option for statistical evaluation. Methods like effect size analysis and co-occurrence network analysis can really benefit from a larger sample size (n>20). Using this study as a baseline, it can help to determine the specific experimental questions to ask and the sample size, among others.

Appendix

Appendix I - Table A1

Table A1 Sample ID and operation information

Region	Location	Digester sludge sample ID	Digester operation					Feed composition		
			Sludge retention time (d)	pH	Temperature (°C)	Efficiency (% volatile solids removed)[1]	Efficiency (% volatile solids removed)[2]	Primary sludge	Activated sludge	Other
Canada	Iona Island, Vancouver	CAII	19	7.1	37	69	65.1	Yes	No	No
	Lions Gate, Vancouver	CALG	39	7.4	55.8	70	70	Yes	No	No
Hong Kong	Shek Wu Hui*	HKSW	25	7.1	36	49	37.7	Yes	Yes	No
	Yuen Long*	HKYL	12	7.1	29	76	53.8	Yes	Yes	No
Japan	Kyushu_HW*	JPHW	23.1	7.5	19	-	-	Yes	Yes	No
	Kansai_HG*	JPHG	29.7	7.2	40.9	-	33	Yes	Yes	No
	Kanto_MR*	JPMR	20	7.6	51.1	-	-	Yes	Yes	No
	Shikoku_MU*	JPMU	63	7.4	34.1	-	-	Yes	Yes	No
	Chugoku_SS*	JPSS	25.3	-	36.5	-	-	Yes	Yes	No
	Kansai_TR*	JPTR	27.8	7.3	42.1	-	22.7	Yes	Yes	No
	Tohoku_TO*	JPTO	30.4	-	38	-	-	Yes	Yes	No
Netherlands	Drenthe	NEDR	-	-	55	-	-	Yes	Yes	No
	Breda	NEBR	28.3	7.7	35.2	-	48.3	Yes	Yes	No
	A. van de Groep	NEAV	29.5	8.1	39.3	71.7	76.3	Yes	Yes	Food waste
United States	Rantoul, Illinois*	USRA	20	7.5	29.5	17.7	21.5	Yes	No	No
	Waukesha, Milwaukee	USWA	-	7.1	34.4	-	57.2	Yes	Yes	No
	Oakland, California	USOA	15	-	51.2	-	65.4	Yes	Yes	No
	San Fransisco, California	USSF	28.1	7.5	38.1	65.9	40.6	Yes	Yes	No
	Tulsa, Oklahoma	USTU	21.1	6.9	33.3	27.9	32.8	Yes	Yes	No

Appendix II - Table A2

Table A2 Relative abundance of bacterial population at phylum level

	Detected Phyla	Total Relative Abundance (%)	High Salinity	Pre-treatment	Thermo-philic	HKSW	<30°C	~40°C	Meso-philic	Food Waste
Official Phyla	Actinobacteria	0.775	1.239	0.827	1.776	0.788	0.667	0.898	0.008	0.020
	Armatimonadetes	0.113	0.398	0.012	0.050	0.175	0.069	0.059	0.137	0
	Bacteroidetes	19.361	30.589	17.378	31.998	28.484	19.319	23.681	3.442	5.646
	Caldiserica	0.957	0.044	0.309	0.586	3.204	1.744	1.137	0.633	0
	Calditrichaeota	0.022	0	0	0	0	0.126	0.015	0.037	0
	Chlamydiae	0.466	0	0	0	0	0	0	3.687	0
	Chloroflexi	11.195	9.628	2.715	0.997	6.880	3.152	2.551	5.983	0.041
	Cyanobacteria	0.051	0.066	0	0.015	0.018	0.233	0.077	0	0
	Dictyoglomi	0.161	0	0	1.229	0	0	0	0.060	0
	Elusimicrobia	0.072	0	0	0	0.018	0.172	0.041	0.349	0
	Fibrobacteres	3.792	0.022	0.111	0.020	0	0.135	0.156	6.484	0
	Firmicutes	6.669	13.103	2.197	24.164	3.852	4.178	4.853	0.906	23.404
	Fusobacteria	0.062	0.022	0.148	0.021	0	0.158	0.103	0.044	0.102
	Gemmatimonadetes	0.025	0	0	0.011	0.035	0.123	0.027	0.004	0
	Kiritimatiellaeota	0.644	0	0.815	0.022	0.508	2.238	1.495	0.077	0.389
	Nitrospirae	0.055	0.022	0	0.120	0	0.068	0	0.228	0
	Planctomycetes	3.215	0.199	4.530	1.320	3.099	2.371	3.153	10.990	0.020
	Proteobacteria	7.376	8.101	10.220	5.150	4.167	18.536	11.982	0.851	0.061
	Spirochaetes	5.532	16.135	3.172	1.402	7.335	4.901	10.366	0.943	0
	Synergistetes	1.262	3.121	0.296	3.417	0.210	0.272	1.001	0.430	0
	Tenericutes	0.605	0.066	0.407	0.981	1.173	0.343	0.828	1.002	1.350

	Thermotogae	1.779	2.346	0.950	5.321	1.155	0.423	2.533	1.481	0.0409
	Verrucomicrobia	2.242	0.044	6.863	0.115	0.753	4.215	2.747	3.198	0.020
Uncultivated and candidate Phyla	Acetothermia	0.236	0	0	0	0	0.041	0.607	1.217	0
	Acidobacteria	1.144	1.262	0.074	0.986	1.593	2.862	1.509	0.865	0
	Aegiribacteria	0.017	0	0	0	0	0.014	0	0.123	0
	Atribacteria	0.201	0.044	0	1.259	0.018	0.041	0.090	0.157	0
	Unclassified Bacteria	6.798	0.089	16.292	0.617	3.676	4.012	1.564	22.488	1.678
	BRC1	0.392	0.111	0.062	0.033	0	0.141	0.109	1.003	0
	CK-2C2-2	0.107	0	0	0	0	0.041	0.097	0.722	0
	Cloacimonetes	7.102	6.884	13.885	4.840	1.943	3.428	18.092	7.703	57.651
	Coprothermobacteraeota	1.119	0	0.025	8.652	0	0	0.175	0.098	0.041
	Dadabacteria	0.004	0	0	0	0	0.027	0.004	0	0
	Deinococcus-Thermus	0.006	0	0	0.016	0	0	0.002	0.030	0
	Dependentiae	0.023	0.044	0	0.021	0.070	0	0.046	0	0
	Epsilonbacteraeota	0.110	0	0.457	0.032	0.018	0.261	0.079	0.033	0
	FCPU426	0.040	0	0	0	0.035	0.061	0.023	0.203	0
	Halanaerobiaeota	0.085	0	0	0.472	0	0	0	0.211	0
	Hydrogenedentes	0.169	0.553	0.025	0.218	0.053	0.229	0.253	0.018	0
	Hydrothermae	0.602	0	0	3.319	0	0	0	1.108	0
	Latescibacteria	0.215	0	0	0	0	1.221	0.328	0.173	0
	Lentisphaerae	0.196	0	0.074	0.082	0	1.030	0.376	0.003	0
	Margulisbacteria	0.007	0	0	0	0	0	0	0.055	0
	Marinimicrobia (SAR406 clade)	0.231	0.398	0	0.001	0.035	0.164	1.240	0.009	0
	MAT-CR-M4-B07	0.003	0	0	0	0	0.022	0	0.003	0
	Modulibacteria	0.003	0	0	0	0	0	0.007	0.014	0
	Omnitrophicaeota	3.125	0	0	0.011	0	1.198	0.240	14.098	0
	Patescibacteria	10.776	5.201	18.131	0.419	30.585	21.491	7.238	3.119	9.452
	Rokubacteria	0.616	0	0	0	0	0	0.003	4.927	0
	WPS-2	0.014	0	0	0.001	0	0.058	0.010	0.041	0

WS1	0.037	0.155	0	0.023	0.018	0.014	0.084	0	0
WS2	0.015	0	0	0.011	0	0.014	0	0.099	0
WS4	0.019	0.022	0	0	0.018	0.055	0.056	0	0
Zixibacteria	0.009	0	0	0	0	0.027	0.002	0.043	0
uncultured	0.025	0	0.025	0.038	0	0.068	0.063	0.004	0

Appendix III - Table A3

Table A2 Species-level identification using BLAST against Sliva 132 NR reference, including (a) selected OTUs related syntrophs that are abundant or phylogenetically close to reference illustrated by Arb; (b) abundant OTUs related to methanogens with relative abundance >1%.

(a)

OTUs	Mothur (Ref=Sliva 132 NR Ref)					BLAST (Ref=SILVA 132 NR Ref)			
	Phylum	Class	Order	Family	Genus	Species	Score	Query cover (%)	E-value	Identity (%)
12	Cloacimonetes	Cloacimonadia	Cloacimonadales	*Cloacimonadaceae*	*Candidatus Cloacimonas*	*Candidatus Cloacamonas acidaminovorans*	2791	100	0	99.74
460	Cloacimonetes	Cloacimonadia	Cloacimonadales	*Cloacimonadaceae*	*Candidatus Cloacimonas*	*Candidatus Cloacimona s acidaminovorans*	1707	87	0	99.78
553	Cloacimonetes	Cloacimonadia	Cloacimonadales	*Cloacimonadaceae*	*Candidatus Cloacimonas*	metagenome	2636	100	0	97.62
152	Firmicutes	Clostridia	Clostridiales	*Peptococcaceae*	*Pelotomaculum*	*Pelotomaculum isophthalicicum*	2266	100	0	94.79
152	Firmicutes	Clostridia	Clostridiales	*Peptococcaceae*	*Pelotomaculum*	*Pelotomaculum isophthalicicum*	2266	100	0	94.79
1583	Firmicutes	Clostridia	Clostridiales	*Peptococcaceae*	*Pelotomaculum*	*Pelotomaculum isophthalicicum*	2244	100	0	94.54
1583	Firmicutes	Clostridia	Clostridiales	*Peptococcaceae*	*Pelotomaculum*	*Pelotomaculum isophthalicicum*	2244	100	0	94.54
476	Firmicutes	Clostridia	Clostridiales	*Syntrophomonadaceae*	*Syntrophomonas*	*Syntrophomonas wolfei*	2130	89	0	93.88
1430	Firmicutes	Clostridia	Clostridiales	*Syntrophomonadaceae*	*Syntrophomonas*	*Syntrophomonas sapovorans*	2599	100	0	98.26

1601	Firmicutes	Clostridia	Clostridiales	*Syntrophomonadaceae*	*Syntrophomonas*	*Syntrophomonas curvata*	2145	100	0	92.87
2292	Firmicutes	Clostridia	Clostridiales	*Syntrophomonadaceae*	*Syntrophomonas*	*Syntrophomonas palmitatica*	2900	100	0	100
2702	Firmicutes	Clostridia	Clostridiales	*Syntrophomonadaceae*	*Syntrophomonas*	*Syntrophomonas cellicola*	2412	95	0	97.52
2099	Firmicutes	Clostridia	Clostridiales	*Syntrophomonadaceae*	*Syntrophothermus*	*Syntrophothermus lipocalidus*	2549	97	0	96.87
19156	Firmicutes	Clostridia	Clostridiales	*Syntrophomonadaceae*	*Syntrophothermus*	*Syntrophothermus lipocalidus*	2519	97	0	96.49
13282	Firmicutes	Clostridia	Thermoanaerobacterales	*Thermoanaerobacteraceae*	*Syntrophaceticu*	*Syntrophaceticus schinkii*	2152	97	0	94.15
13321	Firmicutes	Clostridia	Thermoanaerobacterales	*Thermoanaerobacteraceae*	*Syntrophaceticu*	*Syntrophaceticus schinkii*	2126	97	0	93.75
135	Proteobacteria	Deltaproteobacteria	Deltaproteobacteria Incertae Sedis	*Syntrophorhabdaceae*	*Syntrophorhabdus*	*Syntrophorhabdus aromaticivorans*	2276	100	0	94.4
195	Proteobacteria	Deltaproteobacteria	Deltaproteobacteria Incertae Sedis	*Syntrophorhabdaceae*	*Syntrophorhabdus*	*Syntrophorhabdus aromaticivorans*	2204	96	0	94.36
1954	Proteobacteria	Deltaproteobacteria	Deltaproteobacteria Incertae Sedis	*Syntrophorhabdaceae*	*Syntrophorhabdus*	*Syntrophorhabdus aromaticivorans*	2298	100	0	94.68
5424	Proteobacteria	Deltaproteobacteria	Deltaproteobacteria Incertae Sedis	*Syntrophorhabdaceae*	*Syntrophorhabdus*	*Syntrophorhabdus aromaticivorans*	2453	100	0	96.56
2221	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	*Desulfovibrionaceae*	*Desulfovibrio*	*Desulfovibrio desulfuricans*	2627	100	0	98.91
5338	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	*Desulfovibrionaceae*	*Desulfovibrio*	*Desulfovibrio desulfuricans*	2527	100	0	97.7
5345	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	*Desulfovibrionaceae*	*Desulfovibrio*	*Desulfovibrio vulgaris*	2327	100	0	95.24
5367	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	*Desulfovibrionaceae*	*Desulfovibrio*	*Desulfovibrio desulfuricans*	2663	100	0	99.32
5372	Proteobacteria	Deltaproteobacteria	Desulfovibrionales	*Desulfovibrionaceae*	*Desulfovibrio*	*Desulfovibrio desulfuricans*	2580	100	0	98.25
758	Proteobacteria	Deltaproteobacteria	Desulfuromonadales	*Geobacteraceae*	*Geobacter*	*uncultured Desulfuromonadales bacterium*	2507	100	0	97.23
1454	Proteobacteria	Deltaproteobacteria	Desulfuromonadales	*Geobacteraceae*	*Geobacter*	*uncultured Desulfuromonadales bacterium*	2483	100	0	96.97
2735	Proteobacteria	Deltaproteobacteria	Desulfuromonadales	*Geobacteraceae*	*Geobacter*	*Geobacter lovleyi SZ*	2560	100	0	97.91
1752	Proteobacteria	Deltaproteobacteria	Desulfuromonadales	*Pelobacteraceae*	*Pelobacter*	*Pelobacter propionicus*	2604	100	0	98.44
15	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Smithella*	Smithella sp. D17	2268	100	0	94.24
32	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Smithella*	Smithella sp. D17	2237	100	0	93.82

71	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Smithella*	Smithella sp. D17	2292	100	0	94.49
101	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Smithella*	Smithella sp. M82	2468	100	0	96.64
131	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Smithella*	Smithella sp. D17	2265	100	0	94.15
204	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Smithella*	Smithella sp. M82	2483	100	0	96.77
360	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Smithella*	Smithella sp. D17	2270	100	0	94.22
1803	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Smithella*	*Smithella propionica*	2215	91	0	96.02
737	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Syntrophus*	*Syntrophus gentianae*	2305	100	0	94.68
739	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Syntrophus*	*Syntrophus aciditrophicus*	2300	100	0	94.62
740	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophaceae*	*Syntrophus*	*Syntrophus buswellii*	2302	100	0	94.63
1107	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophobacteraceae*	*Syntrophobacter*	*Syntrophobacter sulfatireducens*	2158	100	0	92.93
1108	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophobacteraceae*	*Syntrophobacter*	*Syntrophobacter fumaroxidans*	2239	100	0	93.98
2900	Proteobacteria	Deltaproteobacteria	Syntrophobacterales	*Syntrophobacteraceae*	*Syntrophobacter*	*Syntrophobacter wolinii*	2248	93	0	95.76
584	Synergistetes	Synergistia	Synergistales	*Synergistaceae*	*Thermanaerovibrio*	*Thermanaerovibrio acidaminovorans DSM 6589*	2660	100	0	99.93
8887	Synergistetes	Synergistia	Synergistales	*Synergistaceae*	*Thermanaerovibrio*	*Thermanaerovibrio acidaminovorans DSM 6589*	2595	100	0	99.1
8898	Synergistetes	Synergistia	Synergistales	*Synergistaceae*	*Thermanaerovibrio*	*Thermanaerovibrio acidaminovorans DSM 6589*	2449	100	0	97.15

(b)

OTUs	Mothur (Ref=Sliva 132 NR Ref)					BLAST (Ref=SILVA 132 NR Ref)				
	Phylum	Class	Order	Family	Genus	Species	Score	Query cover (%)	E-value	Identity (%)

41	Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanobacteriaceae*	*Methanobacterium*	*Methanobacterium beijingense*	2451	99	0	98.35
47	Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanobacteriaceae*	*Methanobacterium*	*Methanobacterium formicicum*	2566	100	0	99.64
21	Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanothermobacteraceae*	*Methanothermobacter*	*Methanobacterium thermaggregans*	2259	100	0	95.62
28	Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanothermobacteraceae*	*Methanothermobacter*	*Methanobacterium thermaggregans*	2237	100	0	95.33
10	Euryarchaeota	Methanobacteria	Methanobacteriales	*Methanothermobacteraceae*	*Methanothermobacter*	*Methanothermobacter tenebrarum strain RMAS*	2512	100	0	98.8
35	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanocorpusculaceae*	*Methanocorpusculum*	*Methanocorpusculum aggregans*	2355	100	0	97.13
32	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanomicrobiaceae*	*Methanoculleus*	*Methanoculleus receptaculi*	2459	96	0	99.78
24	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanomicrobiaceae*	*Methanoculleus*	*Methanoculleus sp. DTU006*	2577	100	0	100
14	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanomicrobiaceae*	*Methanoculleus*	*Methanoculleus sp. SLH121*	2410	100	0	97.85
19	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanomicrobiaceae*	*Methanoculleus*	*Methanoculleus sp. SLH121*	2460	100	0	98.93
25	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanoregulaceae*	*Methanolinea*	*Methanolinea mesophila*	2316	100	0	96.63
1	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanoregulaceae*	*Methanolinea*	*Methanolinea tarda*	2294	100	0	96.34
45	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanoregulaceae*	*Methanoregula*	*Methanoregula boonei 6A8*	2368	100	0	97.28
9	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanospirillaceae*	*Methanospirillum*	*Methanospirillum hungatei JF-1*	2353	100	0	97.13
22	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanospirillaceae*	*Methanospirillum*	*Methanospirillum hungatei JF-1*	2575	100	0	100
48	Euryarchaeota	Methanomicrobia	Methanomicrobiales	*Methanospirillaceae*	*Methanospirillum*	*Methanospirillum hungatei JF-1*	2414	100	0	97.92
4	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosaetaceae*	*Methanothrix*	*Methanothrix concilii GP6*	2407	100	0	97.78
5	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosaetaceae*	*Methanothrix*	*Methanothrix concilii GP6*	2584	100	0	100
7	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosaetaceae*	*Methanothrix*	*Methanothrix concilii GP6*	2468	100	0	98.57

11	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosaetaceae*	*Methanothrix*	*Methanothrix concilii GP6*	2440	100	0	98.21
16	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosaetaceae*	*Methanothrix*	*Methanothrix concilii GP6*	2346	100	0	96.93
20	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosaetaceae*	*Methanothrix*	*Methanothrix concilii GP6*	2435	100	0	98.49
26	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosaetaceae*	*Methanothrix*	*Methanothrix concilii GP6*	2473	100	0	98.57
3	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosaetaceae*	*Methanothrix*	*Methanothrix thermophila*	2507	100	0	98.93
33	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosarcinaceae*	*Methanosarcina*	*Methanosarcina barkeri*	2564	100	0	99.57
34	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosarcinaceae*	*Methanosarcina*	*Methanosarcina barkeri*	2536	100	0	99.22
15	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosarcinaceae*	*Methanosarcina*	*Methanosarcina flavescens*	2431	100	0	97.87
12	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosarcinaceae*	*Methanosarcina*	*Methanosarcina flavescens strain E03.2*	2558	100	0	99.5
6	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosarcinaceae*	*Methanosarcina*	*Methanosarcina flavescens*	2497	100	0	98.72
8	Euryarchaeota	Methanomicrobia	Methanosarcinales	*Methanosarcinaceae*	*Methanosarcina*	*Methanosarcina horonobensis*	2481	100	0	98.57
2	Euryarchaeota	Thermococci	Methanofastidiosales	*Methanofastidiosaceae*	*Candidatus Methanofastidiosum*	Candidatus Methanofastidiosum ADurb.Bin023	2586	100	0	100
18	Euryarchaeota	Thermococci	Methanofastidiosales	*Methanofastidiosaceae*	*Candidatus Methanofastidiosum*	Candidatus Methanofastidiosum ADurb.Bin023	2503	100	0	99
29	Euryarchaeota	Thermococci	Methanofastidiosales	*Methanofastidiosaceae*	*Candidatus Methanofastidiosum*	Candidatus Methanofastidiosum ADurb.Bin023	2407	100	0	97.93
13	Euryarchaeota	Thermoplasmata	Methanomassiliicoccales	*Methanomassiliicoccaceae*	*Methanomassiliicoccus*	*Methanomassiliicoccus luminyensis B10*	2383	100	0	97.5
17	Euryarchaeota	Thermoplasmata	Methanomassiliicoccales	*Methanomassiliicoccaceae*	*Methanomassiliicoccus*	Methanomicrobia archaeon DTU008	2579	100	0	100

Appendix IV - Table A4

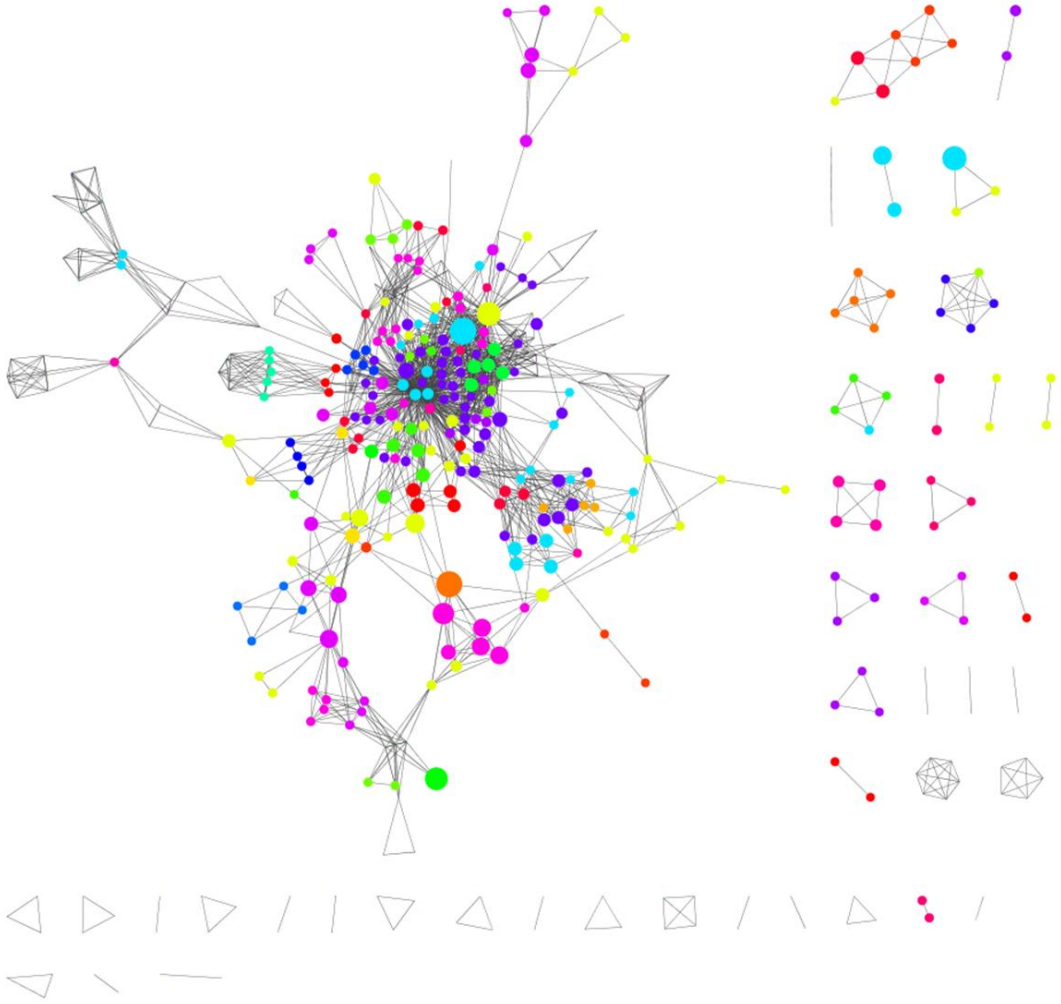
Table A3. Optimal growth condition, substrate utilisation, methanogenic pathways, and origin of the abundant OTUs related to methanogens

Species	OTU Number	Methanogenic Pathway	Temperature ; optimal temperature	NaCl; optimal NaCl	Substrate	Isolated from	Reference
Methanobacterium beijingense	41	Hydrogenotrophic	25-50; 37		H_2/CO_2	Anaerobic digesters	(Ma et al., 2005)
Methanobacterium formicicum	47	Hydrogenotrophic	25-45		Formate, H_2/CO_2	Cow grazing clover pasture	(Kelly et al., 2014)
Methanobacterium thermaggregans	21, 28	Hydrogenotrophic	50-70		H_2/CO_2		(Schupp et al., 2018)
Methanothermobacter tenebrarum strain RMAS	10	Hydrogenotrophic	45-80; 70	0.001-20 g/L	H_2/CO_2	Gas-associated formation water of a natural gas field	(Takahashi et al., 2012)
Methanocorpusculum aggregans	35	Hydrogenotrophic	39		H_2/CO_2	Anaerobic digester	(Joshi et al., 2018; XUN et al., 1989)
Methanoculleus receptaculi	32	Hydrogenotrophic	30-65, 50-55	0-1.3 M; 0.2 M	Formate, H_2/CO_2	Oil production water from oil field	(Cheng et al., 2008)
Methanoculleus sp. DTU006	24	Hydrogenotrophic	35-54		Formate, H_2/CO_2	Metagenomic recovery	(Kougias et al., 2017)
Methanoculleus sp. SLH121	19						N/A
Methanolinea mesophila	25	Hydrogenotrophic	20-40; 37	0-25 g/L	Formate, H_2/CO_2	Rice field soil	(Sakai et al., 2011)
Methanolinea tarda	01	Hydrogenotrophic	35-55; 50	0-1.5%	H2/CO2	Mesophilic anaerobic digester	(Sekiguchi et al., 2008; Yamamoto et al., 2014)
Methanoregula boonei 6A8	45	Hydrogenotrophic	35	0.4-50 mM	H_2/CO_2	Ombrotrophic peat bog	(Bräuer et al., 2015, 2011)
Methanospirillum hungatei JF-1	09, 22, 48	Hydrogenotrophic	20-40; 37		Formate, H_2/CO_2		(Gunsalus et al., 2016; Yu et al., 2013))
Methanothrix concilii GP6	04, 05, 07, 11, 16, 20, 26	Acetoclastic	35-40		Acetate, CO_2 (with DIET)		(Patell and Sprott, 1990; Yang et al., 2019)
Methanothrix thermophila	03	Acetoclastic	N/A; 55-60	Not required	Acetate	Thermal lake mud	(Kamagata et al., 2009)

Methanosarcina barkeri	33, 34	Hydrogenotrophic Acetoclastic Methlotrophic	25-42; 42	Not required	H_2/CO_2, methanol, methylamine and acetate		(Maestrojuan and Boone, 1999)
Methanosarcina flavescens	15, 12, 06	Hydrogenotrophic Acetoclastic Methlotrophic	45	6.8 mM	H_2/CO_2, methanol, mono-, di- and trimethylamine and acetate	Full-scale biogas plant	(Kern et al., 2016)
Methanosarcina horonobensis	8	Acetoclastic Methlotrophic	20-42; 37	0-0.35 M; 0.1 M	Methanol, dimethylamine, trimethylamine, dimethylsulfide and acetate	Groundwater	(Shimizu et al., 2011)
Euryarchaeota archaeon ADurb.Bin023	02, 18, 29	Methylotrophic			Methanol, methylamine	Metagenomic recovery	(Nobu et al., 2016)
Methanomassiliicoccus luminyensis B10	13	Hydrogenotrophic	25-45; 37	0.5-1.0 %	Reducing methanol with H_2	Human faeces	
Methanomicrobia archaeon DTU008	17						N/A

Appendix V - Figure A1

Figure A1. Co-occurrence microbial network constructed using 19 bacterial samples. Each node represents a taxon with its size proportional to the relative abundance and coloured according to their phylum. The network contains 433 nodes, 972 edges, and 58 sub networks.



Acidobacteria	R:255 G:57 B:0 - #FF3900
Actinobacteria	R:255 G:113 B:0 - #FF7100
Atribacteria	R:255 G:170 B:0 - #FFAA00
BRC1	R:255 G:227 B:0 - #FFE300
Bacteroidetes	R:227 G:255 B:0 - #E3FF00
CK-2C2-2	R:170 G:255 B:0 - #AAFF00
Caldiserica	R:113 G:255 B:0 - #71FF00
Chloroflexi	R:57 G:255 B:0 - #39FF00
Cloacimonetes	R:0 G:255 B:0 - #00FF00
Coprothermobacteraeota	R:0 G:255 B:57 - #00FF39
Cyanobacteria	R:0 G:255 B:113 - #00FF71
Epsilonbacteraeota	R:0 G:255 B:170 - #00FFAA
Fibrobacteres	R:0 G:255 B:227 - #00FFE3
Firmicutes	R:0 G:227 B:255 - #00E3FF
Fusobacteria	R:0 G:170 B:255 - #00AAFF
Hydrogenedentes	R:0 G:113 B:255 - #0071FF
Kiritimatiellaeota	R:0 G:57 B:255 - #0039FF
Marinimicrobia_(SAR406_clade)	R:0 G:0 B:255 - #0000FF
Omnitrophicaeota	R:57 G:0 B:255 - #3900FF
Patescibacteria	R:113 G:0 B:255 - #7100FF
Planctomycetes	R:170 G:0 B:255 - #AA00FF
Proteobacteria	R:227 G:0 B:255 - #E300FF
Spirochaetes	R:255 G:0 B:227 - #FF00E3
Synergistetes	R:255 G:0 B:170 - #FF00AA
Tenericutes	R:255 G:0 B:113 - #FF0071
Thermotogae	R:255 G:0 B:57 - #FF0039
Verrucomicrobia	R:255 G:0 B:0 - #FF0000

Reference

Adam, P.S., Borrel, G., Brochier-Armanet, C., Gribaldo, S., 2017. The growing tree of Archaea: New perspectives on their diversity, evolution and ecology. ISME J. 11, 2407–2425. https://doi.org/10.1038/ismej.2017.122

Ahlert, S., Zimmermann, R., Ebling, J., König, H., 2016. Analysis of propionate-degrading consortia from agricultural biogas plants. Microbiologyopen. https://doi.org/10.1002/mbo3.386

Alexiev, A., Coil, D.A., Badger, J.H., Enticknap, J., Ward, N., Robb, F.T., Eisen, J.A., 2014. Complete Genome Sequence of Coprothermobacter proteolyticus DSM 5265. Genome Announc. https://doi.org/10.1128/genomea.00470-14

Angle, J.C., Morin, T.H., Solden, L.M., Narrowe, A.B., Smith, G.J., Borton, M.A., Rey-Sanchez, C., Daly, R.A., Mirfenderesgi, G., Hoyt, D.W., Riley, W.J., Miller, C.S., Bohrer, G., Wrighton, K.C., 2017. Methanogenesis in oxygenated soils is a substantial fraction of wetland methane emissions. Nat. Commun. 8, 1–9. https://doi.org/10.1038/s41467-017-01753-4

Appels, L., Lauwers, J., Degrve, J., Helsen, L., Lievens, B., Willems, K., Van Impe, J., Dewil, R., 2011. Anaerobic digestion in global bio-energy production: Potential and research challenges. Renew. Sustain. Energy Rev. 15, 4295–4301. https://doi.org/10.1016/j.rser.2011.07.121

Balashov, S., Ehrlich, R.L., Palmer, J.N., Workman, A.D., Hammond, J., Cohen, N.A., Ehrlich, G.D., Mell, J.C., Nithinadappauphsupennedu, N.D.A., Workmanamailmedupennedu, A.D.W., 2018. Species-level bacterial community profiling of the healthy sinonasal microbiome using Pacific Biosciences sequencing of full-length 16S rRNA genes. bioRxiv.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J. Comput. Biol. https://doi.org/10.1089/cmb.2012.0021

Barua, S., Dhar, B.R., 2017. Advances towards understanding and engineering direct interspecies electron transfer in anaerobic digestion. Bioresour. Technol. 244, 698–707. https://doi.org/10.1016/j.biortech.2017.08.023

Batstone, D.J., Jensen, P.D., 2010. Anaerobic Processes, in: Treatise on Water Science. https://doi.org/10.1016/B978-0-444-53199-5.00097-X

Ben Hania, W., Bouanane-Darenfed, A., Cayol, J.L., Ollivier, B., Fardeau, M.L., 2016. Reclassification of Anaerobaculum mobile, Anaerobaculum thermoterrenum, Anaerobaculum hydrogeniformans as Acetomicrobium mobile comb. nov., Acetomicrobium thermoterrenum comb. nov. and Acetomicrobium hydrogeniformans comb. nov., respectively, and emendati. Int. J. Syst. Evol. Microbiol. https://doi.org/10.1099/ijsem.0.000910

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., Boutell, J.M., Bryant, J., Carter, R.J., Keira Cheetham, R., Cox, A.J., Ellis, D.J., Flatbush, M.R., Gormley, N.A., Humphray, S.J., Irving, L.J., Karbelashvili, M.S., Kirk, S.M., Li, H., Liu, X., Maisinger, K.S., Murray, L.J., Obradovic, B., Ost, T., Parkinson, M.L., Pratt, M.R., Rasolonjatovo, I.M.J., Reed, M.T., Rigatti, R., Rodighiero, C., Ross, M.T., Sabot, A., Sankar, S. V., Scally, A., Schroth, G.P., Smith, M.E., Smith, V.P., Spiridou, A., Torrance, P.E., Tzonev, S.S., Vermaas, E.H., Walter, K., Wu, X., Zhang, L., Alam, M.D., Anastasi, C., Aniebo, I.C., Bailey, D.M.D., Bancarz, I.R., Banerjee, S., Barbour, S.G., Baybayan, P.A., Benoit, V.A., Benson, K.F., Bevis, C., Black, P.J., Boodhun, A., Brennan, J.S., Bridgham, J.A., Brown, R.C., Brown, A.A., Buermann, D.H., Bundu, A.A., Burrows, J.C., Carter, N.P., Castillo, N., Catenazzi, M.C.E., Chang, S., Neil Cooley, R., Crake, N.R., Dada, O.O., Diakoumakos, K.D., Dominguez-Fernandez, B., Earnshaw, D.J., Egbujor, U.C., Elmore, D.W., Etchin, S.S., Ewan, M.R., Fedurco, M., Fraser, L.J., Fuentes Fajardo, K. V., Scott Furey, W., George, D., Gietzen, K.J., Goddard, C.P., Golda, G.S., Granieri, P.A., Green, D.E., Gustafson, D.L., Hansen, N.F., Harnish, K., Haudenschild, C.D., Heyer, N.I., Hims, M.M., Ho, J.T., Horgan, A.M., Hoschler, K., Hurwitz, S., Ivanov, D. V., Johnson, M.Q., James, T., Huw Jones, T.A., Kang, G.D., Kerelska, T.H., Kersey, A.D., Khrebtukova, I., Kindwall, A.P.,

Kingsbury, Z., Kokko-Gonzales, P.I., Kumar, A., Laurent, M.A., Lawley, C.T., Lee, S.E., Lee, X., Liao, A.K., Loch, J.A., Lok, M., Luo, S., Mammen, R.M., Martin, J.W., McCauley, P.G., McNitt, P., Mehta, P., Moon, K.W., Mullens, J.W., Newington, T., Ning, Z., Ling Ng, B., Novo, S.M., O'Neill, M.J., Osborne, M.A., Osnowski, A., Ostadan, O., Paraschos, L.L., Pickering, L., Pike, Andrew C., Pike, Alger C., Chris Pinkard, D., Pliskin, D.P., Podhasky, J., Quijano, V.J., Raczy, C., Rae, V.H., Rawlings, S.R., Chiva Rodriguez, A., Roe, P.M., Rogers, John, Rogert Bacigalupo, M.C., Romanov, N., Romieu, A., Roth, R.K., Rourke, N.J., Ruediger, S.T., Rusman, E., Sanches-Kuiper, R.M., Schenker, M.R., Seoane, J.M., Shaw, R.J., Shiver, M.K., Short, S.W., Sizto, N.L., Sluis, J.P., Smith, M.A., Ernest Sohna Sohna, J., Spence, E.J., Stevens, K., Sutton, N., Szajkowski, L., Tregidgo, C.L., Turcatti, G., Vandevondele, S., Verhovsky, Y., Virk, S.M., Wakelin, S., Walcott, G.C., Wang, J., Worsley, G.J., Yan, J., Yau, L., Zuerlein, M., Rogers, Jane, Mullikin, J.C., Hurles, M.E., McCooke, N.J., West, J.S., Oaks, F.L., Lundberg, P.L., Klenerman, D., Durbin, R., Smith, A.J., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. https://doi.org/10.1038/nature07517

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. https://doi.org/10.1093/bioinformatics/btu170

Bräuer, S., Cadillo-Quiroz, H., Kyrpides, N., Woyke, T., Goodwin, L., Detter, C., Podell, S., Yavitt, J.B., Zinder, S.H., 2015. Genome of Methanoregula boonei 6A8 reveals adaptations to oligotrophic peatland environments. Microbiol. (United Kingdom) 161, 1572–1581. https://doi.org/10.1099/mic.0.000117

Bräuer, S.L., Cadillo-Quiroz, H., Ward, R.J., Yavitt, J.B., Zinder, S.H., 2011. Methanoregula boonei gen. nov., sp. nov., an acidiphilic methanogen isolated from an acidic peat bog. Int. J. Syst. Evol. Microbiol. 61, 45–52. https://doi.org/10.1099/ijs.0.021782-0

Bremer, H., Dennis, P., 1996. Modulation of chemical composition and other parameters of the cell by growth rate, in: Escherichia coli and Salmonella typhimurium, in: Cellular and Molecular Biology.

Cai, M., Wilkins, D., Chen, J., Ng, S.K., Lu, H., Jia, Y., Lee, P.K.H., 2016. Metagenomic reconstruction of key anaerobic digestion pathways in municipal sludge and industrial wastewater biogas-producing systems. Front. Microbiol. 7, 1–12. https://doi.org/10.3389/fmicb.2016.00778

Calusinska, M., Goux, X., Fossépré, M., Muller, E.E.L., 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. https://doi.org/10.1186/s13068-018-1195-8

Campanaro, S., Treu, L., Kougias, P.G., Zhu, X., Angelidaki, I., 2018. Taxonomy of anaerobic digestion microbiome reveals biases associated with the applied high throughput sequencing strategies. Sci. Rep. 8, 1–12. https://doi.org/10.1038/s41598-018-20414-0

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G., Knight, R., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6, 1621–1624. https://doi.org/10.1038/ismej.2012.8

Carneiro, M.O., Russ, C., Ross, M.G., Gabriel, S.B., Nusbaum, C., DePristo, M.A., 2012. Pacific biosciences sequencing technology for genotyping and variation discovery in human data. BMC Genomics.

Castelle, C.J., Hug, L.A., Wrighton, K.C., Thomas, B.C., Williams, K.H., Wu, D., Tringe, S.G., Singer, S.W., Eisen, J.A., Banfield, J.F., 2013. Extraordinary phylogenetic diversity and metabolic versatility in aquifer sediment. Nat. Commun. 4, 1–10. https://doi.org/10.1038/ncomms3120

Chandler, J.A., Jewell, W.J., Gossett, J.M., 1980. Predicting methane fermentation biodegradability. Symp Biotechnol Energy Prod Conserv 2nd Proc.

Cheng, L., Qiu, T.-L., Deng, Y., Zhang, H., Li, X., Xiao-Bo, Wang, W.-D., 2008. Isolation and

characterization of Methanoculleus receptaculi sp. nov. from Shengli oil field, China. FEMS Microbiol. Lett. 285, 65–71. https://doi.org/10.1111/j.1574-6968.2008.01212.x

Chouari, R., Le Paslier, D., Daegelen, P., Ginestet, P., Weissenbach, J., Sghir, A., 2005. Novel predominant archaeal and bacterial groups revealed by molecular analysis of an anaerobic sludge digester. Environ. Microbiol. 7, 1104–1115. https://doi.org/10.1111/j.1462-2920.2005.00795.x

Chovatia, M., Sikorski, J., Schröder, M., Lapidus, A., Nolan, M., Tice, H., del Rio, T.G., Copeland, A., Cheng, J.F., Lucas, S., Chen, F., Bruce, D., Goodwin, L., Pitluck, S., Ivanova, N., Mavromatis, K., Ovchinnikova, G., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y.J., Jeffries, C.D., Chain, P., Saunders, E., Detter, J.C., Brettin, T., Rohde, M., Göker, M., Spring, S., Bristow, J., Markowitz, V., Hugenholtz, P., Kyrpides, N.C., Klenk, H.P., Eisen, J.A., 2009. Complete genome sequence of Thermanaerovibrio acidaminovorans type strain (Su883 T). Stand. Genomic Sci. 1, 254–261. https://doi.org/10.4056/sigs.40645

Coenye, T., Vandamme, P., 2003. Intragenomic heterogeneity between multiple 16S ribosomal RNA operons in sequenced bacterial genomes. FEMS Microbiol. Lett. https://doi.org/10.1016/S0378-1097(03)00717-1

D H Bergey; William B Whitman; M Goodfellow; Peter Kämpfer; Hans-Jürgen, 2005. Bergey's Manual® of Systematic Bacteriology: Volume Two: The Proteobacteria. New York Springer, ©2005.

Daugherty, M., Polanuyer, B., Farrell, M., Scholle, M., Lykidis, A., De Crécy-Lagard, V., 2002. Complete reconstitution of the human coenzyme A biosynthetic pathway via comparative genomics. J. Biol. Chem. https://doi.org/10.1074/jbc.M201708200

Deppenmeier, U., 2002. The unique biochemistry of methanogenesis. Prog Nucleic Acid Res Mol Biol 71, 223–283.

DESA/UN-WUP, 2018. World Urbanization Prospects: The 2018 Revision. Dep. Econ. Soc. Aff.

Dixon, P., 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x

Dohm, J.C., Lottaz, C., Borodina, T., Himmelbauer, H., 2008. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res. https://doi.org/10.1093/nar/gkn425

Dolfing J, 1987. Acetogenesis, in: Alexander J.B. Zehnder (Ed.), Biology of Anaerobic Microorganisms Volume 1 of Wiley Series in Ecological and Applied Microbiology. Wiley, New York, pp. 417–468.

Earl, J.P., Adappa, N.D., Krol, J., Bhat, A.S., Balashov, S., Ehrlich, R.L., Palmer, J.N., Workman, A.D., Blasetti, M., Sen, B., Hammond, J., Cohen, N.A., Ehrlich, G.D., Mell, J.C., 2018. Species-level bacterial community profiling of the healthy sinonasal microbiome using Pacific Biosciences sequencing of full- length 16S rRNA genes 1–26.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics. https://doi.org/10.1093/bioinformatics/btr381

Eijsink, V.G.H., Horn, S.J., Pope, P.B., Zamanzadeh, M., Frank, J.A., Hagen, L.H., Arntzen, M.Ø., 2016. Quantitative Metaproteomics Highlight the Metabolic Contributions of Uncultured Phylotypes in a Thermophilic Anaerobic Digester. Appl. Environ. Microbiol. 83. https://doi.org/10.1128/aem.01955-16

Ellis, J.T., Tramp, C., Sims, R.C., Miller, C.D., 2012. Characterization of a Methanogenic Community within an Algal Fed Anaerobic Digester. ISRN Microbiol. https://doi.org/10.5402/2012/753892

Embree, M., Nagarajan, H., Movahedi, N., Chitsaz, H., Zengler, K., 2014. Single-cell genome and metatranscriptome sequencing reveal metabolic interactions of an alkane-degrading methanogenic community. ISME J. 8, 757–767. https://doi.org/10.1038/ismej.2013.187

Era7, 2016. Full-length 16S taxonomic profiling with PacBio [WWW Document]. URL https://era7bioinformatics.com/en/page.cfm?id=2716&title=full-lenght-16s-taxonomic-profiling-

with-pacbio

Evans, P.N., Parks, D.H., Chadwick, G.L., Robbins, S.J., Orphan, V.J., Golding, S.D., Tyson, G.W., 2015. Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. Science (80-.). 350, 434–438. https://doi.org/10.1126/science.aac7745

Faust, K., Raes, J., 2016. CoNet app: inference of biological association networks using Cytoscape. F1000Research. https://doi.org/10.12688/f1000research.9050.2

Ferguson, D.J., Gorlatova, N., Grahame, D.A., Krzycki, J.A., 2000. Reconstitution of dimethylamine:coenzyme M methyl transfer with a discrete corrinoid protein and two methyltransferases purified from Methanosarcina barkeri. J. Biol. Chem. https://doi.org/10.1074/jbc.M910218199

Ferguson, D.J., Krzycki, J.A., 1997. Reconstitution of trimethylamine-dependent coenzyme M methylation with the trimethylamine corrinoid protein and the isozymes of methyltransferase II from Methanosarcina barkeri. J. Bacteriol. https://doi.org/10.1128/jb.179.3.846-852.1997

Ferry, J.G., 1993. Methanogenesis: Ecology, Physicology, Biochemistry & Genetics, Methanogenesis. Chapman & Hall, New York.

Ferry, J.G., 1992. Methane from acetate. J. Bacteriol.

Fish, J.A., Chai, B., Wang, Q., Sun, Y., Brown, C.T., Tiedje, J.M., Cole, J.R., 2013. FunGene: The functional gene pipeline and repository. Front. Microbiol. 4, 1–14. https://doi.org/10.3389/fmicb.2013.00291

Fournier, G.P., Gogarten, J.P., 2008. Evolution of acetoclastic methanogenesis in Methanosarcina via horizontal gene transfer from cellulolytic Clostridia. J. Bacteriol. https://doi.org/10.1128/JB.01382-07

Frank, J.A., Pan, Y., Tooming-Klunderud, A., Eijsink, V.G.H., McHardy, A.C., Nederbragt, A.J., Pope, P.B., 2016. Improved metagenome assemblies and taxonomic binning using long-read circular consensus sequence data. Sci. Rep. 6, 1–10. https://doi.org/10.1038/srep25373

Franzén, O., Hu, J., Bao, X., Itzkowitz, S.H., Peter, I., Bashir, A., 2015. Improved OTU-picking using long-read 16S rRNA gene amplicon sequencing and generic hierarchical clustering. Microbiome. https://doi.org/10.1186/s40168-015-0105-6

Fuks, G., Elgart, M., Amir, A., Zeisel, A., Turnbaugh, P.J., Soen, Y., Shental, N., 2018. Combining 16S rRNA gene variable regions enables high-resolution microbial community profiling. Microbiome. https://doi.org/10.1186/s40168-017-0396-x

Gagliano, M.C., Braguglia, C.M., Petruccioli, M., Rossetti, S., 2015. Ecology and biotechnological potential of the thermophilic fermentative Coprothermobacter spp. FEMS Microbiol. Ecol. https://doi.org/10.1093/femsec/fiv018

GMI, 2013. Successful Applications of Anaerobic Digestion From Across the World [WWW Document]. Glob. Methane Initiat.

Goodwin, S., McPherson, J.D., McCombie, W.R., 2016. Coming of age: Ten years of next-generation sequencing technologies. Nat. Rev. Genet. 17, 333–351. https://doi.org/10.1038/nrg.2016.49

Graf, J., 2014. The family rikenellaceae, in: The Prokaryotes: Other Major Lineages of Bacteria and The Archaea. https://doi.org/10.1007/978-3-642-38954-2_134

Guangsheng, C., Plugge, C.M., Roelofsen, W., Houwen, F.P., Stams, A.J.M., 1992. Selenomonas acidaminovorans sp. nov., a versatile thermophilic proton-reducing anaerobe able to grow by decarboxylation of succinate to propionate. Arch. Microbiol. https://doi.org/10.1007/BF00245286

Gunsalus, R.P., Cook, L.E., Crable, B., Rohlin, L., McDonald, E., Mouttaki, H., Sieber, J.R., Poweleit, N., Zhou, H., Lapidus, A.L., Daligault, H.E., Land, M., Gilna, P., Ivanova, N., Kyrpides, N., Culley, D.E., McInerney, M.J., 2016. Complete genome sequence of Methanospirillum hungatei type strain JF1. Stand. Genomic Sci. 11, 1–10. https://doi.org/10.1186/s40793-015-0124-8

Ha, P.T., Lindemann, S.R., Shi, L., Dohnalkova, A.C., Fredrickson, J.K., Madigan, M.T., Beyenal, H., 2017. Syntrophic anaerobic photosynthesis via direct interspecies electron transfer. Nat.

Commun. https://doi.org/10.1038/ncomms13924

Hahnke, S., Langer, T., Koeck, D.E., Klocke, M., 2016. Description of Proteiniphilum saccharofermentans sp. nov., Petrimonas mucosa sp. nov. and Fermentimonas caenicola gen. nov., sp. nov., isolated from mesophilic laboratory-scale biogas reactors, and emended description of the genus Proteiniphilum. Int. J. Syst. Evol. Microbiol. https://doi.org/10.1099/ijsem.0.000902

Handley, K.M., Verberkmoes, N.C., Steefel, C.I., Williams, K.H., Sharon, I., Miller, C.S., Frischkorn, K.R., Chourey, K., Thomas, B.C., Shah, M.B., Long, P.E., Hettich, R.L., Banfield, J.F., 2013. Biostimulation induces syntrophic interactions that impact C, S and N cycling in a sediment microbial community. ISME J. https://doi.org/10.1038/ismej.2012.148

Hania, W. Ben, Godbane, R., Postec, A., Hamdi, M., Ollivier, B., Fardeau, M.L., 2012. Defluviitoga tunisiensis gen. nov., sp. nov., a thermophilic bacterium isolated from a mesothermic and anaerobic whey digester. Int. J. Syst. Evol. Microbiol. https://doi.org/10.1099/ijs.0.033720-0

Hao, L., McIlroy, S.J., Kirkegaard, R.H., Karst, S.M., Fernando, W.E.Y., Aslan, H., Meyer, R.L., Albertsen, M., Nielsen, P.H., Dueholm, M.S., 2018. Novel prosthecate bacteria from the candidate phylum Acetothermia. ISME J. 12, 2225–2237. https://doi.org/10.1038/s41396-018-0187-9

Harismendy, O., Ng, P.C., Strausberg, R.L., Wang, X., Stockwell, T.B., Beeson, K.Y., Schork, N.J., Murray, S.S., Topol, E.J., Levy, S., Frazer, K.A., 2009. Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome Biol. https://doi.org/10.1186/gb-2009-10-3-r32

Hatamoto, M., Imachi, H., Fukayo, S., Ohashi, A., Harada, H., 2007. Syntrophomonas palmitatica sp. nov., an anaerobic syntrophic, long-chain fatty-acid-oxidizing bacterium isolated from methanogenic sludge. Int. J. Syst. Evol. Microbiol. 57, 2137–2142. https://doi.org/10.1099/ijs.0.64981-0

Hebert, P.D.N., Braukmann, T.W.A., Prosser, S.W.J., Ratnasingham, S., deWaard, J.R., Ivanova, N. V., Janzen, D.H., Hallwachs, W., Naik, S., Sones, J.E., Zakharov, E. V., 2018. A Sequel to Sanger: Amplicon sequencing that scales. BMC Genomics. https://doi.org/10.1186/s12864-018-4611-3

Hodgman, C., 2017. BIOS Instant Notes in Bioinformatics, BIOS Instant Notes in Bioinformatics. https://doi.org/10.4324/9780203967249

Holmes, D.E., Shrestha, P.M., Walker, D.J.F., Dang, Y., Nevin, K.P., Woodard, T.L., Lovley, D.R., 2017. Metatranscriptomic Evidence for Direct Interspecies Electron Transfer between Geobacter and Methanothrix Species in Methanogenic Rice Paddy Soils. Appl. Environ. Microbiol. 83. https://doi.org/10.1128/AEM.00223-17

Huber, H., Thomm, M., König, H., Thies, G., Stetter, K.O., 1982. Methanococcus thermolithotrophicus, a novel thermophilic lithotrophic methanogen. Arch. Microbiol. 132, 47–50. https://doi.org/10.1007/BF00690816

Hugenholtz, P., 2002. Exploring prokaryotic diversity in the genomic era. Genome Biol. 3, REVIEWS0003.

Jackson, B.E., Bhupathiraju, V.K., Tanner, R.S., Woese, C.R., McInerney, M.J., 1999. Syntrophus aciditrophicus sp. nov., a new anaerobic bacterium that degrades fatty acids and benzoate in syntrophic association with hydrogen- using microorganisms. Arch. Microbiol. 171, 107–114. https://doi.org/10.1007/s002030050685

Jenicek, P., Celis, C.A., Koubova, J., Pokorna, D., 2011. Comparison of microbial activity in anaerobic and microaerobic digesters. Water Sci. Technol. https://doi.org/10.2166/wst.2011.579

Jia, Y., Wilkins, D., Lu, H., Cai, M., Lee, P.K.H., 2016. Long-term enrichment on cellulose or xylan causes functional and taxonomic convergence of microbial communities from anaerobic digesters. Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.03360-15

Joshi, A., Lanjekar, V., Dhakephalkar, P.K., Dagar, S.S., 2018. Cultivation of multiple genera of

hydrogenotrophic methanogens from different environmental niches. Anaerobe 50, 64–68. https://doi.org/10.1016/j.anaerobe.2018.02.001

Kadnikov, V. V., Mardanov, A. V., Beletsky, A. V., Rakitin, A.L., Frank, Y.A., Karnachuk, O. V., Ravin, N. V., 2019. Phylogeny and physiology of candidate phylum BRC1 inferred from the first complete metagenome-assembled genome obtained from deep subsurface aquifer. Syst. Appl. Microbiol. 42, 67–76. https://doi.org/10.1016/j.syapm.2018.08.013

Kamagata, Y., Nakamura, K., Mikami, E., Endo, G., Oyaizu, H., Kawasaki, H., Yamasato, K., Koga, Y., 2009. Characterization of Three Thermophilic Strains of Methanothrix ("Methanosaeta") thermophila sp. nov. and Rejection of Methanothrix ("Methanosaeta") thermoacetophila. Int. J. Syst. Bacteriol. 42, 463–468. https://doi.org/10.1099/00207713-42-3-463

Kanehisa, M., Sato, Y., Morishima, K., 2016. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. J. Mol. Biol. https://doi.org/10.1016/j.jmb.2015.11.006

Kato, S., Hashimoto, K., Watanabe, K., 2012. Methanogenesis facilitated by electric syntrophy via (semi)conductive iron-oxide minerals. Environ. Microbiol. https://doi.org/10.1111/j.1462-2920.2011.02611.x

Kelly, W.J., Leahy, S.C., Li, D., Perry, R., Lambie, S.C., Attwood, G.T., Altermann, E., 2014. The complete genome sequence of the rumen methanogen Methanobacterium formicicum BRM9. Stand. Genomic Sci. 9, 1–8. https://doi.org/10.1186/1944-3277-9-15

Kerkhof, L., Ward, B.B., 1993. Comparison of Nucleic Acid Hybridization and Fluorometry for Measurement of the Relationship between RNA/DNA Ratio and Growth Rate in a Marine Bacterium. Appl. Environ. Microbiol.

Kern, T., Fischer, M.A., Deppenmeier, U., Schmitz, R.A., Rother, M., 2016. Methanosarcina flavescens sp. nov., a methanogenic archaeon isolated from a full-scale anaerobic digester. Int. J. Syst. Evol. Microbiol. 66, 1533–1538. https://doi.org/10.1099/ijsem.0.000894

Khanal, S.K., Huang, J.-C., 2006. Online Oxygen Control for Sulfide Oxidation in Anaerobic Treatment of High-Sulfate Wastewater. Water Environ. Res. https://doi.org/10.2175/106143006x98804

Kim, M., Ahn, Y.H., Speece, R.E., 2002. Comparative process stability and efficiency of anaerobic digestion; mesophilic vs. thermophilic. Water Res. https://doi.org/10.1016/S0043-1354(02)00147-1

Kim, M., Morrison, M., Yu, Z., 2011. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. J. Microbiol. Methods 84, 81–87. https://doi.org/10.1016/j.mimet.2010.10.020

Knight, R., Vrbanac, A., Taylor, B.C., Aksenov, A., Callewaert, C., Debelius, J., Gonzalez, A., Kosciolek, T., McCall, L.I., McDonald, D., Melnik, A. V., Morton, J.T., Navas, J., Quinn, R.A., Sanders, J.G., Swafford, A.D., Thompson, L.R., Tripathi, A., Xu, Z.Z., Zaneveld, J.R., Zhu, Q., Caporaso, J.G., Dorrestein, P.C., 2018. Best practices for analysing microbiomes. Nat. Rev. Microbiol. https://doi.org/10.1038/s41579-018-0029-9

Koch, M., Dolfing, J., Wuhrmann, K., Zehnder, A.J.B., 1983. Pathways of propionate degradation by enriched methanogenic cultures. Appl. Environ. Microbiol. 45, 1411–1414.

Korlach, J., Bjornson, K.P., Chaudhuri, B.P., Cicero, R.L., Flusberg, B.A., Gray, J.J., Holden, D., Saxena, R., Wegener, J., Turner, S.W., 2010. Real-time DNA sequencing from single polymerase molecules. Methods Enzymol.

Kougias, P.G., Campanaro, S., Treu, L., Zhu, X., Angelidaki, I., 2017. A novel archaeal species belonging to Methanoculleus genus identified via de-novo assembly and metagenomic binning process in biogas reactors. Anaerobe 46, 23–32. https://doi.org/10.1016/j.anaerobe.2017.02.009

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. https://doi.org/10.1093/molbev/msy096

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7 : Molecular Evolutionary Genetics Analysis Version 7 . 0 for Bigger Datasets Brief communication. Mol. Biol. Evol. https://doi.org/10.1093/molbev/msw054

Kurr, M., Huber, R., König, H., Jannasch, H.W., Fricke, H., Trincone, A., Kristjansson, J.K., Stetter, K.O., 1991. Methanopyrus kandleri, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110°C. Arch. Microbiol. 156, 239–247. https://doi.org/10.1007/BF00262992

Letunic, I., Bork, P., 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. https://doi.org/10.1093/nar/gkw290

Li, M., Li, D., Tang, Y., Wu, F., Wang, J., 2017. Cytocluster: A cytoscape plugin for cluster analysis and visualization of biological networks. Int. J. Mol. Sci. https://doi.org/10.3390/ijms18091880

Liu, F., Rotaru, A.E., Shrestha, P.M., Malvankar, N.S., Nevin, K.P., Lovley, D.R., 2012. Promoting direct interspecies electron transfer with activated carbon. Energy Environ. Sci. https://doi.org/10.1039/c2ee22459c

Liu, Q., Ren, Z.J., Huang, C., Liu, B., Ren, N., Xing, D., 2016. Multiple syntrophic interactions drive biohythane production from waste sludge in microbial electrolysis cells. Biotechnol. Biofuels. https://doi.org/10.1186/s13068-016-0579-x

Liu, Y., Balkwill, D.L., Henry, C.A., Drake, G.R., Boone, D.R., 1999. Characterization of the anaerobic propionate- degrading syntrophs Smithella propionica. Int. J. Syst. Bacteriol. 49, 545–556. https://doi.org/10.1099/00207713-49-2-545

Liu, Y., Boone, D.R., 1991. Effects of salinity on methanogenic decomposition. Bioresour. Technol. 35, 271–273.

Lovley, D.R., 2017. Syntrophy Goes Electric: Direct Interspecies Electron Transfer. Annu. Rev. Microbiol. https://doi.org/10.1146/annurev-micro-030117-020420

Lu, Y., Conrad, R., 2005. In situ stable isotope probing of methanogenic Archaea in the rice rhizosphere. Science (80-.). 309, 1088–1090. https://doi.org/10.1126/science.1113435

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, A., Buchner, A., Lai, T., Steppi, S., Jacob, G., Förster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S., Hermann, S., Jost, R., König, A., Liss, T., Lüßbmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., Schleifer, K.H., 2004. ARB: A software environment for sequence data. Nucleic Acids Res. 32, 1363–1371. https://doi.org/10.1093/nar/gkh293

Lv, L., Mbadinga, S.M., Wang, L.Y., Liu, J.F., Gu, J.D., Mu, B.Z., Yang, S.Z., 2015. Acetoclastic methanogenesis is likely the dominant biochemical pathway of palmitate degradation in the presence of sulfate. Appl. Microbiol. Biotechnol. https://doi.org/10.1007/s00253-015-6669-z

Ma, K., Liu, X., Dong, X., 2005. Methanobacterium beijingense sp. nov., a novel methanogen isolated from anaerobic digesters. Int. J. Syst. Evol. Microbiol. 55, 325–329. https://doi.org/10.1099/ijs.0.63254-0

Maeder, D.L., Anderson, I., Brettin, T.S., Bruce, D.C., Gilna, P., Han, C.S., Lapidus, A., Metcalf, W.W., Saunders, E., Tapia, R., Sowers, K.R., 2006. The Methanosarcina barkeri genome: Comparative analysis with Methanosarcina acetivorans and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes. J. Bacteriol. 188, 7922–7931. https://doi.org/10.1128/JB.00810-06

Maestrojuan, G.M., Boone, D.R., 1999. Characterization of Methanosarcina barkeri MST and 227, Methanosarcina mazei S-6T, and Methanosarcina vacuolata Z-76IT. Int. J. Syst. Evol. Bacteriol. 41, 267–274.

Mao, C., Feng, Y., Wang, X., Ren, G., 2015. Review on research achievements of biogas from anaerobic digestion. Renew. Sustain. Energy Rev. 45, 540–555. https://doi.org/10.1016/j.rser.2015.02.032

McCarty, P.L., Rittmann, B.E., 2001. Anaerobic Treatment by Methanogenesis, in: Environmental Biotechnology: Principles and Applications. Mc Graw Hill, pp. 569–627.

McGlynn, S.E., Chadwick, G.L., Kempes, C.P., Orphan, V.J., 2015. Single cell activity reveals direct electron transfer in methanotrophic consortia. Nature. https://doi.org/10.1038/nature15512

McIlroy, S.J., Kirkegaard, R.H., Dueholm, M.S., Fernando, E., Karst, S.M., Albertsen, M., Nielsen, P.H., 2017. Culture-independent analyses reveal novel anaerolineaceae as abundant primary fermenters in anaerobic digesters treating waste activated sludge. Front. Microbiol. https://doi.org/10.3389/fmicb.2017.01134

McInerney, M.J., Sieber, J.R., Gunsalus, R.P., 2009. Syntrophy in anaerobic global carbon cycles. Curr. Opin. Biotechnol. https://doi.org/10.1016/j.copbio.2009.10.001

Mei, R., Nobu, M.K., Narihiro, T., Kuroda, K., Muñoz Sierra, J., Wu, Z., Ye, L., Lee, P.K.H., Lee, P.H., van Lier, J.B., McInerney, M.J., Kamagata, Y., Liu, W.T., 2017. Operation-driven heterogeneity and overlooked feed-associated populations in global anaerobic digester microbiome. Water Res. 124, 77–84. https://doi.org/10.1016/j.watres.2017.07.050

Minoche, A.E., Dohm, J.C., Himmelbauer, H., 2011. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. Genome Biol. https://doi.org/10.1186/gb-2011-12-11-r112

Mir, M.A., Hussain, A., Verma, C., 2016. Design considerations and operational performance of anaerobic digester: A review. Cogent Eng. https://doi.org/10.1080/23311916.2016.1181696

Müller, B., Sun, L., Westerholm, M., Schnürer, A., 2016. Bacterial community composition and fhs profiles of low- and high-ammonia biogas digesters reveal novel syntrophic acetate-oxidising bacteria. Biotechnol. Biofuels. https://doi.org/10.1186/s13068-016-0454-9

Nakamura, K., Oshima, T., Morimoto, T., Ikeda, S., Yoshikawa, H., Shiwa, Y., Ishikawa, S., Linak, M.C., Hirai, A., Takahashi, H., Altaf-Ul-Amin, M., Ogasawara, N., Kanaya, S., 2011. Sequence-specific error profile of Illumina sequencers. Nucleic Acids Res.

https://doi.org/10.1093/nar/gkr344

Narihiro, T., Kamagata, Y., 2013. Cultivating Yet-to-be Cultivated Microbes: The Challenge Continues. Microbes Environ. https://doi.org/10.1264/jsme2.me2802rh

Narrowe, A.B., Angle, J.C., Daly, R.A., Stefanik, K.C., Wrighton, K.C., Miller, C.S., 2017. High-resolution sequencing reveals unexplored archaeal diversity in freshwater wetland soils. Environ. Microbiol. https://doi.org/10.1111/1462-2920.13703

Nguyen, M.D., 2018. Oxidation-Reduction Potential-Based Micro-Aeration Control System for Anaerobic Digestion. University of Hawai‘i at Mānoa.

Nobu, M.K., Narihiro, T., Kuroda, K., Mei, R., Liu, W.T., 2016. Chasing the elusive Euryarchaeota class WSA2: Genomes reveal a uniquely fastidious methyl-reducing methanogen. ISME J. 10, 2478–2487. https://doi.org/10.1038/ismej.2016.33

Nobu, M.K., Narihiro, T., Rinke, C., Kamagata, Y., Tringe, S.G., Woyke, T., Liu, W.T., 2015. Microbial dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor. ISME J. 9, 1710–1722. https://doi.org/10.1038/ismej.2014.256

Olson, M. V., Kaul, R., Barber, R.D., Zhang, L., Harnack, M., Smith, K.S., Ingram-Smith, C., 2011. Complete Genome Sequence of Methanosaeta concilii, a Specialist in Aceticlastic Methanogenesis. J. Bacteriol. 193, 3668–3669. https://doi.org/10.1128/jb.05031-11

Park, J.H., Kang, H.J., Park, K.H., Park, H.D., 2018. Direct interspecies electron transfer via conductive materials: A perspective for anaerobic digestion applications. Bioresour. Technol. 254, 300–311. https://doi.org/10.1016/j.biortech.2018.01.095

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., Tyson, G.W., 2015. CheckM: assessing the quality of microbial genomes recovered from. Cold Spring Harb. Lab. Press Method 1, 1–31. https://doi.org/10.1101/gr.186072.114

Parte, A.C., 2018. LPSN - List of prokaryotic names with standing in nomenclature (Bacterio.net), 20 years on. Int. J. Syst. Evol. Microbiol. 68, 1825–1829. https://doi.org/10.1099/ijsem.0.002786

Patell, G., Sprott, G., 1990. Methanosaeta concilii gen. nov. sp. nov. (“Methanothrix concilii”) and Methanosaeta thermoacetophila nom. rev., comb. nov. Int. J. Syst. Bacteriol. 40, 79–82.

Patell, G.B., Roth, L.A., 1977. Effect of sodium chloride on growth and methane production of methanogens. Can. J. Microbiol. 23, 893–897.

Pattnaik, P., Mishra, S.R., Bharati, K., Mohanty, S.R., Sethunathan, N., Adhya, T.K., 2000. Influence of salinity on methanogenesis and associated microflora in tropical rice soils. Microbiol. Res. 155, 215–220.

Pavan, M.E., Pavan, E.E., Glaeser, S.P., Etchebehere, C., Kämpfer, P., Pettinari, M.J., López, N.I., 2018. Proposal for a new classification of a deep branching bacterial phylogenetic lineage: Transfer of Coprothermobacter proteolyticus and Coprothermobacter platensis to Coprothermobacteraceae fam. nov., within Coprothermobacterales ord. nov., Coprothermobacte. Int. J. Syst. Evol. Microbiol. https://doi.org/10.1099/ijsem.0.002720

Pelletier, E., Kreimeyer, A., Bocs, S., Rouy, Z., Gyapay, G., Chouari, R., Rivière, D., Ganesan, A., Daegelen, P., Sghir, A., Cohen, G.N., Médigue, C., Weissenbach, J., Le Paslier, D., 2008a. “Candidatus Cloacamonas acidaminovorans”: Genome sequence reconstruction provides a first glimpse of a new bacterial division. J. Bacteriol. 190, 2572–2579. https://doi.org/10.1128/JB.01248-07

Pelletier, E., Kreimeyer, A., Chouari, R., Rivie, D., Ganesan, A., Daegelen, P., Sghir, A., Cohen, G.N., Me, C., Weissenbach, J., Paslier, D. Le, 2008b. “ Candidatus Cloacamonas Acidaminovorans ”: Genome Sequence Reconstruction Provides a First Glimpse of a New Bacterial Division □ † 190, 2572–2579. https://doi.org/10.1128/JB.01248-07

Pollard, M.O., Gurdasani, D., Mentzer, A.J., Porter, T., Sandhu, M.S., 2018. Long reads: their purpose and place. Hum. Mol. Genet. https://doi.org/10.1093/hmg/ddy177

Poulsen, L.K., Ballard, G., Stahl, D.A., 1993. Use of rRNA fluorescence in situ hybridization for measuring the activity of single cells in young and established biofilms. Appl. Environ.

Microbiol.

Pritchett, M.A., Metcalf, W.W., 2005. Genetic, physiological and biochemical characterization of multiple methanol methyltransferase isozymes in Methanosarcina acetivorans C2A. Mol. Microbiol. 56, 1183–1194. https://doi.org/10.1111/j.1365-2958.2005.04616.x

Pruesse, E., Peplies, J., Glöckner, F.O., 2012. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28, 1823–1829. https://doi.org/10.1093/bioinformatics/bts252

Qin, Q.-S., Feng, D.-S., Liu, P.-F., He, Q., Li, X., Liu, A.-M., Zhang, H., Hu, G.-Q., Cheng, L., 2017. Metagenomic Characterization of *Candidatus* Smithella cisternae Strain M82_1, a Syntrophic Alkane-Degrading Bacteria, Enriched from the Shengli Oil Field. Microbes Environ. Environ. https://doi.org/10.1264/jsme2.me17022

Qiu, Y.L., Hanada, S., Ohashi, A., Harada, H., Kamagata, Y., Sekiguchi, Y., 2008. Syntrophorhabdus aromaticivorans gen. nov., sp. nov., the first cultured anaerobe capable of degrading phenol to acetate in obligate syntrophic associations with a hydrogenotrophic methanogen. Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.02378-07

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 41, 590–596. https://doi.org/10.1093/nar/gks1219

Rhoads, A., Au, K.F., 2015. PacBio Sequencing and Its Applications. Genomics, Proteomics Bioinforma. 13, 278–289. https://doi.org/10.1016/j.gpb.2015.08.002

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. https://doi.org/10.1038/ismej.2009.2

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584. https://doi.org/10.7717/peerj.2584

Roosaare, M., Vaher, M., Kaplinski, L., Möls, M., Andreson, R., Lepamets, M., Kõressaar, T., Naaber, P., Kõljalg, S., Remm, M., 2017. StrainSeeker: fast identification of bacterial strains from raw sequencing reads using user-provided guide trees. PeerJ. https://doi.org/10.7717/peerj.3353

Rotaru, A.-E., Shrestha, P.M., Liu, F., Shrestha, M., Shrestha, D., Embree, M., Zengler, K., Wardman, C., Nevin, K.P., Lovley, D.R., 2014. A new model for electron flow during anaerobic digestion: direct interspecies electron transfer to Methanosaeta for the reduction of carbon dioxide to methane. Energy Environ. Sci. 7, 408–415. https://doi.org/10.1039/C3EE42189A

Rotaru, A.E., Shrestha, P.M., Liu, F., Markovaite, B., Chen, S., Nevin, K.P., Lovley, D.R., 2014. Direct interspecies electron transfer between Geobacter metallireducens and Methanosarcina barkeri. Appl. Environ. Microbiol. 80, 4599–4605. https://doi.org/10.1128/AEM.00895-14

Sakai, S., Ehara, M., Tseng, I.-C., Yamaguchi, T., Brauer, S.L., Cadillo-Quiroz, H., Imachi, H., Zinder, S.H., 2011. Methanolinea mesophila sp. nov., a hydrogenotrophic methanogen isolated from rice field soil, and proposal of the archaeal family Methanoregulaceae fam. nov. within the order Methanomicrobiales. Int. J. Syst. Evol. Microbiol. 62, 1389–1395. https://doi.org/10.1099/ijs.0.035048-0

Sauer, K., Harms, U., Thauer, R.K., 1997. Methanol:Coenzyme M methyltransferase from methanosarcina barkeri purification, properties and encoding genes of the corrinoid protein MT1. Eur. J. Biochem. https://doi.org/10.1111/j.1432-1033.1997.t01-1-00670.x

Sawatdeenarunat, C., Sung, S., Khanal, S.K., 2017. Enhanced volatile fatty acids production during anaerobic digestion of lignocellulosic biomass via micro-oxygenation. Bioresour. Technol. https://doi.org/10.1016/j.biortech.2017.02.029

Schatz, M.C., Delcher, A.L., Salzberg, S.L., 2010. Assembly of large genomes using second-generation sequencing. Genome Res. https://doi.org/10.1101/gr.101360.109

Schloss, P.D., Jenior, M.L., Koumpouras, C.C., Westcott, S.L., Highlander, S.K., 2016. Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. PeerJ 4, e1869.

https://doi.org/10.7717/peerj.1869

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. 75, 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schupp, B., Paulik, C., Nagy, K., Rittmann, S.K.-M.R., Zwirtmayr, S., Reischl, B., Mauerhofer, L.-M., Pappenreiter, P., Seifert, A.H., Bernacchi, S., Schmider, T., Schuster, B., 2018. Physiology and methane productivity of Methanobacterium thermaggregans. Appl. Microbiol. Biotechnol. 102, 7643–7656. https://doi.org/10.1007/s00253-018-9183-2

Segata, N., Börnigen, D., Morgan, X.C., Huttenhower, C., 2013. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. Nat. Commun. https://doi.org/10.1038/ncomms3304

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., Huttenhower, C., 2011. Metagenomic biomarker discovery and explanation. Genome Biol. https://doi.org/10.1186/gb-2011-12-6-r60

Segota, I., Long, T., 2019. A high-resolution pipeline for 16S-sequencing identifies bacterial strains in human microbiome. bioRxiv. https://doi.org/10.1101/565572

Sekiguchi, Y., Kamagata, Y., Nakamura, K., Ohashi, A., Harada, H., 2000. Syntrophothermus lipocalidus gen. nov., sp. nov., a novel thermophilic, syntrophic, fatty-acid-oxidizing anaerobe which utilizes isobutyrate. Int. J. Syst. Evol. Microbiol. 50, 771–779. https://doi.org/10.1099/00207713-50-2-771

Sekiguchi, Y., Ohashi, A., Harada, H., Imachi, H., Kamagata, Y., Sakai, S., Hanada, S., 2008. Methanolinea tarda gen. nov., sp. nov., a methane-producing archaeon isolated from a methanogenic digester sludge. Int. J. Syst. Evol. Microbiol. 58, 294–301.

https://doi.org/10.1099/ijs.0.65394-0

Sekiguchi, Y., Ohashi, A., Parks, D.H., Yamauchi, T., Tyson, G.W., Hugenholtz, P., 2015. First genomic insights into members of a candidate bacterial phylum responsible for wastewater bulking. PeerJ 3, e740. https://doi.org/10.7717/peerj.740

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res. https://doi.org/10.1101/gr.1239303

Shimizu, S., Upadhye, R., Ishijima, Y., Naganuma, T., 2011. Methanosarcina horonobensis sp. nov., a methanogenic archaeon isolated from a deep subsurface miocene formation. Int. J. Syst. Evol. Microbiol. 61, 2503–2507. https://doi.org/10.1099/ijs.0.028548-0

Shin, J., Lee, S., Go, M., Lee, S.Y., Kim, S.C., Lee, C., Cho, B., 2016. Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing. Nat. Publ. Gr. 1–10. https://doi.org/10.1038/srep29681

Sieber, J.R., McInerney, M.J., Gunsalus, R.P., 2010. Genomic Insights into Syntrophy: The Paradigm for Anaerobic Metabolic Cooperation. Annu. Rev. Microbiol. 66, 429–452. https://doi.org/10.1146/annurev-micro-090110-102844

Singer, E., Bushnell, B., Coleman-Derr, D., Bowman, B., Bowers, R.M., Levy, A., Gies, E.A., Cheng, J.F., Copeland, A., Klenk, H.P., Hallam, S.J., Hugenholtz, P., Tringe, S.G., Woyke, T., 2016. High-resolution phylogenetic microbial community profiling. ISME J. 10, 2020–2032. https://doi.org/10.1038/ismej.2015.249

Solli, L., Håvelsrud, O.E., Horn, S.J., Rike, A.G., 2014. A metagenomic study of the microbial communities in four parallel biogas reactors. Biotechnol. Biofuels. https://doi.org/10.1186/s13068-014-0146-2

Speece, R.E., 1996. Competition and and predominance of methanogens, in: Speece, J.M. (Ed.), Anaerobic Biotechnology for Industrial Wastewater Treatment. Archae Press, Nashville,

Tennessee, pp. 54–57.

Stackebrandt, E., 2014. The family Propionibacteriaceae: Genera other than Propionibacterium, in: The Prokaryotes: Actinobacteria. https://doi.org/10.1007/978-3-642-30138-4_194

Stewart, L.C., Jung, J.-H., Kim, Y.-T., Kwon, S.-W., Holden, J.F., Park, C.-S., 2015. Methanocaldococcus bathoardescens sp. nov., a hyperthermophilic methanogen isolated from a volcanically active deep-sea hydrothermal vent. Int. J. Syst. Evol. Microbiol. 65, 1280–1283. https://doi.org/10.1099/ijs.0.000097

Strang, O., Ács, N., Wirth, R., Maróti, G., Bagi, Z., Rákhely, G., Kovács, K.L., 2017. Bioaugmentation of the thermophilic anaerobic biodegradation of cellulose and corn stover. Anaerobe. https://doi.org/10.1016/j.anaerobe.2017.05.014

Summers, Z.M., Fogarty, H.E., Leang, C., Franks, A.E., Malvankar, N.S., Lovley, D.R., 2010. Direct exchange of electrons within aggregates of an evolved syntrophic coculture of anaerobic bacteria. Science (80-.). https://doi.org/10.1126/science.1196526

Sun, L., Toyonaga, M., Ohashi, A., Tourlousse, D.M., Matsuura, N., Meng, X.Y., Tamaki, H., Hanada, S., Cruz, R., Yamaguchi, T., Sekiguchi, Y., 2016. Lentimicrobium saccharophilum gen. nov., sp. nov., a strictly anaerobic bacterium representing a new family in the phylum bacteroidetes, and proposal of lentimicrobiaceae fam. nov. Int. J. Syst. Evol. Microbiol. https://doi.org/10.1099/ijsem.0.001103

Takahashi, A., Takamizawa, K., Kamagata, Y., Mochimaru, H., Mori, C., Tamaki, H., Nakamura, K., 2012. Methanothermobacter tenebrarum sp. nov., a hydrogenotrophic, thermophilic methanogen isolated from gas-associated formation water of a natural gas field. Int. J. Syst. Evol. Microbiol. 63, 715–722. https://doi.org/10.1099/ijs.0.041681-0

Takai, K., Nealson, K.H., Horikoshi, K., 2004. Methanotorris formicicus sp. nov., a novel extremely thermophilic, methane-producing archaeon isolated from a black smoker chimney in the Central Indian Ridge. Int. J. Syst. Evol. Microbiol. 54, 1095–1100. https://doi.org/10.1099/ijs.0.02887-0

Tan, B., de Araujo e Silva, R., Rozycki, T., Nesbo, C., Foght, J., 2014. Draft Genome Sequences of Three Smithella spp. Obtained from a Methanogenic Alkane-Degrading Culture and Oil Field Produced Water. Genome Announc. 2, 1–2. https://doi.org/10.1128/genomea.01085-14

Taubert, M., Stähly, J., Kolb, S., Küsel, K., 2019. Divergent microbial communities in groundwater and overlying soils exhibit functional redundancy for plant-polysaccharide degradation. PLoS One 14, e0212937. https://doi.org/10.1371/journal.pone.0212937

Thauer, R.K., 1998. Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. Microbiology. https://doi.org/10.1099/00221287-144-9-2377

Thauer, R.K., Kaster, A.K., Seedorf, H., Buckel, W., Hedderich, R., 2008. Methanogenic archaea: Ecologically relevant differences in energy conservation. Nat. Rev. Microbiol. 6, 579–591. https://doi.org/10.1038/nrmicro1931

Timmerman, L., 2015. DNA Sequencing Market Will Exceed $20 Billion, Says Illumina CEO Jay Flatley. Forbes.

Travers, K.J., Chin, C.S., Rank, D.R., Eid, J.S., Turner, S.W., 2010. A flexible and efficient template format for circular consensus sequencing and SNP detection. Nucleic Acids Res. https://doi.org/10.1093/nar/gkq543

Utami, Y.D., Kuwahara, H., Igai, K., Murakami, T., Sugaya, K., Morikawa, T., Nagura, Y., Yuki, M., Deevong, P., Inoue, T., Kihara, K., Lo, N., Yamada, A., Ohkuma, M., Hongoh, Y., 2019. Genome analyses of uncultured TG2/ZB3 bacteria in 'Margulisbacteria' specifically attached to ectosymbiotic spirochetes of protists in the termite gut. ISME J. 13, 455–467. https://doi.org/10.1038/s41396-018-0297-4

van Lier, J.B., Mahmoud, N., Zeeman, G., 2008. Biological wastewater treatment: principles, moddeling and design - Chapter 16: Anaerobic wastewater treatment, Biological Wastewater Treatment : Principles, Modelling and Design. https://doi.org/10.1021/es00154a002

Vanwonterghem, I., Evans, P.N., Parks, D.H., Jensen, P.D., Woodcroft, B.J., Hugenholtz, P., Tyson,

G.W., 2016. Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. Nat. Microbiol. 1, 1–9. https://doi.org/10.1038/nmicrobiol.2016.170

Vasco-Correa, J., Khanal, S., Manandhar, A., Shah, A., 2018. Anaerobic digestion for bioenergy production: Global status, environmental and techno-economic implications, and government policies. Bioresour. Technol. 247, 1015–1026. https://doi.org/10.1016/j.biortech.2017.09.004

Vorholt, J.A., Thauer, R.K., 1997. The active species of "CO2" utilized by formylmethanofuran dehydrogenase from methanogenic Archaea. Eur. J. Biochem. https://doi.org/10.1111/j.1432-1033.1997.00919.x

Wagner, J., Coupland, P., Browne, H.P., Lawley, T.D., Francis, S.C., Parkhill, J., 2016. Evaluation of PacBio sequencing for full-length bacterial 16S rRNA gene classification. BMC Microbiol. 16, 1–17. https://doi.org/10.1186/s12866-016-0891-4

Wang, J., Ren, J., Li, M., Wu, F.X., 2012. Identification of hierarchical and overlapping functional modules in PPI networks. IEEE Trans. Nanobioscience. https://doi.org/10.1109/TNB.2012.2210907

Wawrik, B., Marks, C.R., Davidova, I.A., McInerney, M.J., Pruitt, S., Duncan, K.E., Suflita, J.M., Callaghan, A. V., 2016. Methanogenic paraffin degradation proceeds via alkane addition to fumarate by "Smithella" spp. mediated by a syntrophic coupling with hydrogenotrophic methanogens. Environ. Microbiol. https://doi.org/10.1111/1462-2920.13374

Westerholm, M., Roos, S., Schnürer, A., 2010. Syntrophaceticus schinkiigen. nov., sp. nov., an anaerobic, syntrophic acetate-oxidizing bacterium isolated from a mesophilic anaerobic filter. FEMS Microbiol. Lett. 309, 100–104. https://doi.org/10.1111/j.1574-6968.2010.02023.x

Woese, C.R., Fox, G.E., 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. U. S. A.

Woese, C.R., Kandler, O., Wheelis, M.L., 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. U. S. A.

Wood, C., 2016. Illumina Is the Bedrock of the Genomic Revolution. ARK Invest.

Wu, Y.W., Tang, Y.H., Tringe, S.G., Simmons, B.A., Singer, S.W., 2014. MaxBin: An automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. Microbiome. https://doi.org/10.1186/2049-2618-2-26

Xia, Y., Wang, Y., Fang, H.H.P., Jin, T., Zhong, H., Zhang, T., 2014. Thermophilic microbial cellulose decomposition and methanogenesis pathways recharacterized by metatranscriptomic and metagenomic analysis. Sci. Rep. https://doi.org/10.1038/srep06708

Xia, Y., Yang, C., Zhang, T., 2018. Microbial effects of part-stream low-frequency ultrasonic pretreatment on sludge anaerobic digestion as revealed by high-throughput sequencing-based metagenomics and metatranscriptomics. Biotechnol. Biofuels. https://doi.org/10.1186/s13068-018-1042-y

XUN, L., BOONE, D.R., MAH, R.A., 1989. Deoxyribonucleic Acid Hybridization Study of Methanogenium and Methanocorpusculum Species, Emendation of the Genus Methanocorpusculum, and Transfer of Methanogenium aggregans to the Genus Methanocorpusculum as Methanocorpusculum aggregans comb. nov. Int. J. Syst. Bacteriol. 39, 109–111. https://doi.org/10.1099/00207713-39-2-109

Yamamoto, K., Woyke, T., Cadillo-Quiroz, H., Imachi, H., Tamaki, H., Liu, W.-T., Goodwin, L., Kyrpides, N., Zinder, S.H., Kamagata, Y., 2014. Complete Genome Sequence of Methanolinea tarda NOBI-1T, a Hydrogenotrophic Methanogen Isolated from Methanogenic Digester Sludge. Genome Announc. 2, 1–2. https://doi.org/10.1128/genomea.00876-14

Yang, J., Ma, L., Jiang, H., Wu, G., Dong, H., 2016. Salinity shapes microbial diversity and community structure in surface sediments of the Qinghai-Tibetan Lakes. Sci. Rep. https://doi.org/10.1038/srep25078

Yang, P., Tan, G.-Y.A., Aslam, M., Kim, J., Lee, P.-H., 2019. Metatranscriptomic evidence for classical and RuBisCO-mediated CO2 reduction to methane facilitated by direct interspecies

electron transfer in a methanogenic system. Sci. Rep. 9, 4116. https://doi.org/10.1038/s41598-019-40830-0

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., Euzéby, J., Amann, R., Rosselló-Móra, R., 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12, 635–645. https://doi.org/10.1038/nrmicro3330

You, J., Wu, G., Ren, F., Chang, Q., Yu, B., Xue, Y., Mu, B., 2016. Microbial community dynamics in Baolige oilfield during MEOR treatment, revealed by Illumina MiSeq sequencing. Appl. Microbiol. Biotechnol. https://doi.org/10.1007/s00253-015-7073-4

Youssef, N., Sheik, C.S., Krumholz, L.R., Najar, F.Z., Roe, B.A., Elshahed, M.S., 2009. Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.00592-09

Yu, R.Q., Reinfelder, J.R., Hines, M.E., Barkay, T., 2013. Mercury methylation by the methanogen Methanospirillum hungatei. Appl. Environ. Microbiol. 79, 6325–6330. https://doi.org/10.1128/AEM.01556-13

Yutin, N., Galperin, M.Y., 2013. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. Environ. Microbiol. https://doi.org/10.1111/1462-2920.12173

Zeeman, G., Sanders, W.T.M., Wang, K.Y., Lettinga, G., 1997. Anaerobic treatment of complex wastewater and waste activated sludge - Application of an upflow anaerobic solid removal (UASR) reactor for the removal and pre-hydrolysis of suspended COD, in: Water Science and Technology. https://doi.org/10.1016/S0273-1223(97)00210-2

Zhang, W., Werner, J.J., Agler, M.T., Angenent, L.T., 2014. Substrate type drives variation in reactor microbiomes of anaerobic digesters. Bioresour. Technol.

https://doi.org/10.1016/j.biortech.2013.10.004

Zhang, Z., Schwartz, S., Wagner, L., Miller, W., 2000. A Greedy Algorithm for Aligning DNA Sequences. J. Comput. Biol. 7, 203–214. https://doi.org/10.1089/10665270050081478

Zhou, Z., Pan, J., Wang, F., Gu, J.D., Li, M., 2018. Bathyarchaeota: Globally distributed metabolic generalists in anoxic environments. FEMS Microbiol. Rev. 42, 639–655. https://doi.org/10.1093/femsre/fuy023

Ziels, R.M., Svensson, B.H., Sundberg, C., Larsson, M., Karlsson, A., Yekta, S.S., 2018. Microbial rRNA gene expression and co-occurrence profiles associate with biokinetics and elemental composition in full-scale anaerobic digesters. Microb. Biotechnol. https://doi.org/10.1111/1751-7915.13264