The study of viruses of *Methanosaeta*, a key acetoclastic methanogen for digestion process stability

I-Chieh Chien

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Heidi L. Gough

J. Scott Meschke

David A. C. Beck

Program Authorized to Offer Degree:

Department of Civil and Environmental Engineering

University of Washington

Abstract

The study of viruses of *Methanosaeta*, a key acetoclastic methanogen for digestion process stability

I-Chieh Chien

Chair of the Supervisory Committee:

Professor Heidi L. Gough and Professor J. Scott Meschke

Department of Civil and Environmental Engineering

Stable digestion processes is important for wastewater treatment plants to treat organic solid waste produced from primary and secondary sludge. The stability of the digestion process has greatly improved but process failure still occurs. Toxicants and inadequate operation have been demonstrated to cause failure in some cases. However, in some other cases, there is no obvious explanation for observed upsets. Digestion failure is usually indicated by a decreased rate of methane production, accumulation of acetic acid, and decreased pH, which suggest that the primary consumer of acetate in most mesophilic anaerobic digesters, *Methanosaeta*, has been lost. In natural ecosystems, viruses can

target the most rapid growing populations ("killing the winner" theory) and influence the composition of microbial communities. The most dominant methanogenic Archaea, *Methanosaeta* in mesophilic digesters, appear to be a favorable target for virus attack. This virus-host pair may provide a mechanism for many observed digestion process upsets.

No virus of *Methanosaeta* has been reported previously and the cultivation of *Methanosaeta* is challenging due to the requirement of strict anaerobic conditions and their slow growth rate. The growth of *Methanosaeta concilii* was improved using modified DSM 334 solid media. *Methanosaeta concilii* DSM 6752 colonies appeared on the surface of media (0.5% of gellan) after one month of incubation in serum bottles at 35^{o}C using spread plating. By pour plating (0.5% of gellan), growth of *Methanosaeta* was found after about one week incubation at 35^{o}C. In addition, ammonia toxicity to *Methanosaeta concilii* DSM 6752 in liquid media was also investigated. No inhibition of the methane production when the concentration of ammonia-nitrogen was below 16.8 mg/L. The suppression of methane yield (day 4.8) was 15.6%, 39.6%, and 51.3% at free ammonia-nitrogen of 28.7 mg/L, 47.9 mg/L, and 71.2 mg/L, respectively.

Viruses in two lab-scale reactors (the daily-fed and hourly-fed reactors) dominated by *Methanosaeta* were investigated because *Methanosaeta* viruses were possibly present. All of the observed VLPs were of head-tailed morphology but virus populations between the two reactors were distinct. VBR values

were low in both the daily-fed reactor (0.123) and the hourly-fed reactor (0.093) suggesting that lysogenic or pseudolysogenic life cycles could be the favorable strategy for viruses targeting slow-growing hosts in the methanogenic environment. The presence of many VLPs with 60 nm capsid and 120 nm tail or viruses having the same genome sizes (30 kbp) in both reactors may suggest they target a common host such as *Methanosaeta*, *Proteobacteria* or *Firmicutes*.

The connection between viruses and *Methanosaeta* was studied using a genomic approach based on clustered regularly interspaced short palindromic repeats (CRISPRs). *Methanosaeta* spacers matched viral contigs were found in the two reactors (daily-fed and hourly-fed reactors) and two digesters (West Point and South Plant), suggesting that the *Methanosaeta* viruses might still active in these systems. Presence of various phage proteins in contigs targeted by *Methanosaeta* spacers suggests these sequences are originated from *Methanosaeta* viruses. The connection between viruses and *Methanosaeta* is further confirmed by finding conserved PAMs on viral contigs targeted by both CRISPR/Cas systems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BLAST	basic local alignment search tool	SDS	sodium dodecyl sulfate
Cas	CRISPR-associated	SP	South Plant
CRISPR	clustered regularly interspaced short palindromic repeat	SRSR	short regularly spaced repeat
crRNA	CRISPR RNA	TAN	total ammonia nitrogen
CSTR	continuous stirred tank reactor	TBE	Tris/Borate/EDTA
EDTA	ethylenediaminetetraacetic acid	TCD	thermal conductivity detector
EFM	epifluorescence microscopy	TE	Tris-EDTA
FID	flame ionization detector	TEM	transmission electron microscope
GC	gas chromatograph	UASB	upflow anaerobic sludge blanket
HRT	hydraulic retention time	VBR	virus-to-bacteria ratio
MCE	mixed cellulose ester	VFA	volatile fatty acid
NCBI	national center for biotechnology information	VLPs	virus-like particles
NTA	nitrilotriacetic acid	VSF	virus size-fraction filtrate
OTU	operational taxonomic unit	WP	West Point
PA		WWTP	wastewater treatment plants
PAM	protospacer adjacent motif		
PES	polyethersulfone		
PFGE	pulsed field gel electrophoresis		
qPCR	quantitative polymerase chain reaction		
rDNA	ribosomal DNA		

ACKNOWLEDGEMENTS

I would like to thank my advisor Professor John Ferguson for his guidance and encouragement. I would also like to express my gratitude to my committee members, Professor Heidi Gough, Professor Scott Meschke, Professor David Beck, Professor David Stahl and Professor John Leigh for their useful recommendations on this research.

I acknowledge funding support from Dr. and Mrs. Paul B. Liao Endowed Regental Fellowship, U.S. National Science Foundation, King County Fellowship Program and PDRA2.

Finally, I would like to express appreciation to my parents, my wife Ching-Yu and my son Tzu-Hsun for their support.

Chapter 1 Introduction and Objectives

Anaerobic digestion processes have been widely used to treat organic waste produced from municipal, industrial, and agricultural sources. The processes have many advantages, including reducing waste volume and the cost for waste disposal, stabilizing organic matter, destroying pathogens, controlling order problems, and generating a renewable energy, methane. Recently, anaerobic digesters at municipal waste water treatment plants are getting more attention due to their potential to co-digest organic compounds from other waster streams and to increase methane yield. Many efforts have been made over the last several decades to improve digestion stability, but process failure still occurs. In some cases, actors such as a wide variety of toxicants, organic overload, lack of micronutrients, and poor mixing have been demonstrated to cause the upset and failure of digestion, while in other cases the observed failure has no obvious explanation.

Two poplulations of methanogenic Archaea are responsible for methane production in the digesters. One group is hydrogenotrophic methanogens and the other group is acetoclastic methanogens. Hydrogenotrophic methanogens produce methane from hydrogen and carbon dioxide, and acetoclastic methanogens split acetate into methane and carbon dioxide. In

mesophilic digesters, acetoclastic methanogens are thought to mediate two-thirds of the energy flow and methane production (Jeris and McCarty 1965; Smith and Mah 1966). There are only two known genera of acetoclastic methanogens, *Methanosaeta* and *Methanosarcina*. Having a higher growth rate (*k*) and lower acetate affinity (higher *K*s), *Methanosarcina* usually dominate in environments with high acetate concentrations. In contrast, with lower growth rate and higher acetate affinity, *Methanosaeta* generally dominate when acetate concentrations are low. In anaerobic digesters, acetate concentrations are generally below 15 mM and thus *Methanosaeta* usually predominate in this environment (Raskin, Zheng et al. 1995; Angenent, Zheng et al. 2002; McHugh, Carton et al. 2003; Leclerc, Delgenes et al. 2004; Yu, Lee et al. 2005).

Digestion failure is usually characterized by a decreased rate of methane production, accumulation of acetic and short chain length organic acids, and decreased pH (Kroeker, Schulte et al. 1979), which implies the functional microorganisms involved in the methanogenesis, acetoclastic methanogens, were lost.

Viruses are abundant and ubiquitous biological entities in natural environments. By targeting prokaryotic hosts and lyzing microbial cells, viruses influence composition of microbial communities and controlling nutrient cycles in the oceans (Suttle 2005; Suttle 2007). In addition,

phages are hypothesized to target the most rapidly growing populations ("killing the winner" theory) and thus protect other competitors from extinction in the same environment. Therefore, viruses are also believed to stimulate microbial diversity (Thingstad and Lignell 1997; Thingstad 2000). In mesophilic digesters, the most dominant methanogenic Archaea is *Methanosaeta*, thus, they appear to be a favorable target for virus attack. This virus-host pair may provide a mechanism for many observed digestion process upsets. Therefore, the hypothesis of this study is: *Methanosaeta* viruses are present in *Methanosaeta*-dominated environments such as mesophilic full-scale anaerobic digesters and acetate-fed enrichment reactors.

The first step to assess the hypothesis is to identify *Methanosaeta* viruses. No virus of *Methanosaeta* has been reported previously and the cultivation of *Methanosaeta* is challenging due to the requirement of strict anaerobic conditions and their slow growth rate. Therefore, the specific objectives of this research were: (1) to improve the culture techniques for *Methanosaeta* in solid media (i.e. growing *Methanosaeta* both as colonies and as lawn) and liquid media (i.e. testing potential ammonia toxicity to the activity of *Methanosaeta*), (2) to evaluate two viral screening and isolation approaches (plaque assays and infection tests) for *Methanosaeta*, (3) to investigate the persistent virus-like particles in two *Methanosaeta* dominant reactors, and (4) to investigate the connection between viruses and *Methanosaeta* in reactors using a genomics

approach based on clustered regularly interspaced short palindromic repeats (CRISPRs).

Objective 1: Improve growth of *Methanosaeta* in solid media and evaluate plaque assay approaches for *Methanosaeta*

Before year 2010 (Carbonero, Oakley et al.), the cultivation of *Methanosaeta* colonies on solid media was unsuccessful. The growth of *Methanosaeta* using solid media was only reported in a deep-agar media and the colonies formed as cotton-like flocs after 6 months of cultivation. A modified media reported by Carbonero (2010) was able to grow *Methanosaeta* colonies (< 1 mm in diameter) after incubation of one month. To improve the culture and isolation techniques, approaches to culture *Methanosaeta* using solid media including culture vessels, plating techniques and gellan gum concentrations were evaluated. In addition, a viral challenging method, plaque assay, for *Methanosaeta* was assessed in chapter 4.

Objective 2: Study Ammonia toxicity and assess infection test approaches for *Methanosaeta*

High concentration of ammonia is known to suppress methanogenesis of anaerobic

digestion processes (McCarty and McKinney 1961; Chen, Cheng et al. 2008). Although the inhibition of a pure strain of *Methanosaeta* by ammonia was identified in previous studies, the suppression of methanogenesis at different level of free ammonia is unknown (Sprott and Patel 1986; Steinhaus, Garcia et al. 2007). Thus, the methane production of *Methanosaeta* in media with different concentrations of ammonia was investigated. The result from this investigation was used to improve the culture condition for *Methanosaeta* in liquid media. In addition, a viral challenge and isolation approach in liquid media, infection tests, was evaluated. These results pertaining to the objective 2 are presented in chapter 5.

Objective 3: Investigate virus-like particles in two *Methanosaeta*-dominated reactors

Two lab-scale reactors, the daily-fed and hourly-fed reactors, both dominated by *Methanosaeta* were persistently producing virus-like particles. It was likely that *Methanosaeta* viruses were present in these reactors. Thus morphology, size distribution, genome length, and the concentration of observed virus-like particles were investigated and are shown in chapter 6.

Objective 4: Investigate the connection between viruses and *Methanosaeta* using a genomics

approach based on clustered regularly interspaced short palindromic repeats (CRISPRs)

Clustered regularly interspaced short palindromic repeats (CRISPRs) have been demonstrated to be an adaptive immune system against mobile genetic elements such as viruses (Barrangou, Fremaux et al. 2007). During a virus challenge, a fragment of invader DNA named a spacer can be excised and inserted into the host's CRISPR locus. The virus attempting to infect the host can be identified by searching for homologies between the spacers and sequenced viral nucleotide database. CRISPR loci were found in *Methanosaeta concilii* GP6. In chapter 7, the CRISPRs and Cas (CRISPR-associated) genes of *Methanosaeta concilii* DSM 6752 and DSM 2139 were comparatively analyzed. In addition, the virus-host interactions in the two *Methanosaeta*-dominated reactors were studied using CRISPR-based approach.

Chapter 2 Background

2.1 Anaerobic Digestion

Anaerobic digestion is a reliable technology used worldwide to treat organic waste generated from municipal, industrial and agricultural sources and has the advantages of reducing waste volume, destroying pathogens and eliminating odor problems (McCarty 2001). Because it is an anaerobic process, no electron acceptor (oxygen) needs to be present which could reduce energy requirements for aeration. Moreover, the process converts organic waste into methane gas, a renewable energy source (Appels, Lauwers et al. 2011). Recently, this process is drawing more attention because of its potential to treat different waste streams and to increase methane yield, for example, co-digestion of organic fraction of municipal solid wastes with sewage sludge in anaerobic digesters (Mata-Alvarez, Mace et al. 2000).

An empirical design approach has been applied to establish anaerobic digestion in wastewater treatment process. Most of the time, digestion is stable and reliable. Over several decades, considerable efforts have been made to improve stability of anaerobic digestion process (McCarty 2001; Appels, Baeyens et al. 2008). For example, a wide variety of toxicants have been studied

that are suspected to cause upsets of digestion (Chen, Cheng et al. 2008). Organic overload (Nachaiyasit and Stuckey 1997), lack of micronutrients and poor mixing (Kim, Ahn et al. 2002) have also been demonstrated to cause digestion failure. Also, digestion stability has been found to be influenced by using different feed-loading frequency (Bombardiere, Espinosa-Solares et al. 2007). Nevertheless, upsets and failures of the process still occur with no obvious explanation.

2.2 Microbial Community in Digesters

The reliability of anaerobic digesters in wastewater treatment process greatly relies on better understanding of the resident microbial community. Recently, effort has been made to understand the microbial community structure in anaerobic digestion processes (Ariesyady, Ito et al. 2007; Narihiro and Sekiguchi 2007). Analysis of microbial communities in anaerobic digesters using molecular biological approaches has detected many uncultured microorganisms implying that knowledge of microbial ecology remains vastly unexplored.

The conversion of complex organic matter to methane and carbon dioxide in anaerobic digester is mediated by a variety of different populations from two entirely different biological domains, the Bacteria and the Archaea (Figure 2.1). Initially, large organic compounds (polymers) are

hydrolyzed and degraded into smaller intermediates (e.g. sugars, amino acids, fatty acid, alcohols etc.) by primary fermenting bacteria. These intermediates are then converted to acetate, CO_2 and H_2 by secondary fermenters. Acetate, CO_2 and H_2 are subsequently utilized by the acetoclastic methanogens or hydrogenotrophic methanogens to generate methane and carbon dioxide (Schink 1997). In this complex food web, the fermenting bacterial community is composed of extremely diverse groups, whereas the syntrophic bacterial and archaeal community is less phylogenetically diverse (Ariesyady, Ito et al. 2007; Hatamoto, Imachi et al. 2007; Aller and Kemp 2008; Zhang, Banaszak et al. 2009). Moreover, when the digestion process is functionally stable, the bacterial community composition fluctuated was reported, whereas the archaeal community remains relatively stable (Zumstein, Moletta et al. 2000). This implies that there are diverse Bacteria capable of conducting similar functions in digesters, but certain functional processes can only be performed by specific Archaea with unique biological pathways. For example, the methanogenesis can only be conducted by methanogenic Archaea.

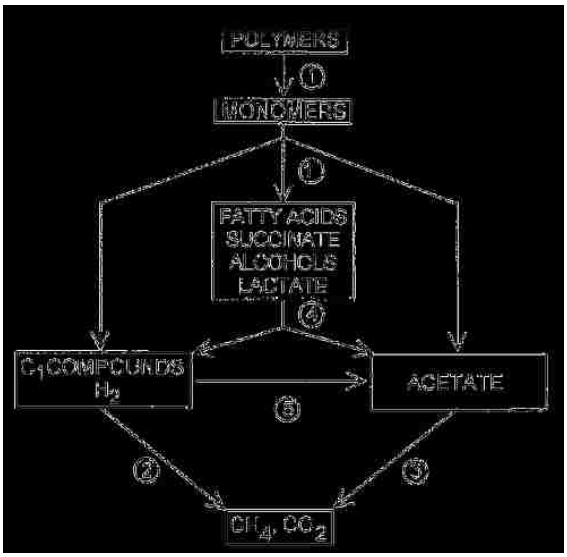


Figure 2.1 Trophic groups of microorganisms involved in anaerobic sludge digestion. Groups of organisms involved: 1) Hydrolytic and primary fermenting bacteria, 2) Hydrogenotrophic methanogens, 3) Acetoclastic methanogens, 4) Secondary fermenting ("syntrophic") bacteria, 5) Homoacetogenic bacteria. (Schink 1997).

2.3 Acetoclastic Methanogens

Acetoclastic methanogens inhabit a variety of natural ecosystems such as paddy field soils (Grosskopf, Janssen et al. 1998), fresh water and marine sediments (Elberson and Sowers 1997; Scholten and Stams 2000) and acidic fens (Steinberg and Regan 2008). They are among the most important methane producers on earth, since two-thirds of the biogenic methane released to atmosphere is derived from methyl group of acetate (Conrad 1999; Smith and Ingram-Smith 2007). Acetoclastic methanogens are also essential in engineered systems such as anaerobic digesters and upflow anaerobic sludge blanket (UASB). *Methanosaeta* usually strongly predominates in

mesophilic sludge treatment digesters (Leclerc, Delgenes et al. 2004) and are thought to mediate two-thirds of the energy flow and methane production (Jeris and McCarty 1965; Smith and Mah 1966). Although syntrophic acetate-oxidizing bacteria, competitors of acetoclastic methanogens, have been shown to convert acetate to hydrogen and carbon dioxide (Ito, Yoshiguchi et al. 2011) when ammonium nitrogen concentrations (NH_4^+-N) are elevated to 5.5 g/L and 6.9 g/L (Westerholm, Dolfing et al. 2011), their dominance in mesophilic digesters was not reported.

Most species of methanogens use hydrogen and carbon dioxide as substrates (hydrogenotrophic methanogens) while only two genera of methanogens, *Methanosaeta* and *Methanosarcina*, are known to utilize acetate (acetoclastic methanogens). Because they use the same substrate, acetate, *Methanosarcina* and *Methanosaeta* compete with each other for substrate by different growth strategies. Having a higher growth rate (*k*) and lower acetate affinity (higher *K*s), *Methanosarcina* usually dominate in environments with high acetate concentrations. In contrast, with lower growth rate and higher acetate affinity, *Methanosaeta* generally dominate when acetate concentrations are low. *Methanosaeta* is ubiquitous and dominant in major biogenic methane-releasing environments and have been postulated to be the predominant methane producer on earth (Smith and Ingram-Smith 2007).

2.4 Ammonia Toxicity

When assessing viral activity based on metabolic activity, it is also important to consider other factors that might cause inhibition. Inhibition of methanogenesis by ammonia has been previously recognized (Debaere, Devocht et al. 1984; Koster and Lettinga 1984; Jarrell, Saulnier et al. 1987; Bhattacharya and Parkin 1989; Hajarnis and Ranade 1993; Lay, Li et al. 1998; Eldem, Ozturk et al. 2004; Sawayama, Tada et al. 2004). Anaerobic processes can benefit from total ammonia nitrogen (TAN, sum of ammonia-nitrogen and ammonium-nitrogen) ranging from 50 and 200 mg/L (McCarty and McKinney 1961; McCarty 1964; Liu and Sung 2002). McCarty and McKinney (1961) and McCarty (1964) also reported that TAN are inhibitory at any concentration between 1500 and 3000 mg/L when pH is above 7.4, and are toxic at any concentration beyond 3000 mg/L at any pH level. TAN ranged from 1700 to 14000 g/L has been reported to suppress 50% of methane production (Chen, Cheng et al. 2008). The variation of the reported inhibitory TAN level may be due to varied experimental conditions such as substrates, inocula, acclimation periods, temperature and pH.

NH_3 (free ammonia/un-ionized ammonia) is implied to be more toxic than NH_4^+ since it is more membrane-permeable (Debaere, Devocht et al. 1984) and can interfere with intracellular ion

exchange (Sprott and Patel 1986). In contrast, some studies showed that ammonium is more toxic than free ammonia to methanogenesis (Koster and Koomen 1988; Lay, Li et al. 1998). In aqueous solution, the concentration of free ammonia is a function of total ammonia concentration (sum of $[NH_3]$ and $[NH_4^+]$), pH and equilibrium constant K_a (Equation 5.1). In addition, K_a is a function of temperature (Equation 5.2).

$$[NH_3] = \frac{[NH_3] + [NH_4^+]}{1 + [H^+]/Ka} \quad (5.1)$$

$$pKa = 0.09018 + 2729.92/T \quad \text{(Emerson, Russo et al. 1975)} \quad (5.2)$$

where Ka = equilibrium constant

T = temperature (K)

When the temperature is at 35°C (Ka = 8.95), the fraction of $[NH_3]$ can be expressed as $1/(1+[H^+]/Ka)$ and is the function of pH (Table 2.1). As shown in Table 5.1, the fraction of free ammonia increases as pH rises. Therefore, more inhibition of methanogenesis may be observed when pH increases and TAN remains constant.

Several studies have considered optimal free ammonia concentrations for methanogenic communities. 55±11 mg/L of free ammonia-nitrogen was reported to be the maximum tolerable

in anaerobic acetate and propionate enrichment cultures at a 40-day solid retention time (Bhattacharya and Parkin 1989). Debaere (1984) suggested that optimal condition of the reactor is achieved when free ammonia-nitrogen was kept below a concentration of 80 to 100 mg/L, and inhibition of methanogenesis was observed when the concentration of NH_3-N was at 150 mg/L.

Table 2.1 Fraction of free ammonia (NH_3) as a function of pH at 35°C

pH	fraction of free ammonia	pH	fraction of free ammonia
6	0.001	7.6	0.043
6.2	0.002	7.8	0.066
6.4	0.003	8	0.101
6.6	0.004	8.2	0.151
6.8	0.007	8.4	0.220
7	0.011	8.6	0.309
7.2	0.017	8.8	0.415
7.4	0.027	9	0.529

Methanosaeta concilii has been reported to be the most sensitive methanogen to ammonia among the tested methanogenic pure cultures. Complete inhibition was found at 2140 mg TAN/L at pH 7.0 (Sprott and Patel 1986). Cell density of *Methanosaeta concilii* GP6 has been monitored in a microreactor with ammonia varying between 250 and 2,500 mg TAN/L at pH 7.6 (Steinhaus, Garcia et al. 2007). Cell density was reduced to 50% when the concentration of TAN reached around 1000 mg/L, and dramatically decreased (~75%) when TAN rose beyond 1900 mg/L.

2.5 Inducing Agents

Viruses can establish a stable relationship with their hosts. This is known as a lysogenic life cycle (Weinbauer 2004). In the lysogenic cycle, viruses (or prophages) remain dormant inside the host cells and their genomes replicate along with the host's. This cycle can be broken by inducing agents that trigger a switch to the lytic life cycle in which virus particles propagate causing the lysis of host cells. A variety of inducing agents have been reported to associated with lysogeny (Ackermann and DuBow 1987). Other inducing agents such as hydrogen peroxide, temperature changes, polyaromatic hydrocarbons, fuel oil, trichloroethylene, and pesticides have been also reported (Wommack and Colwell 2000; Weinbauer 2004).

Archeoviruses have been induced by UV and mitomycin C. Production of *Sulfolobus spindle* virus (SSV1) can be induced by UV treatment or mitomycin C (5-15 μg/mL) (Reiter, Palm et al. 1987; Liu and Huang 2002; Frols, Gordon et al. 2007). *Acidianus* two-tailed virus (ATV) infecting the hyperthermophilic Archaea *Acidianus convivator* could be induced by mitomycin C (0.5 μg/mL), UV, and cold shock. Induction of ATV was observed when temperature shifted from 85^{o}C (optimal temperature) to 75^{o}C under exponential growth. Halophage SNJ1 was induced from *Natrinema* sp. strain F5 by mitomycin C (1 μg/mL) (Mei, Chen et al. 2007). Viruses of

Halobacterium are lytic when environmental salinity decreased (Daniels and Wais 1990). In contrast, the production of virus was not observed when *Methanococcus voltae* A3 was subjected to various stresses including UV, temperature shift from 30^{o}C to 37^{o}C, starvation, growth with mitomycin C, and acridine orange (Wood, Whitman et al. 1989). Thus, while several common microbial stressors have been demonstrated for inducing archeovirus production, these stressors are not universal and must be tested separately for each strain.

2.6 Viruses and Their Connection to Process Stability

Viruses are abundant and ubiquitous biological entities in natural environments. Their abundance exceeds that of Bacteria and Archaea in aquatic environments by about 10-fold (Bergh, Borsheim et al. 1989; Wommack and Colwell 2000; Weinbauer 2004; Suttle 2007). In addition to aquatic environments, they have been reported in many ecosystems, including soils (Ashelford, Day et al. 2003; Williamson, Radosevich et al. 2005), sediments (Filippini and Middelboe 2007; Siem-Jorgensen, Glud et al. 2008) and thermophilic environment (Rice, Stedman et al. 2001). It has been suggested that most observed viruses are bacteriaphages, the viruses of Bacteria and Archaea (Cochlan, Smith et al. 1993; Weinbauer 2004). By targeting prokaryotic hosts and lysing microbial cells, bacteriophages can influence composition of microbial communities and control

nutrient cycles in the oceans (Suttle 2005; Suttle 2007). In addition, phages are thought to target the most rapid growing populations (killing the winner theory) and thus protect other competitors from extinction in the same environment. Therefore, bacteriophages are also believed to stimulate microbial diversity (Thingstad and Lignell 1997; Thingstad 2000).

In contrast to natural ecosystems, knowledge about viral abundance, diversity, and phage-host interactions in engineered biosystems is sparser. In environmental engineering, microbes have been utilized in variety of processes such as activated sludge and anaerobic digestion. Similar to natural ecosystems, abundant phages have been documented (Khan, Satoh et al. 2002; Otawa, Lee et al. 2007; Wu and Liu 2009). Fluctuation of phage concentrations have been reported to correlate with bacterial concentrations in a membrane bioreactor treating industrial wastewater implying a close phage-host interactions (Shapiro, Kushmaro et al. 2010). Wu and Liu (2009) reported that indigenous viruses were abundant and dynamic in the influent, primary settlers, activated sludge, anaerobic sludge and effluent of a municipal wastewater treatment facility. Decreasing of polyphosphate-accumulating bacteria, *Microlunatus phosphovorus*, and increasing of its lytic phages found in a laboratory-scale activated sludge process suggests that phages might be able to influence the performance of phosphate removal (Lee, Otawa et al. 2007). Moreover, phages affiliated with *Siphoviridae*, *Myoviridae*, and *Cystoviridae* were found to propagate in a UASB

methanogenic digester, although their roles and corresponding hosts were not identified (Park, Ikenaga et al. 2007). In summary, phages are abundant in many engineered bioreactors, which indicate that these phages may infect indigenous microbes. More important, phages might have influence on the stability and performance of bioreactors by controlling microbial populations with unique and irreplaceable physiological functions (e.g. polyphosphate accumulating or methanogenic populations).

Digestion failure is usually indicated by a decreased rate of methane production, accumulation of acetic and short chain length organic acids, and decreased pH (Kroeker, Schulte et al. 1979), which implies the functional microorganisms involved in the methanogenesis, acetoclastic methanogens, were lost. *Methanosaeta* generally outnumber *Methanosarcina* and syntrophic acetate-oxidizing bacteria in mesophilic digesters. Therefore, they appear to be a favorable target for phage attack, and this phage-host pair may provide a mechanism for many observed process upsets.

2.7 Viruses of Methanogens

Archaeal viruses have been reported to exhibit exceptionally complex morphotypes, including linear, fusiforms, droplet, and bottle shapes. However, most documented viruses of methanogens

are of head-tail morphology, which is also the most frequently observed morphotype in bacterial domain. Viruses infecting methanoarchaea also carry many genes of bacterial origin suggesting that these viruses likely infect methanogens by interdomain spreading (Prangishvili, Forterre et al. 2006).

Five viruses infecting hydrogenotrophic methanogens were reported (Jordan, Meile et al. 1989; Meile, Jenal et al. 1989; Wood, Whitman et al. 1989; Nolling, Groffen et al. 1993; Pfister, Wasserfallen et al. 1998; Luo, Pfister et al. 2001). Most of the reported viruses belong to family *Caudovirales* and have head-tail morphology (Figure 2.3). A virus discovered by Wood et al. (1989) was an oval-shaped particle (Figure 2.3 a).

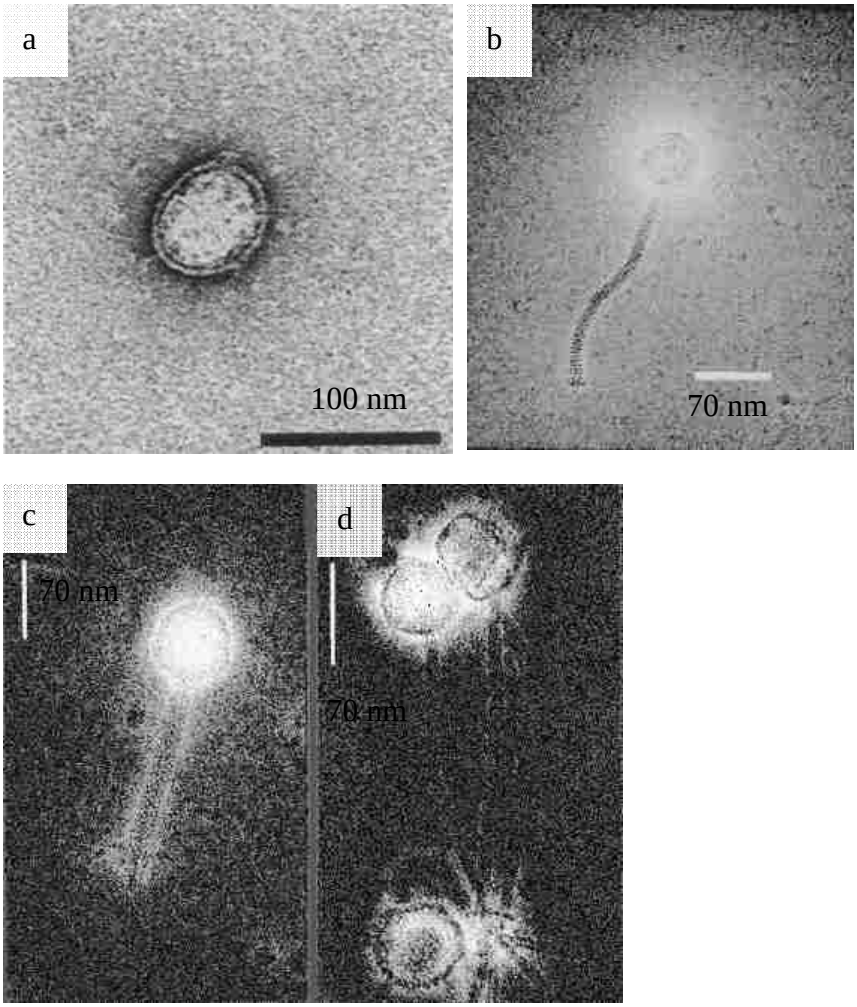


Figure 2.2 Viruses of methanogens

(a) virus-like particle of *Methanococcus voltae* A3 (Wood, Whitman et al. 1989), (b) ψM1: a virus of *Methanobacterium thermoautotrophicum* Marburg (Meile, Jenal et al. 1989), (c) ΦF1: a virus of *Methanobacterium thermoformicicum*:, and (d) ΦF3: a virus of *Methanobacterium thermoautotrophicum* (Nolling, Groffen et al. 1993)

Lytic siphophages ψM1 and ψM2 are the best characterized viruses of methanogens (Figure 2.2 b). ψM2 is a mutant derived from ψM1 (Pfister, Wasserfallen et al. 1998). They having a polyhedral capsid of 55 nm in diameter and a flexible tail of 210 nm in length were found to infect *Methanobacterium thermoautotrophicum* Marburg (Meile, Jenal et al. 1989; Pfister, Wasserfallen

et al. 1998). Nolling et al. (1993) reported that virulent phages ΦF1 and ΦF3 were capable of infecting *Methanobacterium thermoformicicum* (strains 2-245, FTF, FF1, FF3 and CSM3) and *Methanobacterium thermoautotrophicum* (strain AH), respectively. ΦF1 is morphologically similar to *Myoviridae* and has a capsid of 70 nm in diameter and a stiff tail of 160 nm in length (Figure 2.2 c) while ΦF3 is morphologically similar to *Siphoviridae* and has a capsid of 55 nm in diameter and a long flexible tail of 230 nm in length (Figure 2.2 d).

Methanosaeta viruses have not been previously identified probably due to the challenge of culturing their host. *Methanosaeta* is notorious for slow-growing and is resistant to grow on solid media (a key step for the traditional virus isolation method, the plaque assay). Its viruses remain to be found.

2.8 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs)

The recently discovered CRISPR/Cas system (Mojica, Diez-Villasenor et al. 2000; Jansen, van Embden et al. 2002), found to date in most archaeal and 40% of bacterial genomes (Jansen, van Embden et al. 2002; Grissa, Vergnaud et al. 2007b; Makarova, Haft et al. 2011), provide prokaryotes with adaptive immunity against mobile genetic elements (e.g. viruses and conjugative

plasmids) in a sequence-specific manner (Sorek, Kunin et al. 2008; Horvath and Barrangou 2010; Marraffini and Sontheimer 2010). The CRISPR/Cas immune process can be divided into three sequential steps: (1) adaptation, (2) expression, and (3) interference (Makarova, Haft et al. 2011). In the adaptation step, the spacer and a newly generated repeat sequence are integrated into the leader end of the CRISPR locus. In the expression step, long pre-crRNA generated from CRISPR loci are transcribed and further processed into short crRNA monomers, which containing a full spacer and partial adjoining repeat sequences. In the last step, these crRNAs guide the Cas proteins to cleave the complementary invading DNA targets.

CRISPR was first documented by Ishino (1987) in the 3' flanking region of *iap* gene in *Escherichia coli*. This group of repeat elements was subsequently reported in other bacteria and archaea (Mojica, Ferrer et al. 1995; Bult, White et al. 1996; Masepohl, Gorlitz et al. 1996; Sensen, Charlebois et al. 1998). These short, partial palindromic repeat sequences clustered in genomes and were separated by unique non-repetitive sequences. Because these features are distinguishable from other known repeat elements, a novel type of repeat family, Short Regularly Spaced Repeats (SRSRs), was proposed (Mojica, Diez-Villasenor et al. 2000). The denomination was later renamed as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) to address the feature that most repeat sequences were partially palindromic (Jansen, van Embden

et al. 2002). Jansen et al. (2002) analyzed CRISPR *in silico* and reported a common 300-500 bp sequence, the leader sequence, flanking on one side of CRISPR loci. The conservation of leader and repeat sequence were found within one species but not intraspecies.

CRISPR-associated (*cas*) genes were always identified in the genomes having CRISPR sequences, and closely located on either side of CRISPR loci, which suggested a strong association between CRISPR loci and *cas* genes (Jansen, van Embden et al. 2002). Using sequence and structure analysis of Cas proteins, putative function of the CRISPR/Cas system has been proposed, for example, replicon partitioning (Mojica, Ferrer et al. 1995), DNA repair (Makarova, Aravind et al. 2002), and DNA rearrangement (DeBoy, Mongodin et al. 2006). The finding of homologous sequences between spacers and extrachromosomal elements (virus and conjugative plasmid DNA) and the failure for a phage to infect a microbe having spacers matching to this phage raised the hypothesis that CRISPR/Cas might be an immune system in prokaryotes (Bolotin, Ouinquis et al. 2005; Mojica, Diez-Villasenor et al. 2005; Pourcel, Salvignol et al. 2005). The first experimental evidence of CRISPR immunity was demonstrated by Barrangou et al (2007) showing that novel spacer sequences derived from challenged phages were inserted in the CRISPR1 locus of phage-resistant *Streptococcus thermophilus*. In addition, inactivating *cas5* and *cas7* caused the loss of phage resistance and failure to generate the

bacteriophage-insensitive mutant, respectively, indicating that *cas* genes were essential for performing and acquiring the CRISPR immunity (Barrangou, Fremaux et al. 2007). Moreover, a *Staphylococcus epidermidis* CRISPR locus possessing a space matching to *nickase gene* in conjugative plasmid was shown to prevent conjugation and plasmid transformation (Marraffini and Sontheimer 2008).

CRISPR loci consist of repeat and spacer sequences. Generally, the repeat sequence within a locus was highly conserved (Jansen, van Embden et al. 2002). The size of the repeats ranges from 23 to 50 bp (average: 31 bp) (Marraffini and Sontheimer 2010). Based on sequence similarity, CRISPR repeats can be roughly classified into 12 groups (Kunin, Sorek et al. 2007). Repeat sequences in some groups were predicted to form stable RNA secondary structure (stem-loop) but other groups did not have detectable structures (Kunin, Sorek et al. 2007). The size of spacer sequences varies between 17 to 84 bp (average: 36 bp) (Marraffini and Sontheimer 2010). In a CRISPR locus, the average number of repeat-spacer unit was 66 (Marraffini and Sontheimer 2010), but the it was not uncommon for the number to reach 100 in archaea (Garrett, Vestergaard et al. 2011). Appearance of multiple CRISPR loci in a given genome was frequently observed. There are 18 CRISPR loci in the genome of *Methanococcus jannaschii* (Bult, White et al. 1996), which is the highest number of arrays reported so far. CRISPR loci are generally

located on chromosomes but their presence on plasmids was also documented (Lillestol, Shah et al. 2009; Millen, Horvath et al. 2012).

CRISPR loci are generally preceded by low complexity and AT-rich sequences referred to as leaders (Jansen, van Embden et al. 2002; Tang, Bachellerie et al. 2002; Pourcel, Salvignol et al. 2005; Lillestol, Redder et al. 2006). Leader sequences are several hundred base pairs long and are located on one side of the loci (leader end). The position and orientation of the leader sequences are invariable with respect to the orientation of the repeat sequence. The conservation of leader sequence was found within a species but not intraspecies (Jansen, van Embden et al. 2002). Leader sequences are implied to be the region where transcription starts (Lillestol, Redder et al. 2006). Starts of transcription have been experimentally demonstrated in bacteria *Xanthomonas oryzae* (Semenova, Nagornykh et al. 2009), crenarchaeon *Sulfolobus* (Lillestol, Shah et al. 2009), and euryarchaeon *Pyrococcus furiosus* (Hale, Kleppe et al. 2008). The CRISPR promoter might locate within the leader sequences as well (Tang, Bachellerie et al. 2002; Lillestol, Shah et al. 2009). In addition, these sequences might be involved in the insertion of newly acquired spacer sequences in CRISPR loci (Pourcel, Salvignol et al. 2005; Lillestol, Shah et al. 2009; Shah and Garrett 2011), which is one of the three steps of CRISPR/Cas action, the adaptation.

In response to the virus or plasmid challenge, a fragment of invader DNA, protospacer, will be excised and incorporated at or near the first repeat of a CRISPR locus. The selection of protospacers from invader genetic elements remains unclear but has been linked to the protospacer-associated motif (CRISPR motif or PAM) (Mojica, Diez-Villasenor et al. 2005; Marraffini and Sontheimer 2010). The motif, generally several base pairs long, is a conserved region located 2 to 4 bp next to one end of a protospacer in virus or plasmid sequences. For example, there are AGAAW and GGNG motifs for CRISPR1 and CRISPR3, respectively for *Streptococcus thermophilus* (Deveau, Barrangou et al. 2008; Horvath, Romero et al. 2008). Also, there are CC (CRISPR family I), TC (CRISPR family II), and GT (CRISPR family II) motifs in *Sulfolobus* (Lillestol, Shah et al. 2009). In addition, mutation within PAM regions allowed phages to infect *Streptococcus thermophilus* even the CRISPR containing spacer identical to phage genomic sequences, suggesting a critical role of the PAM in CRISPR-mediated resistance. It is also reported that PAMs determine the direction of spacers in CRISPR loci (Mojica, Diez-Villasenor et al. 2005).

The addition of new spacers at the leader end of CRISPR arrays has been reported (Barrangou, Fremaux et al. 2007; Deveau, Barrangou et al. 2008; Horvath, Romero et al. 2008; Erdmann and

Garrett 2012). The insertion typically occurred at the first position (next to leader) while the addition at six of the eight repeats has been describe in locus E of *Sulfolobus solfataricus* P2 (Erdmann and Garrett 2012). Comparing spacer sequences among different *Streptococcus thermophilus* strains, the spacer sequences were more conserved at the trailer end (leader-distal) and were hypervariable at the leader end of the locus (Horvath and Barrangou 2010). This is consistent with the discovery of newer spacers near the leader sequence and older sequences near the trailer end of *Leptospirillum* sp. populations in natural acidophilic biofilms (Andersson and Banfield 2008; Tyson and Banfield 2008). Therefore, spacer sequences not only provide the information of foreign genetic elements but also stored the history of invading events, which provides an alternative tool to study the relationship and evolution between the viruses and their corresponding hosts in several ecosystems (Andersson and Banfield 2008; Snyder, Bateson et al. 2010; Sorokin, Gelfand et al. 2010; Pride, Sun et al. 2011; Stern, Mick et al. 2012).

CRISPR/Cas systems have been recently classified based on multiple criteria, including the phylogenies of most conserved *cas* genes, sequence and organization of CRISPR repeats, and the genome architecture of CRISPR/Cas arrays (Makarova, Haft et al. 2011). Three major types (I, II, and III) along with several subtypes were proposed. The *cas3*, *cas9*, and *cas10* are signature genes of CRISPR/Cas type I, II, and III systems, respectively. Type I and II systems are observed

in both the Bacteria and Archaea while the type II system is exclusive to the Bacteria. For many archaea, finding of multiple and different types of CRISPR/Cas systems located in a single genome was not unusual (Makarova, Haft et al. 2011). Compared to the Bacteria, archaeal CRISPR systems were found to be more complex and diverse (Garrett, Vestergaard et al. 2011).

Chapter 3 Overview of Materials and Methods

3.1 Enrichment Reactors

The study reactors consisted of two semi-continuous stirred tank reactors (CSTR) that were originally established in 2002, as described by Conklin et al (2006). The reactors were inoculated with anaerobic digester sludge from the West Point Treatment Plant in Seattle, WA, fed aseptic Reduced Anaerobic Mineral Media (RAMM) supplemented with vitamins and acetate (234mM) as the sole carbon and energy source, and operated at 30-34°C. One reactor was fed in a single daily dose (daily-fed reactor) and the other in hourly increments (hourly-fed reactor). The daily-fed reactor was re-inoculated in 2007 and 2008 due to community population shift from *Methanosarcina* to *Methanosaeta*. But reestablishment of *Methanosarcina* population did not succeed. Table 3.1 summarizes the operating conditions for the reactors during the current study, which was conducted between 2007 and 2013. During the study period, the reactors were fed 150 ml/day sterile feed and the same volume was manually wasted daily from each reactor. Due to a lower operating reactor volume in the daily-fed reactor during the study period, this reactor had a lower hydraulic retention time (HRT) (13.3 days) than the hourly-fed reactor (20 days).

Gas production by the reactors was continually recorded by wet test meters (Precision Scientific, Chicago, IL). The concentrations of methane and carbon dioxide in headspace gas were analyzed by Carle Series 100A gas chromatograph equipped with a thermal conductivity detector (GC-TCD, Chandler Engineering, Tulsa, OK). Helium was used as the carrier gas (20 mL/min). A 6 foot long packed Haysep® Q80/100 column (Supelco; Bellefonte, PA) was used to separate nitrogen, methane and carbon dioxide. The injector, oven and detector temperatures were set at 110ºC. The concentration of acetate in the reactor effluent was analyzed using a Shimadzu GC-2010 equipped with flame ionization detector (GC-FID, Shimadzu Scientific Instruments, Columbia, MD). and a DB-FFAP column (122-3232, Agilent Technologies, Santa Clara, CA). One μL of sample was injected by an autosampler system (AOC-20i, Shimadzu Scientific Instruments, Columbia, MD). Helium was used as carrier gas. The injector and detector temperatures were set at 260ºC. The oven began at 80°C for one minute, then ramped to 110°C at a rate of 3°C per minute. The temperature was then ramped to 200°C at a rate of 40°C per minute, and then held at 200°C for six minutes.

Table 3.1 Enrichment reactor characteristics

Parameter	Daily-fed reactor	Hourly-fed reactor
reactor volume	2 L	3 L
vol. of media fed per day	150 ml	150 ml
hydraulic retention time	13.3 days	20 days
date of most recent inoculation	March-26-2008	2002
pH	7.3	7.4
CH_4 % in headspace	54%	56%

daily-fed reactor has been inoculated with anaerobic sludge from a West Point digester on January/27/2008

3.2 Feed Media

The feed media for the enrichments was modified from the Revised Anaerobic Mineral Medium developed by Shelton and Tiedje (Shelton and Tiedje 1984) (Table 3.2). The concentrations of nickel and cobalt are increased to 6 and 2.4 μM, respectively, due to their beneficial effect to methanogenesis (Speece 1996; Kida, Shigematsu et al. 2001; Conklin, Stensel et al. 2006). Other modifications are shown in Table 3.1.

To prepare the media, all compounds except the vitamin, $Na_2S{\cdot}9H_2O$, resazurin, and acetic acid were added to DI water. After autoclaving at 121^{o}C for 50 min, the headspace gas was purged with nitrogen gas for 15 min. The container was capped with a rubber stopper and cooled to room temperature. Vitamin stock solution, $Na_2S{\cdot}9H_2O$, and resazurin were added while purging

with nitrogen gas. Finally, acetic acid was added after color change of the resazurin indicated that reduced conditions were achieved. If the media was not reduced (color of redox indicator remained pink), additional $Na_2S \cdot 9H_2O$ was added. The media was pumped into bleach sterilized bags (under nitrogen purging). These bags were connected to reactors and stored in the refrigerator.

Table 3.2 Feed media for enrichments

Compound	Final concentration (mM)	Note
$CH_3COO(NH_4)$	92.4	modified
$CH_3COO(Na) \cdot 3H_2O$	5	modified
KH_2PO_4	2	
K_2HPO_4	2	
$Ca(CH_3COO)_2$	0.5	modified
$MgCl_2 \cdot 6H_2O$	0.5	
$FeCl_2 \cdot 4H_2O$	0.1	
Trace Element Stock Solution	10 mL	Table 3.3
Cysteine Stock Solution	2.5 mL	modified
Viamin Stock Solution	50 mL	Table 3.4
$Na_2S \cdot 9H_2O$	0.22	reductant
resazurin	0.2	redox indicator
CH_3COOH	135.6	modified

modified: the components were modified by Anne Conklin (2006) from original recipe (Shelton and Tiedje 1984)

The recipes of trace element and vitamin stock solutions are listed in Table 3.3 and Table 3.4.

The vitamin (100X) and trace element stock solutions (500X) were sterilized by filtration (0.22 μm) and autoclave, respectively. The vitamin solution was stored at 4^{o}C in the dark.

Table 3.3 Trace element stock solution for feed media

Compound	Stock concentration (500X)(μM)	Note
$NiCl_2 \cdot 6H_2O$	6	modified
$MnCl_2 \cdot 4H_2O$	2.52	
$CoCl_2 \cdot 6H_2O$	2.39	modified
$ZnCl_2$	1.38	
Na_2SeO_3	0.29	
H_3BO_3	0.21	
$CuCl_2$	0.03	
$NaMoO_4 \cdot 2H_2O$	0.01	

modified: the components were modified by Anne Conklin (2006) from original recipe (Shelton and Tiedje 1984)

Table 3.4 Vitamin solution for feed media

Compound	Stock concentration (100X)(μg/L)
4-aminobenzoic acid	50
d-biotin (H)	50
cyanocobalamin (B12)	1
folic acid (M)	20
1,4-napthoquinone hydrate 97%	50
nicotinamide	50
D-pantothenic acid Hemicalcium Salt	50
pyridoxine 98%	100
Riboflavin (B2)	50
thiamine hydrochloride (B1)	50
DL-6-8-thioctic acid amide	50

3.3 Gassing Station and Oxygen Scrubbing Device

A gassing station was used for preparing the media, exchanging headspace gas of serum bottles, purging syringes and needles for gas sampling, filling gas bags for enrichment culture wasting, and preparing anoxic stock solutions such as DI water, antibiotics, and acetate. The gases used are nitrogen (99.998%, NI 4.8-T; 99.999%, NI-5.0UH-T), carbon dioxide (CD-M-50), and 20% carbon dioxide and 80% nitrogen (BG NICD20C-K). Trace oxygen in these gases was removed by two types of gas purifiers. One of them was titanium scrubbing solution and the other was copper-based oxygen removal tower. For the titanium scrubbing solution, 5 mL of titanium (III) chloride solution (20%, stabilized in HCl, ST43-500), 6.25 g sodium bicarbonate, and 2.2 g of trisodium citrate were dissolved in 500 mL of DI water. Gases were bubbled through the scrubbing solution to remove trace-level oxygen contamination. The oxygen removal tower was comprised of heat-reduced copper through which the gases were passed (detail description of a oxygen removal tower, see (Robb and Place 1995)).

3.4 Anaerobic Chamber

Many experiments such as infection tests, plaque assays, and media preparation were set up in an

anaerobic glove box (Type A, Coy Lab Products Inc., Grass Lake, MI). Labwares were transferred into the chamber through a vacuum airlock to reduce introduction of oxygen. Nitrogen and hydrogen gas mix (H_2: CO_2: N_2 = 5%: 5%: 90%) were used for the gas exchange in the airlock. In the chamber, hydrogen provided by gas mix could react with oxygen in the presence of heated palladium catalyst. The reaction by product, water, was adsorbed by desiccant to control the humidity inside the chamber. The palladium and desiccant were routinely regenerated by heating at 200^{o}C for 2 h and stored in a 105^{o}C oven.

3.5 Transmission Electron Microscopy (TEM)

VLPs in the samples were viewed using transmission electron microscope (TEM, Philips CM100 TEM, 100kV). TEM samples were prepared in general accordance with the methods described by Ackermann (Ackermann 2009). Samples were filtered through 0.2 µm low protein binding polyethersulfone (PES) syringe filters (Whatman, Clifton, NJ). Filtrate was placed in a 10 ml-polycarbonate tube (Seton Scientific, CA) and ultracentrifuged at 171,500×g and 4°C for 2 h (Beckman Coulter 70.1 Ti rotor, Fullerton, CA). The viral pellet was resuspended in 1 ml of 0.1 M ammonium acetate and centrifuged again at 79,302×g and 4°C for 1 h. The pellet was resuspended into 0.2 ml of 0.1 M ammonium acetate and stored at 4°C.

For TEM grid preparation, 3 µl of poly-L-lysine was placed on a 300 mesh carbon-stabilized formvar-coated copper grid (TedPella Inc, Reading, CA) for 1 min. Then the poly-L-lysine was removed by wicking using Whatman no. 1 filter paper. The viral suspension (5 µl) was placed on the grid for 2 min. Finally, 1µl of negative stain, phosphotungstic acid (PTA, 2%, pH 7.2) was added for 1 min, and the grid was wicked dry. Prepared grids were placed in a desiccator at least overnight until examination by TEM.

3.6 Epifluorescence Microscopy (EFM)

Filtrate was harvested from reactors or infection tests using Whatman 0.2 µm PES filters for enumerating VLPs. VLPs were enumerated using the EFM methods described by Noble and Fuhrman (Noble and Fuhrman 1998) with minor modification. Samples were preserved in 10 mM sodium pyrophosphate, 2% formaldehyde and 94.5 mM of NaCl (final concentrations); preservation solutions used were prefiltered using a 0.02 µm Whatman Anotop™ syringe filter. Fixed solutions were held on ice for 15 min and filtered through Whatman 0.2 µm PES filters. Two µl of 1,000× SYBR Gold (Invitrogen Molecular Probes, Eugene, OR) was added to the filtrate to stain viral nucleic acid. After vortexing, samples were incubated on ice in the dark for at least 15 min. The stained VLPs were mounted on 0.02 µm pore-sized Anodisc^{TM} aluminum

oxide-membrane filters (6809-6002, Whatman, Clifton, NJ). After drying the filter on a Kimwipe® tissue (Kimberly-Clark Corp., Roswell, GA) in the dark, the filter was placed on a glass slide and covered with a cover slip with freshly made mounting solution (50% glycerol and 50% PBS with 0.1% p- phenylenediamide). Slides were stored at -20°C until enumeration. Nucleic acid stained VLPs were directly enumerated using a Leica DM-LB microscope (Buffalo Grove, IL). For enumeration of VLPs, at least 15 images per filter were captured by digital camera (Leica DC 300F) using Image-Pro Plus 5.1 (Media Cybernetics, Inc., Bethesda, MD). Fields were randomly selected, avoiding the area near the supporting ring of the filter. Images were processed by IrfanView (version 4.27) (converted to grey images). VLPs were enumerated using ImageJ (version 1.43u).

3.7 Glycerol Stock of Methanogens

Methanosaeta concilii DSM 6752, DSM 2139, and GP6 were preserved in 20% glycerol. 100 ml of 40% glycerol solution were prepared in a 160-mL serum bottle and purged with nitrogen for 30 min. The glycerol solution was autoclaved and transferred into an anaerobic chamber for dispensing. 20% culture stock was made by 1:1 (v/v) mixing of late log phase methanogens and 40% glycerol solution. The stock was sealed in 10-mL serum bottle and stored at −80^{o}C.

Chapter 4 Improving Culture Methods for *Methanosaeta* Using Solid Media

4.1 Introduction

The growth of *Methanosaeta* in solid media is challenging probably due to the lack of proper solid media, the requirement of strictly anaerobic culture condition and its slow growth rate (the doubling time is 4.8 days) (Janssen 2003). Studies of *Methanosaeta* have been mostly conducted in liquid media. The purification and isolation of *Methanosaeta* from environments using dilution to extinction usually took several months (Zehnder, Huser et al. 1980; Ma, Liu et al. 2006; Mizukami, Takeda et al. 2006). To improve the culture and isolation techniques and to assess plaque assay approaches for *Methanosaeta*, the methods to culture *Methanosaeta* using solid media were evaluated in this chapter.

Methanosaeta-like cells from paddy field soils have been shown to grow in a deep-agar medium and the colonies formed as cotton-like flocs after 6 months of cultivation (Mizukami, Takeda et al. 2006). This same research group also reported that no growth of *Methanosaeta* on roll-tube agar was found after one year of cultivation (Mizukami, Takeda et al. 2006). Recently, a modified cultivation method to grow *Methanosaeta* on solid media was published (Carbonero, Oakley et al.

2010) incorporating several changes of the media (DSM 334). First, the original reductants, cysteine-HCl and sulfide, were replaced by 10 mL/L of a 1.25% (w/v) Titanium (III) citrate solution due to the possible inhibitory effect of H_2S and cysteine on methanogenic cultures (Brauer, Yashiro et al. 2006). Second, the chelating agent nitrilotriacetic acid (NTA) was substituted with citrate due to its potential inhibitory effect on *Methanosaeta*. Third, an alternative solidifying agent, gellan gum, was used for the solid media because it is thermally stable, chemically inert, more economical, and more visibly clearbetter optical clear (Kang, Veeder et al. 1982; Rule and Alexander 1986). Using the improved media, small white/cream colonies were observed inside and on the surface of 1% (w/v) gellan gum after 3 to 5 days of cultivation and about 1 mm diameter colonies formed on 1% solid plates after one month of incubation (Carbonero, Oakley et al. 2010).

The plaque assay is a traditional method used to purify, isolate, and quantify viruses. Among various modifications of this method, one of the most frequently used procedures is the double-agar layer plaque assay (Kropinski, Mazzocco et al. 2008). In principle, viruses are mixed with the host of interest in molten soft agar. This agar is then poured on a layer of base agar precast in a Petri plate. During incubation, the host will grow as a lawn in the top agar. The viruses lyse the contacted hosts and continuously propagate to kill the host cells in their vicinity,

forming visible clear zones called plaques. Each plaque is considered to have originated from a single virus, and therefore this approach can be used to purify and isolate the viruses of the host. Additionally, by counting the number of plaques on the plate, the titer of the virus can also be determined. To apply this method, the host culture must be amenable to grow as a lawn on solid media.

In this chapter, the attempts to grow *Methanosaeta* both as colonies and as lawn were described. In addition, plaque assay approaches for *Methanosaeta* was also assessed.

4.2 Materials and Methods

4.2.1 *Methanosaeta* Cultures and Media

In this study, two *Methanosaeta concilii* strains, DSM 6752 and DSM 2139, purchased from German collection of microorganisms and cell cultures (DSMZ) were used. DSM 6752 was a pure strain and DSM 2139 is described by DSMZ as a highly enriched culture. However, no bacteria were observed in DSM 2139 using confocal microscopy during periodic examinations. To culture the strains, the suggested DSMZ media, DSM 334, was utilized. The recipe of the

media is shown in Table 4.1. Trace element solution and vitamin solution were sterilized by autoclaving and 0.2 μm filtration, respectively. For preparation of the media, all compounds in Table 4.1, except $Na_2S \cdot 9H_2O$, resazurin and the vitamin solution were added to distilled water in a flask and autoclaved. After autoclaving, the media was purged under 20% of CO_2 and 80% of N_2 while it was still hot. The flask was then capped with a sterile black rubber stopper and cooled to room temperature. $Na_2S \cdot 9H_2O$, resazurin and vitamin solution were added using anaerobic techniques (under the CO_2/ N_2 purge). 37% of HCl was added into media until the pH reached 7.0. If the media was colorless, the flask was capped by a stopper. If the media was pink, the media was gradually dosed with additional $Na_2S \cdot 9H_2O$ until it turned colorless. Lab tape was used to secure the stopper before the flask was transferred into an anaerobic glove box for media distribution.

Modified DSM 334 (Carbonero, Oakley et al. 2010) was utilized for solid media in plaque assays. The reductants, cysteine-HCl, and $Na_2S \cdot 9H_2O$, were replaced by 10 mL/L of a 1.25% (w/v) titanium (III) citrate solution. An iron chelating agent, nitrilotriacetic acid (NTA), was substituted with 67 mM of sodium citrate for preparing the trace element solution. For the solidifying agent, agar was replaced by Gelzan (Sigma-Aldrich; St. Louis, MO). The Gelzan concentration used in this study ranged from 0.1% (w/v) to 1.0% (w/v). In addition, 30 mg/L (final concentration) of

ampicillin and vancomycin were used in the modified DSM 334 media to inhibit bacterial growth.

Table 4.1 DSM 334 media

Media	**DSM 334**	**modified DSM 334**	
Compound	Concentration	Concentration	Unit
KH_2PO_4	0.300	0.300	g/L
NaCl	0.600	0.600	g/L
$MgCl_2 \cdot 6H_2O$	0.100	0.100	g/L
$CaCl_2 \cdot 2H_2O$	0.080	0.080	g/L
NH_4Cl	1.000	1.000	g/L
$KHCO_3$	4.000	4.000	g/L
Cysteine-HCl·H_2O	0.300	--	g/L
$Na_2S \cdot 9H_2O$	0.300	--	g/L
Titanium (III) citrate	--	10.000	mL/L
Resazurin	1.000	1.000	mg/L
Trace element solution	10.000	10.000	mL/L
Vitamin solution	10.000	10.000	mL/L
Trace element solution			
Compound	Concentration	Concentration	Unit
Nitrilotriacetic acid (NTA)	12.800	--	g/L
Sodium citrate dihydrate	--	19.7	g/L
$FeCl_3 \cdot 6H_2O$	1.350	1.350	g/L
$MnCl_2 \cdot 4H_2O$	0.100	0.100	g/L
$CoCl_2 \cdot 6H_2O$	0.024	0.024	g/L
$CaCl_2 \cdot 2H_2O$	0.100	0.100	g/L
$ZnCl_2$	0.100	0.100	g/L
$CuCl_2 \cdot 2H_2O$	0.025	0.025	g/L
H_3BO_3	0.010	0.010	g/L
$Na_2MoO_4 \cdot 2H_2O$	0.024	0.024	g/L
NaCl	1.000	1.000	g/L
$NiCl_2 \cdot 6H_2O$	0.120	0.120	g/L
$Na_2SeO_3 \cdot 5H_2O$	0.026	0.026	g/L
Vitamin solution			
Biotin	2.000	2.000	mg/L
Folic acid	2.000	2.000	mg/L
Pyridoxine-HCl	10.000	10.000	mg/L
Thiamine-HCl·$2H_2O$	5.000	5.000	mg/L
Riboflavin	5.000	5.000	mg/L
Nicotinic acid	5.000	5.000	mg/L
D-Ca-pantothenate	5.000	5.000	mg/L
Vitamin B_{12}	0.100	0.100	mg/L
p-Aminobenzoic acid	5.000	5.000	mg/L
Lipoic acid	5.000	5.000	mg/L

4.2.2 Methane Measurement

Methane concentrations produced by methanogens were analyzed using a SRI 8610C GC-FID with a Supelco Alumina Sulfate Plot column (50 m, 0.53 mm i.d.). Gas samples (100 μL at bottle internal pressure) were taken using a 1-mL BD syringe (BD309602) and a 25-gauge needle (BD305125) (Becton, Dickinson and Company; Franklin Lakes, NJ) connected by a Teflon Mininert valve (Supelco; Bellefonte, PA). The bottles were allowed to adjust to the ambient temperature before sampling. Before injection into GC-FID (valve closed), 25-gauge needles were replaced with 30-gauge needles (BD 305128) because 30-gauge needles matched the diameter of the GC-FID injection port while 25-gauge needles were easier to handle when penetrating the blue rubber stopper (2048-11800, Bellco Glass; Vineland, NJ). 10 mL of methane standards ranging from 0.2% to 100% were prepared in 20-mL Tedlar gas bags (Jensen Inert Products; Coral Springs, FL). Each standard was prepared by mixing methane and nitrogen gases in the gas bag (final volume was 10 mL). Triplicate injections were made for each standard curve and the standard curves were generated around every four months. A typical standard curve is shown in Figure 4.1. The overall standard curve (Figure 4.1 a) consists of two linear sections which are used to calculation methane concentration when peak area equal or below 2500 (Figure 4.1 b) and above 2500 (Figure 4.1 c).

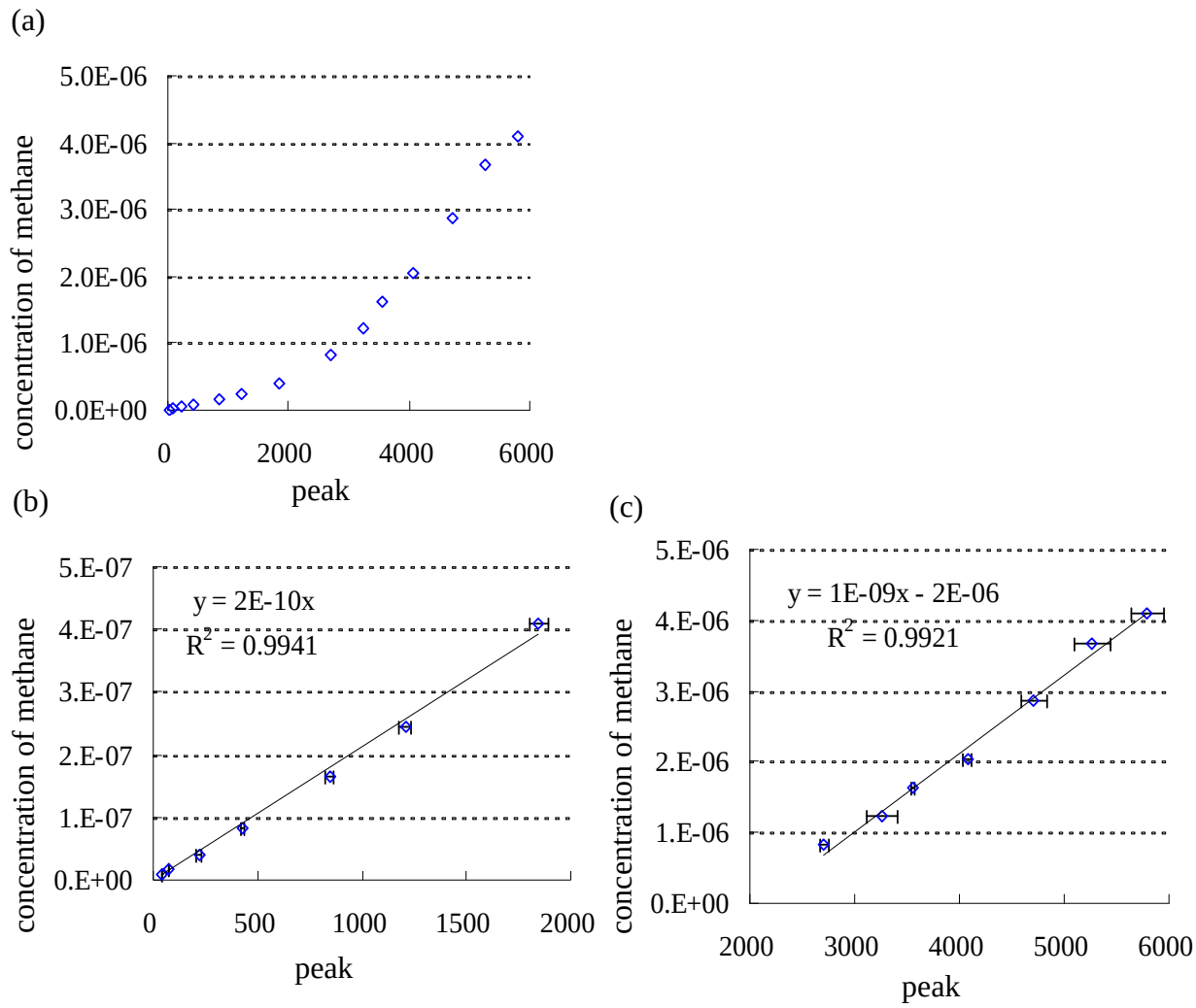


Figure 4.1 Standard curve of methane using GC-FID
(a) overall standard curve of methane. (b) the standard curve of methane when peak area ≤ 2500. (c) the standard curve of methane when peak area > 2500. Error bars represent avedevs of triplicate injections.

4.2.3 Virus Size-fraction Filtrate (VSF)

Virus size-fraction filtrates (VSFs) were harvested from enriched reactors or anaerobic digesters for plaque assays or infection tests. In addition, VSFs were also collected from the infection tests

during serial transfers. To collect the VSFs from enrichment reactors, two steps were performed: (1) the cultures in the reactors were anoxically wasted into nitrogen-purged serum bottles through 18-gauge needles (BD1838, Becton, Dickinson and Company; Franklin Lakes, NJ); (2) the cultures in the bottles were withdrawn by syringes and needles and filtered through 0.2 µm pore-size low protein binding PES syringe filters (6780-2502, Whatman, Clifton, NJ) into sterile anoxic serum bottles. For the VSFs from infection tests, the filtrates were generated by step (2) described above. For the VSFs originating from full-scale digesters, pretreatment such as centrifuging and prefiltering was necessary. Anaerobic sludge was placed in 50-mL conical tubes and centrifuged at 4^{o}C and 4150 rpm for 30 min in a Sorvall Legend RT tabletop centrifuge equipped with a swing-bucket rotor (75006441, Thermo Scientific, Waltham, MA). The supernatant was collected and vacuum filtered through 1µm glass fiber filters. The filtrates were then filtered through 0.45 and 0.20 µm PES syringe filters (6780-2502 and 6780-2504, Whatman, Clifton, NJ). The VSFs were transferred to anoxic serum bottles and stored at 4^{o}C before use.

4.2.4 Plaque Assay

Modified DSM 334 media was used for the plaque assay. For solid media preparation, all reagents except resazurin, titanium (III) citrate solution, and vitamin solution were dissolved in

distilled water in a flask with a stir bar. The media was autoclaved at 121^{o}C and 15 psi for 25 min. After autoclaving, the media was cooled under the purge of CO_2/N_2 gas mix (1:4) and was kept warm on a hot plate. Resazurin, titanium (III) citrate solution, antibiotics (ampicillin and vancomycin), and vitamin solution were added into the media and the pH was adjusted to 7.0 using 37% HCl. When the media was reduced (colorless at pH 7.0), it was immediately transferred to the anaerobic glove box and distributed into 15-mL or 50-mL polystyrene centrifuge tubes (352099 and 352074, Corning, Tewksbury, MA) pre-warmed to 50^{o}C using a dry bath. Plating techniques such as pour plating and spread plating were utilized for plaque assays. To assess the growth of *Methanosaeta* on solid media, streak plating was also used. Several media containers were tested for the plaque assays including plastic and glass Petri dishes, serum bottles (W012492, Wheaton; Millville, NJ), Balch tubes (2048-00150, Bellco Glass; Vineland, NJ), cell culture flask (430168, Corning, Tewksbury, MA) and Wolfe bottles (Bellco Glass; Vineland, NJ) (Hermann, Noll et al. 1986).

4.3 Results and Discussion

4.3.1 Gellan Gum Concentration

In Carbonero's study (2010), it took one week to obtain *Methanosaeta concilii* (DSM 6752) colonies in media with 0.1% gellan (semisolid) and one month to have colonies on plates with 1% gellan (solid). However, 0.1% semisolid plates were almost impossible to be manipulated for plaque assays. Therefore, multiple gellan concetrations were tested to reduce the time for observed *Methanosaet* sp. growth while preserving gellan strength for plaque assays. Gellan concentrations ranging from 0.12% to 0.8% were tested. Depending on the plating techniques (e.g. pouring, spreading, and streaking), the type of containers (e.g. Petri dishes, Balch tubes, and serum bottles), and the orientation of containers during incubation (e.g. tilted or inverted to avoid moisture condensing on the surface of media), gellan concetrations from 0.15 % to 0.5% were used in this study.

4.3.2 Cultivation Containers and Plating Techniques

Initially, Petri dishes and cell culture flasks were selected due to easy manipulation (Figure 4.2 a and d). The plates and flasks were placed in the glove box at room temperature and in a BBL GasPak^{TM} 100 Anaerobic jar (260626 Becton, Dickinson and Company; Franklin Lakes, NJ) (Figure 4.2 b) incubated at 35^{o}C. In both cases, the media in plates and flasks turned pink within one week and no colonies were found on the media after three months. This indicated that the

media may have been exposed to oxygen under these conditions. Therefore, 160-mL glass serum bottles and Balch tubes were tested with *Methanosaeta concilii* DSM 6752. By pour plating *Methanosaeta* using media with 0.2% of gellan in Balch tubes, the media turned turbid in one week at 35^{o}C (Figure 4.3 a). In addition, the growth of *Methanosaeta* was confirmed by detecting increased methane in headspace gas and observing *Methanosaeta*-like cells when these samples were examined under the microscope (Figure 4.3 b). By spread plating *Methanosaeta* on the media with 0.5% of gellan at 35^{o}C in serum bottles, white colonies (1-3 mm in diameter) were formed (Figure 4.3 c). Methane was also detected and the *Methanosaeta*-like cells were shown in Figure 4.3 d. Nevertheless, the media in serum bottles or Balch tubes turned pink occasionally, which was probably due to the leak from the old septa. Wolfe bottles had the advantages of large surface area and being gas-tight; however, they were not routinely used because of the lack of availability of these bottles.

Both spread and pour plating were able to grow *Methanosaeta* in solid media. By spread plating, white colonies formed on the surface of the media after about one month of incubation. However, the culture did not grow as a lawn. By pour plating, white cloudy cultures grew inside the media and these cultures formed faster than spread plating (after one week) and produced turbidity, which could be used for the testing of plaque assay.

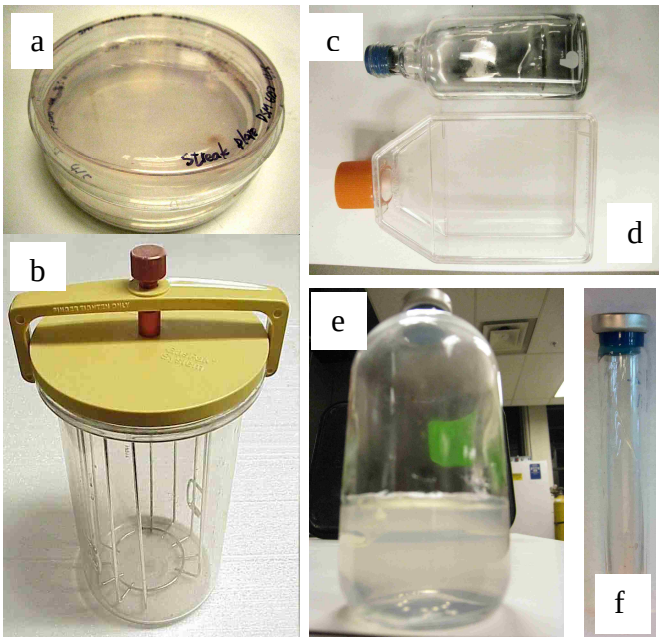


Figure 4.2 Cultivation containers used in plaque assays.
(a) Petri dish. (b) BBL GasPak™ 100 Anaerobic jar. (c) Wolfe bottle. (d) cell culture flasks (e) 160 mL serum bottle. (f) Balch tube.

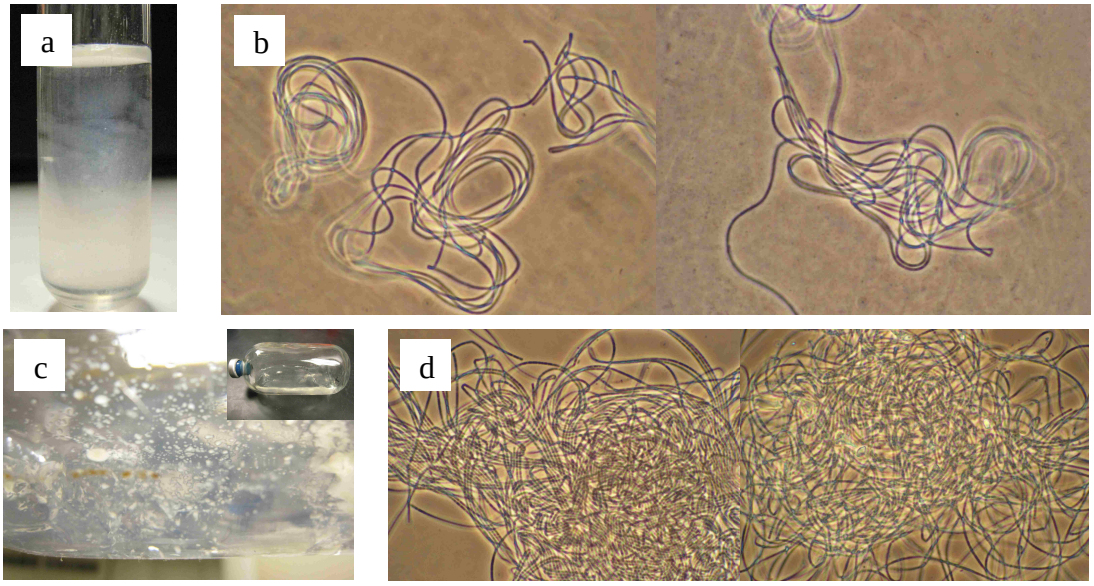


Figure 4.3 Cultivation of *Methanosaeta concilii* (DSM 6752) using serum bottles and Balch tubes.
(a) Pour plating *Methanosaeta concilii* in the media with 0.2% of gellan in the Balch tube. (b) *Methanosaeta*-like cells from the tube. (c) Colonies of *Methanosaeta concilii* on the surface of media (0.5% gellan). (d) *Methanosaeta*-like cells from the colonies.

4.3.3 Fragmentation of *Methanosaeta*

Methanosaeta usually grow as a long thread (up to hundred micrometers) (Figure 4.4) when cultured without any physical shear force such as shaking or stirring. In order to disperse *Methanosaeta* more homogeneously in the media, several methods such as shaking and ultrasound were tested. Figure 4.4 shows the results of fragmentation of cultures. Both shaking and ultrasound were able to shear long threads into small fragments after 70 min. However, single cells of *Methanosaeta* were rarely observed. The fragmented and untreated cultures were then inoculated into Balch tubes using media with different gellan concentration to further evaluate the growth rate by measuring the methane production rate in Balch tubes (Figure 4.5). Negative correlation was found between methane generation rates and gellan concentrations. Consistently, the fragmented cultures produced less methane than those untreated. This result suggested that sheared cells might be damaged during the treatment. No difference of the evenness of *Methanosaeta* flocs in tubes was observed between sheared or untreated culture. The media with 0.2% gellan was still too soft to handle. Therefore, untreated *Methanosaeta* pour plating using the media with 0.5% gellan was used for the assessment of plaque assays.

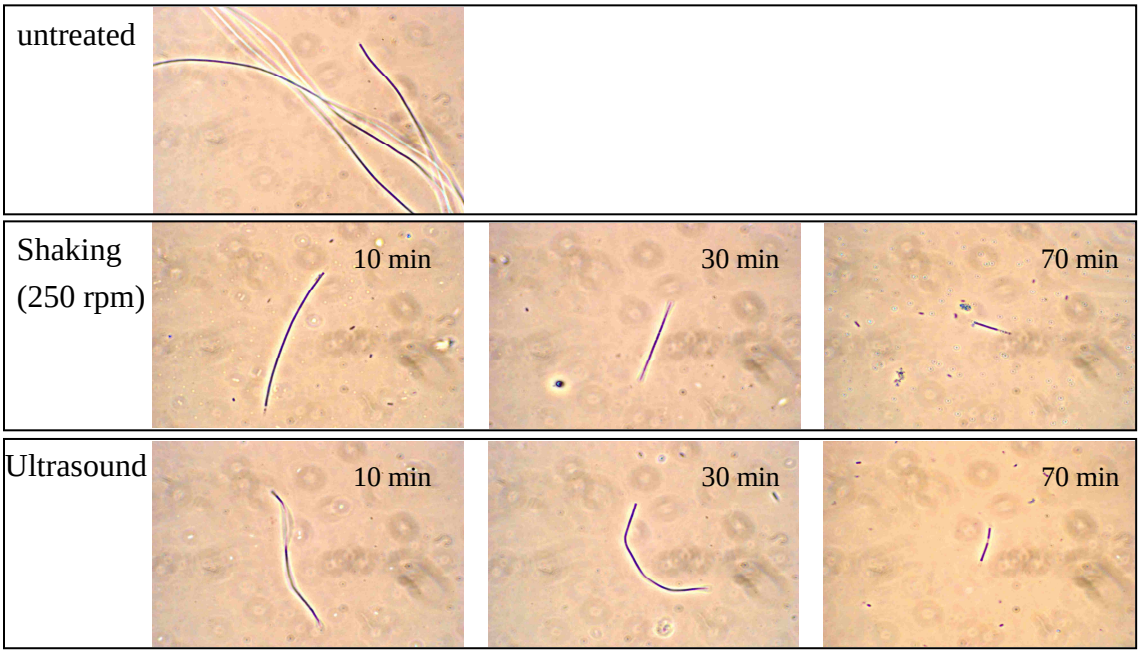


Figure 4.4 Fragmentation of *Methanosaeta concilii* (DSM 6752) by shaking and ultrasonic. Images were observed at 1000X magnification (oil immersion).

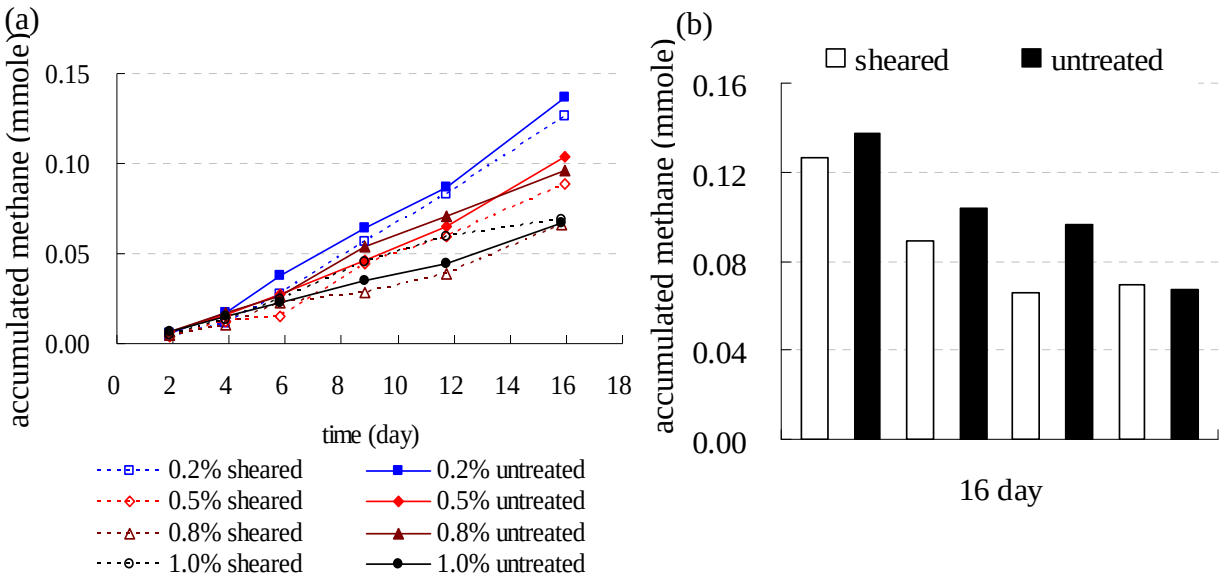
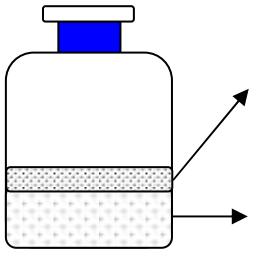


Figure 4.5 Effect of gellan concentrations and fragmentation of cultures to the growth rate. (a) methane generation for ultrasonic and untreated *Methanosaeta concilii* (DSM 6752) in Balch tubes. (b) methane yield at day 16. 0.1 mL of cultures were inoculated into 5 mL of media (expected final methane yield is 0.4 mmole).

4.3.4 Plaque Assays

Typical plaque assays performed in this study were shown in figure 4.6 and table 4.2. Double-layer plaque assays were used. 20-40 mL of media (0.1-0.5% of gellan) was poured as the bottom layer (providing additional nutrients for the growth of *Methanosaeta* inoculated in the top layer and diffusing the metabolites out from the top layer). 5-10 mL of media (0.1-0.5% of gellan), *Methanosaeta concilii* and VSFs were mixed and poured as the top layer. In this layer, plaques would form if infectious viruses were present. VSFs from the daily-fed enrichment, the hourly-fed enrichment, and upset full-scale digesters were evaluated in the plaque assays. Serial diluted VSFs were spotted on the top layer or mixed with *Methanosaeta* and pour plated. VSFs and *Methanosaeta* were incubated in liquid media for 2 hours at 35^{o}C with shaking before pour plating. In most tests, no plaques were observed. Several plaque-like clear zones were found in viral assays. Material from these clear zones was analyzed by EFM, however, no VLP was found. Although *Methanosaeta* was able to grow in the modified media, the cell concentration in gellan was not very dense, which made the observation of plaques difficult. Concentration of sodium acetate in the bottom layer was increased (from 6.8 g/L to 24.6 g/L) in order to increase cell density but a change in turbidity was not obvious.



160 mL serum bottle

Top layer:
media (0.5%) 5-10 mL, DSM 6752, VSFs
pour plating or spotting

Bottom layer
media (0.1-0.5%) 20-40 mL
supplying more media, diluting generated metabolites

Figure 4.6 Experimental bottle setup for double-layer plaque assays for *Methanosaeta concilii*.

Table 4.2 Experimental design for plaque assays for *Methanosaeta concilii*

	Top layer			Bottom layer
bottle	media 0.5 % gellan	*Methanosaeta* (DSM 6752)	Virus filtrate (Daily reactor)	media 0.1 % gellan
1	10 ml	1 ml	100 µl	20 ml
2	10 ml	1 ml	--	20 ml
3	10 ml	1 ml	100 µl	40 ml
4	10 ml	1 ml	--	40 ml

note: 1. Serial dilution of virus size-fraction filtrates (VSFs): 10^{-1}, 10^{-2}, 10^{-3} and 10^{-4}

2. VSFs and *Methanosaeta* were mixed and shaken for 2 hours at 35°C before pour plating

3. Samples were duplicated (10^{-2}, 10^{-3} and 10^{-4})

4.4 Conclusions

The growth of *Methanosaeta concilii* was observed using modified DSM 334 solid media in glass containers such as serum bottles and Balch tubes. *Methanosaeta concilii* colonies appeared on the surface of media (0.5% of gellan) after one month of incubation in serum bottles at 35°C by spread plating but these colonies did not spread as a lawn. By pour plating (0.5% of gellan),

growth of *Methanosaeta* was found by observing formation of turbidity and detecting methane in headspace gas of serum bottles after about one week incubation at 35^{o}C. In the future, growth of *Methanosaeta* on solid media can be further evaluated using different anaerobic techniques. For example, Petri dishes with pressure-controlled stainless steel containers or gas-tight bags (Nakamura, Tamaki et al. 2011) could better avoid oxygen contamination, since Petri dishes have the advantage of large surface area and flat transparent covers that make the observation of growth much easier than serum bottles (thick, round glass bottles).

By pour plating, a double-layer plaque assay was tested in serum bottles to screen for viruses of *Methanosaeta* but plaques were not observed. The modified DSM 334 media improved culturing but was still not sufficient for plaque assays. For plaque assays in the future, inducing (mitomycin C and UV) and enriched lytic virus might be worth incorporating into the tests.

Chapter 5 Ammonia Toxicity and the Assessment of Infection Tests for *Methanosaeta*

5.1 Introduction

During the initial assessment of viral impact on two pure *Methanosaeta* strains cultured in liquid media, and challenged by VSFs from enrichment reactors and full-scale digesters, methane suppression results did not agree with TEM results and negative controls (i.e. methane yields remained inhibited while virus concentrations decreased). Inhibitors were investigated and ammonia was shown to be the inhibitor.

For anaerobic digesters, treatment processes can benefit from total ammonia nitrogen (TAN, sum of ammonia-nitrogen and ammonium-nitrogen) ranging from 50 and 200 mg/L (McCarty and McKinney 1961; McCarty 1964; Liu and Sung 2002). But inhibition of methanogenesis was observed when TAN was elevated. 50% methane suppression was reported in a wide range of TAN (1700 to 14000 g/L) (Chen, Cheng et al. 2008). For *Methanosaeta concilii*, the ammonia toxicity level was reported in only two studies. In one of the studies, methane generation was completely inhibited at 2140 mg TAN/L at pH 7.0 (Sprott and Patel 1986). In the other study, cell

density of *Methanosaeta concilii* GP6 reduced to 50% when the concentration of TAN reached about 1000 mg/L, and dramatically decreased (~75%) when TAN rose beyond 1900 mg/L (Steinhaus, Garcia et al. 2007). Although the inhibition of *Methanosaeta* by ammonia was identified in previous studies, the suppression of methanogenesis at different level of free ammonia is unknown. Therefore, the methane production of *Methanosaeta* in media with different free ammonia concentrations was investigated in this chapter.

When the ammonia toxicity was controlled, improved infection test approaches were evaluated. The assessment of infection tests was summarized in the second section of this chapter.

5.2 Materials and Methods

5.2.1 Acetate Solution and Anoxic DI water

Acetate solution was made by dissolving 0.539 g ammonia acetate and 0.952 g sodium acetate trihydrate in distilled water to 100 mL. The solution was placed into 160-mL serum bottles and purged with nitrogen. The bottles were capped with blue rubber stoppers and sealed with 20 mm aluminum crimps (224178-01, Wheaton; Millville, NJ). The bottle were autoclaved at 121^{o}C for

25 min. Anoxic DI water, used as controls in infection tests, was also prepared as described above.

5.2.2 Collection of Virus Size-fraction Filtrates from Upset Digesters

Virus size-fraction filtrates (VSFs) from anaerobic digesters, which had been upset, were screened for the virus of *Methanosaeta*. For example, Filtrate 2009 was obtained from anaerobic digesters at the Port Angeles wastewater treatment plant (WWTP) when the digesters became upset in Nov 2009. Two digesters (West and East) were experiencing different levels of upset when the sludge was collected. The West digester had greater symptoms of upset and had volatile fatty acid (VFA) 2500-3000 mg/L as acetic acid, alkalinity 1700-2300 mg/L as $CaCO_3$, and a pH of 5.8. The upset of the east digester was much less pronounced and had VFA's 380-850 mg/L, alkalinity 2800-3000 mg/L, and pH of 6.8. Although the East digester was less upset, it was foaming severely. Filtrate 2009 was used to challenge methanogens from enrichments (daily-fed and hourly-fed reactors) and *Methanosaeta concilii* DSM 6752 and DSM 2139. After digester samples arrived at the laboratory, VSFs were prepared immediately. The method to process VSFs was described in 4.2.3.

5.2.3 Concentrating VLPs

VLPs in VSFs were concentrated by a centrifugal ultrafiltration device, Amicon Ultra-15, (UFC903024, Millipore, Billerica, MA). 15 mL of VSFs were added to the filtration unit and centrifuged at 4000×g for 10 min. After centrifuge, filtrate was discarded. If the total volume of VSFs was more than 15 mL, the process was repeated until all of the VSFs were filtered. The final volume of VSFs (retentate of ultrafiltration) was reduced to 150-250 μL by adjusting the time of centrifugation.

5.2.4 Infection Test Setup

Two strains of *Methanosaeta*, DSM 6752 and DSM 2139, were challenged by VSFs in serum bottles. A flowchart of the infection test approach is illustrated as Figure 5.1. In the infection test, *Methanosaeta*, 0.14 M acetate media and virus size-fraction filtrates (VSFs) were added to serum bottles by needle and syringe (in control bottles, VSFs were replaced by anoxic water or media). Methane yield in the headspace of serum bottles was measured by GC-FID during the incubation. Bottles producing less methane than the controls were suspected to be inoculated with *Methanosaeta* viruses. The original VSFs harvested from enrichments or anaerobic digesters contained a variety of viruses and each passage of this VSF was expected to dilute those unrelated

viruses out of bottles. VSFs collected from bottles having methane suppression were serially transferred to fresh *Methanosaeta* cultures. If the VSFs did not inhibit methane production during serial transfer, the set of infection tests was abandoned. If the methane yield was suppressed by VSFs during serial transfer, controls to inactivate viruses (such as autoclave or UV treatment) or to remove viruses from filtrate (such as 0.02 μm filtration) were performed to assess whether the inhibition was caused by viruses. The morphology of virus-like particles (VLPs) in the VSFs was viewed by transmission electron microscope (TEM).

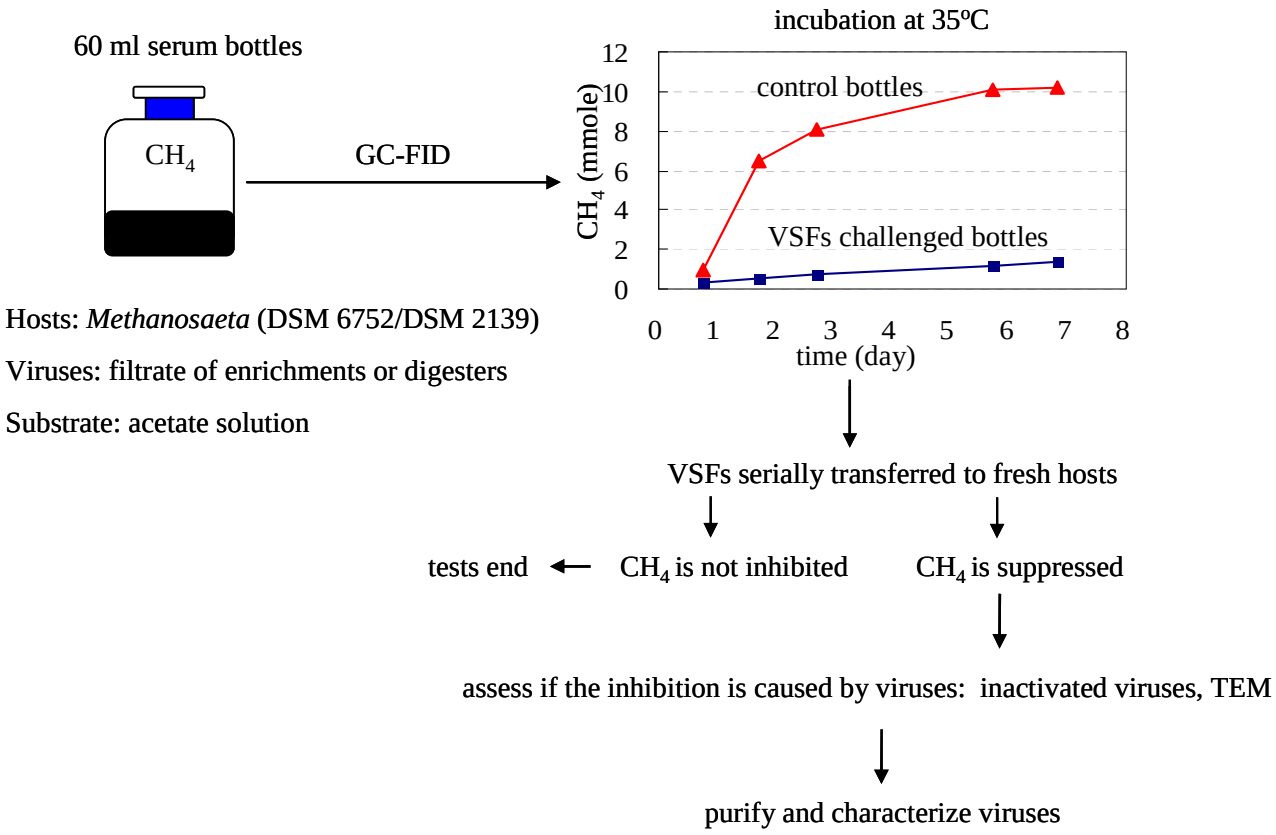


Figure 5.1 Flowchart of planned infection test approach with idealized results scenarios

5.3 Results and Discussion

5.3.1 Investigating the Inhibition of Methanogenesis

Inoculation and Serial Passage

Two lab *Methanosaeta*-dominated enrichments were used for infection test to assess the first transfer of Filtrate 2009 (see Table 5.1). The result of the first transfer of Filtrate 2009 is shown in Figure 5.2. Compared to DI water controls, bottles (177 and 180) inoculated with VSF from the east digester produced less methane in both enrichment cultures. In contrast, bottles (176 and 179) inoculated with the VSF from the larger upset west digester generated more methane than that of DI water control, which might be due to the adding of more acetate in the VSFs from the digester. Because methane generation was suppressed in bottles containing VSFs from the East digester, this filtrate (named Filtrate 2009) was considered to potentially have infective virus particles. Therefore, VSF was harvested from bottles 177 and 180 to see if inhibition of methanogenesis could be transferred.

Table 5.1 Experimental design for challenging *Methanosaeta*-dominated enrichments with Filtrate 2009

Serum	Source of methanogens:	Source of VSFs	note

bottle no.			
176 A,B	daily-fed enrichment: 5 mL	west digester: 3.5 mL	infection challenge (sick)
177 A,B	daily-fed enrichment: 5 mL	east digester: 3.5 mL	infection challenge (sick)
178 A,B	daily-fed enrichment: 5 mL	DI water: 3.5 mL	negative control
179 A,B	hourly-fed enrichment: 5 mL	west digester: 3.5 mL	infection challenge (sick)
180 A,B	hourly-fed enrichment: 5 mL	east digester: 3.5 mL	infection challenge (sick)
181 A,B	hourly-fed enrichment: 5 mL	DI water: 3.5 mL	negative control

DI: deionized water (anoxic)

1.5 mL acetate feed was added to all bottles for a total bottle volume of 10 mL

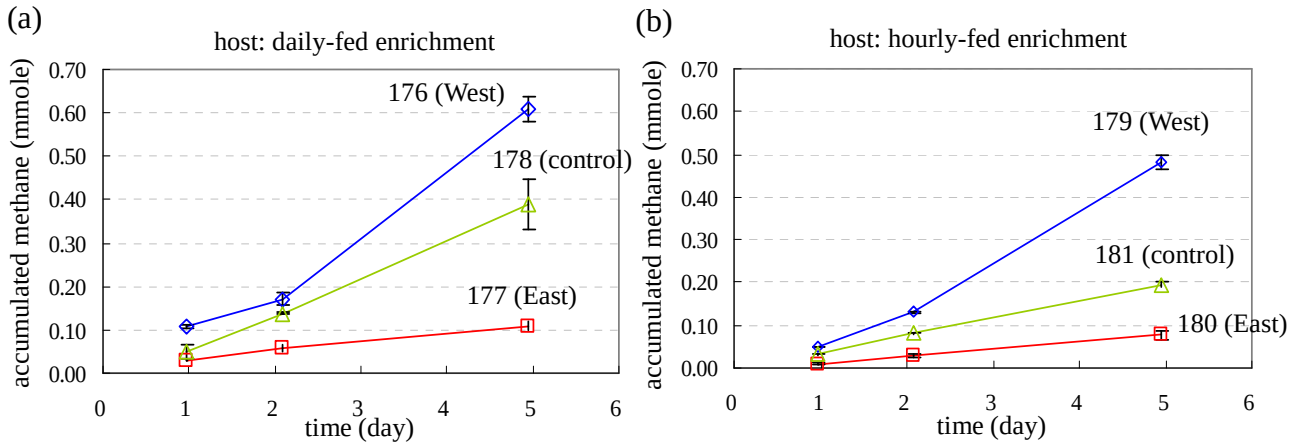


Figure 5.2 Challenging *Methanosaeta*-dominant enrichments with Filtrate 2009.

Hosts in the challenging tests were daily-fed enrichment (a) and hourly-fed enrichment (b)

Filtrate 2009 was transferred one more time (second transfer) in the enrichments to verify the ability to suppress the generation of methane (Table 5.2). The result is illustrated in Figure 5.3. As in the previous infection test, enrichment receiving Filtrate 2009 produced less methane than its corresponding DI control. At the end of the test, 75% and 95% methane suppression were observed for the daily-fed and the hourly-fed enrichments, respectively. Again, the VSF was

collected, and further transfers were done with a pure culture of *Methanosaeta concilii* (DSM 6752).

Table 5.2 Experimental design for transferring of Filtrate 2009 in enrichments

Serum bottle no.	Source of methanogens:	Source of VSFs	note
182 A,B	daily-fed enrichment: 5 mL	177A: 3.5 mL	filtrate transfer
183 A,B	daily-fed enrichment: 5 mL	177B: 3.5 mL	filtrate transfer
184 A,B	daily-fed enrichment: 5 mL	DI water: 3.5 mL	negative control
185 A,B	hourly-fed enrichment: 5 mL	180A: 3.5 mL	filtrate transfer
186 A,B	hourly-fed enrichment: 5 mL	180B: 3.5 mL	filtrate transfer
187 A,B	hourly-fed enrichment: 5 mL	DI water: 3.5 mL	negative control

DI: deionized water (anoxic)

1.5 mL acetate feed was added to all bottles for a total bottle volume of 10 mL

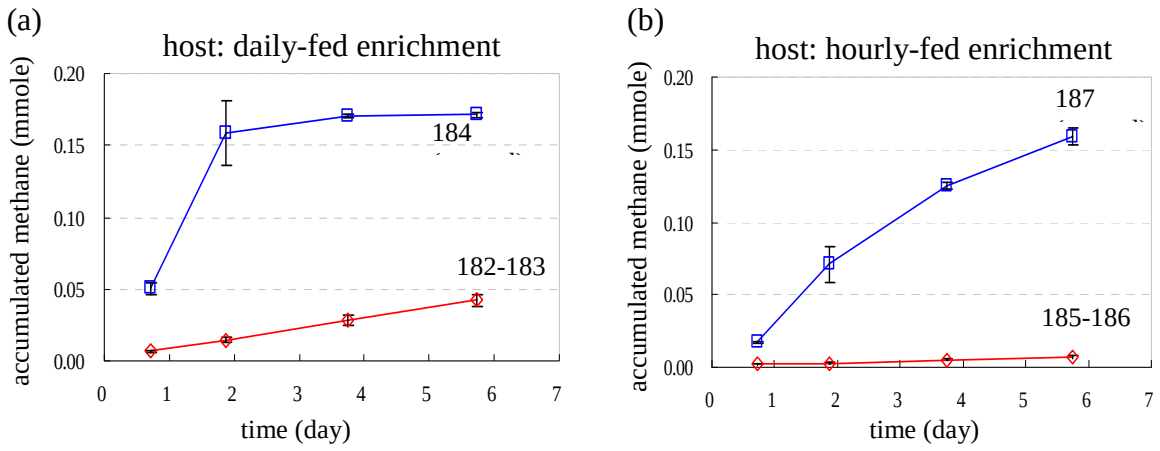


Figure 5.3 Transfer of Filtrate 2009 in enrichments

Hosts for the transfer were daily-fed enrichment (a) and hourly-fed enrichment (b)

Filtrate 2009 was then serially passed (9 transfers) into *Methanosaeta concilii* (DSM 6752) to confirm its ability to suppress methane generation. The experiment design and typical results are shown in Table 5.10 and Figure 5.12. Generally, at the end of an infection experiment, methane yield from the culture receiving Filtrate 2009 was inhibited about 75 % compared to the DI water control.

Table 5.3 Experimental design for transferring of Filtrate 2009 in *Methanosaeta concilii* (DSM 6752)

Serum bottle no.	Source of methanogens: *Mst. concilii* DSM 6752	Source of VSFs	note
192 A,B	*Mst. concilii*: 5 mL	182A: 3.5 mL	filtrate transfer
193 A,B	*Mst. concilii*: 5 mL	182B: 3.5 mL	filtrate transfer
194 A,B	*Mst. concilii*: 5 mL	183A: 3.5 mL	filtrate transfer
195 A,B	*Mst. concilii*: 5 mL	183B: 3.5 mL	filtrate transfer
196 A,B	*Mst. concilii*: 5 mL	DI water: 3.5 mL	negative control

DI: deionized water (anoxic)

1.5 mL acetate feed was added to all bottles for a total bottle volume of 10 mL

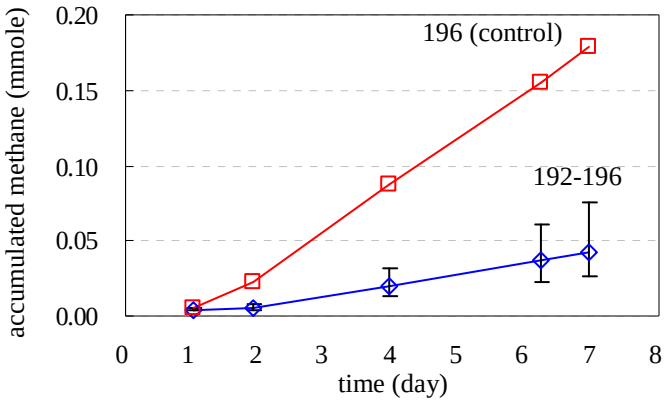


Figure 5.4 Transfer of Filtrate 2009 in *Methanosaeta concilii* (DSMZ 6752)

TEM Micrographs of VLPs in Filtrate 2009

During the serial passage of Filtrate 2009, VLPs in the filtrate were examined periodically using TEM. The contaminant viruses (i.e. viruses of other hosts), originating from anaerobic digesters and the enrichments, were expected to be diluted in the system during serial passage and leave a majority of VLPs associated with the available hosts, *Methanosaeta concilii* (DSMZ 6752). The morphology of VLPs in the filtrate at second transfer and at fourth transfer are shown in Figures 5.13 and 5.14. These images indicated that all three families of tailed phage were present in the Filtrate 2009.

The morphology of virus-like particles in the filtrate at 8^{th} transfer and at 11^{th} transfer is shown in Figure 5.15 (the dilution factors at 8^{th} and 11^{th} transfer were 1.8×10^{-3} and 7.9×10^{-5}). Tailed phages were not found in the filtrate, instead, spherical and oval-shaped particles ranging from 30 nm to 50 nm were most frequently observed. VLPs consistent with this size and morphology were not observed during serial passages, which may indicate that the inhibition of methanogenesis during infection tests was not caused by *Methanosaeta* viruses.

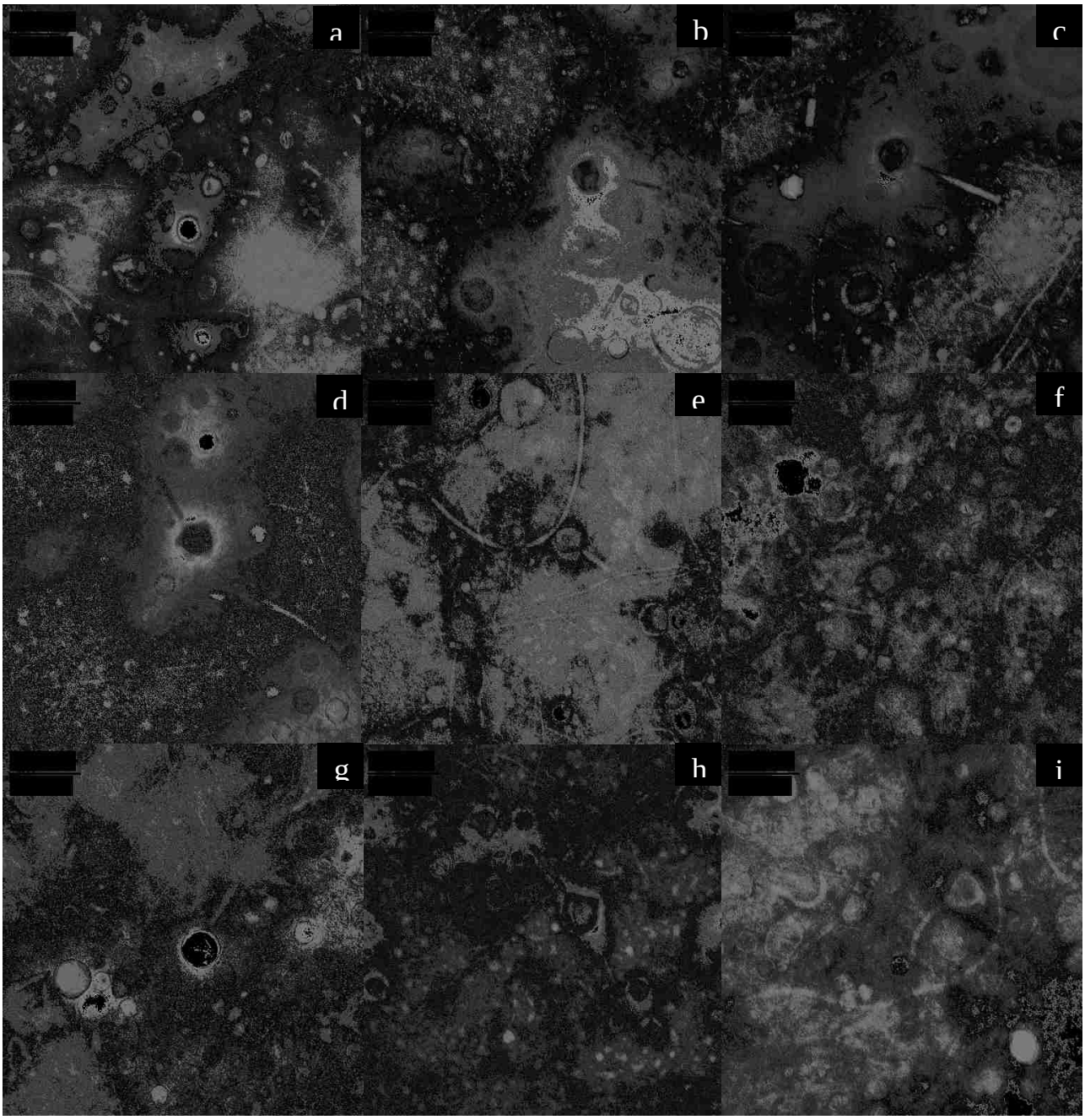


Figure 5.5 TEM micrographs of filtrate 186 (2nd transfer of filtrate 2009)
In the Filtrate 2009, tailed phage including *Myoviridae* (tail contractile; e.g. 5.5 (d)), *Siphoviridae* (tail long, noncontractile; e.g. 5.5 (c)) and *Podoviridae* (tail short; e.g. 5.5 (f)) were observed.

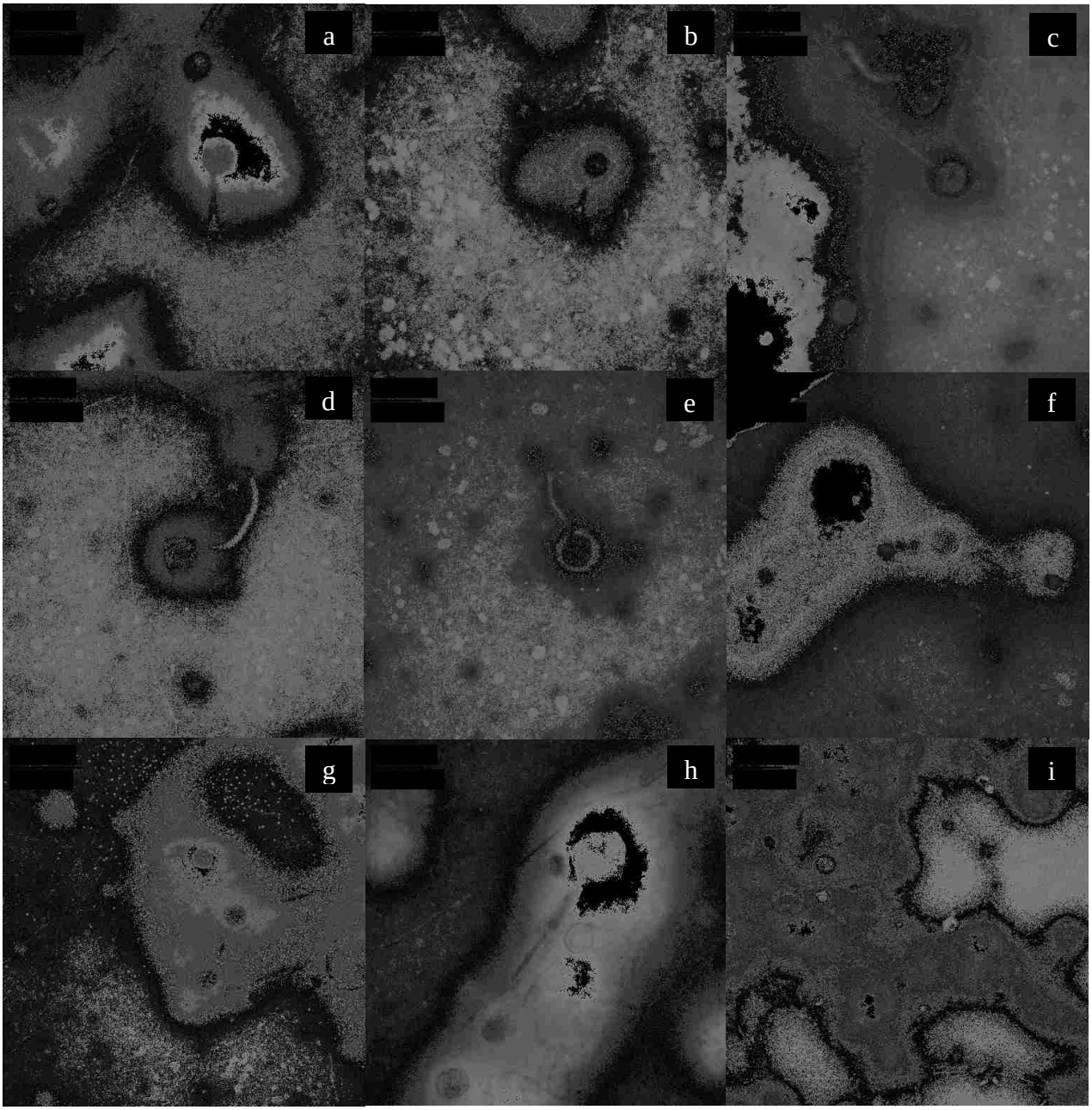
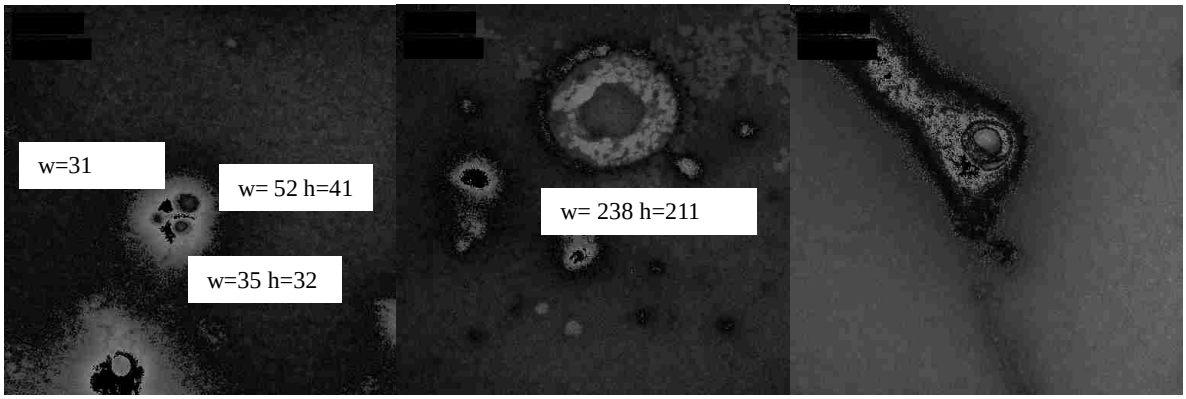


Figure 5.6 TEM micrographs of filtrate 204 (4th transfer of filtrate 2009).
In the Filtrate 2009, tailed phage including *Myoviridae* (tail contractile; e.g. 5.6 (h)), *Siphoviridae* (tail long, noncontractile; e.g. 5.6 (c)) and *Podoviridae* (tail short; e.g. 5.6 (b)) were present.

(a)



(b)

Figure 5.7 TEM micrographs of filtrate.
(a) 8th transfer and (b) 11th transfer

Inactivating Viruses by Autoclave

In order to examine whether infective viruses caused the inhibition of methane generation, viruses in Filtrate 2009 were inactivated by autoclaving. The experiment is described in Table 5.4 and the result are shown in Figure 5.8. *Methanosaeta* inoculated with autoclaved filtrate (bottle 205) produced less methane than *Methanosaeta* with Filtrate 2009. Compared to the DI water control, 68 % of methane yield was inhibited in bottles with Filtrate 2009, and 88 % of methane yield was suppressed in bottles with autoclaved Filtrate 2009. It also implied that the inhibitor may increase or was more toxic after heat treatment.

Table 5.4 Experimental design to inactivated viruses in Filtrate 2009 by autoclave

Serum bottle no.	Source of methanogens: *Mst. concilii* DSM 6752	Source of VSFs	Note
202 A,B	5 mL	194 A: 3.5 ml	filtrate transfer
203 A,B	5 mL	194 B: 3.5 ml	filtrate transfer
204 A,B	5 mL	195 B: 3.5 ml	filtrate transfer
205 A,B	5 mL	autoclaved* 195 A: 3.5 ml	to inactive viruses
206 A,B	5 mL	DI water: 3.5 ml	negative control

*: autoclaved at 121°C for 30 min (three cycles)

DI: deionized water (anoxic)

1.5 mL acetate feed was added to all bottles for a total bottle volume of 10 mL

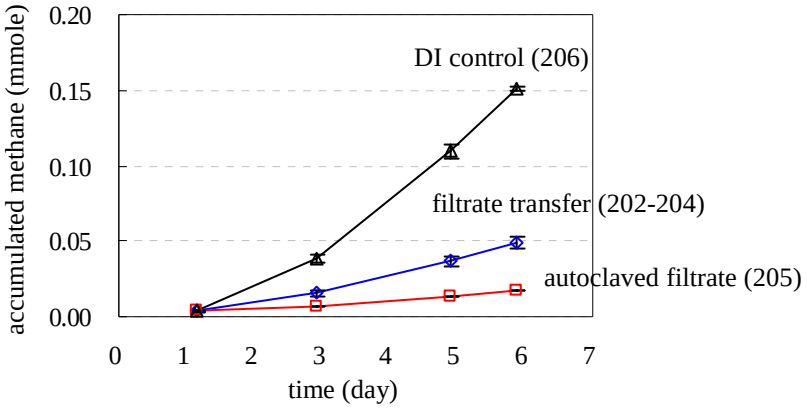


Figure 5.8 Inactivated viruses in Filtrate 2009 by autoclave

Removal of VLPs by 0.02 µm pore size Filtration

0.02 µm filters were used to remove VLPs to examine if *Methanosaeta concilii* receiving 0.02 µm filtrate could show recovery of methane generation (Table 5.5). Filtrate was obtained by two consecutive filtrations: 0.45 µm polyethersulfone (PES) syringe filter was first applied to remove

large particles (e.g. *Methanosaeta* cells) and 0.02 µm Anodisc filter membranes were used to remove particles larger than 20 nm (e.g. VLPs, generally ranging from 30 to 200 nm). In addition, to evaluate the possible impact of oxygen exposure during the filtration process, DI water was also filtered, using the same process (bottle 229). The results are illustrated in Figure 5.9. Methane accumulated in headspace was no different for *Methanosaeta* inoculated with Filtrate 2009 or with 0.02µm-filtered Filtrate 2009. The result of this test indicated that VLPs were not cause of the observed inhibition of methane generation and instead a chemical inhibitor that could be passed through a 0.02 µm filter was suspected.

Table 5.5 Experimental design to remove VLPs from Filtrate 2009 using 0.02 µm pore size filters

Serum bottle no.	Source of methanogens: *Mst. Concilii* (DSM 6752)	Source of VSFs	Note
226 A,B	5 mL	214 A: 3.5 mL	filtrate transfer
227 A,B	5 mL	214 B: 3.5 mL	filtrate transfer
228 A,B	5 mL	215 A, B (0.02 µm Anodisc): 3.5 mL	remove virus
229 A,B	5 mL	DI filtrate (0.02 µm Anodisc): 3.5 mL	process control
233 A,B	5 mL	DI water: 3.5 mL	negative control

DI: deionized water (anoxic)

1.5 mL acetate feed was added to all bottles for a total bottle volume of 10 mL

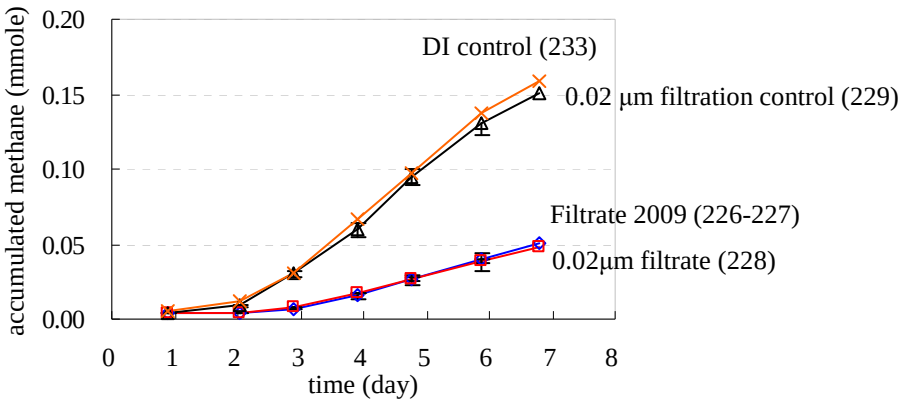


Figure 5.9 Removing VLPs from Filtrate 2009 using 0.02 µm pore size filter

Examination of Other Inhibitors - Syringe Filter and Filtration Process

The concern of eluting potential toxic substances from filter membrane was examined by using filters made from different materials. The experimental designs are shown in Tables 5.6. As shown in Figure 5.10, the amount of final methane yield was similar no matter whether Filtrate 2009 was harvested by filtering through 0.22 µm mixed cellulose esters (MCE) membrane (bottle 208) or 0.2 µm polyethersulfone (PES) membrane (bottle 207 C). This result indicated that the material of filters has no effect on the methane inhibition.

To further examine the influence of the particle removal process, centrifugation instead of filtration was used to collect VLPs. The experimental designs are shown in Tables 5.7. In Figure 5.11, centrifugation (bottle 218) was used in addition to filtration (bottle 214-215) to remove bacterial cells from previous infection tests. The final methane yield remains inhibited no matter

whether the host cells were removed by filtration (bottle 214-215) or centrifugation (bottle 218). This result implied that the process of filtration was not related to methane suppression. In addition, because concentrates may be exposed to oxygen during the process of centrifugation, DI controls (bottle 220 and 221) were used to evaluate the impact of oxygen. 37% and 27% of methane yield were suppressed (compared to bottle 225) by oxygen in bottles 220 and bottle 221, respectively. This result suggested that the inhibitory activity in bottles 218 and 219 was not only caused by oxygen exposure during the process (twice the inhibition (80%) than oxygen exposure (37%)) but also by unknown inhibitors.

Table 5.6 Experimental design to evaluate the process of filtration: materials of PES and MCE filters

Serum bottle no.	Source of methanogens: *Mst. Concilii* (DSMZ 6752)	Source of VSFs	Note
207 A	5 mL	202 A (0.45 μm PES): 3.5 mL	filtrate transfer
207 B	5 mL	202 A (0.45 μm PES): 3.5 mL	filtrate transfer
207 C	5 mL	203 A (0.2 μm PES): 3.5 mL	filter evaluation
208 A,B	5 mL	202 B (0.22 μm MCE): 3.5 mL	filter evaluation
212 A,B	5 mL	DI water: 3.5 mL	negative control

polyethersulfone (PES); mixed cellulose esters (MCE)

DI: deionized water (anoxic)

1.5 mL acetate feed was added to all bottles for a total bottle volume of 10 mL

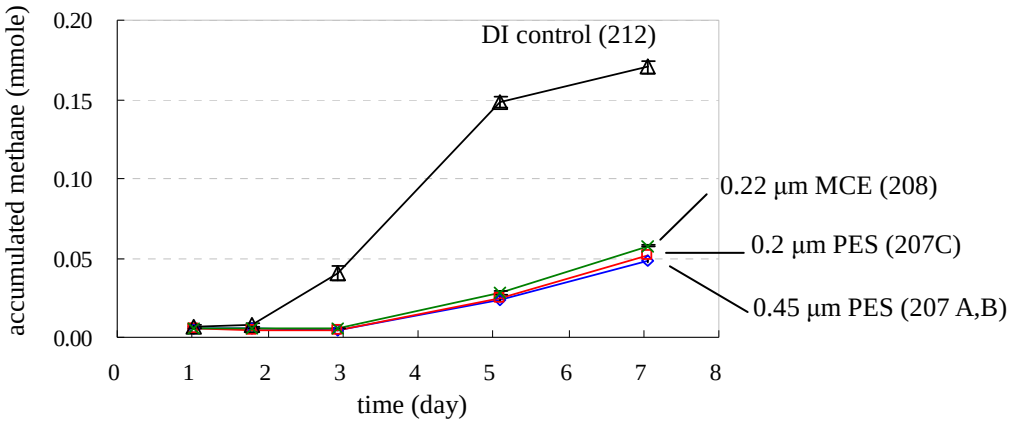


Figure 5.10 Evaluating the process of filtration: materials of PES and MCE filters

Table 5.7 Experimental design to evaluate the filtration process

Serum bottle no.	Source of methanogens: *Mst. Concilii* (DSMZ 6752)	Source of VSFs	Note
214 A,B	5 mL	208 A: 3.5 mL	filtrate transfer
215 A,B	5 mL	208 B: 3.5 mL	filtrate transfer
218	5 mL	207 B (3500 RPM 20 min): 3.5 mL	centrifuge
219	5 mL	207 B (3500 RPM 20 min, 0.45 µm PES): 3.5 mL	centrifuge
220	5 mL	DI (3500 RPM 20 min): 3.5 mL	process control
221	5 mL	DI (3500 RPM 20 min, 0.45 µm PES): 3.5 mL	process control
225 A,B	5 mL	DI water: 3.5 mL	negative control

DI: deionized water (anoxic)

1.5 mL acetate feed was added to all bottles for a total bottle volume of 10 mL

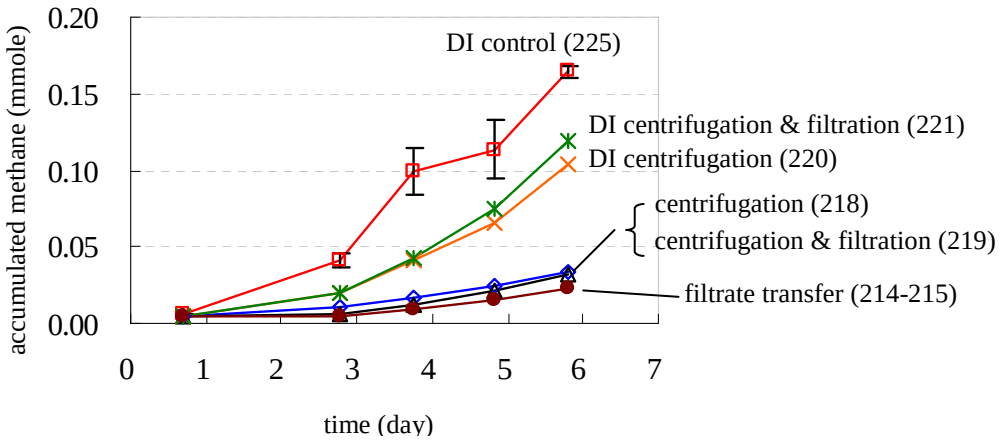


Figure 5.11 Evaluating the filtration process

Examination of Other Inhibitors - Transfer Filtrate of DI Water Control

The transfer of filtrate from earlier DI water controls (bottle 225) was also conducted (Table 5.8), and methane yield was partially inhibited (Figure 5.12). Compared to DI water control, methane yield was inhibited 42 % in bottle 231 (DI first transfer), 53 % in bottle 230 (DI 2nd transfer) and 69 % in bottle 226 and 227 (Filtrate 2009). The results suggest that an inhibitory agent was formed during the control tests with DSM 6752.

Table 5.8 Transfer the filtrate of previous DI water control

Serum bottle no.	Source of methanogens: *Mst. Concilii* (DSMZ 6752)	Source of VSFs	Note
226 A,B	5 mL	214 A: 3.5 mL	filtrate transfer
227 A,B	5 mL	214 B: 3.5 mL	filtrate transfer
230 A,B	5 mL	224 A (DI control 2nd transfer): 3.5 mL	DI control 2nd transfer
231 A,B	5 mL	225 A (DI control 1st transfer): 3.5 mL	DI control 1st transfer
233 A,B	5 mL	DI water: 3.5 mL	DI water control

DI: deionized water (anoxic)

1.5 mL acetate feed was added to all bottles for a total bottle volume of 10 mL

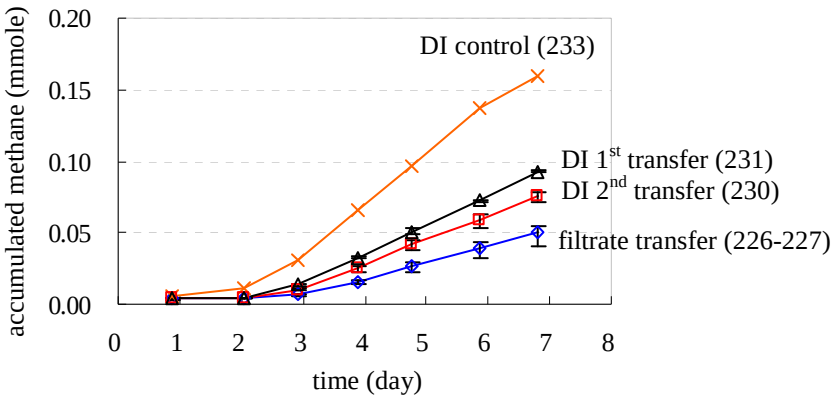


Figure 5.12 Transfer the filtrate of previous DI water control

Filtrate 2009 harvested from a sick digester at the Port Angeles WWTP was initially considered to have potential infective *Methanosaeta* viruses due to its consistent suppression of methane yield of *Methanosaeta* in enrichment and *Methanosaeta concilii* (DSM 6752) during serial passage. However, virus-caused methane suppression was not supported by the TEM observations and inactivated filtrate controls (autoclaved filtrate and 0.02 µm filtrate). Methane generation was further inhibited with the addition of autoclaved filtrate. The aforementioned results showed that the methane impairment was not related to infective viruses. Several test results showed that the inhibitor was not eluted from filters and not generated during the filtration process. Methane yield in bottles that received the filtrate of previous DI controls was partially inhibited. This implied that the inhibitor could originate from the culture of *Methanosaeta concilii*, media, or acetate stock solution since these were the only substances added into the bottle. The above tests results were summarized in Table 5.9.

Table 5.9 Summary of the investigation of the inhibitor in the infection tests

Results of infection tests	Implication of the inhibitor	Figure
Autoclaved VSFs increased inhibitory	inhibitor increase or more toxic after treatment	5.8
0.02 μm filtered VSFs remains inhibitory	soluble or particle less than 20 nm	5.9
Inhibitory VSFs harvested by both MCE or PES filter	not eluted from the material of filter	5.10
Inhibitory remained when filtration was substituted by centrifugation	not generated during the filtration process	5.11

5.3.2 Evaluation of Potential Ammonia Toxicity

pH Adjustment

pH is an important variable for ammonia toxicity (Table 2.1). A range of initial pH was tested to evaluate its influence on methane generation during infection tests. Two identical groups of bottles were prepared (Table 5.17). pH of one group was kept with the initial condition (about 8.0) and pH of the other group was adjusted to about 7.2. Methane production during the tests is illustrated in Figure 5.21. As expected, the methane yield of the bottles without pH adjustment was inhibited (Figure 5.21 a and c). The pH of the bottles seeded with autoclaved VSFs increased by 0.12 (autoclaved VSFs of the daily-fed enrichment) and 0.06 (autoclaved VSFs of the hourly-fed enrichment). When the pH of these bottles was adjusted to 7.2, methane production was partially recovered. Moreover, the methane generation in pH-adjusted DI controls did not increase when compared to unadjusted DI controls.

Table 5.10 Infection tests with and without pH adjustment

Source of methanogens: *Mst. Concilii* (DSMZ 2139)	Source of VSFs: 3.5 mL	Initial pH	Initial NH_3-N (mg/L)
5 mL	VSF-Daily	7.95	66.7
5 mL	VSF-Hourly	7.92	62.6
5 mL	VSF-Daily (autoclaved)	8.07	85.4
5 mL	VSF-Hourly (autoclaved)	7.98	71.0
5 mL	DI water	7.91	23.3
5 mL	VSF-Daily	7.2	12.8
5 mL	VSF-Hourly	7.13	10.9
5 mL	VSF-Daily (autoclaved)	7.26	14.7
5 mL	VSF-Hourly (autoclaved)	7.23	13.7
5 mL	DI water	7.27	5.7

DI water: deionized water (anoxic)

1.5 mL acetate feed was added to all bottles for a total bottle volume of 10 mL

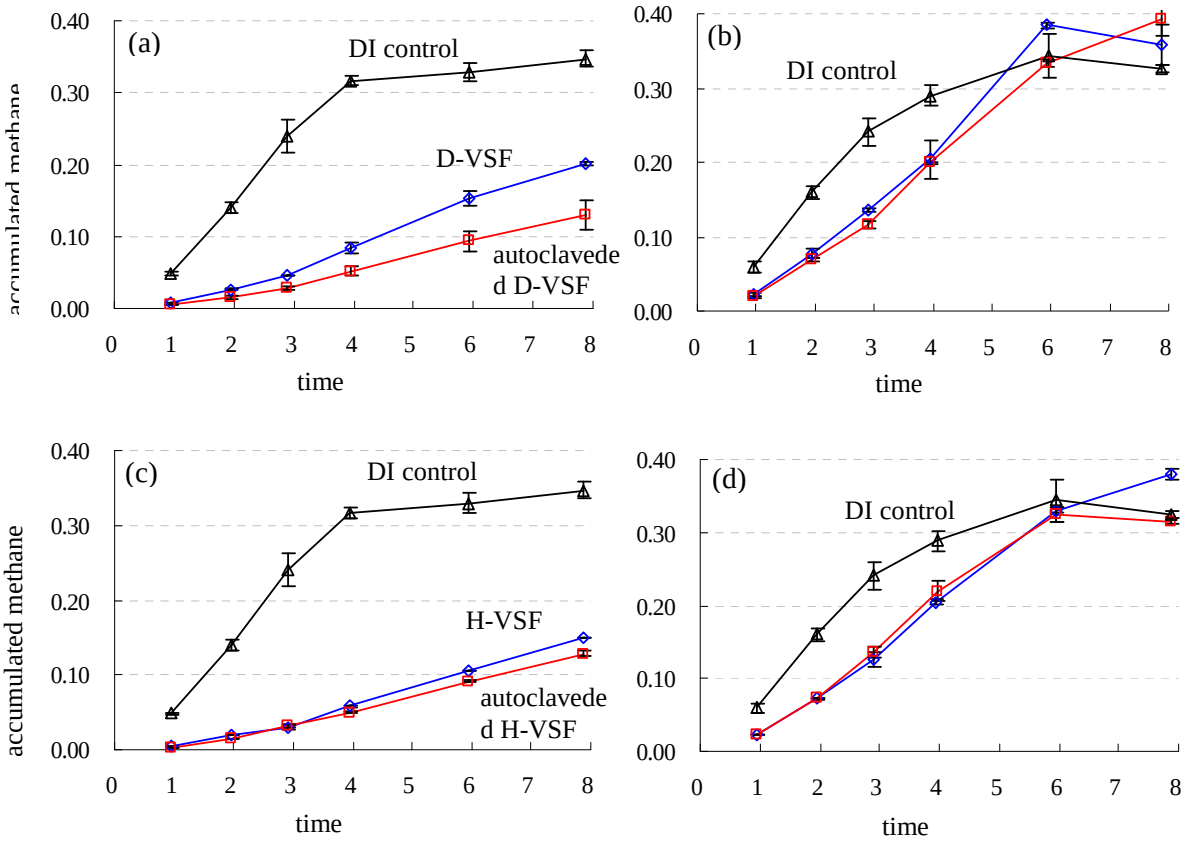


Figure 5.13 Infection tests with and without pH adjustment
Methanosaeta concilii DSM2139 receiving VSF of daily-fed enrichment without (a) and with pH adjustment (b). *Methanosaeta concilii* DSM2139 receiving VSF of hourly-fed enrichment without (c) and with pH adjustment (d)

Ammonia Toxicity Tests

Free ammonia (NH_3) has been demonstrated to be toxic to *Methanosaeta concilii*, and the speciation of ammonia nitrogen (NH_3/NH_4^+) is influenced by pH (see Sections 2.4). Therefore, the correlation between the levels of methane inhibition and the concentration of free ammonia was evaluated. Initial free ammonia concentrations ranging from 12.8 to 86.4 mg/L (ammonia-N: 10.5 mg/L to 71.2 mg/L) were tested (Figure 5.14). On day 4.8, no suppression of methane

generation was observed when NH_3-N concentration were below 16.8 mg/L (10.5 and 16.8 mg/L). Methane yield was reduced to 84% at a concentration of 28.7 mg/L ammonia-N. Methane production declined to 60% and 49% with ammonia-N at 47.9 mg/L and 71.2 mg/L, respectively. A past study has reported that 55.2 mg/L of NH_3-N (120 mM NH_4Cl at 35^oC, pH 7 for 21 days) suppressed methane yield of *Methanosaeta concilii* to 80% (Sprott and Patel 1986) (Figure 5.15). The inhibition reported by Sprott and Patel was more pronounced than the finding in this study (49% methane inhibited at 70.8 mg/L of NH_3-N). The difference can be attributed to the difference in culture media (total ammonia, sodium concentration etc.), pH (7.9 in this test), and the tested strains (DSM 6752 in this study).

Free ammonia concentration and ammonia-nitrogen were calculated for DSM 334 media for pH ranging from 7.0 to 8.5 (Table 5.18). In addition, the experimental observed percentage of suppression of methanogenesis and corresponding ammonia-nitrogen is also summarized. When pH of the media was below 7.2, the influence of ammonia-nitrogen on the methane production was negligible. When pH of the media increases above 7.5, notable inhibition occurs. Combined with the finding of optimal growth pH at 7.62-7.71 (Steinhaus, Garcia et al. 2007), a suggested pH range for *Methanosaeta concilii* DSM 6752 should be kept between 7.4 to 7.6. Additionally, NH_4Cl concentration in the media can be reduced 50% (if reducing nitrogen source of the media

does not has negative effect on methanogenesis) to culture *Methanosaeta concilii* DSM 6752 at its optimal growth pH 7.6 with negligible ammonia toxicity.

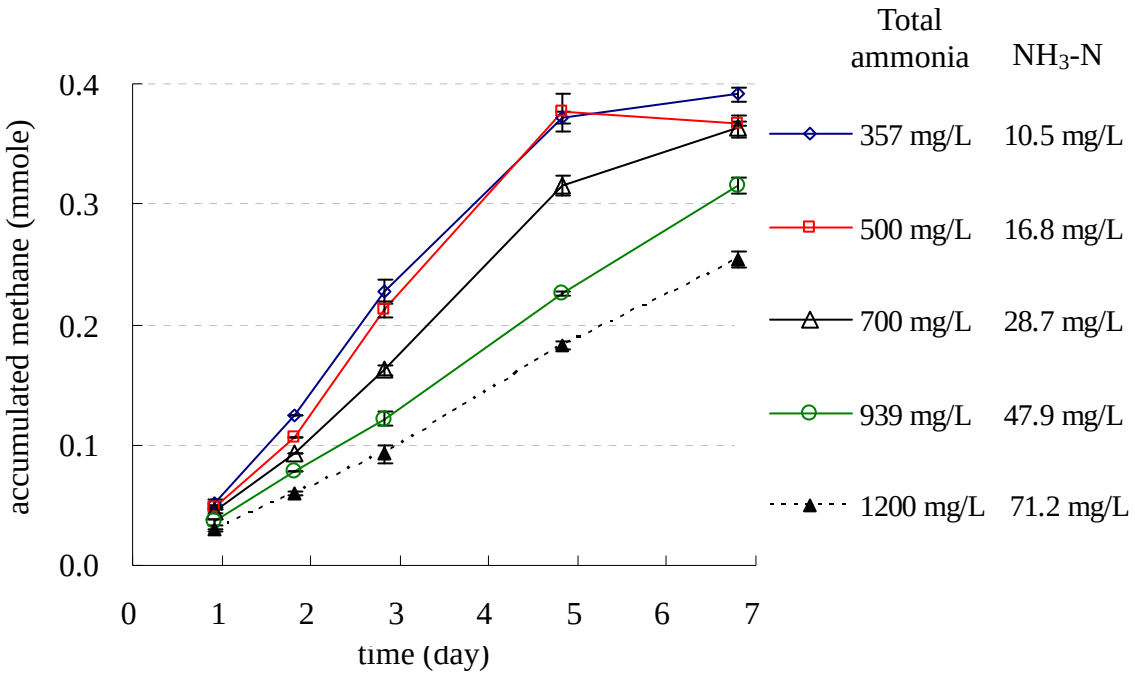


Figure 5.14 Influence of ammonia toxicity on methane yield of *Methanosaeta concilii* DSM 6752

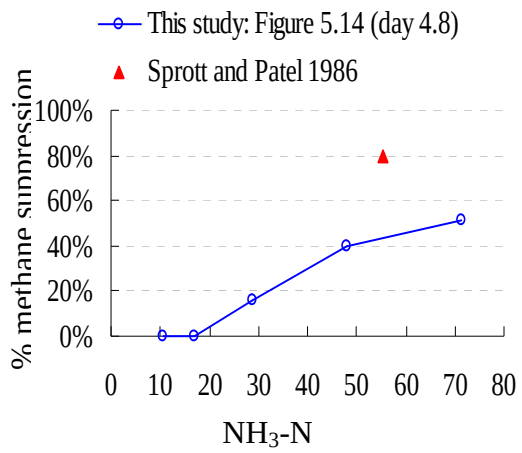


Figure 5.15 Correlation between ammonia nitrogen concentrations (mg/L) and methane suppressions (%)

Table 5.11 Ammonia concentration in DSM 334 media

pH	[NH_3] (mM)	NH_3-N (mg/L)	Ammonia toxicity test in this study	
			NH_3-N (mg/L)	Suppression of methanogenesis (%)
7.0	0.21	9.1		
7.1	0.26	11.5	10.5	not observed
7.2	0.33	14.4	16.8	not observed
7.3	0.41	18.0		
7.4	0.51	22.6		
7.5	0.64	28.2	28.7	15.6%
7.6	0.80	35.2		
7.7	1.00	43.9	47.9	39.6%
7.8	1.24	54.5		
7.9	1.53	67.4	71.2	51.3%
8.0	1.89	83.1		
8.1	2.31	102.0		
8.2	2.82	124.4		
8.3	3.42	150.7		
8.4	4.11	181.1		
8.5	4.90	215.7		

Ammonia Toxicity in the Infection Tests

In infection tests, ammonia originated from three sources: (1) media which carried the host, *Methanosaeta*, (2) VSFs and (3) acetate stock solution (ammonia acetate: 0.07 M) (Table 5.12). The experiment in this section was modified by removing ammonia from acetate stock solution and VSFs in order to see if the infection tests could be improved. To remove ammonia from acetate stock solution, acetic acid was used instead of the mixture of ammonia acetate and sodium acetate (bottle 4-6). To further reduce ammonia in VSFs, viruses in VSFs were concentrated by centrifugal ultrafiltration (Amicon Ultra-15), which reduced the volume and total ammonia by about 93% (from 3.5 mL to 250μL) (bottle 7-8). VSFs used were from anaerobic digesters of West Point WWTP (WP) and South Plant (SP). The pH of the individual bottles were measured, and corresponding free ammonia concentration was calculated (Table 5.12). As shown in Figure 5.16, without reducing ammonia in infection tests (bottle 1-3), methane generation of *Methanosaeta concilii* (DSM 6752) received VSFs from WP and SP was inhibited 90% and 99% (corresponding NH_3 concentrations 131.6 and 153.3 mg/L, respectively) at day 5.8 (compared to control, bottle 6). When acetic acid was utilized as the substrate, the corresponding NH_3 concentrations were decreased to 34.4 and 41.9 mg/L, respectively (bottle 4 and 5). The inhibitions were reduced to 51% (38% recovery) and 67% (32% recovery), respectively. When ammonia concentrations in VSFs were lower by concentrating the VLPs

(bottles 7 and 8), the corresponding NH_3 concentrations were reduced to 3.8 and 3.9 mg/L, respectively, and the inhibitions were not observed. The influence of pH (7.27-8.18) and ion concentration of media (e.g. VSF in bottle 1 and DI in bottle 3) on the growth of *Methanosaeta concilii* should be evaluated in detail to distinguish their impact on methanogenesis from that of ammonia toxicity. However, the current result still clearly indicated that the impairment of methanogenesis in infection tests was caused by ammonia. VLPs in VSFs did not suppress methane yield of *Methanosaeta concilii*, suggesting the infective *Methanosaeta concilii* virus was not present.

Table 5.12 Study of ammonia toxicity to *Methanosaeta concilii* DSM 6752 in the infection tests

Bottle	*Mst. concilii* DSM 6752	VSF/DI 3.5 mL	0.14 M acetate 1.5 mL	Initial pH	NH_3-N (mg/L)
1	5 mL	WP-VSF	NaAc/NH_4Ac	8.18	131.6
2	5 mL	SP-VSF	NaAc/NH_4Ac	8.26	153.3
3	5 mL	DI	NaAc/NH_4Ac	7.96	28.6
4	5 mL	WP-VSF	HAc	7.63	34.4
5	5 mL	SP-VSF	HAc	7.72	41.9
6	5 mL	DI	HAc	7.27	3.0
7	5 mL	WP-VSF: 230 µL + DI	HAc	7.27	3.8
8	5 mL	SP-VSF: 250 µL + DI	HAc	7.27	3.9

VSFs used were from anaerobic digesters of West Point WWTP (WP) and South Plant (SP)
NaAc: sodium acetate; NH_4Ac: ammonium acetate; HAc: acetic acid
DI water: deionized water (anoxic)

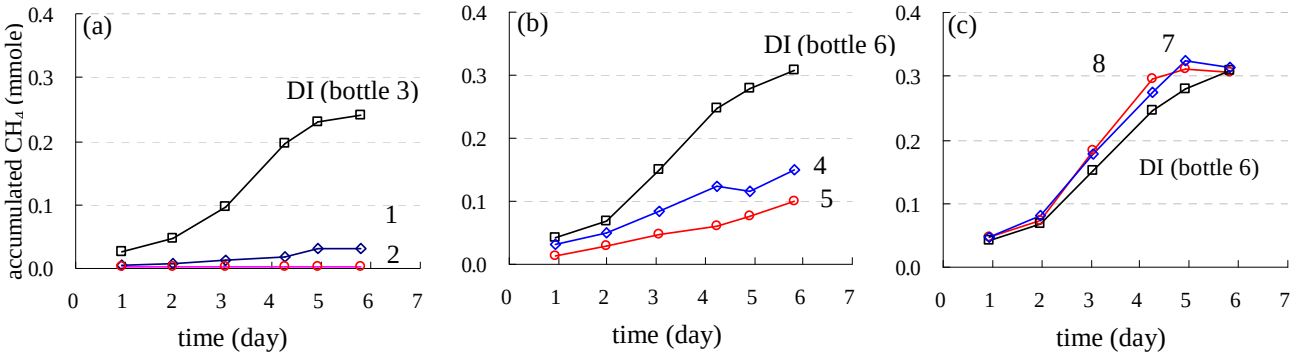


Figure 5.16 Study of ammonia toxicity to *Methanosaeta concilii* DSM 6752 in the infection tests (a) original recipe of the infection tests, (b) removing ammonia in acetate solution using acetic acid, (c) reducing concentration of ammonia using acetic acid and Amicon Ultra-15

Improving Infection tests

From the results of previous experiments, it is clear that controlling ammonia (especially free ammonia) concentrations is critical in infection tests. One approach is to reduce the total ammonia/ammonium in the systems. Therefore, in infection tests, the acetate stock solution which combined ammonia acetate and sodium acetate was substituted with DSM 334 media (Table 4.1) (less ammonium in DSM 334 media: ammonium concentrations are1 g/L and 5.4 g/L in the DSM 334 media and the original acetate solution). In addition, VLPs in VSFs from either enriched reactors or anaerobic digesters were concentrated by centrifugal ultrafiltration. Another approach is to control pH during the tests such as with carbonate buffering. Both these modifications were used to screen for *Methanosaeta* viruses in VSFs from the reactors and digesters. Carbonate buffering was tested by evaluating the influence of different initial

headspace CO_2 concentrations (Table 5.13 and Figure 5.17). Compared to one another, no difference was found when 10% or 15% of CO_2 was present into the bottles. Minor suppression was observed in bottles with CO_2 concentration above 20%. Therefore, 10% CO_2 was selected for subsequent infection tests. In addition, 20 mL of fresh media was spiked in bottles when all acetate was consumed, and the feasibility of adding extra media was examined. The result showed no negative effect of adding fresh media. The addition of more fresh media could support more growth cycles for *Methanosaeta*, and therefore could produce more infective viruses (if they are present in VSFs) to make observable impact on methane yield.

Table 5.13 Effect of initial CO_2 headspace concentrations on methane generation of *Methanosaeta concilii*

Bottle	DSM 2139	DSM 334	CO2	Initial pH	End pH	Initial NH_3-N (mg/L)	End NH_3-N (mg/L)
9	4 ml	16 ml	10%	7.56	7.68	32.2	42.0
10	4 ml	16 ml	15%	7.35	7.53	20.2	30.2
11	4 ml	16 ml	20%	7.22	7.38	15.1	21.6
12	4 ml	16 ml	25%	7.04	7.24	10.0	15.8

the tests were conducted in 160 mL serum bottles

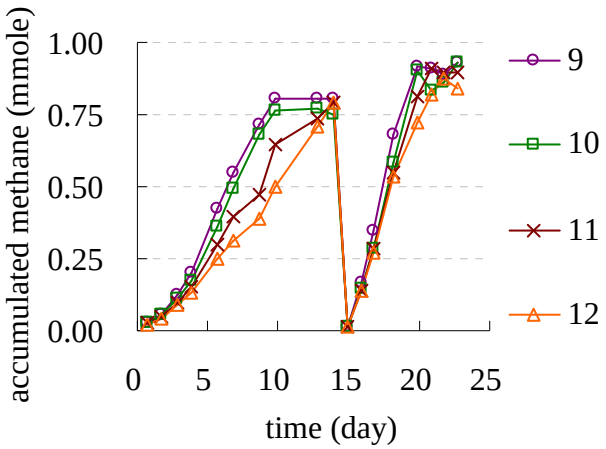


Figure 5.17 Effect of different initial CO_2 headspace concentrations on methane generation of *Methanosaeta concilii*

5.3.3 Assessment of Improved Infection Tests for Screening *Methanosaeta* viruses

The improved infection test was used to screen for *Methanosaeta* viruses in VSFs from enrichments and anaerobic digesters as described in Table 5.14. The experiments were conducted in 240-mL serum bottles. Both strains of *Methanosaeta concilii* were used as tested hosts. Initially, the inocula (hosts) were 20% (4mL in 20 mL, bottles 8 and 9) but were reduced to 10% (bottle 13-17) to increase virus-to-host ratio. Also, the input of VSFs was increased from 3.5 mL to a concentrate harvested from 90 mL (increasing about 26 times by centrifugal ultrafiltration). Final volume of VSFs varied from 150 to 225 μL. 10% CO_2 was added in the headspace of the serum bottles. The results of the infection tests are shown in Figure 5.18. No obvious suppression

of methane yield was found in bottles receiving VSFs (Figure 5.18 a, b and c). Intact VLPs were observed by TEM (Figure 5.18 d)indicating the ultrafiltration process was experimentally feasible.

Table 5.14 Challenging *Mst. concilii* with VSFs of enrichments using improved infection tests

Bottle	*Mst. concilii*	Media	VSF	CO2	Note
8 A,B	DSM 2139 4 mL	16 ml	H-fed 90 mL (225 µL)	10%	H-fed
9 A,B	DSM 2139 4 mL	16 ml	- -	10%	control
13 A,B	DSM 2139 1.65 mL	15 ml	D-fed 90 mL (150µL)	10%	D-fed
14 A,B	DSM 2139 1.65 mL	15 ml	- -	10%	control
15 A,B	DSM 6752 1.65 mL	15 ml	D-fed 90 mL (165µL)	10%	D-fed
16 A,B	DSM 6752 1.65 mL	15 ml	H-fed 90 ml (150µL)	10%	H-fed
17 A,B	DSM 6752 1.65 mL	15 ml	- -	10%	control

Experiments were conducted in 240 mL serum bottles (~220 mL headspace)

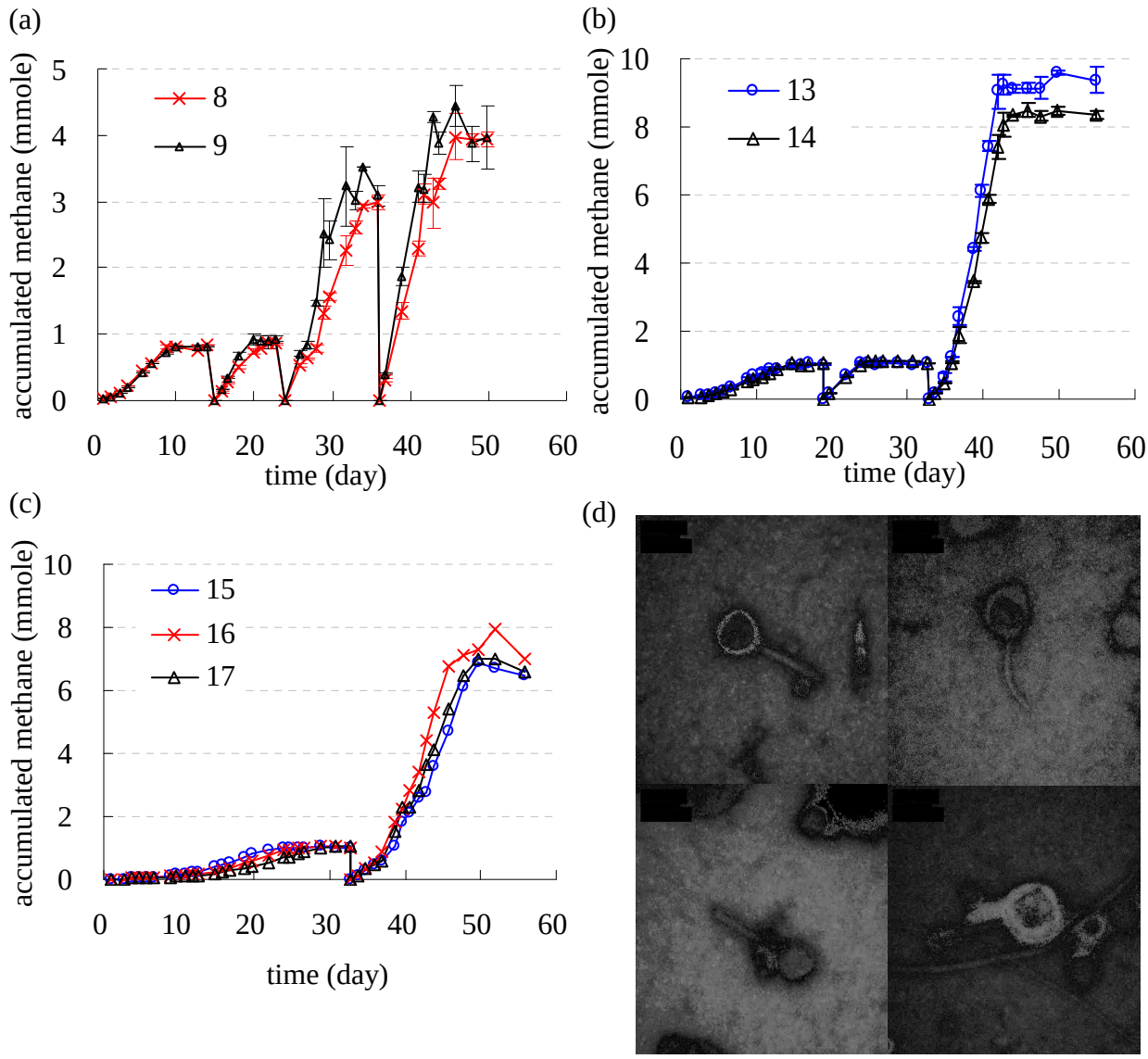


Figure 5.18 Challenging *Mst. concilii* with VSFs of enrichments using improved infection tests *Mst. concilii* DSM 2139 challenged by VSFs from the hourly-fed enrichment (a) and daily-fed enrichment (b). *Methanosaeta concilii* DSM 6752 challenged by VSFs from both enrichments (c). TEM of VLPs in VSFs after centrifugal ultrafiltration (d). Scale bar is 80 nm.

The experimental design of a typical infection tests using VSFs from anaerobic digesters is illustrated in Table 5.15. During the cultivation (36 days), no difference in methane production was found between cultures seeded with VSFs (bottle 18 and 20) (Figure 5.19) and the

corresponding controls (bottles 19 and 21). Table 5.16 summarized the VSFs that have been used (results not shown) in infection tests but none of the VSFs have caused the impairment of methanogenesis of *Methanosaeta*.

Table 5.15 *Mst. concilii* challenged with VSFs of an anaerobic digester using improved infection tests

Bottle	*Mst. concilii*	DSM 334	VSF	CO2	Note
18 A,B	DSM 2139 1.65 mL	15 mL	WP 60 mL (150 uL)	10%	West Point WWTP
19	DSM 2139 1.65 mL	15 mL	- -	10%	control
20 A,B	DSM 6752 1.65 mL	15 mL	WP 60 mL (150 uL)	10%	West Point WWTP
21	DSM 6752 1.65 mL	15 mL	- -	10%	control

VSFs collected from a digester at West Point (WP) WWTP
100 mL of media was spiked at day 22

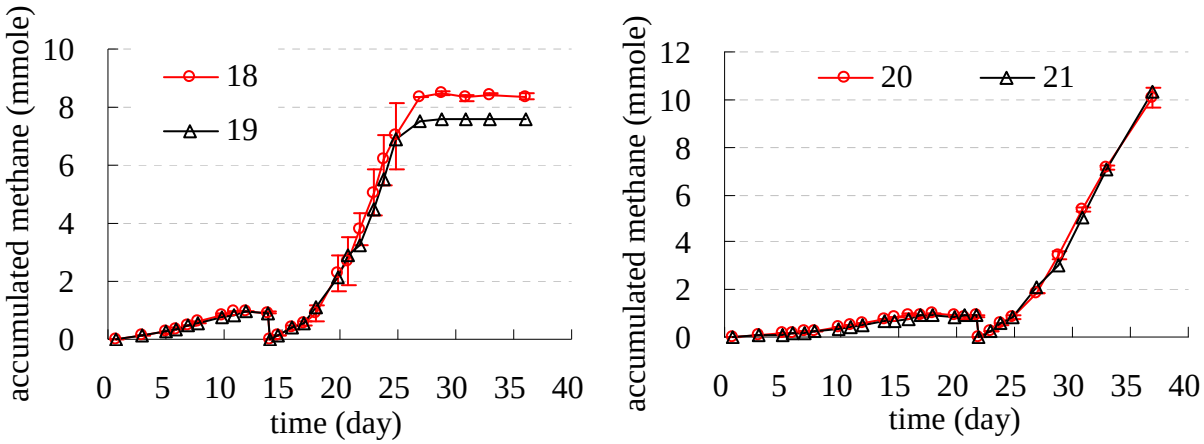


Figure 5.19 *Mst. concilii* challenged with VSFs of an anaerobic digester using improved infection tests

Table 5.16 Summary of the VSFs that have been used in infection tests

VSFs	Date	Note
D-fed enrichment	November 2011/ April 2012	healthy reactor
H-fed enrichment	November 2011/ April 2012	healthy reactor
West Point digester	November 2011/ April 2012	healthy digester
South Plant digester	April 2012	healthy digester
Port Angeles digester	May 2012	recovered from upset after 3 month
Brightwater digester	July 2012	upset digester

Tested hosts were *Methanosaeta concilii* DSM 6752 and 2139

Since the ammonia concentration has been reduced below toxic level in the infection tests, *Methanosaeta concilii* has not been suppressed by the VSFs collected from enriched methanogenic reactors or digesters of WWTPs. However, it remained unclear if a virus of *Methanosaeta concilii* might be present and reproduced by lysogenic pathway. It has been shown that the virus-to-bacteria ratio is increased in nutrient-rich, more productive environments (Wommack and Colwell 2000), which suggests that the viruses of fast growing hosts might prefer a lytic life cycle. Under anaerobic conditions, limited energy is gained by slow-growing *Methanosaeta* when acetate is converted to methane. Therefore, the lysogenic pathway might be favored by the *Methanosaeta* virus.

Test of Inducing Agents and Induction Experiments

In order to perform induction experiments, inducing agents including temperature (heat shock),

mitomycin C, and UV were tested. In addition, two short-chain volatile fatty acids (intermediate products of digestion), propionic and butyric acid, were also examined (Table 5.17). The effect of these inducing agents on the growth of *Methanosaeta concilii* is shown in Figure 5.20. 60^{o}C heat shock for 20 min dramatically inactivated the cells, while no negative effect was found using 50^{o}C and 55^{o}C. For the mitomycin C, concentration at 15 μg/mL caused 80% inhibition, and partial suppression of methane generation was found in bottle containing 5 μg/mL of mitomycin C. For the UV treatment, 80% to 90% reduction of methane yield was observed when cells were exposed to UV at least 10s. Partial suppression of methanogenesis was found in bottles that received propionic and butyric acid during the cultivation, but the level of inhibition was not correlated to the tested concentration range. The purpose of this test was to select suitable concentration or intensity of inducing agents (i.e. cause partial impairment but avoid severe damage to host cells). Therefore, heat shock 55^{o}C for 20 min, 5 μg/mL mitomycin C, UV exposure for 10 s, and 30 mM propionate and 12 mM butyrate were selected for further study.

Table 5.17 Test of inducing agents

Bottle	Inducing agent	Note
T 50	temperature	50°C, 20 min
T 55	temperature	55°C, 20 min
T 60	temperature	60°C, 20 min
C 1	mitomycin C	1 μg/mL
C 5	mitomycin C	5 μg/mL
C 15	mitomycin C	15 μg/mL
C 0	deoxygenated DI	process control
UV 10	UV	10s
UV 30	UV	30s
UV 40	UV	40 s
UV 0	--	40 s (UV off, process control)
VFA 1	VFA	propionate: 7.5 mM; butyrate 3 mM
VFA 2	VFA	propionate: 15 mM; butyrate 6 mM
VFA 3	VFA	propionate: 30 mM; butyrate 12 mM
VFA 0	Deoxygen DI	process control
control	--	no inducing agent

2 mL of *Methanosaeta concilii* DSM 6752 was added to 8 mL of DSM334 media in 60 mL serum bottles

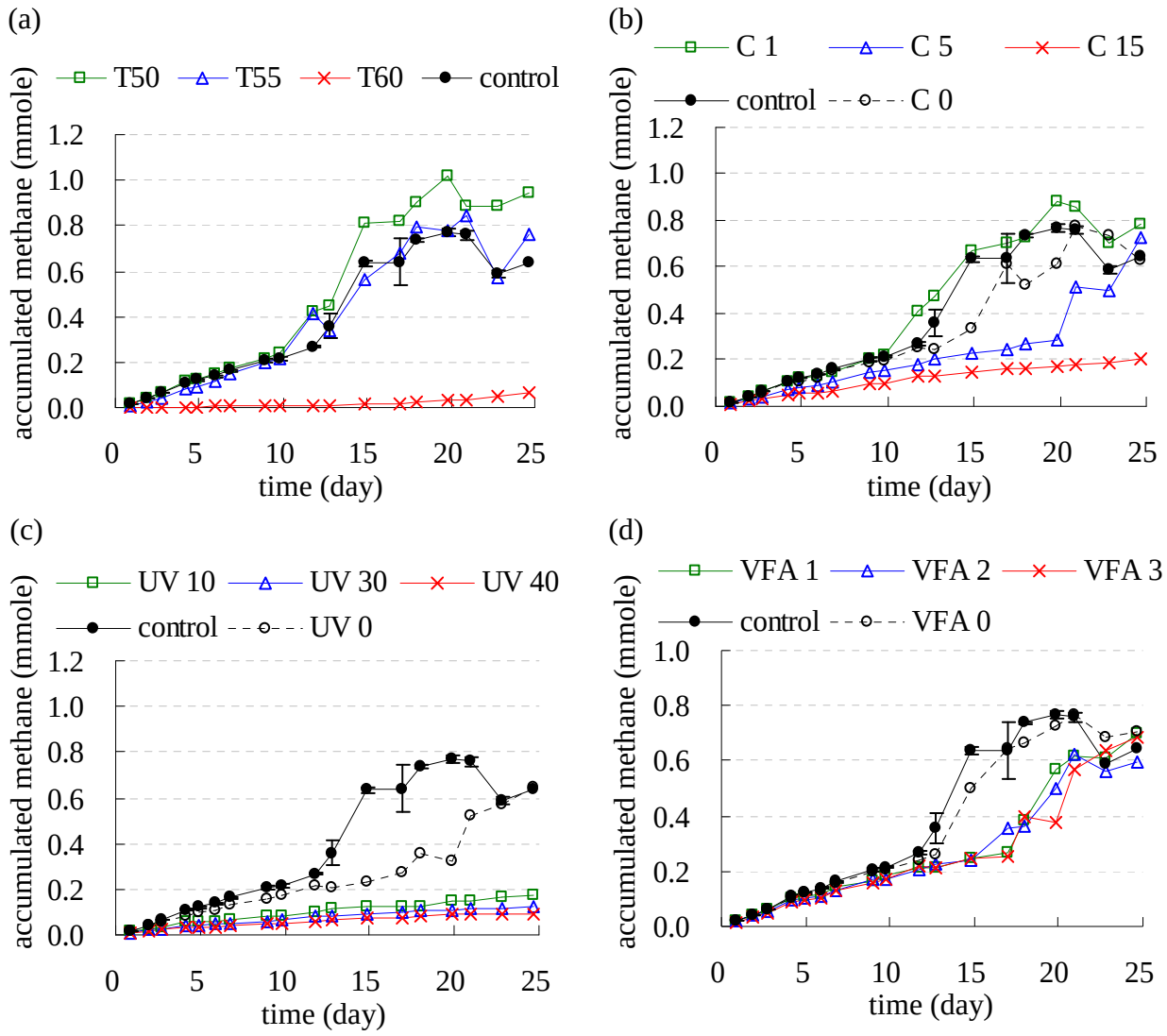


Figure 5.20 Effect of inducing agents on the growth of *Methanosaeta concilii* DSM 6752
2 mL of *Methanosaeta concilii* DSM 6752 was inoculated in 8 mL of media for the tests. (a) heat shock at 50, 55 and 60°C for 20 min. (b) concentration of mitomycin C: 1, 5 and 15 µg/mL. (c) UV exposure for 10, 30 and 40 s. (d) concentration of propionic acid/butyric acid: 7.5/3, 15/6 and 30/12 mM.

The induction tests were performed on *Methanosaeta concilii* DSM 6752 that had been seeded with VSFs from the daily-fed (Figure 5.21 a) and hourly-fed (Figure 5.21 c) enrichments (1st transfer). 4 mL *Methanosaeta concilii* was added in 16 mL media (20% inocula). Host cells were

incubated until early log phase prior to introduction of inducing agents (day 7). Samples were taken before (day 7, before induction) and after induction (day 14) for EFM to measure VLPs concentrations (Figure 5.21 b and d). Methane yield was partially inhibited in bottles treated by UV. For bottles receiving the rest of inducing agents, the suppression of *Methanosaeta concilii* was less (Figure 5.21) than when testing of inducing agents (Figure 5.20) possibly due to the higher concentration of more active *Methanosaeta concilii* in the induction experiment. Comparing concentration of VLPs before and after induction, no obvious increasing was observed (Figure 5.28 b and d). The results imply the decline of methane generation might be directly caused by inducing agents but not the induction of viruses from *Methanosaeta*.

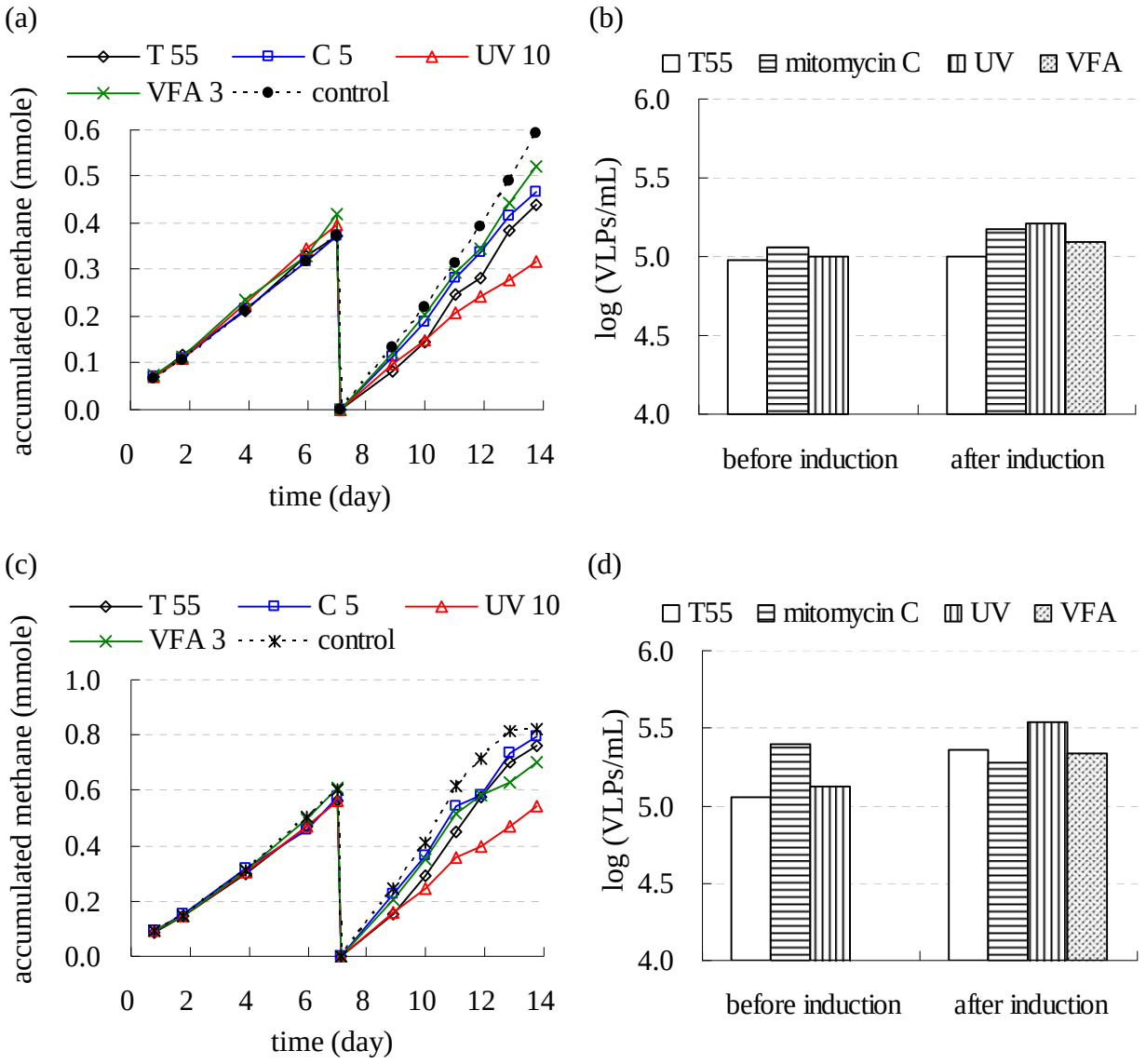


Figure 5.21 Induction experiments of *Methanosaeta concilii* DSM 6752 with enrichment VSFs (a) induction test using VSFs of daily-fed enrichment. (b) concentration of VLPs before (day 7) and after (day 14) induction for test (a). (c) induction test using VSFs of hourly-fed enrichment. (d) concentration of VLPs before (day 7) and after (day 14) induction for test (c).

5.4 Conclusions

Free ammonia concentrations above 16.8 mg/L are toxic to *Methanosaeta concilii* DSM 6752 and the ammonia toxicity can be avoided by reducing TAN concentrations or/and using carbonate buffering to control pH of the media.

For *Methanosaeta concilii* DSM 6752, there was no inhibition of the methane production when the concentration of ammonia-nitrogen was below 16.8 mg/L. The suppression of methane yield (day 4.8) was 15.6%, 39.6%, and 51.3% at free ammonia-nitrogen of 28.7 mg/L, 47.9 mg/L, and 71.2 mg/L, respectively. A past study has reported that 55.2 mg/L of NH_3-N (120 mM NH_4Cl at 35^{o}C, pH 7 for 21 days) suppressed methane yield of *Methanosaeta concilii* to 80% (Sprott and Patel 1986). Difference in level of inhibition was possibly due to the use of different media, strains and anaerobic technique. Based on the calculation of free ammonia-nitrogen and reported optimal growth pH in DSM 334 media (TAN: 1000 mg/L) (Steinhaus, Garcia et al. 2007), the suggested pH for cultivation of *Methanosaeta concilii* DSM 6752 should kept between 7.4 and 7.6 to maintain free ammonia-nitrogen below 35.2 mg/L (minor suppression of methanogenesis). Additionally, NH_4Cl concentration in the media can be reduced 50% (if reducing nitrogen source of the media does not have negative effect on methanogenesis) to culture *Methanosaeta concilii*

DSM 6752 at its optimal growth pH 7.6 with negligible ammonia toxicity (NH_3-N: 17.6 mg/L).

The virus challenge experiment performed in liquid media (the infection test) was improved. The concentrations of free ammonia in infection tests were reduced by using DSM 334 media, concentrated VSFs and 10% carbon dioxide as headspace gas (to reduce the fraction of toxic ammonia-nitrogen). The virus-to-host ratio was increased by using concentrated VSFs and less host inocula. In addition, the addition of more fresh media was shown to support more growth cycles for *Methanosaeta*, and therefore could produce more infective viruses (if they are present in VSFs) to make an observable impact on methane yield. VSFs collected from the daily-fed enrichment, the hourly-fed enrichment, and upset full-scale digesters were used in improved infection tests but none of the VSFs have caused the impairment of methanogenesis of *Methanosaeta*.

Chapter 6 Characterizing microbial and viral communities in enrichment reactors

6.1 Introduction

In this chapter, persistent virus populations that inhabit laboratory methanogenic reactors receiving a sterile feed with acetate were characterized (description of reactors, see section 3.1). Virus morphologies were visualized using transmission electron microscopy (TEM). Sizes of VLPs were measured to show particle distribution patterns. Virus concentrations were enumerated using epifluorescence microscopy (EFM). Pulsed field gel electrophoresis (PFGE) was used to reveal genome sizes of viruses. In addition, concentrations of two acetoclastic methanogen genera and bacteria were determined using quantitative PCR (qPCR). The prokaryotic community structures of two reactors were investigated using 16S rDNA clone libraries.

The scope of this chapter is (1) to investigate the occurrence of viruses in two *Methanosaeta*-dominated reactors, (2) to examine the morphology, size distribution and genome length of observed viruses, and (3) to determine the concentration of viruses in relation to

acetoclastic methanogens and bacteria in two acetate-fed methanogenic reactors.

6.2 Materials and Methods

6.2.1 Autofluorescence Microscopy

The reactor effluents were examined by laser scanning confocal microscopy (Leica SP5 II, Leica Microsystems GmbH, Wetzlar, Germany) under the following conditions: excitation (405 nm laser) and emission (430-500 nm), which are used to visualize methanogens using their natural autofluorescent characteristics (Doddema and Vogels 1978).

6.2.2 Pulsed Field Gel Electrophoresis (PFGE)

Cellular materials in 1 L of sample from each reactor were initially separated from viruses by centrifuging at 3696×g and 4^{o}C for 30 min in a Sorvall Legend RT tabletop centrifuge equipped with a swinging-bucket rotor (Cat. No. 75006441) (Thermo Scientific, Waltham, MA). Supernatant was collected and serially filtered through 0.45 and 0.20 µm pore size low protein binding polyethersulfone (PES) syringe filters (6780-2502 and 6780-2504, Whatman, Clifton, NJ).

Viruses in the filtrates were concentrated using a series of centrifugal ultrafiltration units (Centricon UFC703008, Amicon Ultra-15 UFC903024 and Amicon Ultra-0.5 UFC503008; Millipore, Billerica, MA) following the manufacturer's instructions. In addition, ultrafilter retentates were washed twice in SM buffer (100 mM NaCl, 8 mM $MgSO_4 \cdot 7H_2O$, 50 mM Tris-Cl; without gelatin) and the final volume adjusted to 50 µl. The concentrate was treated with 1U of DNase I (AM2222, Life Technologies, Grand Island, NY) at 37°C for 1 h to degrade free DNA followed by 65°C for 15 min and the addition of 10 mM of EDTA (final concentration) to inactivate the DNase I.

Equal volumes of concentrate and 2% low-melting agarose (A9414, Sigma-Aldrich, St. Louis, MO) were mixed by gentle pipetting and transferred to plug molds. After solidification, plugs were incubated overnight in extraction buffer (100 mM EDTA, pH 8.0, 1% SDS and 1 mg/ml proteinase K (4333793, Life Technologies, Grand Island, NY)) at room temperature with gentle mixing on a rocker. The plugs were rinsed three times with 1X TE buffer for 30 min and stored at 4°C. Plugs and the DNA marker (N3551S, New England Biolabs, Ipswich, MA) were placed into wells of a 1% agarose gel (162-0137, BioRad, Hercules, CA) prepared in 0.5X TBE (wells were sealed with 1% agarose). PFGE was performed with CHEF-DR II System (BioRad, Hercules, CA) using the following conditions: 0.5X TBE, 6 V/cm, 15°C for 24 hours, switch times ramped from 1-25

seconds. The gel was stained with SYBR Green I (Molecular Probes, Eugene, OR) for 30 min and the image was acquired by a UVP GDS-8000 System (UVP, Upland, CA).

6.2.3 DNA Extraction

Genomic DNA from the reactors was isolated by UltraClean Soil DNA Isolation Kit (Mo Bio Laboratories, Carlsbad, CA). DNA extraction was performed following the manufacturers' protocol except the samples were disrupted using a FastPrep®-24 bead beater (MP Biomedicals, Solon, OH) at setting 5 m/s for 20 s.

6.2.4 Quantitative PCR (qPCR)

Quantification of 16S rRNA genes of *Methanosaeta*, *Methanosarcina* and Bacteria were performed using an Eppendorf Mastercycler® ep realplex and RealMasterMix kit (Eppendorf, Hauppauge, NY). Standards were prepared using PCR products of 16S rRNA genes of *Methanosaeta concilii* (DSM 6752), *Methanosarcina barkeri* (DSM 804) and *Sphingopyxis* TrD1 (GenBank® accession number: JN940802), which were cloned using TOPO TA Cloning Kit

(Invitrogen Molecular Probes, Grand Island, NY). Plasmids were isolated by QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA) and quantified using a NanoDrop 1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE). Plasmids were linearized by restriction enzyme EcoRI (R6011, Promega Co., Madison, WI). Reaction mixes contained 2 µl DNA templates, 4.5 µl RealMasterMix and individual forward and reverse primers adjusted with water to 10 µl. Primers used for qPCR are summarized in Table 6.1. For qPCR of *Methanosaeta*, thermal cycling conditions were as follows: initial denaturation at 95°C for 10 min, 50 cycles with denaturation for 10 s at 95°C, annealing for 20 s at 61.8°C, extension for 20 s at 68°C. The resulting standard curve for *Methanosaeta* had an efficiency of 98% (R^2=0.9826) with a method detection limit of $1x10^4$ 16S rRNA gene copies per PCR reaction. For qPCR of *Methanosarcina*, thermal cycling conditions were: initial denaturation at 95°C for 10 min, 50 cycles with denaturation for 10 s at 95°C, annealing for 20 s at 64°C, extension for 20 s at 68°C. The resulting standard curve for *Methanosarcina* had an efficiency of 94% (R^2=0.9982) with a method detection limit of $1x10^1$ 16SrRNA gene copies per PCR reaction. qPCR cycle conditions for bacteria were: initial denaturation at 95°C for 10 min, 50 cycles with denaturation for 15 s at 95°C, annealing for 30 s at 55.4°C, extension for 20 s at 68°C. Melting curves generated at the end of the qPCR reactions were routinely examined to verify the correct amplification The resulting standard curve for Bacteria had an efficiency of 84% (R^2=0.9826) with a method detectionlimit of $3x10^3$.

Table 6.1 qPCR primers for *Methanosaeta*, *Methanosarcina* and Bacteria

Target Organisms	Primer	Sequence	Reference
Primers used for qPCR			
Methanosaeta	MS1b	5'-CCGGCCGGATAAGTCTCTTGA-3'	(1)
	SAE835R	5'-GACAACGGTCGCACCGTGGCC-3'	(1)
Methanosarcina	MB1b	5'-CGGTTTGGTCAGTCCTCCGG-3'	(1)
	SAR835R	5'-AGACACGGTCGCGCCATGCCT-3'	(1)
Bacteria	1114F	5'-CGGCAACGAGCGCAACCC-3'	(2)
	1275R	5'-CCATTGTAGCACGTGTGTAGCC-3'	(2)
Primers used for cloning reactions			
Bacteria	8F	5'-AGAGTTTGATCCTGGCTCAG-3'	(3)
Archaea	21F	5'-TACCTTGTTACGACTT-3'	(3)
Universal	1492R	5'-TTCCGGTTGATCCYGCCGGA-3'	(3)

Reference of primers: (1) (Shigematsu, Tang et al. 2003), (2) (Denman and McSweeney 2006), (3) (Stackebrandt and Goodfellow 1991; DeLong 1992)

6.2.5 Clone Libraries and Phylogenetic Analysis

16S rRNA genes were amplified with forward primers specific for *Bacteria*,) and *Archaea* combined with a Universal reverse primer, which are listed in Table 6.1. Clone libraries were constructed by using TOPO TA Cloning Kit (Invitrogen Molecular Probes, Grand Island, NY). 16S rRNA genes in clones were sequenced at the High Throughput Genomics Center at University of Washington from one end (*Bacteria*: 8F, *Archaea*: 21F or 1492R). Sequences were trimmed and assembled using Sequencher® (version 4.9, Gene Codes Corporation, Ann Arbor, MI), and sequences having similarity >99% were defined as an OTU (Operational Taxonomic Unit). Taxonomic affiliation of OTU was determined by RDP Classifier (Wang, Garrity et al. 2007). All

sequences were deposited in the GenBank database under accession numbers KC961776 - KC961942.

6.2.6 Size Measurement of VLPs

The head diameter and tail length of VLPs were measured by ImageJ (Version 1.43u). Micrographs were screened, and clear images of VLPs were selected for the measurement.

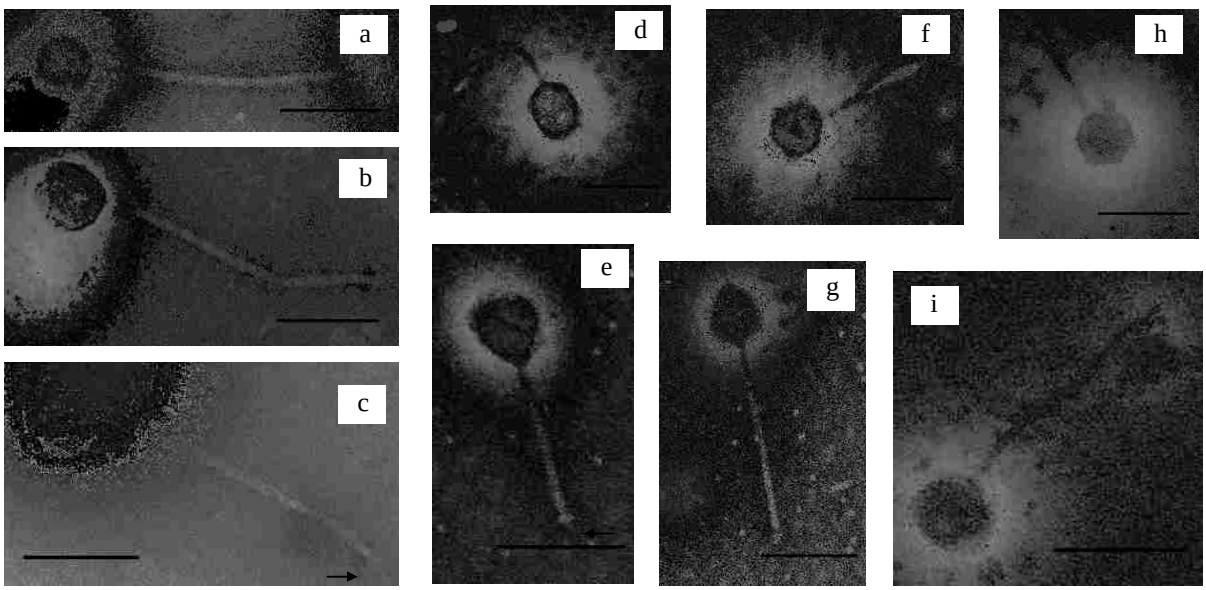
6.3 Results

6.3.1 Morphology of VLPs

TEM images of VLPs from the reactors are shown in Figure 6.1. All of the observed VLPs were head-tailed morphologies, consistent with the descriptions of the order *Caudovirales* (Ackermann 2007). Within this morphology, two distinct variations were observed: isometric capsids with long noncontractile tails (characteristic of the family *Siphoviridae*; ex: Figure 6.1 a, b, c, e, g, i and o) in both reactors, and isometric capsids with long contractile tails (characteristic of the family *Myoviridae*; ex: Figure 6.1 j, p, q, r and s) only in the hourly reactor. One example of an additional

morphology was observed (Figure 6.1 d) that had a moderately elongated capsid. The only described family within the order *Caudovirales* not observed in this study was *Podoviridae*, which is characterized by short tails. Various facultative structures of VLPs were evident in the TEM micrographs (for a review of the description of facultative structures see (Ackermann 1998)). For example, tail sheaths (Figure 6.1 j, k and m), base plates (Figure 6.1 e, n and p), tail spikes (Figure 6.1 e) and tail fibers (Figure 6.1 k, m, r, and s) were observed as indicated by arrows in the figures.

Daily-fed reactor



Hourly-fed reactor

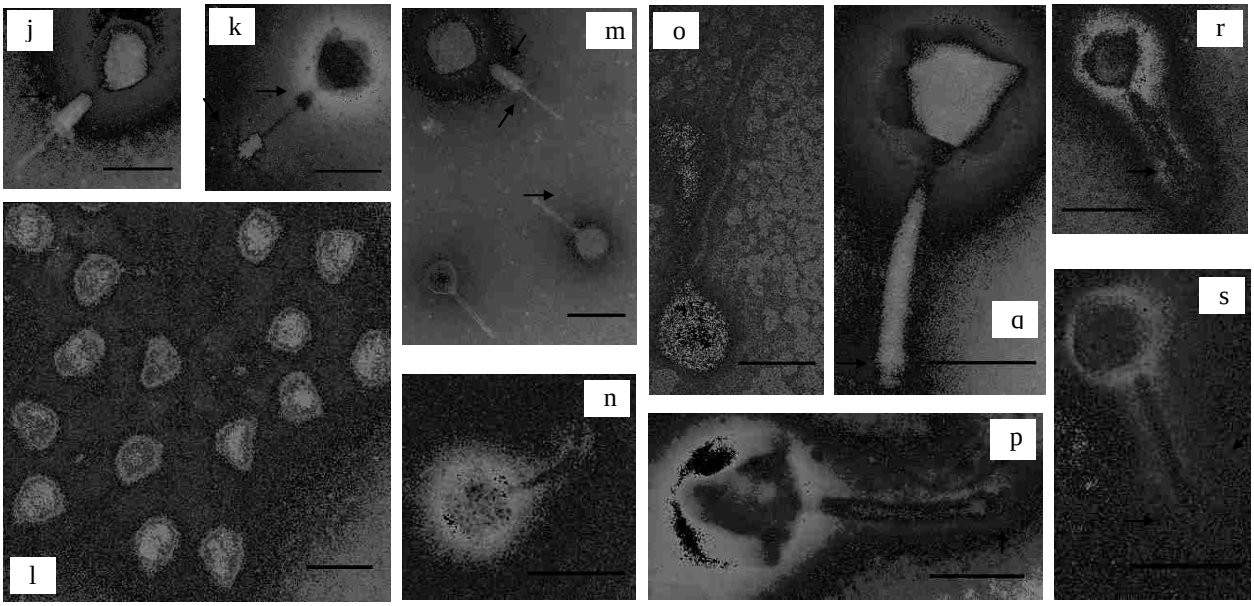


Figure 6.1 Selected transmission electron microscope micrographs of virus-like particles in the reactors

VLPs in the daily-fed reactor (a-i) and the hourly-fed reactor (j-s). VLPs morphologically similar to *Siphoviridae* (a, b, c, e, g, i and o) and to *Myoviridae* (j, p, q, r, and s) were observed. Scale bar represents 100 nm.

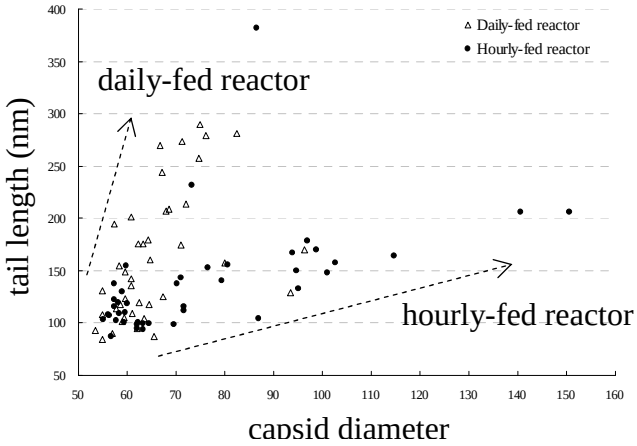
6.3.2 Size distribution of VLPs

The sizes of capsids and tails of 41 VLPs from the daily-fed reactor and 42 VLPs from the hourly-fed reactor were measured. Size distribution of VLPs varied between two reactors (Figure 6.2 a) with smaller capsids and longer tails observed in the daily-fed reactor, and larger capsids and shorter tails in the hourly-fed reactor (Table 6.2).

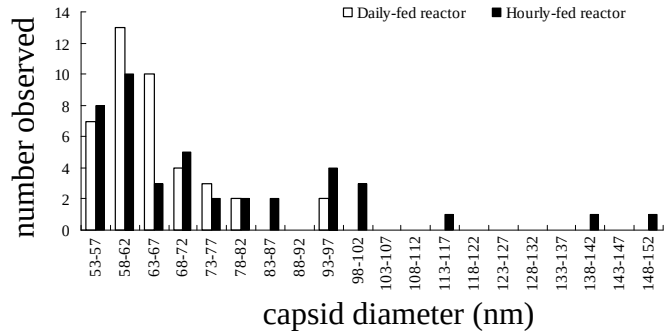
In the daily-fed reactor, the size distribution of isometric capsid diameters showed a distribution peak in the 58 nm - 62 nm group (Figure 6.2 b) while no clear peak was found in the distribution of tail lengths (Figure 6.2 c). Long flexible noncontractile tails were observed. Four of these ranged between 200 nm and 220 nm and 7 tails were longer than 240 nm.

In the hourly-fed reactor, the size distribution of head diameters also showed a peak in the 58 nm - 62 nm group (Figure 6.2 b). VLPs with extremely large capsid were also observed (e.g. Figure 6.1 p: 140 nm and Figure 6.1 q: 150 nm). Most tail lengths ranged from 80 nm to 240 nm with one exception of a VLP with a 380 nm tail (Figure 6.2 c and Figure 6.1 o). The distribution of tail lengths formed two peaks, one with tail length ranging from 100 nm to 119 nm and the other with tail length between 140 nm and 159 nm.

(a)



(b)



(c)

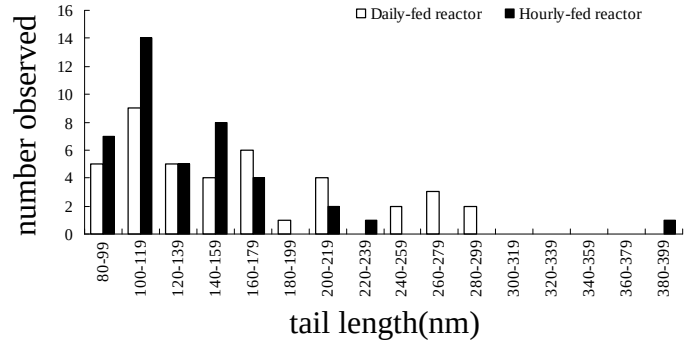


Figure 6.2 Distribution of measured capsid diameters and tail lengths
41 virus-like particles (VLPs) from the daily-fed reactor and 42 VLPs from the hourly-fed reactor were measured. (a) XY scatter chart, (b) size of capsids: X-axis shows capsid diameters and Y-axis shows number observed, (c) tail lengths: X-axis shows tail lengths and Y-axis shows number observed.

Table 6.2 Dimensions of VLPs observed in the reactors in comparison to previously documented phage

Reactor/phage families	Head diameter (nm)			Tail length (nm)			Reference
	average	range	STD[b]	average	range	STD[b]	
daily-fed reactor	66	54 - 96	9.6	162	84 - 289	61.4	This study
hourly-fed reactor	76	55 - 151	22.6	137	87 - 382	51.4	This study
Siphoviridae[a]	55	40 - 97	--	191	79 - 593	--	(Ackermann 1998)
Myoviridae[a]	85	53 - 160	--	167	80 - 485	--	(Ackermann 1998)

[a] 251 total phages examined

[b] STD, standard deviation

6.3.3 Concentration of VLPs using EFM

The average concentrations of VLPs in the reactors quantified using EFM are shown in Figure 6.3 and Table 6.3. Variation of VLP concentrations was low over the course of a week and before and after automated feeding. The concentration of VLPs were similar in the hourly-fed reactor ($8.4 \pm 4.3 \times 10^7$ VLPs/ml) than in the daily-fed reactor ($7.1 \pm 1.5 \times 10^7$ VLPs/ml). Variation shows the deviation among multiple samples collected from the reactors at different times, and no significant difference indicated between the two (T-test, $P = 0.05$).

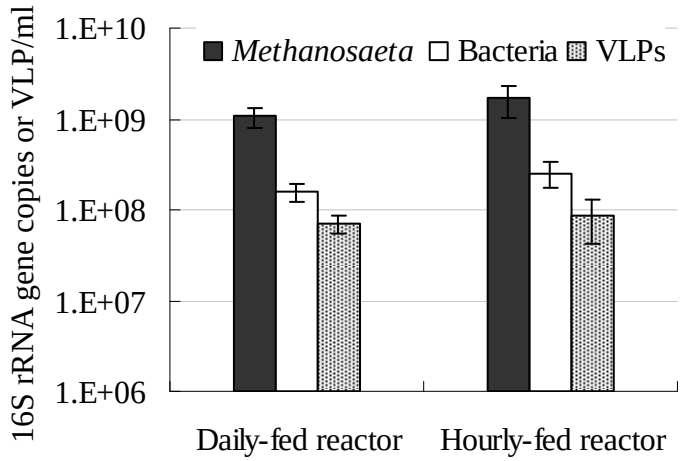


Figure 6.3 Concentrations of *Methanosaeta*, Bacteria and virus-like particle (VLP) in reactors Values are the average of 6 samples for *Methanosaeta* and Bacteria, and 10 (daily-fed) and 7 (hourly-fed) samples for the VLP. Standard deviations were indicated by error bar.

Table 6.3 Concentrations virus-like particle (VLP), *Methanosaeta* and Bacteria in reactors

Reactor	VLPs (VLPs/mL)	*Methanosaeta* (cells/mL)	Bacteria (cells/mL)	Prokaryotic (cells/mL)	Virus-to-Prokaryotic cell Ratio (VBR)
daily-fed reactor	$7.1 \pm 1.5\times10^{7}$	5.3×10^{8}	4.1×10^{7}	5.7×10^{8}	0.12
hourly-fed reactor	$8.4 \pm 4.3\times10^{7}$	8.4×10^{8}	6.7×10^{7}	9.0×10^{8}	0.09

6.3.4 Genome size of VLPs

Viral genome size distributions for the daily-fed and hourly-fed reactors are shown in the PFGE results (Figure 6.4). Different banding patterns were observed between the two reactors. Three bands (two major and one minor) were discovered in each reactor. The major (brightest) bands in

the daily-fed and hourly-fed reactors were observed at approximately 80 and 85 kbp, respectively. The second brightest band was observed at approximately 35 kbp for both reactors. A third dim band in each reactor was observed at a higher molecular weight range (>240 kbp). The ranges of genome size for *Siphoviridae* and *Myoviridae* are ~14 kbp to ~135 kbp and ~24 kbp to 316 kbp, respectively (Hyman and Abedon 2012). Thus, the viruses have genomes >240 kbp in both reactors might belong to *Myoviridae* although *Myoviridae* was not found in the daily-fed reactor using TEM. Additional background smearing observed in the hourly-fed reactor, especially around 145 kbp may suggest DNA from multiple virus species with similar genome sizes.

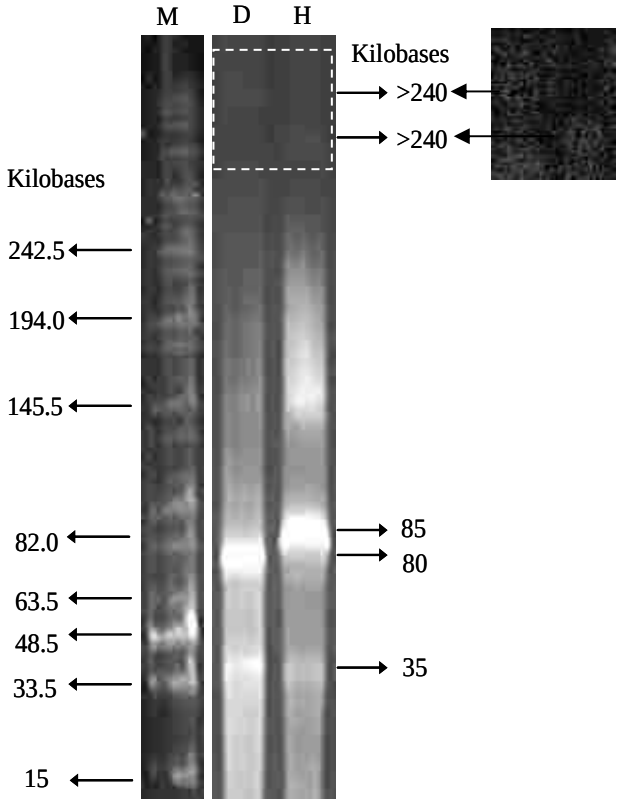


Figure 6.4 Pulsed field gel electrophoresis (PFGE) of viral genomes from reactors

6.3.5 Performance of the Acetate-fed Reactors

Both the daily-fed and hourly-fed reactors had neutral pH and consistent methane production prior to sampling. Monitoring of acetate concentrations, gas production and gas composition in the reactors showed that nearly all of the acetate was converted to methane (calculations data not shown). Neither reactor had any upsets or failure during the course of the study.

6.3.6 Dominant Reactor Microorganisms

Three independent methods were used to demonstrate the dominance of *Methanosaeta* in the reactors including microscopy, qPCR and sequence recovery. Filamentous cells composed of several single flat-end rods (0.8-0.9 × 1.8-3.3 μm) were dominant in both reactors (Figure 6.5 a, c, d and f). These rods were non-motile and generally grew as long threads (3-50 μm). Autofluorescence characteristic of methanogenic Archaea was confirmed in the dominant microbial populations in the samples (Figure 6.5 b and e). The rod-shapes observed are consistent with *Methanosaeta* as described elsewhere (Zehnder, Huser et al. 1980; Janssen 2003; Mizukami, Takeda et al. 2006). Coccoidal morphotypes were observed with much lower frequency, some of which were autofluorescent.

Methanosaeta qPCR signals dominated the reactors (Figure 6.3, Table 6.3), while *Methanosarcina* was not detected (detection limit in the extract sample of $3x10^2$ copies/ml of reactor effluent). *Methanosaeta* 16S rRNA gene copy concentrations were ~7× greater than total bacterial gene copy concentrations. *Methanosaeta* concentrations were statistically equivalent in the two reactors, while bacterial concentrations were higher in the hourly-fed reactor (T-test, P = 0.05).

Phylogenetic affiliation of archaeal and bacterial 16S rRNA genes from the two reactors as determined by clone sequencing are summarized in Table 6.4. In corroboration of the qPCR and microscopic evidence, the most frequently detected archaeal phylogenetic group was *Methanosaeta* in both reactors. One clone (out of 73 archaeal clones) in the daily-fed reactor aligned with the genus *Methanosphaera* (a Coccoidal-shaped microbe). No clone sequences from the group *Methanosarcina* were detected.

For the bacterial domain, the most frequently observed clones in both reactors were from the phyla *Proteobacteria*, *Firmicutes*, and *Bacteroidetes* (Table 6.4). While other groups were consistent between the two reactors, more clones for the WWE1 (for a description of phylum WWE1 see (Chouari, Le Paslier et al. 2005)) were obtained from the hourly-fed reactor. Of the *Proteobacteria*,

the *Arcobacter* was represented by 35 clones, consisting of 45.7% of the *Proteobacteria* in the daily-fed and 29.8% in the hourly-fed reactors (data not shown).

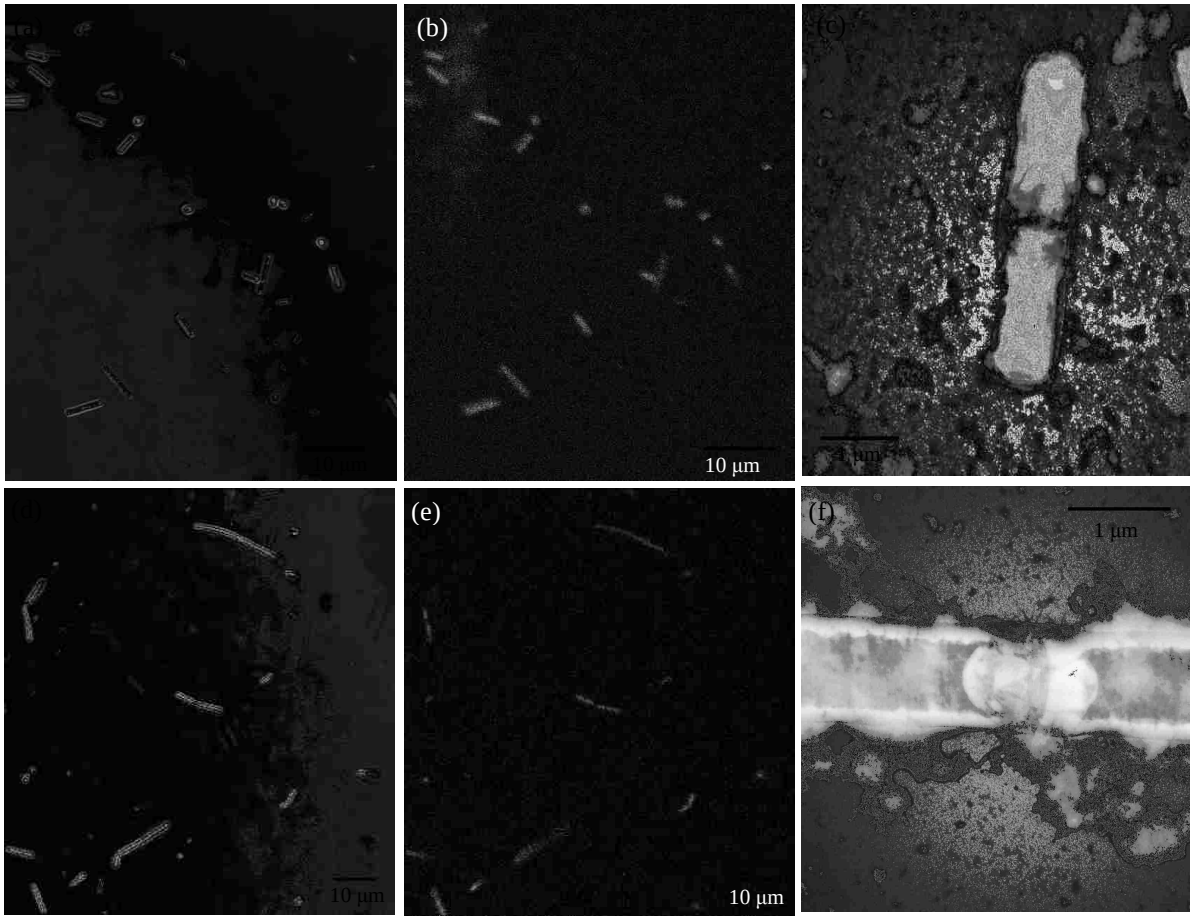


Figure 6.5 Micrographs of enriched cultures in the daily-fed (a-c) and hourly-fed reactor (d-f) (a) and (d) are transmitted light bright field images. (b) and (e) are fluorescence micrographs. (c) and (f) are transmission electron microscope micrographs

Table 6.4 Phylogenetic analysis of 16S rRNA gene in the daily-fed and hourly-fed reactors

Clone Library	Taxon[a]	No. of OTU[b]	No. of clone	Daily-fed (%)	Hourly-fed (%)
Archaea-21F	*Methanosaeta*	4	36	96.4	100.0
	Methanosphaera	1	1	3.6	-
Archaea-1492R	*Methanosaeta*	4	36	100.0	100.0
Bacteria	*Proteobacteria*	6	43	47.8	44.7
	Firmicutes	7	21	19.6	25.5
	Bacteroidetes	5	13	13.0	14.9
	WWE1	2	7	2.2	12.8
	minor groups[c]	6	9	17.4	2.1

[a]Taxon was determined by RDP classifier (Wang, Garrity et al. 2007); Archaea was grouped at the genus level and Bacteria was group at phylum level

[b]OTU, Operational Taxonomic Unit

[c]Clones with either 2 or less representatives in the library were from the groups *Spirochaetes*, *Chloroflexi*, *Lentisphaerae*, SR1 and unclassified bacteria.

6.4 Discussion

6.4.1 VLPs in Enriched Reactors

VLPs have been previously reported in anaerobic digesters (Park, Ikenaga et al. 2007; Wu and Liu 2009). However to the best of our knowledge, this is the first time that VLPs associated specifically with acetate-fed methanogenic consortia enriched in *Methanosaeta* have been characterized. Anaerobic digester sludge from municipal wastewater treatment plants have had higher reported VLP concentrations (e.g. 2.4×10^{10} VLPs/ml (Wu and Liu 2009)) likely associated with higher microbial diversity in these systems compared to the study enrichment reactors. Still,

the prevalence of *Siphoviridae* observed in the study reactors was consistent with Ackermann's 2007 survey that showed this virus family to be the most common tailed-phage documented in published electron micrograph images irrespective of the habitat. *Siphovirdae* were also the most abundant identified morphotype (16%) in a previously studied methanogenic digester treating wastewater from a beer brewery (Park, Ikenaga et al. 2007). Observation of this common viral family is particularly interesting because although some archeoviruses have been reported to exhibit exceptionally complex morphotypes (e.g. linear, fusiforms, droplet and bottle shapes) (Prangishvili, Forterre et al. 2006; Pina, Bize et al. 2011), currently characterized viruses of methanogens are of head-tailed morphologies (Jordan, Meile et al. 1989; Meile, Jenal et al. 1989; Nolling, Groffen et al. 1993).

Several observations demonstrate that virus populations between two reactors differed. First, VLPs morphologically similar to *Myoviridae* were only found in the hourly-fed reactor (Figure 6.1). Second, different dominant VLP capsid sizes and tail lengths were observed in the two reactors (Figure 6.2 a), a parameter thought to be highly conserved and uniform for individual virus species (Ackermann 1998). Third, PFGE banding patterns suggested that the viruses in the two reactors had different genome sizes (Figure 6.4, e.g. virus with 80 kbp genome in the daily-fed reactor and virus with 85 kbp in the hourly-fed reactor). Different virus populations may have

developed in the reactors due to several contributing factors such as feeding schedule, HRT, re-inoculation of the daily-fed reactor, natural population drift in each reactor caused by unique "virus-host arm races".

6.4.2 Enriched Cultures and Microbial Communities

One genus of the acetoclastic methanogens, *Methanosaeta*, was dominant in both reactors. Higher periodic acetate concentrations and shorter HRT in the daily-fed reactor was designed to favor *Methanosarcina* due to its higher growth rate (*k*) and lower acetate affinity (higher *K*s) than *Methanosaeta* and *Methanosarcina* was historically dominant in the daily-fed reactor as previously reported (Conklin, Stensel et al. 2006). However, dominance shifted from toward *Methanosaeta* by year 2007. Re-inoculation of the daily-fed reactor with sludge from municipal anaerobic digesters failed to reestablish *Methanosarcina*, as evidenced by qPCR, microscopic, and sequencing results presented in this work. The loss of *Methanosarcina* in the daily-fed reactor remains unexplained, but allowed comparison of VLP communities in two reactor systems that shared a dominant microbial population.

While the Bacteria were less prevalent than the *Methanosaeta,* their phylogeny still reveal

interesting insights into the reactor microbial communities. The phyla *Proteobacteria*, *Firmicutes* and *Bacteroidetes* detected in the reactors have previously been reported as core microbial groups in several full-scale mesophilic anaerobic digesters (Riviere, Desvignes et al. 2009) and in mesophilic bovine serum albumin-fed reactors (Tang, Shigematsu et al. 2005). One of the OTU detected in this study (containing 35 clones) within phylum *Proteobacteria* was assigned to genus *Arcobacter* and was most frequently detected in both the daily-fed (45.7%) and hourly-fed (29.8%) reactors (data not shown). The closest related sequence (NCBI accession number: GQ136513.1, 100% sequence coverage and similarity) to this OTU was a clone also found in an acetate enriched digester sample, implying that an anaerobic acetate enriched niche may be a favorable habit for *Arcobacter*. Sequences assigned to WWE1 were observed more frequently in the hourly-fed reactor, while the sequences affiliated with minor groups were more dominant in the daily-fed reactor, indicating different bacterial communities between the two reactors. Most of the bacterial 16S rRNA gene sequences recovered did not closely affiliate to any known isolated culture (similarity <97%) but were similar to the clones found in the anaerobic digesters (Tang, Shigematsu et al. 2005; Shigematsu, Tang et al. 2006; Riviere, Desvignes et al. 2009; Ito, Yoshiguchi et al. 2011; Krakat, Schmidt et al. 2011), methanogenic consortium and landfill leachate (Limam, Bouchez et al. 2010) suggesting these anoxic systems may contain many bacterial species with yet unconfirmed metabolic roles.

6.4.3 VLPs and Enriched Cultures

Observations of persistent VLP populations in the reactors suggest that they are actively propagating in the system. Because viruses are obligate parasites, they only reproduce when their corresponding hosts (Bacteria or methanogenic Archaea in this study) are present. VLPs were confirmed to be entering the reactors with the feed, which was prepared aseptically and evaluated by EFM to confirm that VLPs were not present. VLPs that entered the reactors with the anaerobic digestion sludge used to inoculate the reactors is not mathematically not predicted to have persisted in the reactors due to the extended time since last re-inoculation (>3 years prior to the study) in comparison to the systems' retention times. Thus, replication of the observed VLP was the most likely explanation for observation of VLP persistence.

While replication of the virus population was supported by the data, the virus-to-bacteria ratio (VBR) values were lower in the study reactors than is often reported. The ratio of VLP concentrations to the prokaryotic cells (*Methanosaeta* and bacteria) concentration was 0.123 in the daily-fed reactor and 0.093 in the hourly-fed reactor (assuming 2 and 3.8 16S rRNA genes copies per organism for *Methanosaeta* (Barber, Zhang et al. 2011) and Bacteria (Fogel, Collins et al. 1999), respectively. In contrast to the study calculation, when VBR has been used to study

relationships between viruses and bacteria in many natural aquatic ecosystems, the number of virus is higher than the number of bacteria with typical VBR ranging from 3 to 10 in aquatic ecosystems (Wommack and Colwell 2000). Assumptions used to estimate microorganism concentrations from 16S rRNA copy number may partially contribute to lower VBR values, particularly because the numbers of 16SrRNA genes in a chromosome can vary among species of the same genus and some methanogens have recently been reported to contain more than one chromosome (Hildenbrand, Stock et al. 2011) (though not yet studied for slow-growing *Methanosaeta)*. However, lower VBR values similar to this study's results have been reported in an oligotrophic lake (0.03-0.7) (Tapper and Hicks 1998) and in rhizosphere soil (0.04) (Ashelford, Day et al. 2003). VBR values in Archaea-dominant environments have not been reported.

Finding of low VLPs concentration in two *Methanosaeta*-dominant reactors studied here may due to several reasons. First, viruses in reactors might prefer lysogenic or pseudolysogenic (Los and Wegrzyn 2012) life cycles, in which viruses persist inside host cells and thus were not counted using the study methods. Second, it has been shown that the VBR is higher in nutrient-rich, more productive environments (Wommack and Colwell 2000). However, the acetate-fed methanogenic reactors in this study are not classified as productive environments for viruses. Under anaerobic conditions, limited energy is gained by slow growing *Methanosaeta* when acetate is converted to

methane, and only limited substrate (e.g. vitamins) was available for the growth of bacteria (although metabolic roles of bacteria in reactors is unknown). Third, a positive correlation between VBR and host community diversity has been previously observed (Tuomi, Fagerbakke et al. 1995), suggesting that the VBR may be an indicator of host community diversity. For the reactors in the current study, *Methanosaeta* was found to be dominant in the daily-fed and hourly-fed reactors and low diversity of microbial communities was observed, which may correlate with the low VBR observed in two enriched reactors.

Despite persistent VLP presence in both *Methanosaeta* dominant reactors, upset or failure of these methanogenic reactors was not observed over the study period. This might be explained by the present of temperate viruses, chronic infections or co-evolution of virus and host populations in reactors. First, limited energy conditions and a slow growing host under anaerobic conditions may be favorable for lysogenic or pseudolysogenic life cycles. Second, some crenarchaeal viruses have been found to be chronically produced without host cells lysis (Prangishvili and Garrett 2005; Prangishvili, Vestergaard et al. 2006; Pina, Bize et al. 2011). Although viruses of acetoclastic methanogens have not yet been reported, this hypothesis seems less likely because currently identified head-tail morphotypes, such as those observed in this study, have previously been reported as lytic (Meile, Jenal et al. 1989; Nolling, Groffen et al. 1993; Luo, Pfister et al. 2001).

Finally, an evolutionary arms race between viruses and hosts could result in minor changes in population structure without the changing of community stability (Andersson and Banfield 2008). Nevertheless, this theory requires more investigation for methanogenic habitats.

6.5 Conclusions

Persistent VLPs were present in the *Methanosaeta*-dominated daily-fed and hourly-fed reactors. Virus populations between the two reactors were distinct which was supported by three observations: (1) VLPs morphologically similar to *Myoviridae* were only found in the hourly-fed reactor; (2) the different size distribution of VLPs between two reactors, suggesting that distinct virus species existed in the individual reactors; (3) different PFGE banding patterns also indicate that certain virus species were unique in the individual reactors.

All of the observed VLPs were of head-tailed morphology. VLPs morphologically belonging to *Siphoviridae* were found in both reactors, and VLPs morphologically similar to *Myoviridae* were only found in the hourly-fed reactor. The numbers of VLPs was 18% higher in the hourly-fed reactor ($8.4 \pm 4.3 \times 10^7$ VLPs/ml) than in the daily-fed reactor ($7.1 \pm 1.5 \times 10^7$ VLPs/ml). Major viral genome size of the daily-fed reactor and hourly-fed reactors was at about 80 and 85 kbp,

respectively. Viruses with 35 kbp and >240 kbp genomes were found in both reactors.

VBR values were low in both the daily-fed reactor (0.123) and the hourly-fed reactor (0.093). This result may imply that lysogenic or pseudolysogenic life cycles could be the favorable strategy for viruses targeting slow-growing hosts in the methanogenic environment. Besides, low VBR might also due to low diversity of microbial communities in reactors.

The presence of many VLPs with 60 nm capsid and 120 nm tail or viruses having the same genome sizes (30 kbp) in both reactors may suggest they target a common host such as *Methanosaeta*, *Proteobacteria* or *Firmicutes*. Since both reactors were dominated by *Methanosaeta*, viruses infecting *Methanosaeta* are likely to propagate in the systems.

For the bacterial domain, both reactors were dominated by *Proteobacteria*, *Firmicutes*, and *Bacteroidetes*. Of the *Proteobacteria*, the *Arcobacter* was represented by 35 clones, consisting of 45.7% of the *Proteobacteria* in the daily-fed and 29.8% in the hourly-fed reactors.

Chapter 7 The study of the connection between viruses and *Methanosaeta* using a genomic approach based on clustered regularly interspaced short palindromic repeats (CRISPRs)

7.1 Introduction

In this chapter, interactions between viruses and *Methanosaeta* were studied using molecular biology approach based on Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) (literature of CRISPRs reviewed in section 2.8). CRISPR has been shown to be a tool to study the relationship between viruses and hosts in many ecosystems (Andersson and Banfield 2008; Snyder, Bateson et al. 2010; Sorokin, Gelfand et al. 2010; Pride, Sun et al. 2011; Stern, Mick et al. 2012). Using this tool, the present of *Methanosaeta* viruses in two *Methanosaeta*-dominated reactors (the daily-fed and hourly-fed reactors) and three full-scale digesters (West Point, South Plant, and Port Angeles) was investigated. Initially, comparative analyses of CRISPR/Cas systems of three *Methanosaeta concilii* strains were conducted. For the identification of CRISPR spacer in microbial genomes, two different tools (CRISPRFinder and Crass) were compared. Isolated spacer sequences were then used to identify the putative *Methanosaeta* viruses in constructed viral nucleotide database. These viral sequences were used as queries to search for

homologies in NCBI database. Finally, recognition sites for the CRISPR/Cas systems on targeted viral sequences were investigate.

7.2 Materials and Methods

7.2.1 Overview of the experimental approaches

The experimental approaches are summarized in Figure 7.1.The experiments to collect, concentrate, and purify viruses for metagenomic sequencing were done in conjunction with a former Master's student, Scott Pease (MS in ENVH, 2012). The data and results of viral metagenomes analyzed by Pease are located in Appendix A. Microbial metagenomes from the daily-fed and hourly-fed reactors were sequenced individually. The spacer sequences in the CRISPR loci were extracted from microbial metagenomes, and these spacer sequences were blasted against the viral metagenomes to identify the corresponding viruses. In addition, pure strains of *Methanosaeta concilii* (DSM 6752 and DSM 2139) were also sequenced. Combined with a *Methanosaeta concilii* GP6 genome retrieved from NCBI, variation of the CRISPR and CRISPR-associated (*cas*) genes among the three strains were investigated. Additionally, viral metagenomes of three full-scale anaerobic digesters were sequenced. The presence of viruses of

Methanosaeta concilii in these digesters was studied using the extracted spacer sequences.

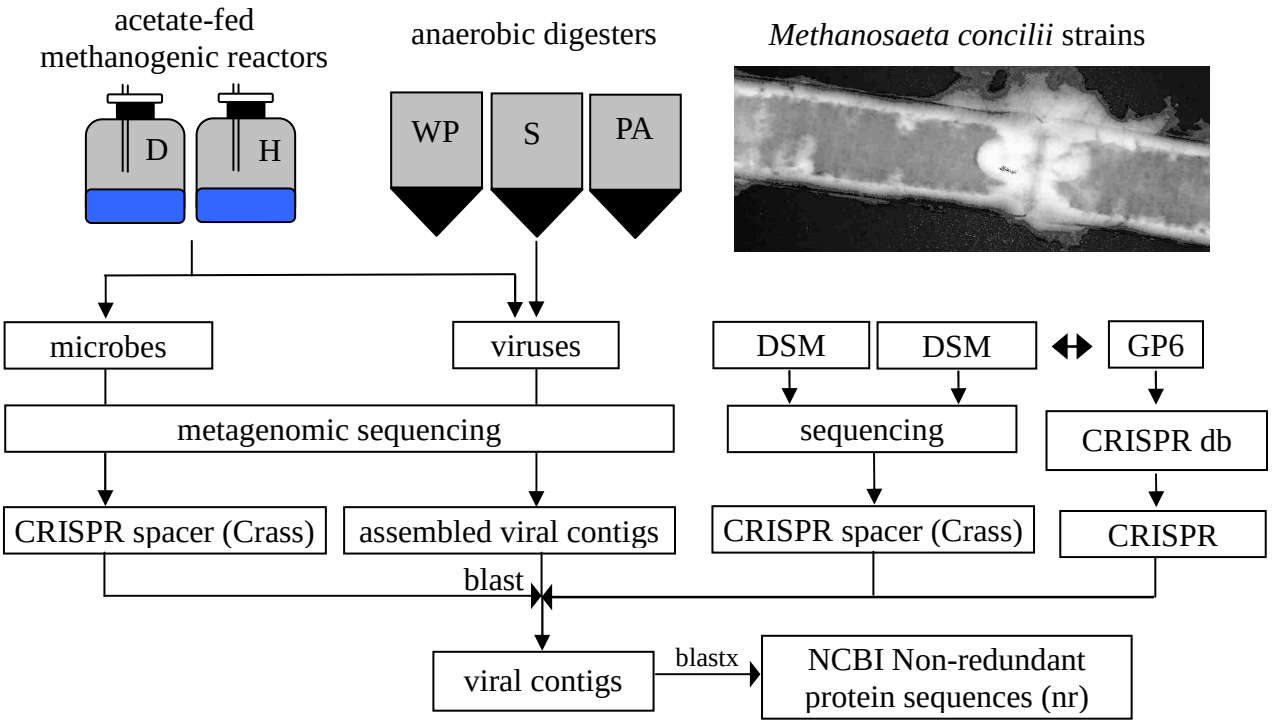


Figure 7.1 Flowchart of the experimental approach

7.2.2 Sequencing Enrichments and *Methanosaeta concilii* Genomes

Sample sequencing was performed using an Ion torrent personal genome machine (PGM) with a 316 chip and the long reads protocol (Functional Genomics Laboratory, University of Washington). DNA of the daily-fed and hourly-fed enrichments was barcoded and sequenced with a 316 chip. Likewise, barcoded DNA of *Methanosaeta concilii* DSM 6752 and DSM 2139 were pooled and sequenced. Reads were quality filtered using the clip function in Nesoni (v. 58;

Victorian Bioinformatics Consortium) to trim reads to a length with a minimum quality score of 10 and discard reads <24 bp. Assembly parameters were optimized using VelvetOptimiser v. 2.2.0 (Zerbino and Birney 2008) and were applied to contig assembly using MetaVelvet (Namiki, Hachiya et al. 2012).

7.2.3 *Methanosaeta concilii* CRISPR/Cas System Comparison

The whole genome of *Methanosaeta concilii* GP6 was retrieved from NCBI and used as a template for the comparative analysis. The variations of CRISPR/Cas regions in strain DSM 6752 and DSM 2139 were identified using the *breseq* (http://barricklab.org/breseq) (Barrick, Yu et al. 2009) and were visualized using Integrative Genomics Viewer (IGV) (Thorvaldsdottir, Robinson et al. 2013).

7.2.4 CRISPR and Spacer Identification

Since the discovery of CRISPR in many genomes of bacteria and archaea, several tools have been developed to identify these unique repeat elements, including the CRISPR Recognition Tool (CRT) (Bland, Ramsey et al. 2007), CRISPRFinder (Grissa, Vergnaud et al. 2007a), and

Crass (Skennerton, Imelfort et al. 2013). CRISPRFinder is so far the most widely used tool to detect CRISPR. Crass is a recently developed tool to identify CRISPR for metagenomic data without the need of assembly, which brings the advantage of avoiding the sequence error generated during assembling. In addition, Crass uses a new algorithm to increase the sensitivity, specificity and speed of CRISPR detection.

In this study, CRISPR and spacer sequences were identified and extracted using CRISPRfinder (http://crispr.u-psud.fr/Server/) (Grissa, Vergnaud et al. 2007a) and Crass v. 0.3.6 (http://ctskennerton.github.io/crass/) (Skennerton, Imelfort et al. 2013). Contigs and filtered reads were used as input for CRISPRfinder and Crass, respectively. Identified CRISPRs were grouped by the repeat sequences. Homologies of repeat sequences were searched in the NCBI nr/nt database and CRISPRdatabase (http://crispr.u-psud.fr/crispr/) (Grissa, Vergnaud et al. 2007b) in order to find the suspected microbes having the corresponding CRISPR.

7.2.5 Spacer Similarity Search

The spacers obtained were blasted against several viral metagenomes using ViroBLAST (Deng, Nickle et al. 2007) with default settings, aside from the word size, 7. The viral metagenomes

used were the daily-fed reactor, hourly-fed reactor, West Point digester, South Plant digester, and Port Angeles digester. For Crass, maximum mismatch between spacer and protospacer (viral) sequences was 2 nt and the E-value was ≤ 7E-7. Homologous sequences or proteins of protospacer-containing viral contigs were searched in the nucleotide collection (nt/nr) database using blastn and the non-redundant protein sequences (nr) database using blastx, respectively. For the search of non-redundant protein sequences, E-value was ≤ 6E-4.

7.2.6 Analysis of Protospacer Adjacent Motif (PAM)

Protospacer adjacent motifs (or protospacer-associated motif) of *Methanosaeta* CRISPR systems were investigated by aligning the invading sequences. Nucleotide conservation near protospacer was examined. The conserved motifs of each loci were identified and illustrated using WebLogo (Crooks, Hon et al. 2004).

7.3 Results and Discussion

7.3.1 Sequencing and Assembly Statistics

Results of sequencing and assembly of enrichments and *Methanosaeta concilii* are summarized in Table 7.1. The Ion torrent 316 chip contains about 6.35 millions wells. Therefore, under

perfect library preparation and sequencing conditions, there should be 6.35 millions reads per sequencing run. In reality, fewer total reads can be obtained due to empty wells, wells with only primer signal, wells with positive control beads, and wells with double signal (polyclonal reads). The number of combined raw reads of two enrichment samples was 4,237,650 (67% of total addressable wells), which is higher than that of combined DSM 6752 and 2139 samples (2,579,244 reads, 41% of total addressable wells). Excluding the reads without barcodes, the final raw reads were 4,213,832 (66%) for enrichments, and 2,564,818 (40%) for *Methanosaeta concilii* strains. Better sequencing results of enrichment samples were also indicated by longer mean read length. In addition, estimated average genome coverage of strain DSM 6752 and DSM 2139 are about 40× (calculated using the *Methanosaeta concilii* GP6 genome length: 3.03 Mb).

N50 is a sequence length-weighted median value widely used to address the quality of assembly. It represents at least 50% of bases in contigs longer than N50. Not surprisingly, the N50 is longer in two pure strains than the enrichments. N50 of the hourly-fed enrichment is larger than that of the daily-fed enrichment which might indicate a less diverse microbial community in the hourly-fed reactor. However, this remains to be confirmed by other approaches, and the bias introduced by DNA extraction should not be excluded.

Table 7.1 Sequencing and assembly statistics of enrichments and *Methanosaeta concilii*

		Enrichments and *Methanosaeta concilii*			
	Statistic	Daily	Hourly	6752	2139
Pre-assembly	Raw reads	2,127,272	2,086,560	1,259,401	1,305,417
	Mean raw read length	188.02	188.64	176.66	172.04
	Raw bases	399,975,626	393,610,055	222,484,454	224,584,415
	Filtered reads	2,070,722	2,036,139	1,178,771	1,215,114
	Mean filtered length	107.97	110.79	100.06	100.35
	Filtered bases	223,578,268	225,581,678	117,947,470	121,941,370
	Percent reads pass	97.34	97.58	93.60	93.08
	Percent bases pass	55.90	57.31	53.01	54.30
Assembled	N50	914	1,348	3,620	3,540
	Longest contig	11,037	13,342	23,120	14,400
	Contigs	7,880	5,662	1,567	1,621
	Contigs > 1k	941	939	773	809
	Bases in contigs > 1k	1,871,679	2,117,487	2,497,469	2,562,059
	Assembled bases	4,011,863	3,621,101	2,821,136	2,896,328
	GC	51.27	51.41	51.38	51.30

7.3.2 CRISPR/Cas of *Methanosaeta concilii*

CRISPR/Cas Systems of *Methanosaeta concilii* GP6

CRISPR/Cas systems of *Methanosaeta concilii* GP6 retrieved from NCBI and CRISPRdatabase were summarized in Tables 7.2 and 7.3, and illustrated in Figure 7.2. Six putative CRISPR loci were found in the *Methanosaeta concilii* GP6 genome, and all loci were located on the chromosome. CRISPR locus 1 contained only one spacer, and there was no *cas* gene in the vicinity of this region. This locus is probably not a CRISPR, so it is excluded in this study. Locus 2 and locus 3 are clustered in a close region, and have identical repeat sequences. A different

repeat sequence is shared by loci 4, 5, and 6, and these loci are clustered. The predicted RNA secondary structures of repeats are shown in Figure 7.3. The repeat sequence of CRISPR/Cas system 1 forms a stable secondary structure (stem-loop structure), and a weaker folding RNA molecule could be generated by the repeat of system 2. The amount of spacers ranges between 8 and 17 in most of the loci, while locus 6 contains 46 spacers. The spacer size is between 31 to 41 bp. The identical sequences (231 bp) were only located in one end (upstream) of loci 2 and 3 suggests they might be the leader sequences. No conserved sequence was detected in either flanking regions of loci 4, 5, and 6.

Both CRISPR clusters are physically linked to *cas* genes (Table 7.3 and Figure 7.2). In system 1, *cas3* (the closest) is located 825 bp upstream from the first repeat of locus 2. In system 2, the *cas3* (the closest) is 6291 bp upstream from the end of locus 4, except *cas6* spotted 1149 bp away from locus 6. The two most conserved *cas* genes, *cas1* and *cas2*, are found in both systems. Both CRISPR/Cas clusters are classified as type I systems because of the appearance of the type-specific *cas3* gene. In addition, CRISPR/Cas system 1 is further classified into subtype I-B due to *cas8b*. Gene *cas4* appears in both systems. *cas5h* and *cas7* are found in the system 1, and *cas6* and *csa1* are exclusive in the system 2. Moreover, many transposase families are found in both systems (gray arrows in Figure 7.2).

The most conserved and universal Cas proteins, Cas1 and Cas2, contain metal-dependent DNA and RNA endonuclease, respectively, and are suggested to participate in the adaption step (spacer acquisition and integration). Also, Cas4 is implied to be involved in the adaption process due to possessing RecB nuclease domain, and is fused to Cas1 in some type I subsystems (Garrett, Vestergaard et al. 2011; Makarova, Haft et al. 2011). Cas3 is a large protein that carries a separated helicas domain and HD endonulease domain, and might be responsible for the cleavage of invader DNA during the interference step. The RNA endonuclease activity in Cas5, Cas6, and Cas7 suggests they could promote the crRNA maturation during processing step. Cas8 is a large subunit of Cascade complex, and might be involved in both adaptation and interference steps.

Table 7.2 CRISPRs of *Methanosaeta concilii* GP6

Locus	Start Position	End Position	# spacers	DR consensus	Repeat (bp)	Spacer (bp)	Locus (bp)
1	2745630	2745727	1	TTTCTCGTGGAACTTGCTCTTGATGT	26	46	97
2	2747357	2747982	8	GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG	37	34 - 40	625
3	2749613	2750460	11	GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG	37	33 - 41	847
4	2851895	2852843	12	GTTTCAATTCCCTTTTCATCGGGATAGGCTCTGCAAT	37	36 - 41	948
5	2854178	2855494	17	GTTTCAATTCCCTTTTCATCGGGATAGGCTCTGCAAT	37	36 - 40	1316
6	2856874	2860348	46	GTTTCAATTCCCTTTTCATCGGGATAGGCTCTGCAAT	37	31 - 41	3474

Table 7.3 *cas* genes of *Methanosaeta concilii* GP6

System/Classification[(a)]	Gene	Start	End	Strand (+/-)	Gene size (bp)
system 1: locus 2-3 and Cas proteins CRISPR-Cas type I-B	*cas3*	2751286	2753754	-	2468
	cas5h	2753729	2754442	-	713
	cas7	2754450	2755370	-	920
	cas8b	2755373	2757256	-	1883
	cas4	2757370	2758005	-	635
	cas2	2758007	2758294	-	287
	cas1	2758297	2759421	-	1124
system 2: locus 2-3 and Cas proteins CRISPR-Cas type I	*csa1*	2840113	2841054	-	941
	cas4	2841087	2841680	-	593
	cas2	2841781	2842059	-	278
	cas1	2842091	2843104	-	1013
	cas3	2843201	2845603	+	2402
	cas6	2861498	2862475	-	977

[a]: classification of CRISPR-Cas (Makarova, Haft et al. 2011)

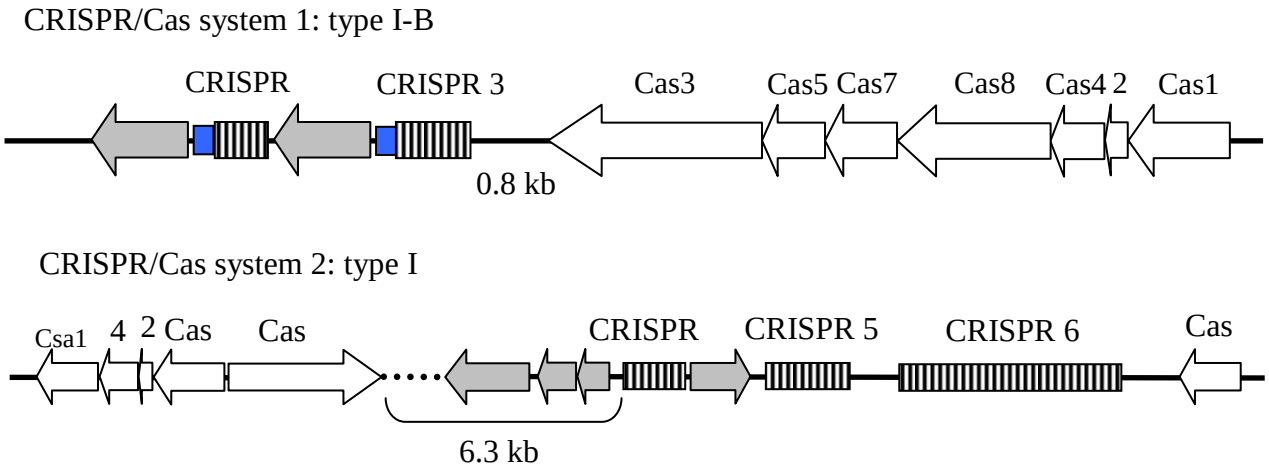
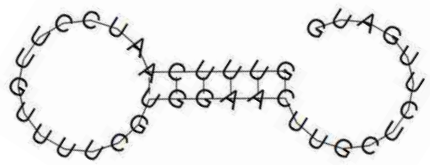
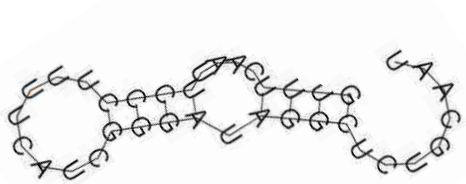


Figure 7.2 Organization of *Methanosaeta concilii* GP6 CRISPR/Cas systems. Transposase families are shown as grey arrows.

CRISPR 1 (locus 2 and 3): GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG



CRISPR 2 (locus 4, 5, and 6):



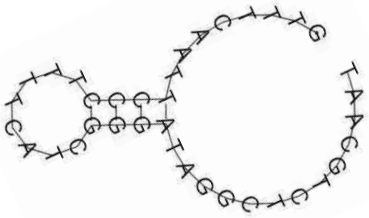


Figure 7.3 Predicted RNA secondary structure of two CRISPR repeat sequence

Two CRISPR/Cas systems were found in the *Methanosaeta concilii* GP6 genome. System 1 is classified as type I-B, and system 2 belongs to type I. To date, the type I system is the most abundant group found in the archaeal genome (Makarova, Haft et al. 2011). The presence of multiple CRISPR/Cas systems in a given archaeon is also frequently reported (Lillestol, Redder et al. 2006; Hale, Kleppe et al. 2008; Garrett, Vestergaard et al. 2011; Terns and Terns 2011). Based on sequence similarity, CRISPR repeats can be classified into at least 12 groups. Some groups can form stable stem-loop RNA secondary structures, while others have no predicted structures (Kunin, Sorek et al. 2007). A stem-loop structure may act as a binding structure for

conserved Cas proteins (Kunin, Sorek et al. 2007), while binding via sequence-specific manner was also reported in *Sulfolobus* (Peng, Brugger et al. 2003). In addition, several groups of repeats clearly correspond to distinct CRISPR-Cas subtype indicating close association between repeats and Cas proteins. Together, this might suggest that the two CRISPR/Cas systems in *Methanosaeta concilii* GP6 adopt different processing mechanisms to mediate immunity.

Comparative Analyses of CRISPR/Cas Systems of *Methanosaeta concilii*

To compare CRISPR/Cas systems among three strains of *Methanosaeta concilii*, sequencing reads were aligned against the template strain (GP6). The results are shown in Table 7.4 and Figure 7.4. For CRISPR/Cas system 1, DSM 6752 has repeat element and contains partial *cas* genes. In contrast, the repeat sequence and most of the *cas* cassette are missing in DSM 2139. For CRISPR/Cas system 2, repeats and cas cassette are not detected in DSM 6752 while they are intact in DSM 2139.

Two CRISPR/Cas systems in strain GP6 probably are both active with respect to the appearance of repeat-spacer units and *cas* genes responsible for all three immunity processing steps. It is noteworthy that the distance between CRISPR locus and cas genes is about 6.3 kb in CRISPR/Cas system 2. Most reported CRISPRs are less than 1kb away from *cas* genes

(Marraffini and Sontheimer 2010), but active CRISPR loci of *Sulfolobus* were also documented when they were several kb (3.5-10 kb) away from *cas* genes (Lillestol, Shah et al. 2009; Erdmann and Garrett 2012). However, there is no detectable leader sequence in system 2. For crenarchaeal genus *Sulfolobus*, lack of a leader sequence has been shown to prevent the extension of CRISPR locus, although the transcripts were detected (Lillestol, Shah et al. 2009; Erdmann and Garrett 2012). These results suggest that both *Methanosaeta concilii* GP CRISPR/Cas systems might be actively transcribed but the addition of new spacer-repeat units could only happens in the system 1.

Table 7.4 Missing coverage in CRISPR/Cas regions

Strain	CRISPR/Cas system	Size (bp)	Description
DSM 6752	1	1927–2156	MCON_3254
		31–109	transposase, IS116/IS110/IS902 family
		35–49	transposase, IS116/IS110/IS902 family
		2140–2303	MCON_3259–MCON_3260
		109–127	hypothetical protein/hypothetical protein
		1359	[MCON_3264]–[cas3]
		1101	[cas3]–[cas5]
		964	[cas5]–[cas7]
		2519	[cas7], cas8b, [cas4]
	2	21389	21 genes: csa1-MCON_3406
DSM 2139	1	11913–11996	MCON_3254–[cas4]
		1639	[cas4], cas2, cas1
	2	1629	MCON_3381
		1195	transposase and inactivated derivatives, IS1 family
		797–904	[MCON_3387]

(a) CRISPR/Cas system 1

(b) CRISPR/Cas system 2

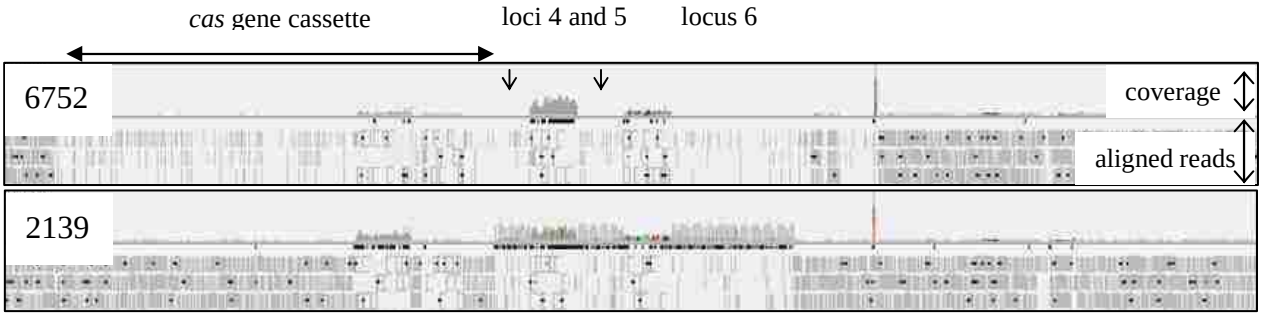


Figure 7.4 Mapping reads to CRISPR/Cas systems
(a) CRISPR/Cas system 1, (b) CRISPR/Cas system 2

Some of the CRISPR/Cas systems for DSM 6752 and DSM 2139 are either truncated or lost. For DSM 6752, the system 1 might have been defective because most of the *cas* genes are absent, whereas the whole system 2 was lost. For DSM 2139, the whole system 1 was missing while the system 2 is intact. The variation of CRISPR/Cas modules of different strains has been documented; for instance, four of six CRISPR/Cas modules were lost in a variant strain of *S. solfataricus* P2 (Redder and Garrett 2006). Besides, *S. islandicus* HVE10/4 has two CRISPR/Cas modules while *S. islandicus* REY15A carries only one CRISPR/Cas module (She, Singh et al. 2001; Reno, Held et al. 2009). In addition, the position of CRISPR/Cas module in the genome

was not conserved for different *S. islandicus* strains (Shah and Garrett 2011). The above observation probably can be explained by complex genome rearrangements in the CRISPR regions. It has been reported that CRISPR systems constantly locate in the region rich in transposable elements (Garrett, Vestergaard et al. 2011; Guo, Brugger et al. 2011; Shah, Vestergaard et al. 2011; You, Liu et al. 2011). The loss of CRISPR/Cas module in *S. solfataricus* P2 was implied by a single recombination event between two bordering IS elements (Redder and Garrett 2006). Bordering IS elements were implied to produce transposons containing whole CIRSPR/Cas modules and to trigger recombination thus leading to the loss of CRISPR systems (Shah, Vestergaard et al. 2011). Loss of CRISPR/Cas systems through genome recombination in *Methanosaeta* is possible because several *transposase* families were found in or adjacent to the *Methanosaeta concilii* CRISPR/Cas systems (Figure 7.2). Moreover, it might be a selective advantage for the DSM 6752 and DSM 2139 to lose the CRISPR/Cas systems because the absence of infective viruses during laboratory transferring.

7.3.3 Repeats and Spacers of Enrichments and *Methanosaeta concilii*

CRISPR Identification

Table 7.5 and 7.6 summarize the CRISPR of *Methanosaeta concilii* and enrichments identified

using CRISPRFinder. CRISPRs were grouped on the basis of repeat sequence similarity. As shown in Table 7.5, there are 5 and 7 repeat groups detected for DSM 6752 and DSM 2139, respectively. The repeat sequence of *Methanosaeta concilii* GP6 CRISPR/Cas system 1 was found in DSM 6752 (group 1), while the repeat sequence of system 2 was only detected in DSM 2139 (group 1), which was consistent with the genome mapping result in the previous section. For other groups, no homology was found in NCBI nucleotide collection (nt/nr) database and CRISPRdatabase. There are 24 spacers extracted from system 1 of DSM 6752, and 5 spacers isolated from system 2 of DSM 2139. Many spacers were identified in other unknown repeat groups, for example, there are 19 spacers in group 2 of DSM 6752, and 13 spacers in group 6 of DSM 2139. However, it is unclear whether both strains carry novel CRISPRs in their genomes.

For the CRISPRs in enrichments, 5 groups were found in the daily-fed reactor, while only 2 groups were detected in the hourly-fed reactor. Among these 7 groups, 3 groups were identical to the repeats of *Methanosaeta concilii* GP6. Group 1 and group 2 of the daily-fed reactor belong to GP6 system 1 and 2, respective. Group 1 of the hourly-fed reactor is homologous to the repeat of GP6 system 1. Again, the repeat sequences of the other groups have no homology in the public database, which also indicates that the microbes in this anaerobic ecosystem are rarely characterized. For the daily-fed reactor, 10 of 18 spacers are from the GP6 CRISPRs (9 of system

1; 1 of system 2). For the hourly-fed reactor, 23 of 27 spacers are isolated from the repeat of GP6 CRISPR system 1.

Table 7.5 CRISPR of *Methanosaeta concilii* DSM 6752 and DSM 2139 (CRISPRFinder)

Sample	Group	CRISPR repeat sequence	Repeat (bp)	# spacers	Repeat homology
Methanosaeta concilii DSM 6752	1	CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	14	
		CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	2	*Methanosaeta*
		CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAACT	38	2	*concilii* GP6
		CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	4	CRISPR/Cas system
		CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	2	1
	consensus	CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	24	
	2	GTTTCAATTGGGCCACTTCTTAGCAGAAATGGATAG	36	8	
		GTTTCAATTGGGCCACTTCTTAGCAGAAATGGATAG	36	1	
		GTTTCAATTGGGCCACTTCTTAGCAGAAATGGATAG	36	3	
		GTTTCAATTGGGCCACTTCTTAGCAGAAATGGATAG	36	4	
		CAATTGGGCCACTTCTTAGCAGAAATGGATAG	32	3	
	consensus	GTTTCAATTGGGCCACTTCTTAGCAGAAATGGATAG	36	19	
	3	GGCTCTCGGAGAGATCGGCGATCCGAAAGCCATCGATCC	39	3	
		GGCTCTCGGAGAGATCGGCGATCCGAAAGCCATCGATCC	39	1	
	consensus	GGCTCTCGGAGAGATCGGCGATCCGAAAGCCATCGATCC	39	4	
	4	GTTTCAATTGGGCCACGTCCTTTCAGACATGGATAG	36	3	
		CTATCCATGTCTGAAAAGACGTGGCCCAATTGAAAC	36	2	
		CTATCCATGTCTGAAAGGACGTGGCCCAATTGAAAC	36	3	
	consensus	CTATCCATGTCTGAAAGGACGTGGCCCAATTGAAAC	36	8	
	5	CTATCCATGTCTGCGAAGACGTGGCCCAATTGAAAC	36	7	
			Σ	62	
Methanosaeta concilii DSM 2139	1	GTTTCAATTCCCTTTTCATCGGGATAGGCTCTGCAAT	37	2	*Methanosaeta*
		TGTTTCAATTCCCTTTTCATCGGGATAGGCTCTGCAAT	38	2	*concilii* GP6
		GTTTCAATTCCCTTTTCATCGGGATAGGCTCTGCAAT	37	1	CRISPR/Cas system
	consensus	GTTTCAATTCCCTTTTCATCGGGATAGGCTCTGCAAT	37	5	2
	2	GTTTCAATTGGGCCACGTCCTTTCAGACATGGATAG	36	4	
		GTTTCAATTGGGCCACGTCCTTTCAGACATGGATAG	36	4	
		GTTTCAATTGGGCCACGTCCTTTCAGACATGGATAGG	37	2	
	consensus	GTTTCAATTGGGCCACGTCCTTTCAGACATGGATAG	36	10	
	3	CCTATCCATGTCTGAAAAGACGTGGCCCAATTG	33	3	
	4	CTATCCATTTCTGCTAAGAAGTGGCCCAATTG	32	4	
		CTATCCATTTCTGCTAAGAAGTGGCCCAATTG	32	2	
	consensus	CTATCCATTTCTGCTAAGAAGTGGCCCAATTG	32	6	
	5	ATCCATTTCTGCGAAGAAGTGGCCCAATTGAAACAC	36	3	
	6	GTTTCAATTGGGCCACGTCTCTTCAGACATGGATAG	36	8	
		GTTTCAATTGGGCCACGTCTCTTCAGACATGGATAG	36	5	
	consensus	GTTTCAATTGGGCCACGTCTCTTCAGACATGGATAG	36	13	
	7	AAGAGCATGTTCTCGGAGATGTG	23	5	
			Σ	45	

Table 7.6 CRISPR of the daily-fed and hourly-fed enrichments (CRISPRFinder)

Sample	Group	CRISPR repeat sequence	Repeat (bp)	# spacers	Repeat homology
Daily-fed enrichment	1	CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAACC	38	2	
		CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	1	*Methanosaeta concilii* GP6 CRISPR/Cas system 1
		CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	2	
		CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAA	36	3	
		CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC		1	
	consensus	CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	9	
	2	GTTTCAATTCCCTTTTCATCGGGATAGGCTCTGCAAT	37	1	GP6 system 2
	3	GTTGCAACGAACTTGAAAACCCGATAGGGATTGAA	35	2	
		ACGTTGCAACGAACTTGAAAACCCGATAGGGATTGAAAC	39	1	
	consensus	GTTGCAACGAACTTGAAAACCCGATAGGGATTGAAAC		3	
	4	CAAGGATTGCCGTCATTGCAGTTCTTTTGGCTATGAGT	38	3	
	5	CCATCAAGAGCAAGTTCCACGAAAA	25	2	
			Σ	18	
Hourly-fed enrichment	1	GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG	37	3	
		GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG	37	7	
		GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATGG	38	3	*Methanosaeta concilii* GP6 CRISPR/Cas system 1
		GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG	37	3	
		TGTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG	38	2	
		GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG	37	2	
		GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG	37	3	
	consensus	GTTTCAATCCTTGTTTTCGTGGAACTTGCTCTTGATG	37	23	
	2	GTTTCAATCCCTATCGGGTTTTCAAGTTCGTTGCAAC	37	1	
		GTTTCAATCCCTATCGGGTTTTCAAGTTCGTTGCAAC	37	1	
		GTTTCAATCCCTATCGGGTTTTCAAGTTCGTTGCAA	36	2	
	consensus	GTTTCAATCCCTATCGGGTTTTCAAGTTCGTTGCAAC	36	4	
			Σ	27	

In the 4 samples investigated (2 pure strains and 2 reactors), more spacers were observed from system 1 than from system 2 (57 spacers from system 1and 8 spacers belong to system 2). This result is contrary to the data found in *Methanosaeta concilii* GP6 genome, in which there are 20 and 75 spacers in system 1 and 2, respectively. If only considering the spacers in reactors (where the viral sequences are present), the majority of the spacers belong to CRISPR system 1 (32 of 33). This observation suggests that CRISPR system 1 is more active than CRISPR system 2 in the environment with persistent viruses, with respect of the spacer acquisition. When challenged

with virus mixture, no extension was found in locus F of *S. solfataricus* P2 CRISPR subfamily I (Erdmann and Garrett 2012), which lacks a leader sequence (extension was found in the other 3 loci in the same CRISPR system). This might explain that more spacers were found in leader-containing system 1 than in leader-absent system 2. However, caution should be awared that the information of sequences flanking CRISPR loci is only available for strain GP6. Therefore, it remains unclear whether the lack of the leader sequence of *Methanosaeta* CRISPR system 2 is common in all strains or just a particular case in GP6.

CRISPR results generated from Crass are listed in Table 7.7. The number of repeat group for DSM 6752, DSM 2139, the daily-fed reactor, and the hourly-fed reactor are 4, 2, 12, and 9, respectively. The numbers are different compared to those found by the previous method, CRISPRFinder (5, 7, 5, and 2). GP6 CRISPR systems 1 and 2 were observed in DSM 6752 and DSM 2139, respectively. These two systems were found in both reactors using Crass but system 2 was not detected in the hourly-fed reactor using the CRISPRFinder. Sequences of the other repeats were blasted against the nr/nt database using megablast program (word size 7). Group 46 is homologous to CRISPR of *Candidatus* Cloacamonas acidaminovorans str. Evry (WWE1 candidate division). Group 28 of the daily-fed reactor and group 1 of the hourly-fed reactor were identical, and homology of the repeat was found in 4 microbes: *Methylomonas methanica*

MC09/*Ramlibacter tataouinensis* TTB310/*Pseudomonas stutzeri* A1501/*Chlorobium luteolum* DSM 273. The number of spacers found in reactors dramatically increased. There were 99, 45, 262, and 263 spacers found in DSM 6752, DSM 2139, the daily-fed reactor, and the hourly-fed reactor, respectively (62, 45, 18, and 27 found using CRISPRFinder). Again, in the reactors, more spacers were identified in CRISPR system 1 (202) than system 2 (38).

Table 7.7 CRISPR of two pure strains and two reactors (Crass)

Sample	GID	Repeat concensus	Repeat Length	# pacers	SP Length	
6752	G1	CTATCCATATCCATTTCTGCTAAGAAGTGGCCCAATTGAAAC	42	15	32	
	G2	CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	62	36	GP6 system 1
	G6	CGTTTCAATTGGGCCACGACTTCTCAGTCATGGATAG	37	7	37	
	G8	CTATCCATGTCTGAAAGGACGTGGCCCAATTGAAAC	36	15	36	
	Σ	4		99		
2139	G1	CTATCCATATCCATTTCTGCAAAGAAGTGGCCCAATTGAAAC	42	28	34	
	G5	ATTGCAGAGCCTATCCCGATGAAAAGGGAATTGAAAC	37	17	38	GP6 system 2
	Σ	2		45		
Daily	G1	CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	96	36	GP6 system 1
	G2	ATTGCAGAGCCTATCCCGATGAAAAGGGAATTGAAAC	37	25	38	GP6 system 2
	G3	AGGTCGATCCCCGCGCATGCGGGGTAGCC	29	7	31	
	G5	GTTGCAACGAACTTGAAAACCCGATAGGGATTGAAAC	37	73	36	
	G12	GTCAGAATCAATAAATGCCCGCCAGGGCATTAAGAC	36	8	36	
	G15	GTTTCAATTCCTTATAGGTAGGCTACAAAC	30	3	36	
	G28	GTCGCGCCCCACGCGGGCGCGTGGATTGAAAC	32	15	34	++
	G30	CGGTTCATCCCCGCGTGAGCGGGGAACACG	30	10	31	
	G46	GTTTCAATCCTTGATTTAATGGACTGTAGAGTAAAAAC	38	6	35	**
	G48	ATTTCCATGGCTGAAAAGTCATGGCCACATTGAAGC	36	8	37	
	G50	GTTGTGGTTTGATAAGAAATTGGAATAGAAGACATTA	37	8	29	
	G63	AGTCGCCCGCTCACGAGCGGGCGTGGATTGAAACCA	36	3	31	
	Σ	12		262		
Hourly	G1	GTCGCGCCCCACGCGGGCGCGTGGATTGAAAC	32	25	34	++
	G2	CATCAAGAGCAAGTTCCACGAAAACAAGGATTGAAAC	37	106	36	GP6 system 1
	G5	GTTGCAACGAACTTGAAAACCCGATAGGGATTGAAAC	37	79	36	
	G6	GTACCCAGCTTACCTATAAGGAATTGAAAC	30	8	35	
	G7	GTTTCAATTGGGCCACTTCTTAGCAGAAATGGATAGA	37	20	36	
	G12	ATTGCAGAGCCTATCCCGATGAAAAGGGAATTGAAAC	37	13	38	GP6 system 2
	G17	AATGTCTTCTATTCCAATTTCTTATCAAACCACAAC	36	6	30	
	G20	CGGTTCATCCCCGCGTGTGCGGGGAACGCGGGGAACAC	38	3	23	
	G27	CGGTCCACCCCCACGCGTGTGGGGATAAC	29	3	33	
	Σ	9		263		

**: homologous repeat found in *Candidatus* Cloacamonas acidaminovorans str. Evry (WWE1 candidate division)

++: homologous repeat found in *Methylomonas methanica* MC09/*Ramlibacter tataouinensis* TTB310/*Pseudomonas stutzeri* A1501/*Chlorobium luteolum* DSM 273

Methods Comparison

The CRISPR identification methods were compared with respect to the number of repeats and spacers detected. The result is shown in Figure 7.5. A clear trend, that the Crass approach is more

sensitive in more complex data than CRIPSRFinder except for the spacers in DSM 6752 (i.e. more repeats and spacers were found in reactors than in pure strains using Crass) was observed.

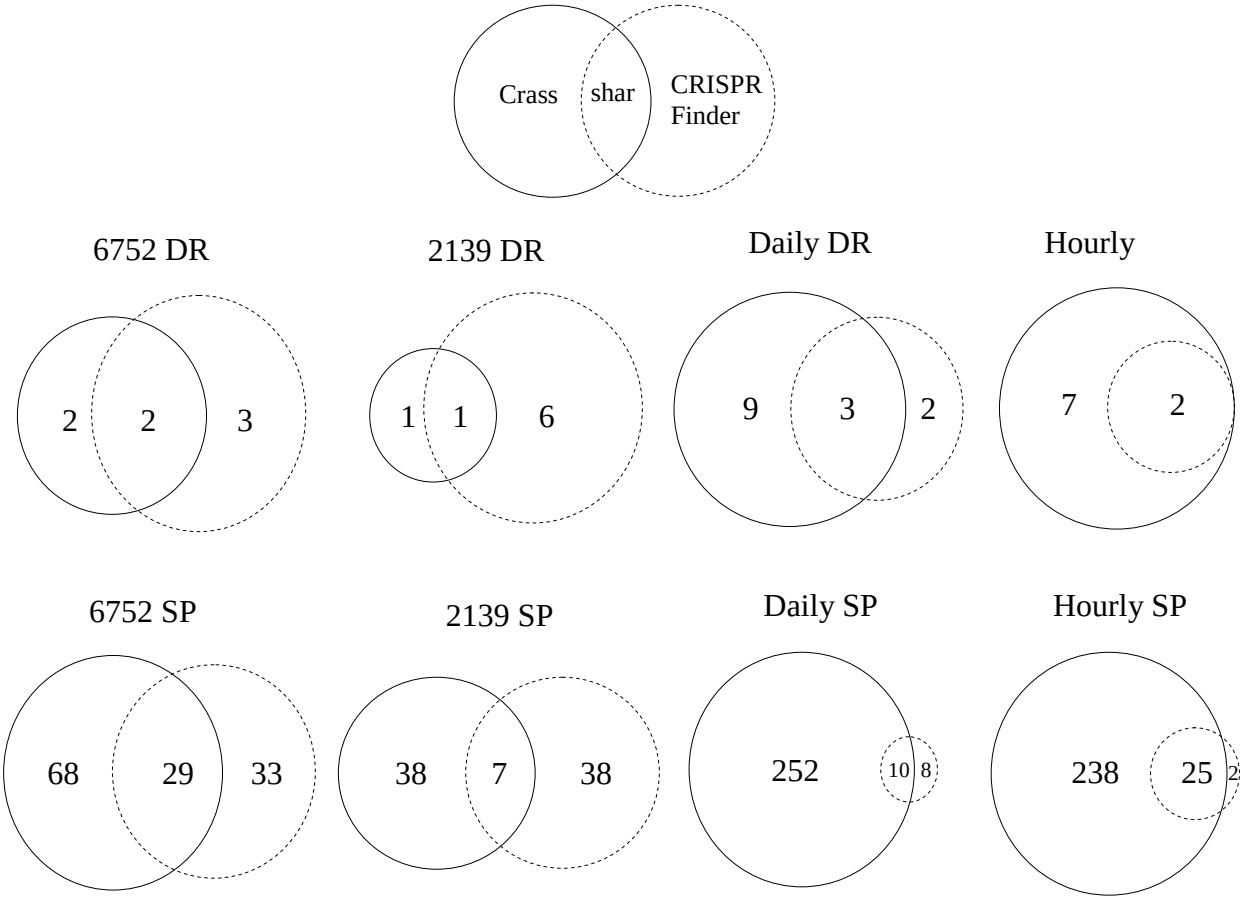


Figure 7.5 Comparison of two CRISPR identification methods (Crass and CRISPRFinder)

It is noteworthy that the current data is caused not only by different tools, but also the distinct data formats used (i.e. reads before (Crass) and after assembling (CRISPRFindr)). The effect of these two factors on the present result is indistinguishable, but Crass has been demonstrated to be more sensitive and specific than other available tools (Skennerton, Imelfort et al. 2013) when

same metagenomes and data format (reads or contigs) were used. In conclusion, when studying the CRISPR, it is better to use Crass for unassembled reads, especially for data obtained from the complex microbial communities.

7.3.4 Studying the Relationship between Viruses and Hosts using CRISPR

Spacer sequences identified by Crass were used as queries to blast against several viral metagenomic libraries, including two reactors (daily-fed and hourly-fed reactors) and three full-scale digesters (West Point, South Plant, and Port Angeles). In addition, the spacer sequences of GP6 were also used. The viral contigs containing a short segment of DNA (protospacers) matching to spacer sequences were retrieved. If several viral contigs were tagged by the same spacer, these contigs were assembled (if possible) into longer sequences. Finally, these protospacer-containing contigs were blasted against the nucleotide collection (nt/nr) database using blastn as well as the non-redundant protein sequences (nr) database using blastx to find homologous sequences or proteins, and hopefully to reveal the putative viruses of *Methanosaeta concilii*. The spacers harvested exclusively from CRISPRFinder did not have any homology in the database, so they were not included.

Protospacer-containing Viral Contigs (*Methanosaeta concilii* CRISPR)

Protospacer-containing viral contigs were identified if contigs contained significant homologies to spacer sequences (100% coverage; mismatch < 2 nt). The result is summarized in Table 7.8. For the *Methanosaeta concilii* strains (DSM 6752, DSM 2139, and GP6), about 10% of the spacers matched the viral contigs containing protospacers. For the spacers extracted from the reactors, the percentages of matched spacers increased to 30%. It is not surprising to have higher percentage of matches for the reactors, since both the spacers and viral contigs were isolated either from the same environment or from related samples. However, even for spacers recovered from the reactors, many of the spacers did not match to any viral contig (70%). It is likely that the spacers were acquired from other mobile genetic elements, such as conjugative plasmids. Since plasmids were likely removed during the virus purification processes (CsCl density gradient or the DNase treatment), they would not be expected to be found in the viral metagenomic libraries. In addition, a fraction of the putative viral populations might be lost due to the selective pressure causing by the present of immune *Methanosaeta* (CRISPR-mediated) in the systems.

Table 7.8 Blast results (*Methanosaeta concilii* spacers against viral metagenomic libraries)

sample	# spacer	match			no match
		overall	contigs with annotated gene	unknown viral contigs	
DSM 6752	99	9 (9%)	3 (3%)	6 (6%)	90 (91%)
DSM 2139	45	5 (11%)	2 (4%)	3 (7%)	40 (89%)
GP6	87	9 (10%)	1 (1%)	8 (9%)	78 (90%)
Daily	121	40 (33%)	3 (2%)	37 (31%)	81 (67%)
Hourly	119	32 (27%)	1 (1%)	31 (26%)	87 (73%)

To understand which systems contained putative viruses of *Methanosaeta*, protospacer-containing contigs were sorted by individual viral metagenomic library (Figure 7.6). All viral metagenomes except Port Angeles have contigs that match *Methanosaeta* spacer sequences. Although CRISPR/Cas systems could provide *Methanosaeta* immunity against the viruses when matched spacers are present, viruses of *Methanosaeta* might still active in these systems for several reasons. First, when searching putative viruses, mismatch was allowed between spacer and protospacer. It has been reported that viruses with a single nucleotide mutation in the protospacer region could escape CRISPR-mediate immunity (Deveau, Barrangou et al. 2008). Second, even with perfectly matched spacers, viruses can still circumvent the CRISPR/Cas system by mutating in PAM regions (Deveau, Barrangou et al. 2008; Semenova, Nagornykh et al. 2009). Third, the CRISPR loci were rapidly evolved and diverse in the level of strain (Zheng, Roberts et al. 2004; Andersson and Banfield 2008; Tyson and Banfield 2008),

suggesting that viruses could propagate by infecting a portion of the host population.

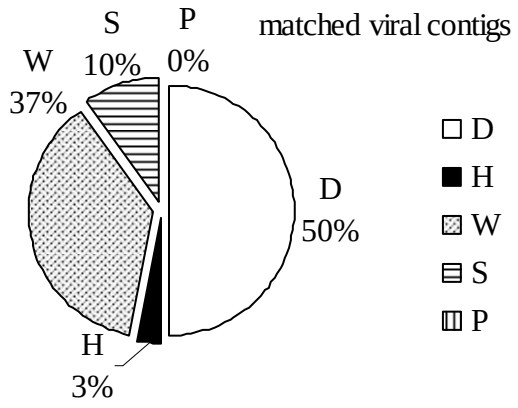


Figure 7.6 Distribution of protospacer-containing viral contigs specific to *Methanosaeta* spacers

Protospacer-containing Viral Contigs (other microorganisms)

Viral contigs matched to spacers isolated from CRISPRs of other microbes were also obtained (Table 7.9). This suggests that some of the observed VLPs infecting microbial hosts in enrichments were still produced. For example, viruses might propagate in the daily-fed reactor by infection a microbe belonging to WWE1 candidate division (Figure 7.9: repeat group 46). There are 16 and 17 viral contigs tagged by the spacers from the daily-fed and hourly-fed reactors, respectively. In addition to the reactors, these viral contigs were also found in the full-scale digesters, suggesting their corresponding hosts are prevalent in the full-scale anaerobic digesters.

Table 7.9 Spacer-matched viral contigs from other microbial CRISPRs

Sample	Repeat group	# matched viral contig	matched viral metagenome				
			D	H	W	S	P
D	5	11	7		2	2	
	46*	3	2			1	
	3	1					1
	28**	1				1	
H	5	9	6	1	1	1	
	1**	6			2	4	
	7	1			1		
	20	1				1	

*: homologous repeat found in *Candidatus* Cloacamonas acidaminovorans str. Evry (WWE1 candidate division)

**: homologous repeat found in *Methylomonas methanica* MC09/*Ramlibacter tataouinensis* TTB310/*Pseudomonas stutzeri* A1501/*Chlorobium luteolum* DSM 273

Homologous Sequences of Viral Contigs

Spacer-tagged viral contigs were searched for homologies in the NCBI database. First, matched viral nucleotide sequences were used as queries to search similar sequences in nucleotide databases using blastn, but no significant homologous sequence was found. Therefore, the viral nucleotide sequences were translated into protein sequences and used as queries to search the non-redundant protein database using blastx. The identified homologous sequences are summarized in Table 7.10. There are 15 homologous sequences found containing either annotated functions or predicted hypothetical proteins. Among them, 8 proteins were targeted by spacers from *Methanosaeta* CRISPR, and 7 proteins were matched to spacers from other CRISPRs. Although many viral contigs were targeted by CRISPR spacers, most of the matched

viral contigs have not been annotated yet. For example, 30% of reactor *Methanosaeta* spacers matched to viral contigs, but only less than 3% of the spacers linked to annotated viral sequences (Table 7.8), which suggests that most viruses inhabiting in the methanogenic environment have not yet been characterized.

By analyzing the genes in contigs, the viral contigs targeted by *Methanosaeta* spacers could reveal their origins. One of the invaders belongs to viruses, due to the identification of a viral structure protein (a capsid protein of deep-sea thermophilic phage D6E) and several viral enzymes including the viral terminase, methyltransferase, and recombinase (Table 7.10). The phage D6E was isolated from its host, *Geobacillus* sp., cultured at 65^{o}C. It is a myovirus with a 60 nm diameter capsid and 120 nm tail (Figure 7.7), and has a 49,335 bp dsDNA genome (Wang and Zhang 2010). Although D6E belongs to the most frequently observed head-tail phage, its genomic sequence is unique among other known phages. In addition, the gene cluster of transcription and replication of D6E is highly similar to a thermophilic siphovirus GVE2, despite their low sequence similarity. Moreover, the D6E has an extensive mosaic genome, which has been reported in other mesophilic phages. Similarities were observed between the phage D6E and the viruses from the studied reactors and digesters: (1) these viruses both inhabit in the anaerobic environments with elevated temperature (mesophilic and thermophilic), (3) viral

genomes have low sequence similarities to known phages (Wang and Zhang 2010) (low significant taxonomic matches for studied viral contigs, see appendix Table A.2), and (3) VLPs with similar morphology were also observed in the hourly-fed reactor (Figure 7.7 and Figure 6.1 j).

In addition to the phage structure protein, several viral enzymes were also targeted by *Methanosaeta* CRISPR spacers for example, terminase, methyltransferase, and recombination protein. The phage terminase (large subunit) is a late protein and is in charge of translocating DNA in viral genome packaging. Methyltransferases are responsible for DNA modification probably to avoid host restriction systems or regulation.

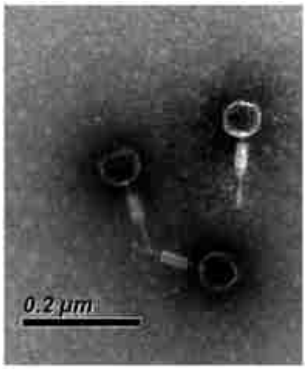


Figure 7.7 The deep-sea thermophilic phage D6E (Wang and Zhang 2010)

Table 7.10 Homologous sequences of spacer-matched viral contigs

Spacer ID	Identical CRISPR	# matched viral contig	# mismatch[b]			Blastx: Non-redundant protein sequences (nr)[c]
			0	1	2	
DG1SP1407	GP6 system I	6	D W S	D W S		capsid protein [Deep-sea thermophilic phage D6E][d]
DG1SP1215	GP6 system I	6	D		D W(2) S(2)	capsid protein [Deep-sea thermophilic phage D6E][d]
HG2SP971	GP6 system I	2		W S		phage recombination protein [*Methanomethylovorans hollandica*]
6752G2SP6	GP6 system I	1	W			putative phage protein, terminase, large subunit [*Paenibacillus alvei*]
6752G2SP272	GP6 system I	1			D	hypothetical protein ABC2837 [*Bacillus clausii* KSM-K16]
6752G1SP580	unknown	2		D		adenine DNA methyltransferase [*Bordetella* phage BPP-1]
2139G1SP791	unknown	1	D			transposase, IS891/IS1136/IS1341 [*Methanococcoides burtonii*]
2139G5SP149	GP6 system II	1	W			hypothetical protein MCON_3328 [*Methanosaeta concilii GP6*]
DG5SP856	unknown	1			D	prophage lambdach01 portal protein [*Caldicellulosiruptor owensensis*]
DG5SP1060	unknown	1			D	hypothetical protein [Environmental Halophage eHP-38_00075]
DG28SP23	4 bacteria[a]	1			S	phage terminase-like protein, large subunit [*Spirochaeta africana*]
HG1SP125	4 bacteria[a]	5		W S (4)		phage-related hypothetical protein[*Sphaerochaeta pleomorpha*]
HG5SP621	unknown	2		D H		hypothetical protein MCON_3328 [*Methanosaeta concilii GP6*]
HG5SP143	unknown	2		D	W	prophage lambdach01 portal protein [*Caldicellulosiruptor owensensis*]
HG5SP959	unknown	1			D	hypothetical protein [*Methanobacterium formicicum*]

(a): *Ramlibacter tataouinensis* TTB310; *Methylomonas methanica* MC09; *Pseudomonas stutzeri* A1501; *Chlorobium luteolum* DSM 273

(b): D: daily-fed reactor, H: hourly-fed reactor, W: West Point WWTP, S: South WWTP,

(c): Maximium E value: 6E-6

(d): (d): The closest relative sequence of the capsid protein [Deep-sea phage D6E] is MCON_1983 [*Mst. concilii* GP6] (Max identity: 57%; E value: 8e-110)

In contrast to contigs with annotated viral DNA, several contigs contained predicted proteins of prokayotic origin, and thus raise several questions. First, are these true viral sequences (e.g. prophage or microbial genes captured and carried by viruses) or microbial genomes contaminated during virus particle purification? If they are microbial DNA, why they are targeted by CRISPR? Why does the *Methanosaeta* CRISPR spacer sequence matches to a gene of *Methanosaeta concilii* GP6 (hypothetical protein MCON_3328 of *Methanosaeta concilii* GP6) (Table 7.10)? An interesting finding shows that the closest similar sequence of the capsid protein (ADE87491.1, Deep-sea phage D6E) is the hypothetical protein, MCON_1983, found in the genome of *Methanosaeta concilii* GP6. Therefore, this hypothetical protein might be a relic of prophage, and perhaps the case is the same for the hypothetical protein MCON_3328. Moreover, identical sequences were observed between *Streptococcus thermophilus* spacers (7%) and its own genome which suggests that CRISPRs might involve in microbial regulatory systems (Horvath, Romero et al. 2008).

Similarly, the contigs matched by other microbial CRISPR spacers reveal their viral origin. Viral proteins (portal protein, terminase) were identified in all but one contig, which contain hypothetical protein MCON_3328 of *Methanosaeta concilii* GP6. To sum up, current evidence suggests that unidentified *Methanosaeta* viruses are propagating in the investigated systems (except for the hourly-fed reactor). In addition to *Methanosaeta* in reactors, several microbes appear to be the hosts of other viruses.

Protospacer Adjacent Motif (PAM)

Protospacer adjacent motifs have been identified in many systems (Horvath, Romero et al. 2008;

Lillestol, Shah et al. 2009; Mojica, Diez-Villasenor et al. 2009). This motif probably serves as a recognition site for the CRISPR/Cas system to select target sequences. To search the PAMs of *Methanosaeta* CRISPR/Cas systems, spacer-matched viral contigs were aligned, and results are shown in Figure 7.8. About 65 contigs were aligned for CRISPR system 1, and a TTG motif immediately upstream of the protospacers was revealed. Similarly, a CC motif located1 nt upstream of the protospacers was identified for *Methanosaeta* CRISPR system 2. The appearance of these motifs suggests the protospacers are not randomly selected.

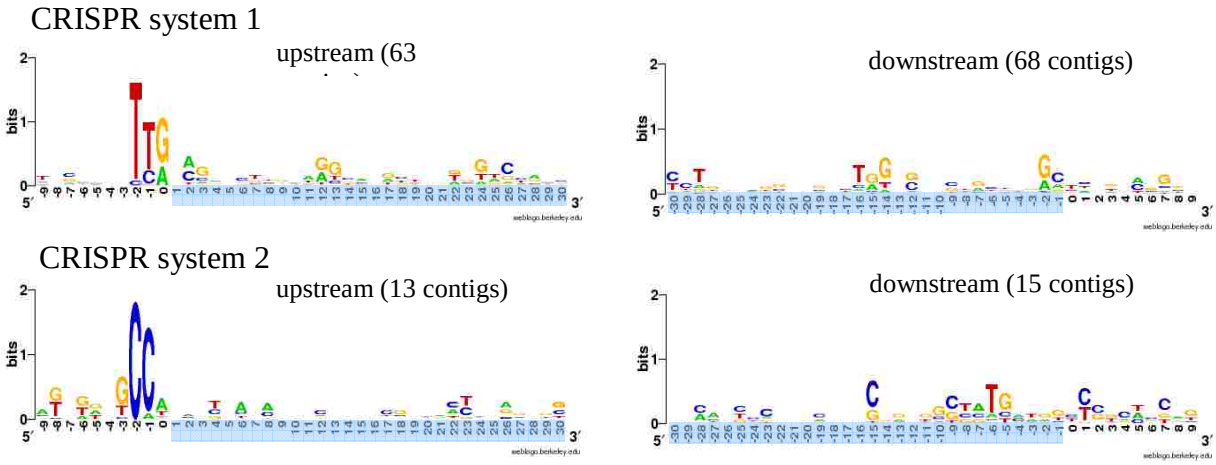


Figure 7.8 Conserved proto-spacer adjacent motifs (PAMs) of *Methanosaeta* CRISPR/Cas systems

7.4 Conclusions

Two CRISPR/Cas systems were found in *Methanosaeta concilii* GP6 genome. System 1 is classified as type I-B, and system 2 belongs to type I. These two CRISPR/Cas systems in strain GP6 probably are both active with respect to the appearance of repeat-spacer units and *cas* genes responsible for all three immunity processing steps. However, the lack of a leader sequence in systems 2 suggests it might not be functional.

CRISPR/Cas systems of DSM 6752 might be defective because most of the *cas* genes are absent, and the whole system 2 is lost. For DSM 2139, system 1 is totally missing while the system 2 is intact. The observed variation of CRISPR/Cas systems among strains probably can be explained by complex genome rearrangements in the CRISPR regions via several *transposase* families.

When identifying and isolating spacers of CRISPR, it would be better to use Crass for unassembled reads, especially for the data obtained from the complex microbial communities. Using this approach, GP6 CRISPR 1 and 2 systems were observed in both reactors. In addition, CRISPR of the other microbes were also detected.

Higher percentage of matching between spacers and viral contigs was found when spacers were isolated from reactors rather than from pure strains (DSM 6752, DSM 2139, and GP6). This is not surprising since both the spacers and viral contigs were isolated either from the same environments or from related samples. However, even for spacers recovered from the reactors, many of the spacers did not match to any viral contig (70%). It is likely that the spacers were

acquired from other mobile genetic elements, such as conjugative plasmids. In addition, it is possible that a fraction of the putative viral populations were lost due to the selective pressure causing by the presence of immune *Methanosaeta* in the systems. *Methanosaeta* spacers matched viral contigs were found in the two reactors (daily-fed and hourly-fed reactors) and two digesters (West Point and South Plant), suggesting that the *Methanosaeta* viruses might propagate in these systems.

30% of reactor *Methanosaeta* spacers matched to viral contigs, but only less than 3% of the spacers linked to annotated viral sequences, indicating that most viruses that inhabit in the methanogenic environment have not yet been characterized. Presence of various phage proteins in contigs targeted by *Methanosaeta* spacers suggests these sequences are originated from *Methanosaeta* viruses. This relationship is further confirmed by finding conserved PAMs on viral contigs targeted by both CRISPR/Cas systems. In addition to *Methanosaeta* in reactors, several microbes were the hosts of other viruses.

Chapter 8 Conclusions

This study is the first to identify viruses of *Methanosaeta* in anaerobic digesters and laboratory reactors. In the course of this study, culturing techniques for *Methanosaeta concilii* in both solid and liquid media have been greatly improved and applied into the investigation of *Methanosaeta* viruses. Despite the fact that viruses of *Methanosaeta* remain recalcitrant to culture, molecular results clearly suggest *Methanosaeta* viruses propagate in two laboratory *Methanosaeta*-dominated reactors and two full-scale anaerobic digesters.

Media and methods for culturing *Methanosaeta* have been improved. By substituting potential toxicants (sodium sulfide, cysteine, NTA, and agar) for modified compounds (titanium (III) citrate, sodium citrate and gellan gum), the growth of *Methanosaeta* was observed after one week using pour plating and one month using spread plating in serum bottles (0.5% gellan, 35^{o}C). Ammonia toxicity to *Methanosaeta* can be avoided with concentration of NH_3-N in media below 16.8 mg/L. This can be achieved by reducing input TAN and controlling pH (e.g. using carbonate buffer system). Based on the improved culturing methods, two viral isolation approaches for *Methanosaeta*, plaque assays and infection tests, were evaluated but no virus was found.

Persistent VLPs were present in the *Methanosaeta*-dominated daily-fed and hourly-fed reactors suggesting these VLPs are likely to infect their corresponding hosts in reactors such as *Methanosaeta*. This implication is supported by detecting putative *Methanosaeta* viruses in both reactors using metagenomic and CRISPR approaches. In studied reactors, *Methanosaeta* viruses might be head-tailed morphology and possibly belong to family *Siphoviridae*. Due to low VBR

found in reactors, lysogenic or pseudolysogenic life cycles might be the favorable strategy for *Methanosaeta* viruses to target slow-growing hosts in the methanogenic environment.

In addition to enrichment reactors, putative *Methanosaeta* viruses were also detected in two full-scale anaerobic digesters at West Point and South Plant WWTPs. The detection of *Methanosaeta* viruses using the CRISPR approach was supported by finding viral proteins including capsid, terminase, methyltransferase, and recombinase on spacer-matched contigs and by identifying a conserved TTG motif and a CC motif *Methanosaeta* in CRISPR system 1 and 2, respectively.

In this study, culture methods for *Methanosaeta* were improved and the present of putative *Methanosaeta* viruses in reactors and digesters was shown. To understand the correlation between *Methanosaeta* viruses and the stability of digestion processes, further research is required. PCR primers or probes targeting spacer-matched viral contig (putative *Methanosaeta* viruses) could be designed and used to monitor specific viral population dynamics before, during and after digester failure. Response of *Methanosaeta* to viral attack through out the upset can be monitored by studying the change of CRISPR spacers near leader regions using PCR-sequencing approaches. Method developments for monitoring viral dynamics and CRISPR spacer acquisitions can be tested in viral challenge experiments such as infection tests. In addition, for *Methanosaeta* viruses, identifying inducing agents which trigger viral particle propagation might be a useful factor to improve stability of digester processes.

Reference

Ackermann, H. W. (1998). "Tailed bacteriophages: the order caudovirales." Adv Virus Res **51**: 135-201.

Ackermann, H. W. (2007). "5500 Phages examined in the electron microscope." Arch Virol **152**(2): 227-243.

Ackermann, H. W. (2009). "Basic phage electron microscopy." Methods Mol Biol **501**: 113-126.

Ackermann, H. W. and M. S. DuBow (1987). "Viruses of prokaryotes. In: General Properties of Bacteriophages." CRC Press, Inc. Boca Raton, Fla. **Vol. I**.

Aller, J. Y. and P. F. Kemp (2008). "Are Archaea inherently less diverse than Bacteria in the same environments?" Fems Microbiology Ecology **65**(1): 74-87.

Andersson, A. F. and J. F. Banfield (2008). "Virus population dynamics and acquired virus resistance in natural microbial communities." Science **320**(5879): 1047-1050.

Angenent, L. T., D. D. Zheng, S. H. Sung and L. Raskin (2002). "Microbial community structure and activity in a compartmentalized, anaerobic bioreactor." Water Environment Research **74**(5): 450-+.

Appels, L., J. Baeyens, J. Degreve and R. Dewil (2008). "Principles and potential of the anaerobic digestion of waste-activated sludge." Progress in Energy and Combustion Science **34**(6): 755-781.

Appels, L., J. Lauwers, J. Degreve, L. Helsen, B. Lievens, K. Willems, J. Van Impe and R. Dewil (2011). "Anaerobic digestion in global bio-energy production: Potential and research challenges." Renewable & Sustainable Energy Reviews **15**(9): 4295-4301.

Ariesyady, H. D., T. Ito and S. Okabe (2007). "Functional bacterial and archaeal community structures of major trophic groups in a full-scale anaerobic sludge digester." Water Research **41**(7): 1554-1568.

Ashelford, K. E., M. J. Day and J. C. Fry (2003). "Elevated abundance of bacteriophage infecting bacteria in soil." Appl Environ Microbiol **69**(1): 285-289.

Barber, R. D., L. Zhang, M. Harnack, M. V. Olson, R. Kaul, C. Ingram-Smith and K. S. Smith (2011). "Complete genome sequence of Methanosaeta concilii, a specialist in aceticlastic methanogenesis." J Bacteriol **193**(14): 3668-3669.

Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero and P. Horvath (2007). "CRISPR provides acquired resistance against viruses in prokaryotes." Science **315**(5819): 1709-1712.

Barrick, J. E., D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh, D. Schneider, R. E. Lenski and J. F. Kim (2009). "Genome evolution and adaptation in a long-term experiment with Escherichia coli." Nature **461**(7268): 1243-U1274.

Bergh, O., K. Y. Borsheim, G. Bratbak and M. Heldal (1989). "High abundance of viruses found in aquatic environments." Nature **340**(6233): 467-468.

Bhattacharya, S. K. and G. F. Parkin (1989). "The effect of ammonia on methane fermentation processes." Journal Water Pollution Control Federation **61**(1): 55-59.

Bland, C., T. L. Ramsey, F. Sabree, M. Lowe, K. Brown, N. C. Kyrpides and P. Hugenholtz (2007). "CRISPR Recognition Tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats." BMC Bioinformatics **8**.

Bolotin, A., B. Ouinquis, A. Sorokin and S. D. Ehrlich (2005). "Clustered regularly interspaced

short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin." Microbiology-Sgm **151**: 2551-2561.
Bombardiere, J., T. Espinosa-Solares, M. Domaschko and M. Chatfield (2007). "Thermophilic anaerobic digester performance under different feed-loading frequency." Appl Biochem Biotechnol **137-140**(1-12): 765-775.
Brauer, S. L., E. Yashiro, N. G. Ueno, J. B. Yavitt and S. H. Zinder (2006). "Characterization of acid-tolerant H-2/CO2-utilizing methanogenic enrichment cultures from an acidic peat bog in New York State." Fems Microbiology Ecology **57**(2): 206-216.
Bult, C. J., O. White, G. J. Olsen, L. X. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. M. Geoghagen, J. F. Weidman, J. L. Fuhrmann, D. Nguyen, T. R. Utterback, J. M. Kelley, J. D. Peterson, P. W. Sadow, M. C. Hanna, M. D. Cotton, K. M. Roberts, M. A. Hurst, B. P. Kaine, M. Borodovsky, H. P. Klenk, C. M. Fraser, H. O. Smith, C. R. Woese and J. C. Venter (1996). "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii." Science **273**(5278): 1058-1073.
Carbonero, F., B. B. Oakley and K. J. Purdy (2010). "Improving the isolation of anaerobes on solid media: The example of the fastidious *Methanosaeta*." Journal of Microbiological Methods **80**(2): 203-205.
Chen, Y., J. J. Cheng and K. S. Creamer (2008). "Inhibition of anaerobic digestion process: A review." Bioresource Technology **99**(10): 4044-4064.
Chouari, R., D. Le Paslier, C. Dauga, P. Daegelen, J. Weissenbach and A. Sghir (2005). "Novel major bacterial candidate division within a municipal anaerobic sludge digester." Applied and Environmental Microbiology **71**(4): 2145-2153.
Cochlan, W. P. W., J.
Steward, G. F., D. C. Smith and F. Azam (1993). "Spatial distribution of viruses, bacteria and chlorophyll a in neritic, oceanic and estuarine environments " Mar. Ecol. Prog. Ser. **92**: 77-87.
Conklin, A., H. D. Stensel and J. Ferguson (2006). "Growth kinetics and competition between Methanosarcina and Methanosaeta in mesophilic anaerobic digestion." Water Environ Res **78**(5): 486-496.
Conrad, R. (1999). "Contribution of hydrogen to methane production and control of hydrogen concentrations in methanogenic soils and sediments." Fems Microbiology Ecology **28**(3): 193-202.
Crooks, G. E., G. Hon, J. M. Chandonia and S. E. Brenner (2004). "WebLogo: A sequence logo generator." Genome Research **14**(6): 1188-1190.
Daniels, L. L. and A. C. Wais (1990). "Ecophysiology of Bacteriophage S5100 Infecting *Halobacterium cutirubrum*." Applied and Environmental Microbiology **56**(11): 3605-3608.
Debaere, L. A., M. Devocht, P. Vanassche and W. Verstraete (1984). "Influence of high NaCl salt levels on methanogenic associations." Water Research **18**(5): 543-548.
DeBoy, R. T., E. F. Mongodin, J. B. Emerson and K. E. Nelson (2006). "Chromosome evolution in the Thermotogales: Large-scale inversions and strain diversification of CRISPR sequences." Journal of Bacteriology **188**(7): 2364-2374.
DeLong, E. F. (1992). "Archaea in coastal marine environments." Proc Natl Acad Sci U S A **89**(12): 5685-5689.
Deng, W., D. C. Nickle, G. H. Learn, B. Maust and J. I. Mullins (2007). "ViroBLAST: a

stand-alone BLAST web server for flexible queries of multiple databases and user's datasets." Bioinformatics **23**(17): 2334-2336.
Denman, S. E. and C. S. McSweeney (2006). "Development of a real-time PCR assay for monitoring anaerobic fungal and cellulolytic bacterial populations within the rumen." FEMS Microbiol Ecol **58**(3): 572-582.
Deveau, H., R. Barrangou, J. E. Garneau, J. Labonte, C. Fremaux, P. Boyaval, D. A. Romero, P. Horvath and S. Moineau (2008). "Phage response to CRISPR-Encoded resistance in Streptococcus thermophilus." Journal of Bacteriology **190**(4): 1390-1400.
Doddema, H. J. and G. D. Vogels (1978). "Improved identification of methanogenic bacteria by fluorescence microscopy." Appl Environ Microbiol **36**(5): 752-754.
Elberson, M. A. and K. R. Sowers (1997). "Isolation of an aceticlastic strain of Methanosarcina siciliae from marine canyon sediments and emendation of the species description for Methanosarcina siciliae." International Journal of Systematic Bacteriology **47**(4): 1258-1261.
Eldem, N. O., I. Ozturk, E. Soyer, B. Calli and O. Akgiray (2004). "Ammonia and pH inhibition in anaerobic treatment of wastewaters, part I: Experimental." Journal of Environmental Science and Health Part a-Toxic/Hazardous Substances & Environmental Engineering **39**(9): 2405-2420.
Emerson, K., R. C. Russo, R. E. Lund and R. V. Thurston (1975). "Aqueous Ammonia Equilibrium Calculations: Effect of pH and Temperature." Journal of the Fisheries Research Board of Canada **32**(12): 2379-2383.
Erdmann, S. and R. A. Garrett (2012). "Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms." Molecular Microbiology **85**(6): 1044-1056.
Filippini, M. and M. Middelboe (2007). "Viral abundance and genome size distribution in the sediment and water column of marine and freshwater ecosystems." FEMS Microbiol Ecol **60**(3): 397-410.
Fogel, G. B., C. R. Collins, J. Li and C. F. Brunk (1999). "Prokaryotic Genome Size and SSU rDNA Copy Number: Estimation of Microbial Relative Abundance from a Mixed Population." Microb Ecol **38**(2): 93-113.
Frols, S., P. M. K. Gordon, M. A. Panlilio, C. Schleper and C. W. Sensen (2007). "Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays." Virology **365**(1): 48-59.
Garrett, R. A., G. Vestergaard and S. A. Shah (2011). "Archaeal CRISPR-based immune systems: exchangeable functional modules." Trends in Microbiology **19**(11): 549-556.
Grissa, I., G. Vergnaud and C. Pourcel (2007a). "CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats." Nucleic Acids Research **35**: W52-W57.
Grissa, I., G. Vergnaud and C. Pourcel (2007b). "The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats." BMC Bioinformatics **8**.
Grosskopf, R., P. H. Janssen and W. Liesack (1998). "Diversity and structure of the methanogenic community in anoxic rice paddy soil microcosms as examined by cultivation and direct 16S rRNA gene sequence retrieval." Appl Environ Microbiol **64**(3): 960-969.
Guo, L., K. Brugger, C. Liu, S. A. Shah, H. J. Zheng, Y. Q. Zhu, S. Y. Wang, R. K. Lillestol, L. M. Chen, J. Frank, D. Prangishvili, L. Paulin, Q. X. She, L. Huang and R. A. Garrett (2011). "Genome Analyses of Icelandic Strains of Sulfolobus islandicus, Model Organisms for Genetic and Virus-Host Interaction Studies." Journal of Bacteriology **193**(7): 1672-1680.

Hajarnis, S. R. and D. R. Ranade (1993). "Revival of ammonia inhibited cultures of *Methanobacterium bryantii* and *Methanosarcina barkeri*." Journal of Fermentation and Bioengineering **76**(1): 70-72.
Hale, C., K. Kleppe, R. M. Terns and M. P. Terns (2008). "Prokaryotic silencing (psi) RNAs in Pyrococcus furiosus." Rna-a Publication of the Rna Society **14**(12): 2572-2579.
Hatamoto, M., H. Imachi, A. Ohashi and H. Harada (2007). "Identification and cultivation of anaerobic, syntrophic long-chain fatty acid-degrading microbes from mesophilic and thermophilic methanogenic sludges." Applied and Environmental Microbiology **73**(4): 1332-1340.
Hermann, M., K. M. Noll and R. S. Wolfe (1986). "Improved agar bottle plate for isolation of methanogens or other anaerobies in a defined gas atmosphere." Applied and Environmental Microbiology **51**(5): 1124-1126.
Hildenbrand, C., T. Stock, C. Lange, M. Rother and J. Soppa (2011). "Genome Copy Numbers and Gene Conversion in Methanogenic Archaea." Journal of Bacteriology **193**(3): 734-743.
Horvath, P. and R. Barrangou (2010). "CRISPR/Cas, the Immune System of Bacteria and Archaea." Science **327**(5962): 167-170.
Horvath, P., D. A. Romero, A.-C. Coute-Monvoisin, M. Richards, H. Deveau, S. Moineau, P. Boyaval, C. Fremaux and R. Barrangou (2008). "Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus." Journal of Bacteriology **190**(4): 1401-1412.
Hyman, P. and S. T. Abedon (2012). "Smaller Fleas: Viruses of Microorganisms." Scientifica **2012**: 23.
Ishino, Y., H. Shinagawa, K. Makino, M. Amemura and A. Nakata (1987). "Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product." Journal of Bacteriology **169**: 5429-5433.
Ito, T., K. Yoshiguchi, H. D. Ariesyady and S. Okabe (2011). "Identification of a novel acetate-utilizing bacterium belonging to Synergistes group 4 in anaerobic digester sludge." ISME J **5**(12): 1844-1856.
Jansen, R., J. D. A. van Embden, W. Gaastra and L. M. Schouls (2002). "Identification of genes that are associated with DNA repeats in prokaryotes." Molecular Microbiology **43**(6): 1565-1575.
Janssen, P. H. (2003). "Selective enrichment and purification of cultures of Methanosaeta spp." J Microbiol Methods **52**(2): 239-244.
Jarrell, K. F., M. Saulnier and A. Ley (1987). "Inhibition of methanogenesis in pure cultures by ammonia, fatty acids and heavy metals and protection against heavy metal toxicity by sewage sludge." Canadian Journal of Microbiology **33**(6): 551-554.
Jeris, J. S. and P. L. McCarty (1965). "The biochemistry of methane fermentation using C14 tracers." Journal Water Pollution Control Federation **37**(2): 178-192.
Jordan, M., L. Meile and T. Leisinger (1989). "Organization of Methanobacterium thermoautotrophicum bacteriophage psi M1 DNA." Mol Gen Genet **220**(1): 161-164.
Kang, K. S., G. T. Veeder, P. J. Mirrasoul, T. Kaneko and I. W. Cottrell (1982). "Agar-like polysaccharide produced by a *Pseudomonas* species - production and basic properties." Applied and Environmental Microbiology **43**(5): 1086-1091.
Khan, M. A., H. Satoh, H. Katayama, F. Kurisu and T. Mino (2002). "Bacteriophages isolated from activated sludge processes and their polyvalency." Water Res **36**(13): 3364-3370.
Kida, K., T. Shigematsu, J. Kijima, M. Numaguchi, Y. Mochinaga, N. Abe and S. Morimura (2001). "Influence of Ni2+ and Co2+ on methanogenic activity and the amounts of

coenzymes involved in methanogenesis." Journal of Bioscience and Bioengineering **91**(6): 590-595.
Kim, M., Y. H. Ahn and R. E. Speece (2002). "Comparative process stability and efficiency of anaerobic digestion; mesophilic vs. thermophilic." Water Res **36**(17): 4369-4385.
Koster, I. W. and E. Koomen (1988). "Ammonia inhibition of the maximum growth rate (μm) of hydrogenotrophic methanogens at various pH-levels and temperatures." Applied Microbiology and Biotechnology **28**(4-5): 500-505.
Koster, I. W. and G. Lettinga (1984). "The influence of ammonium nitrogen on the specific activity of pelletized methanogenic sludge." Agricultural Wastes **9**(3): 205-216.
Krakat, N., S. Schmidt and P. Scherer (2011). "Potential impact of process parameters upon the bacterial diversity in the mesophilic anaerobic digestion of beet silage." Bioresource Technology **102**(10): 5692-5701.
Kroeker, E. J., D. D. Schulte, A. B. Sparling and H. M. Lapp (1979). "Anaerobic treatment process stability." Journal Water Pollution Control Federation **51**(4): 718-727.
Kropinski, A. M., A. Mazzocco, T. E. Waddell, E. Lingohr and R. P. Johnson (2008). Enumeration of Bacteriophages by Double Agar Overlay Plaque Assay. **501:** 69-76.
Kunin, V., R. Sorek and P. Hugenholtz (2007). "Evolutionary conservation of sequence and secondary structures in CRISPR repeats." Genome Biology **8**(4).
Lay, J. J., Y. Y. Li and T. Noike (1998). "The influence of pH and ammonia concentration on the methane production in high-solids digestion processes." Water Environment Research **70**(5): 1075-1082.
Leclerc, M., J. P. Delgenes and J. J. Godon (2004). "Diversity of the archaeal community in 44 anaerobic digesters as determined by single strand conformation polymorphism analysis and 16S rDNA sequencing." Environmental Microbiology **6**(8): 809-819.
Lee, S. H., K. Otawa, M. Onuki, H. Satoh and T. Mino (2007). "Population dynamics of phage-host system of Microlunatus phosphovorus indigenous in activated sludge." J Microbiol Biotechnol **17**(10): 1704-1707.
Lillestol, R. K., P. Redder, R. A. Garrett and K. Brugger (2006). "A putative viral defence mechanism in archaeal cells." Archaea **2**(1): 59-72.
Lillestol, R. K., S. A. Shah, K. Brugger, P. Redder, H. Phan, J. Christiansen and R. A. Garrett (2009). "CRISPR families of the crenarchaeal genus Sulfolobus: bidirectional transcription and dynamic properties." Molecular Microbiology **72**(1): 259-272.
Limam, R. D., T. Bouchez, R. Chouari, T. Li, I. Barkallah, A. Landoulsi and A. Sghir (2010). "Detection of WWE2-related Lentisphaerae by 16S rRNA gene sequencing and fluorescence in situ hybridization in landfill leachate." Can J Microbiol **56**(10): 846-852.
Liu, D. X. and L. Huang (2002). "Induction of the Sulfolobus shibatae virus SSV1 DNA replication by mitomycin C." Chinese Science Bulletin **47**(11): 923-927.
Liu, T. and S. Sung (2002). "Ammonia inhibition on thermophilic aceticlastic methanogens." Water Science and Technology **45**(10): 113-120.
Los, M. and G. Wegrzyn (2012). Pseudolysogeny. Advances in Virus Research, Vol 82: Bacteriophages, Pt A. M. Lobocka and W. T. Szybalski. **82:** 339-349.
Luo, Y. N., P. Pfister, T. Leisinger and A. Wasserfallen (2001). "The genome of archaeal prophage Psi M100 encodes the lytic enzyme responsible for autolysis of Methanothermobacter wolfeii." Journal of Bacteriology **183**(19): 5788-5792.
Ma, K., X. L. Liu and X. Z. Dong (2006). "*Methanosaeta harundinacea* sp nov., a novel acetate-scavenging methanogen isolated from a UASB reactor." International Journal of

Systematic and Evolutionary Microbiology **56**: 127-131.
Makarova, K. S., L. Aravind, N. V. Grishin, I. B. Rogozin and E. V. Koonin (2002). "A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis." Nucleic Acids Research **30**(2): 482-496.
Makarova, K. S., D. H. Haft, R. Barrangou, S. J. J. Brouns, E. Charpentier, P. Horvath, S. Moineau, F. J. M. Mojica, Y. I. Wolf, A. F. Yakunin, J. van der Oost and E. V. Koonin (2011). "Evolution and classification of the CRISPR-Cas systems." Nature Reviews Microbiology **9**(6): 467-477.
Marraffini, L. A. and E. J. Sontheimer (2008). "CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA." Science **322**(5909): 1843-1845.
Marraffini, L. A. and E. J. Sontheimer (2010). "CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea." Nature Reviews Genetics **11**(3): 181-190.
Marraffini, L. A. and E. J. Sontheimer (2010). "Self versus non-self discrimination during CRISPR RNA-directed immunity." Nature **463**(7280): 568-U194.
Masepohl, B., K. Gorlitz and H. Bohme (1996). "Long tandemly repeated repetitive (LTRR) sequences in the filamentous cyanobacterium Anabaena sp PCC 7120." Biochimica Et Biophysica Acta-Gene Structure and Expression **1307**(1): 26-30.
Mata-Alvarez, J., S. Mace and P. Llabres (2000). "Anaerobic digestion of organic solid wastes. An overview of research achievements and perspectives." Bioresource Technology **74**(1): 3-16.
McCarty, P. L. (1964). " Anaerobic waste treatment fundamentals, Part III. Toxic materials and their control." Public Works **95**(12): 91-94.
McCarty, P. L. (2001). "The development of anaerobic treatment and its future." Water Sci Technol **44**(8): 149-156.
McCarty, P. L. and R. E. McKinney (1961). "Salt toxicity in anaerobic treatment." Journal of Water Pollution Control Federation **33**(4): 399-415.
McHugh, S., M. Carton, T. Mahony and V. O'Flaherty (2003). "Methanogenic population structure in a variety of anaerobic bioreactors." Fems Microbiology Letters **219**(2): 297-304.
Mei, Y. J., D. Chen, D. C. Sun, Y. Yang, P. Shen and X. D. Chen (2007). "Induction and preliminary characterization of a novel halophage SNJ1 from lysogenic Natrinema sp F5." Canadian Journal of Microbiology **53**(9): 1106-1110.
Meile, L., U. Jenal, D. Studer, M. Jordan and T. Leisinger (1989). "Characterization of psiM1, a virulent phage of *Methanobacterium thermoautotrophicum* Marburg." Archives of Microbiology **152**(2): 105-110.
Millen, A. M., P. Horvath, P. Boyaval and D. A. Romero (2012). "Mobile CRISPR/Cas-Mediated Bacteriophage Resistance in Lactococcus lactis." PLoS One **7**(12).
Mizukami, S., K. Takeda, S. Akada and T. Fujita (2006). "Isolation and characteristics of methanosaeta in paddy field soils." Biosci Biotechnol Biochem **70**(4): 828-835.
Mizukami, S., K. Takeda, S. Akada and T. Fujita (2006). "Isolation and characteristics of *Methanosaeta* in paddy field soils." Bioscience Biotechnology and Biochemistry **70**(4): 828-835.
Mojica, F. J. M., C. Diez-Villasenor, J. Garcia-Martinez and C. Almendros (2009). "Short motif sequences determine the targets of the prokaryotic CRISPR defence system." Microbiology-Sgm **155**: 733-740.
Mojica, F. J. M., C. Diez-Villasenor, J. Garcia-Martinez and E. Soria (2005). "Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements."

Journal of Molecular Evolution **60**(2): 174-182.
Mojica, F. J. M., C. Diez-Villasenor, E. Soria and G. Juez (2000). "Biological significance of a family of regularly spaced repeats in the genomes of Archaea, bacteria and mitochondria." Molecular Microbiology **36**(1): 244-246.
Mojica, F. J. M., C. Ferrer, G. Juez and F. Rodriguez-Valera (1995). "Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning." Molecular Microbiology **17**(1): 85-93.
Nachaiyasit, S. and D. C. Stuckey (1997). "The effect of shock loads on the performance of an anaerobic baffled reactor (ABR) .1. Step changes in feed concentration at constant retention time." Water Research **31**(11): 2737-2746.
Nakamura, K., H. Tamaki, M. S. Kang, H. Mochimaru, S. T. Lee and Y. Kamagata (2011). "A Six-well Plate Method: Less Laborious and Effective Method for Cultivation of Obligate Anaerobic Microorganisms." Microbes and Environments **26**(4): 301-306.
Namiki, T., T. Hachiya, H. Tanaka and Y. Sakakibara (2012). "MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads." Nucleic Acids Research **40**(20).
Narihiro, T. and Y. Sekiguchi (2007). "Microbial communities in anaerobic digestion processes for waste and wastewater treatment: a microbiological update." Curr Opin Biotechnol **18**(3): 273-278.
Noble, R. T. and J. A. Fuhrman (1998). "Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria." Aquatic Microbial Ecology **14**(2): 113-118.
Nolling, J., A. Groffen and W. M. Devos (1993). "phiF1 and phiF3, two novel virulent, archaeal phages infecting different thermophilic strains of the genus *Methanobacterium*." Journal of General Microbiology **139**: 2511-2516.
Otawa, K., S. H. Lee, A. Yamazoe, M. Onuki, H. Satoh and T. Mino (2007). "Abundance, diversity, and dynamics of viruses on microorganisms in activated sludge processes." Microb Ecol **53**(1): 143-152.
Park, M. O., H. Ikenaga and K. Watanabe (2007). "Phage diversity in a methanogenic digester." Microb Ecol **53**(1): 98-103.
Peng, X., M. Brugger, B. Shen, L. M. Chen, Q. X. She and R. A. Garrett (2003). "Genus-specific protein binding to the large clusters of DNA repeats (short regularly spaced repeats) present in Sulfolobus genomes." Journal of Bacteriology **185**(8): 2410-2417.
Pfister, P., A. Wasserfallen, R. Stettler and T. Leisinger (1998). "Molecular analysis of Methanobacterium phage psiM2." Mol Microbiol **30**(2): 233-244.
Pina, M., A. Bize, P. Forterre and D. Prangishvili (2011). "The archeoviruses." FEMS Microbiol Rev **35**(6): 1035-1054.
Pourcel, C., G. Salvignol and G. Vergnaud (2005). "CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies." Microbiology-Sgm **151**: 653-663.
Prangishvili, D., P. Forterre and R. A. Garrett (2006). "Viruses of the Archaea: a unifying view." Nature Reviews Microbiology **4**(11): 837-848.
Prangishvili, D. and R. A. Garrett (2005). "Viruses of hyperthermophilic Crenarchaea." Trends in Microbiology **13**(11): 535-542.
Prangishvili, D., G. Vestergaard, M. Haring, R. Aramayo, T. Basta, R. Rachel and R. A. Garrett (2006). "Structural and genomic properties of the hyperthermophilic archaeal virus ATV

with an extracellular stage of the reproductive cycle." Journal of Molecular Biology **359**(5): 1203-1216.
Pride, D. T., C. L. Sun, J. Salzman, N. Rao, P. Loomer, G. C. Armitage, J. F. Banfield and D. A. Relman (2011). "Analysis of streptococcal CRISPRs from human saliva reveals substantial sequence diversity within and between subjects over time." Genome Research **21**(1): 126-136.
Raskin, L., D. D. Zheng, M. E. Griffin, P. G. Stroot and P. Misra (1995). "Characterization of microbial communities in anaerobic bioreactors using molecular probes." Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology **68**(4): 297-308.
Redder, P. and R. A. Garrett (2006). "Mutations and rearrangements in the genome of Sulfolobus solfataticus P2." Journal of Bacteriology **188**(12): 4198-4206.
Reiter, W. D., P. Palm, S. Yeats and W. Zillig (1987). "Gene expression in archaebacteria: physical mapping of constitutive and UV-inducible transcripts from the Sulfolobus virus-like particle SSV1." Molecular & General Genetics **209**(2): 270-275.
Reno, M. L., N. L. Held, C. J. Fields, P. V. Burke and R. J. Whitaker (2009). "Biogeography of the Sulfolobus islandicus pan-genome." Proceedings of the National Academy of Sciences of the United States of America **106**(21): 8605-8610.
Rice, G., K. Stedman, J. Snyder, B. Wiedenheft, D. Willits, S. Brumfield, T. McDermott and M. J. Young (2001). "Viruses from extreme thermal environments." Proc Natl Acad Sci U S A **98**(23): 13341-13345.
Riviere, D., V. Desvignes, E. Pelletier, S. Chaussonnerie, S. Guermazi, J. Weissenbach, T. Li, P. Camacho and A. Sghir (2009). "Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge." ISME J **3**(6): 700-714.
Robb, F. T. and A. R. Place (1995). Archaea: a laboratory manual. Thermophiles, Cold Spring Harbor Laboratory Press.
Rule, P. L. and A. D. Alexander (1986). "Gellan gum as a substitute for agar in *Leptospiral* media." Journal of Clinical Microbiology **23**(3): 500-504.
Sawayama, S., C. Tada, K. Tsukahara and T. Yagishita (2004). "Effect of ammonium addition on methanogenic community in a fluidized bed anaerobic digestion." Journal of Bioscience and Bioengineering **97**(1): 65-70.
Schink, B. (1997). "Energetics of syntrophic cooperation in methanogenic degradation." Microbiol Mol Biol Rev **61**(2): 262-280.
Scholten, J. C. M. and A. J. M. Stams (2000). "Isolation and characterization of acetate-utilizing anaerobes from a freshwater sediment." Microbial Ecology **40**(4): 292-299.
Semenova, E., M. Nagornykh, M. Pyatnitskiy, I. I. Artamonova and K. Severinov (2009). "Analysis of CRISPR system function in plant pathogen Xanthomonas oryzae." Fems Microbiology Letters **296**(1): 110-116.
Sensen, C. W., R. L. Charlebois, C. Chow, I. G. Clausen, B. Curtis, W. F. Doolittle, M. Duguet, G. Erauso, T. Gaasterland, R. A. Garrett, P. Gordon, I. H. de Jong, A. C. Jeffries, C. Kozera, N. Medina, A. De Moors, J. van der Oost, H. Phan, M. A. Ragan, M. F. Schenk, Q. X. She, R. K. Singh and N. Tolstrup (1998). "Completing the sequence of the Sulfolobus solfataricus P2 genome." Extremophiles **2**(3): 305-312.
Shah, S. A. and R. A. Garrett (2011). "CRISPR/Cas and Cmr modules, mobility and evolution of adaptive immune systems." Research in Microbiology **162**(1): 27-38.
Shah, S. A., G. Vestergaard and R. A. Garrett (2011). CRISPR/Cas and CRISPR/Cmr Immune

Systems of Archaea.
Shapiro, O. H., A. Kushmaro and A. Brenner (2010). "Bacteriophage predation regulates microbial abundance and diversity in a full-scale bioreactor treating industrial wastewater." Isme Journal **4**(3): 327-336.
She, Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. C. Y. Chan-Weiher, I. G. Clausen, B. A. Curtis, A. De Moors, G. Erauso, C. Fletcher, P. M. K. Gordon, I. Heikamp-de Jong, A. C. Jeffries, C. J. Kozera, N. Medina, X. Peng, H. P. Thi-Ngoc, P. Redder, M. E. Schenk, C. Theriault, N. Tolstrup, R. L. Charlebois, W. F. Doolittle, M. Duguet, T. Gaasterland, R. A. Garrett, M. A. Ragan, C. W. Sensen and J. Van der Oost (2001). "The complete genome of the crenarchaeon Sulfolobus solfataricus P2." Proceedings of the National Academy of Sciences of the United States of America **98**(14): 7835-7840.
Shelton, D. R. and J. M. Tiedje (1984). "General method for determining anaerobic biodegradation potential." Applied and Environmental Microbiology **47**(4): 850-857.
Shelton, D. R. and J. M. Tiedje (1984). "General method for determining anaerobic biodegradation potential." Appl Environ Microbiol **47**(4): 850-857.
Shigematsu, T., Y. Tang, H. Kawaguchi, K. Ninomiya, J. Kijima, T. Kobayashi, S. Morimura and K. Kida (2003). "Effect of dilution rate on structure of a mesophilic acetate-degrading methanogenic community during continuous cultivation." J Biosci Bioeng **96**(6): 547-558.
Shigematsu, T., Y. Tang, Y. Mizuno, H. Kawaguchi, S. Morimura and K. Kida (2006). "Microbial diversity of mesophilic methanogenic consortium that can degrade long-chain fatty acids in chemostat cultivation." Journal of Bioscience and Bioengineering **102**(6): 535-544.
Siem-Jorgensen, M., R. N. Glud and M. Middelboe (2008). "Viral dynamics in a coastal sediment: seasonal pattern, controlling factors and relations to the pelagic-benthic coupling." Marine Biology Research **4**(3): 165-U120.
Skennerton, C. T., M. Imelfort and G. W. Tyson (2013). "Crass: identification and reconstruction of CRISPR from unassembled metagenomic data." Nucleic Acids Research **41**(10).
Smith, K. S. and C. Ingram-Smith (2007). "Methanosaeta, the forgotten methanogen?" Trends Microbiol **15**(4): 150-155.
Smith, P. H. and R. A. Mah (1966). "Kinetics of acetate metabolism during sludge digestion." Appl Microbiol **14**(3): 368-371.
Snyder, J. C., M. M. Bateson, M. Lavin and M. J. Young (2010). "Use of Cellular CRISPR (Clusters of Regularly Interspaced Short Palindromic Repeats) Spacer-Based Microarrays for Detection of Viruses in Environmental Samples." Applied and Environmental Microbiology **76**(21): 7251-7258.
Sorek, R., V. Kunin and P. Hugenholtz (2008). "CRISPR - a widespread system that provides acquired resistance against phages in bacteria and archaea." Nature Reviews Microbiology **6**(3): 181-186.
Sorokin, V. A., M. S. Gelfand and Artamonova, II (2010). "Evolutionary Dynamics of Clustered Irregularly Interspaced Short Palindromic Repeat Systems in the Ocean Metagenome." Applied and Environmental Microbiology **76**(7): 2136-2144.
Speece, R. E. (1996). "Anaerobic Biotechnology for Industrial Wastewaters." Archaea Press, Nashville.
Sprott, G. D. and G. B. Patel (1986). "Ammonia toxicity in pure cultures of methanogenic bacteria." Systematic and Applied Microbiology **7**(2-3): 358-363.
Stackebrandt, E. and M. Goodfellow (1991). Nucleic acid techniques in bacterial systematics.

Chichester; New York, Wiley.
Steinberg, L. M. and J. M. Regan (2008). "Phylogenetic Comparison of the Methanogenic Communities from an Acidic, Oligotrophic Fen and an Anaerobic Digester Treating Municipal Wastewater Sludge." Applied and Environmental Microbiology **74**(21): 6663-6671.
Steinhaus, B., M. L. Garcia, A. Q. Shen and L. T. Angenent (2007). "A portable anaerobic microbioreactor reveals optimum growth conditions for the methanogen Methanosaeta concilii." Applied and Environmental Microbiology **73**(5): 1653-1658.
Stern, A., E. Mick, I. Tirosh, O. Sagy and R. Sorek (2012). "CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome." Genome Research **22**(10): 1985-1994.
Suttle, C. A. (2005). "Viruses in the sea." Nature **437**(7057): 356-361.
Suttle, C. A. (2007). "Marine viruses - major players in the global ecosystem." Nature Reviews Microbiology **5**(10): 801-812.
Tang, T. H., J. P. Bachellerie, T. Rozhdestvensky, M. L. Bortolin, H. Huber, M. Drungowski, T. Elge, J. Brosius and A. Huttenhofer (2002). "Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus." Proceedings of the National Academy of Sciences of the United States of America **99**(11): 7536-7541.
Tang, Y., T. Shigematsu, S. Morimura and K. Kida (2005). "Microbial community analysis of mesophilic anaerobic protein degradation process using bovine serum albumin (BSA)-fed continuous cultivation." J Biosci Bioeng **99**(2): 150-164.
Tapper, M. A. and R. E. Hicks (1998). "Temperate viruses and lysogeny in Lake Superior bacterioplankton." Limnology and Oceanography **43**(1): 95-103.
Terns, M. P. and R. M. Terns (2011). "CRISPR-based adaptive immune systems." Current Opinion in Microbiology **14**(3): 321-327.
Thingstad, T. F. (2000). "Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems." Limnology and Oceanography **45**(6): 1320-1328.
Thingstad, T. F. and R. Lignell (1997). "Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand." Aquatic Microbial Ecology **13**(1): 19-27.
Thorvaldsdottir, H., J. T. Robinson and J. P. Mesirov (2013). "Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration." Briefings in Bioinformatics **14**(2): 178-192.
Tuomi, P., K. M. Fagerbakke, G. Bratbak and M. Heldal (1995). "Nutritional enrichment of a microbial community: The effects on activity, elemental composition, community structure and virus production." Fems Microbiology Ecology **16**(2): 123-134.
Tyson, G. W. and J. F. Banfield (2008). "Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses." Environmental Microbiology **10**(1): 200-207.
Wang, Q., G. M. Garrity, J. M. Tiedje and J. R. Cole (2007). "Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy." Applied and Environmental Microbiology **73**(16): 5261-5267.
Wang, Y. Q. and X. B. Zhang (2010). "Genome Analysis of Deep-Sea Thermophilic Phage D6E." Applied and Environmental Microbiology **76**(23): 7861-7866.
Weinbauer, M. G. (2004). "Ecology of prokaryotic viruses." Fems Microbiology Reviews **28**(2): 127-181.
Westerholm, M., J. Dolfing, A. Sherry, N. D. Gray, I. M. Head and A. Schnurer (2011).

"Quantification of syntrophic acetate-oxidizing microbial communities in biogas processes." Environmental Microbiology Reports **3**(4): 500-505.
Williamson, K. E., M. Radosevich and K. E. Wommack (2005). "Abundance and diversity of viruses in six Delaware soils." Appl Environ Microbiol **71**(6): 3119-3125.
Wommack, K. E. and R. R. Colwell (2000). "Virioplankton: viruses in aquatic ecosystems." Microbiol Mol Biol Rev **64**(1): 69-114.
Wood, A. G., W. B. Whitman and J. Konisky (1989). "Isolation and characterization of an archaebacterial viruslike particle from Methanococcus voltae A3." J Bacteriol **171**(1): 93-98.
Wu, Q. L. and W. T. Liu (2009). "Determination of Virus abundance, diversity and distribution in a municipal wastewater treatment plant." Water Research **43**(4): 1101-1109.
You, X. Y., C. Liu, S. Y. Wang, C. Y. Jiang, S. A. Shah, D. Prangishvili, Q. X. She, S. J. Liu and R. A. Garrett (2011). "Genomic analysis of Acidianus hospitalis W1 a host for studying crenarchaeal virus and plasmid life cycles." Extremophiles **15**(4): 487-497.
Yu, Y., C. Lee and S. Hwang (2005). "Analysis of community structures in anaerobic processes using a quantitative real-time PCR method." Water Science and Technology **52**(1-2): 85-91.
Zehnder, A. J., B. A. Huser, T. D. Brock and K. Wuhrmann (1980). "Characterization of an acetate-decarboxylating, non-hydrogen-oxidizing methane bacterium." Arch Microbiol **124**(1): 1-11.
Zehnder, A. J. B., B. A. Huser, T. D. Brock and K. Wuhrmann (1980). "Characterization of an acetate-decarboxylanting, non-hydrogen-oxidizing methane bacterium." Archives of Microbiology **124**(1): 1-11.
Zerbino, D. R. and E. Birney (2008). "Velvet: Algorithms for de novo short read assembly using de Bruijn graphs." Genome Research **18**(5): 821-829.
Zhang, H., J. E. Banaszak, P. Parameswaran, J. Alder, R. Krajmalnik-Brown and B. E. Rittmann (2009). "Focused-Pulsed sludge pre-treatment increases the bacterial diversity and relative abundance of acetoclastic methanogens in a full-scale anaerobic digester." Water Research **43**(18): 4517-4526.
Zheng, Y., R. J. Roberts and S. Kasif (2004). "Identification of genes with fast-evolving regions in microbial genomes." Nucleic Acids Research **32**(21): 6347-6357.
Zumstein, E., R. Moletta and J. J. Godon (2000). "Examination of two years of community dynamics in an anaerobic bioreactor using fluorescence polymerase chain reaction (PCR) single-strand conformation polymorphism analysis." Environmental Microbiology **2**(1): 69-78.

Appendix A An Analysis of Viral Metagenomes in Acetate-fed Anaerobic Reactors

Results

Taxonomic Profile

After quality trimming and assembly, reactor samples contained 1,846-2,481 contigs, each, while anaerobic digester samples contained 36,900-54,031 contigs each, with N50s of 161-319 bp and mean coverage depths of 3.4-13.1 for all samples (Table A.1). The percentage of contigs with significant taxonomic matches to the M5NR database (E-value $\leq 10^{-5}$) on MG-RAST was lower for the reactor samples (10.2-10.5%), than for the WWTP samples (25.7-32.1%). In the taxonomic classification of contigs matching the M5NR database, viral matches comprised 4.1-5.8% of samples (0.4-1.9% of all contigs) (Table A.2). The taxonomic classification was dominated by matches to bacteria (73.8-82.9%). The reactor samples had a greater percentage of matches to archaea (5.2-7.2%) than the anaerobic digester samples (1-1.5%). Comparison of viromes to 16S rRNA gene databases indicated minimal prokaryotic contamination (0 matches for the reactors, 52 for South WWTP, and 60 for West Point WWTP). Estimates of alpha diversity in each virome were computed using the relative abundances of annotated species in the M5NR database. This estimate indicated 178, 323, 836, and 788 species in the daily, hourly, West Point, and South viromes, respectively. The rarefaction curves indicate that additional sampling may be beneficial in order to better characterize the viral communities of these systems (Figure A.1).

Viral sequences were further classified into families using the same search parameters. Both the

reactors and the digesters were dominated by the order *Caudovirales* (Figure A.2), consistent with TEM image analyses. *Siphoviridae* tended to be the most abundant (36-49%), followed by *Myoviridae* (12-50%) and *Podoviridae* (0-12%). A small number of matches to viruses from other orders were observed for the South WWTP, similar to the anaerobic digester of Tamaki *et al*. (2011). Viral sequences from the reactors were further examined for matches to known methanogen phages or prophages by comparison to the ACLAME database (E-value ≤ 0.001). All matches were to the TerL gene of phage psiM2 of *Methanothermobacterium autotrophicus*. There were five significant matches for the daily fed reactor, three of which were the top hit. The hourly fed reactor had one significant match, which was not the top

Table A.1 Sequence and contig statistics

	Daily Fed Reactor	**Hourly Fed Reactor**	**West Point WWTP**	**South WWTP**
No. of raw sequences	572,672	2,554,267	2,578,357	2,074,807
Mean raw read length (bp)	215	135	149	233
No. of filtered sequences	453,538	2,121,452	2,146,160	2,008,500
Mean filtered length (SD)	*62 (35)*	84 (43)	91 (48)	110 (54)
No. of contigs	1,846	2,481	54,031	36,900
Mean coverage depth	6.7	13.1	3.4	6.1
Contig N50	316	161	202	319
Total bases in contigs	571,331	458,033	12,945,497	11,185,187
Longest contig	3,425	1,744	3,628	6,943
No of contigs >1000 bp	37	9	141	689
Bases in contigs >1000 bp	61,759	11,414	184,502	1,019,123

Table A.2 Functional and taxonomic assignment percentages and distribution of taxonomic assignments

	Daily Fed Reactor	Hourly Fed Reactor	West Point WWTP	South WWTP	Tamaki et al. (2011)
Functional assignment[a]					
Unknown (%)	98.1	96.8	95.8	94.6	95.1
Known (%)	1.9	3.2	4.2	5.4	4.9
Taxonomic assignment[b]					
Unknown (%)	89.5	89.8	74.3	67.9	73.7
Known (%)	10.5	10.2	25.7	32.1	26.3
Taxonomic classification of known[b]					
Viruses (%)	4.1	5.6	5.4	5.8	5.3
Bacteria (%)	80.4	73.8	78.8	82.9	81.8
Archaea (%)	7.2	5.2	1.5	1	2
Eukaryotes (%)	0	2	0.1	0.4	1.1
Unassigned/Unclassified/ Other (%)	8.2	13.5	14.1	9.9	9.7

a. Based on BLAT to the SEED Subsystems database (E < 1e-5).
b. Based on BLAT to the M5NR database (E < 1e-5).

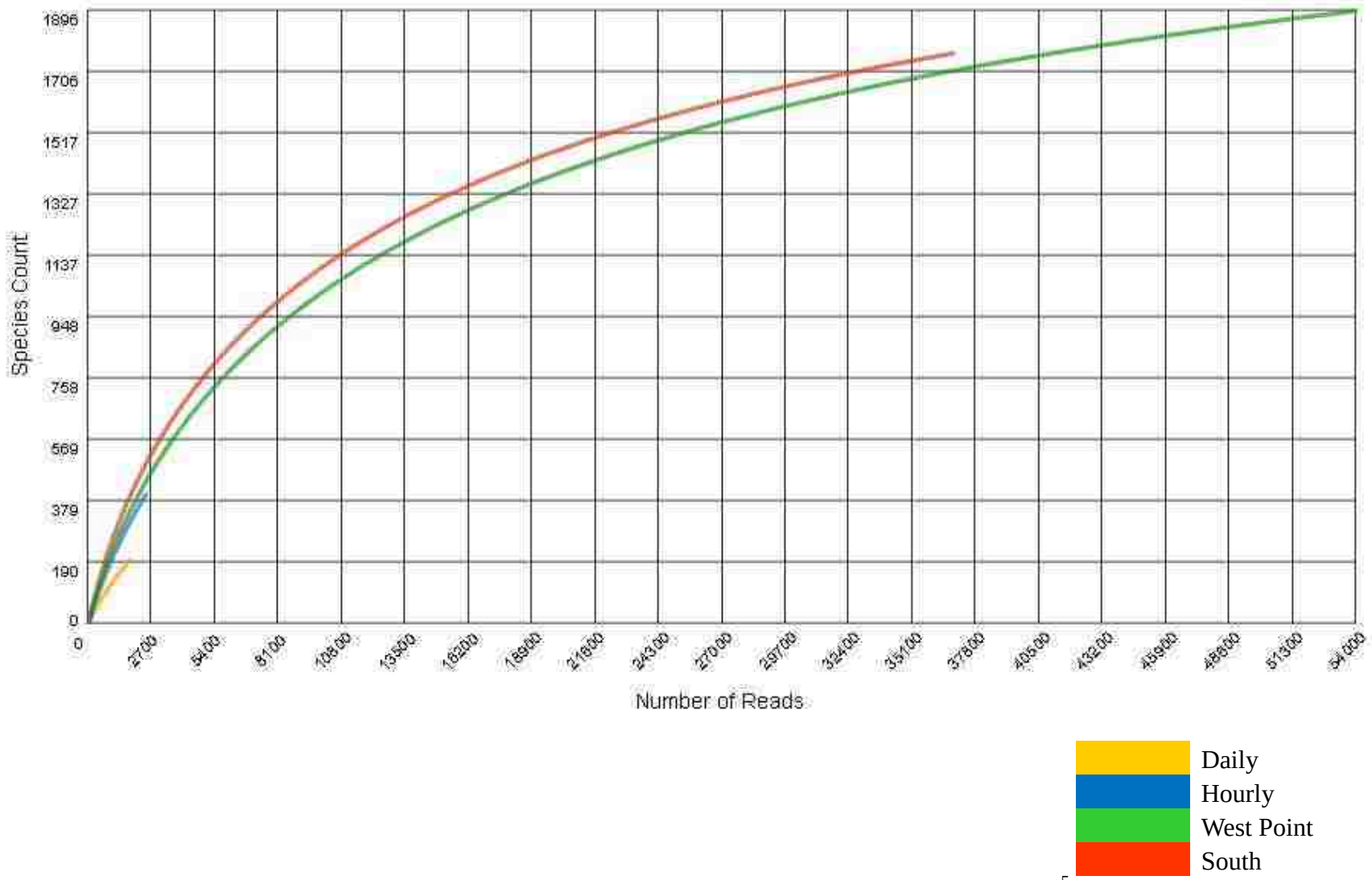


Figure A.1 Rarefaction curves of assembled samples based on BLAT to M5NR database (E-value $\leq 10^{-5}$)

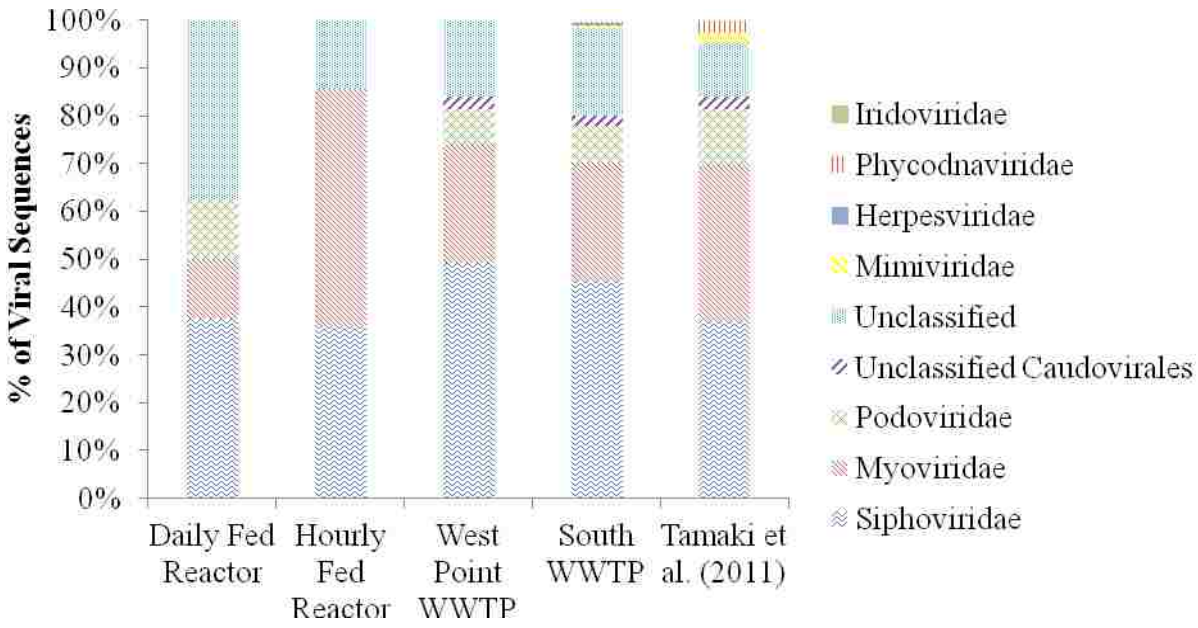


Figure A.2 Comparison of virus family abundance based on BLAT to the M5NR database (E < 1e-5)

Functional Profile

A comparison of the assembled viromes to the SEED subsystems database on MG-RAST (E-value $\leq 10^{-5}$) provided functional annotation for 1.9-5.4% of contigs (Table A.1). The COG database was searched through MG-RAST as well (E-value $\leq 10^{-5}$). COG functions related to DNA methylation were among the most abundant, with DNA modification methylase (COG0863) accounting for 6.5-14.7% of reactor COGs and 25.0-25.7% of anaerobic digester COGs. Transposase and inactivated derivatives (COG0675) were the most abundant in the reactor viromes, comprising 17.6-22.0% of assigned COGs. Aside from phage-related COGs, other abundant COGs were related to transcriptional regulation (COG1475) and replication, recombination and repair, such as single-stranded DNA-binding proteins (COG0629) and helicases (COG0305; COG0553).

Functional Comparisons

A comparison of COG classifications (E-value $\leq 10^{-5}$) among the reactors and anaerobic digesters indicates that they have similar distributions of functions (Figure A.3, Figure A.4), with the hourly fed reactor sample appearing to be slightly divergent from the other samples. The hourly fed sample appears to have more genes related to nucleotide and carbohydrate transport and metabolism and fewer related to replication, recombination, and repair. In comparing the collective COG classifications of the reactors and digesters from this study with 42 other viromes, there are several significant differences (Figure A.5). The reactors and digesters had significantly lower abundance of genes related to signal transduction, cytoskeleton, post-translational modification, lipid transport and metabolism, inorganic ion transport and metabolism, secondary metabolite synthesis, amino acid transport and metabolism, and energy production and conversion (P-value ≤ 0.002).

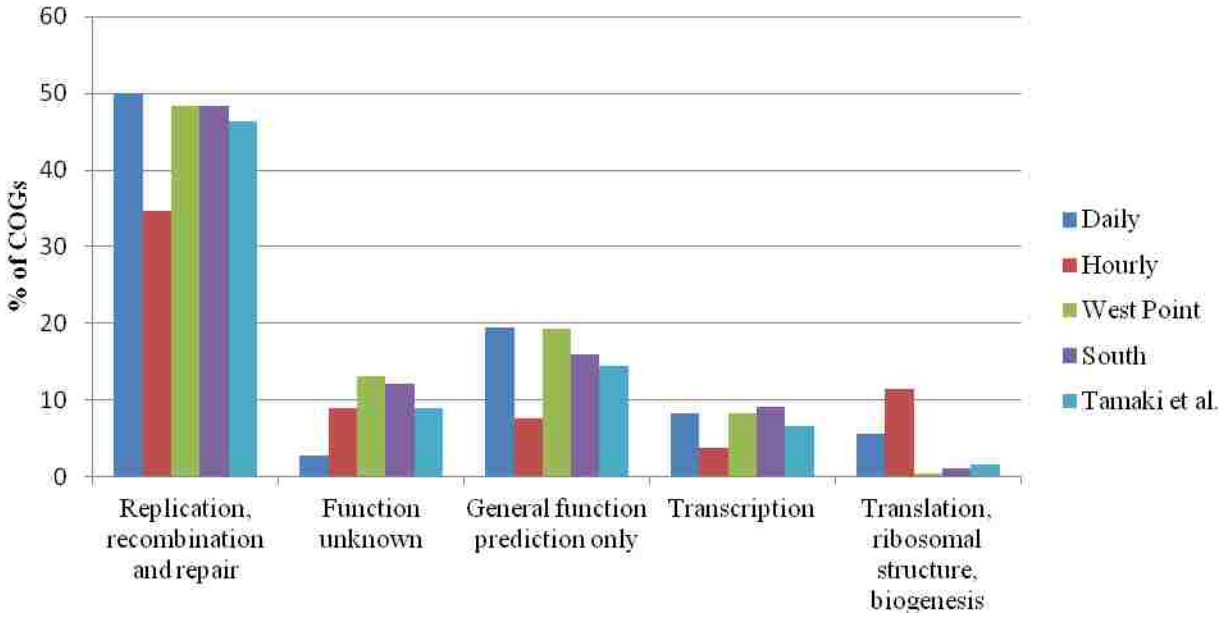


Figure A.3 Comparison of relatively abundant categories of functional profile among the samples from this study and the anaerobic digester of Tamaki *et al*. (2011) based on level two COG assignments (E-value $\leq 10^{-5}$).

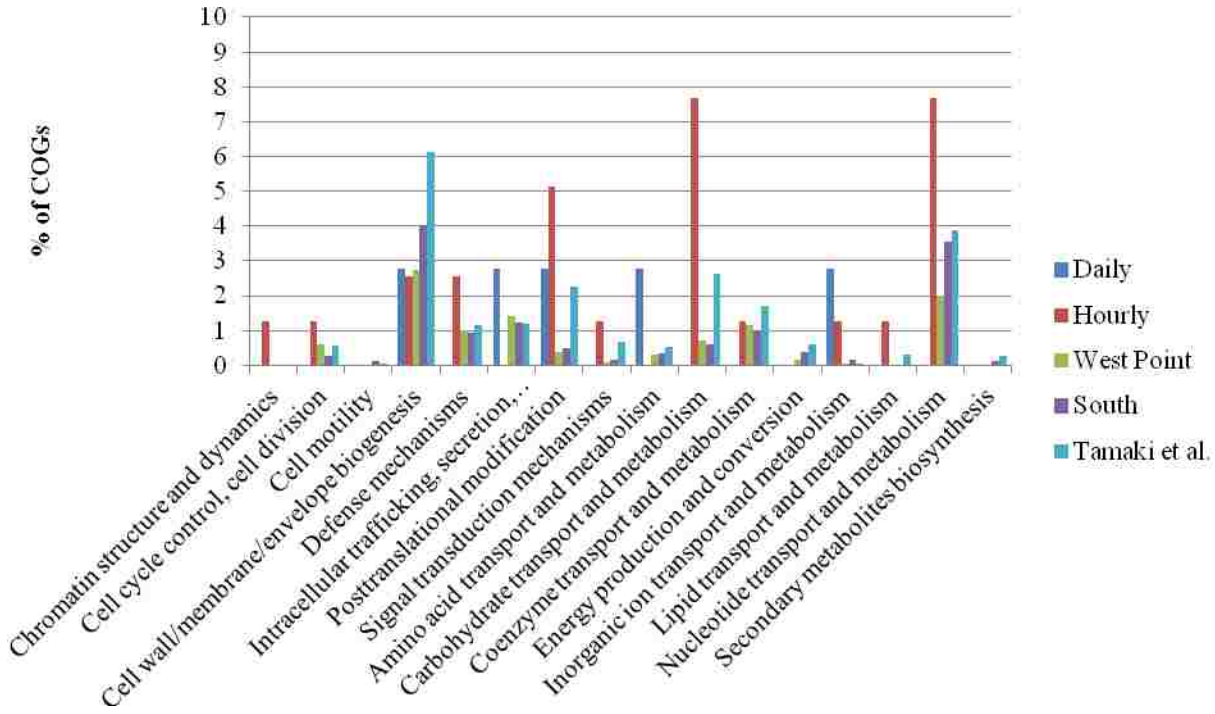


Figure A.4 Comparison of relatively less abundant functional profile among the samples from this study and the anaerobic digester of Tamaki *et al*. (2011) based on level two COG assignments (E-value $\leq 10^{-5}$).

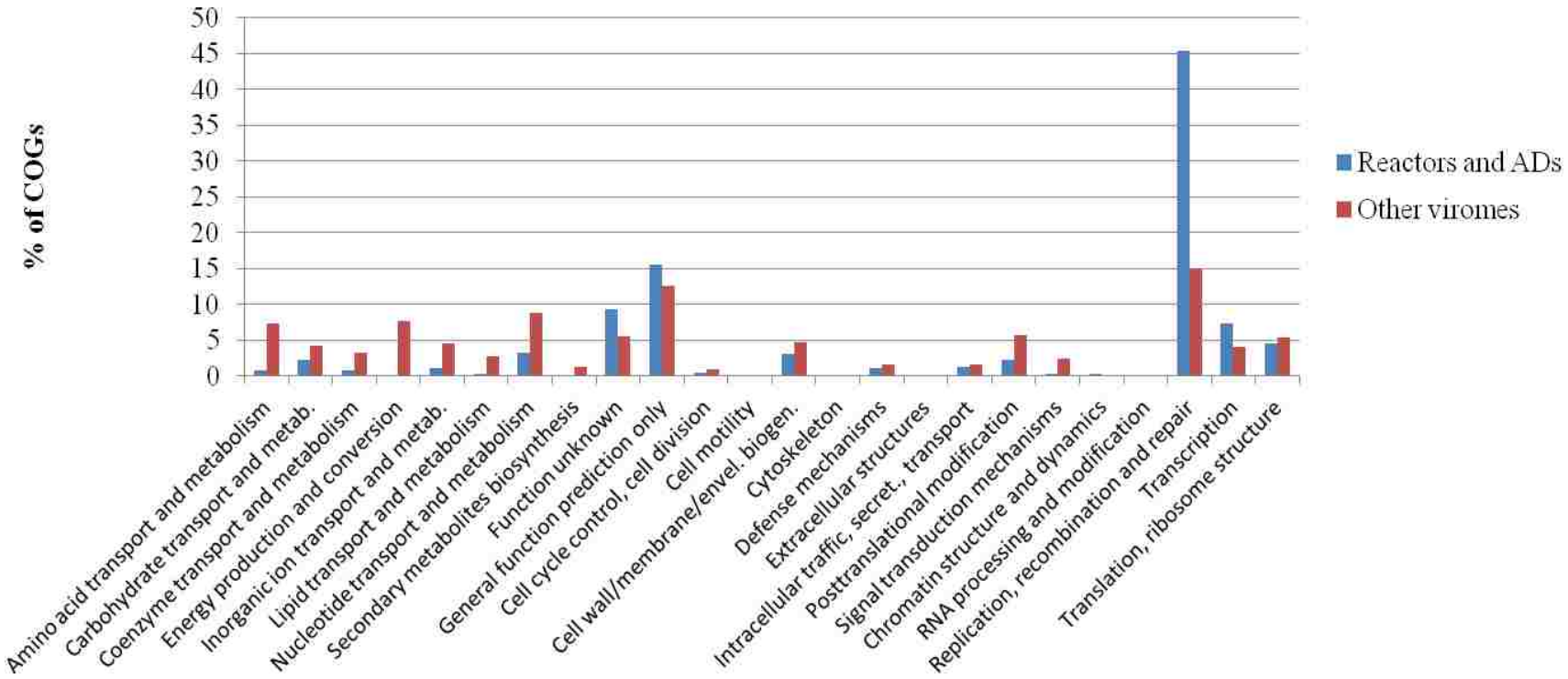


Figure A.5 Comparison of functional profile among the samples from this study and a collection of publicly available metagenomes based on level two COG assignments (E-value $\leq 10^{-5}$). Asterisk denotes statistical significance on normalized data (P-value ≤ 0.002).

CRISPR Spacer Matches to Viromes

CRISPRs in archaea and bacteria serve as indications of viral infection history. A BLASTN comparison (E-value ≤ 0.001) of the anaerobic reactor viromes to the CRISPRdb, containing 2669 identified CRISPRs in 1700 archaea and bacteria, was conducted in order to identify unknown viruses by matches to spacers from prokaryotic species either known or likely to be in the reactors. The significance criteria selected appeared to be sufficiently selective to omit most matches to species unlikely to be in the system while still retaining most matches to species likely to be in the system. The hourly-fed reactor had 14 significant matches to spacers. Of these, three were to *Methanosaeta concilii*, three were to syntrophic bacteria, and two were to *Methanococcales* found in marine geothermal sediment. The daily-fed reactor had 27 significant matches to spacers. Of these, twenty were to *M. concilii*, two of which had 100% identity across the entire spacer; one was to syntrophic bacteria; and two were to other sludge species (Table A.3). To ensure that these matches were not derived from contaminating sequences, the viromes were compared to the direct repeats using the same parameters. The hourly-fed and daily-fed viromes had one and two matches, respectively, none of which were plausible.

Table A.3 Significant matches to spacers in CRISPRdb by BLASTN (E-value ≤ 0.001)

Hourly Fed	
Species Hit	Number of Hits
Methanosaeta concilii GP-6	3
Syntrophic species	3
Methanothermococcus okinawensis IH1	1
Methanocaldococcus vulcanius M7	1
Other	6
Daily Fed	
Species Hit	Number of Hits
Methanosaeta concilii GP-6	20
Syntrophic species	1
Other sludge species	2
Other	4

Discussion

The alpha diversity estimates were greater for the anaerobic digesters than for the reactors. This difference in diversity reflects the differences in energy and nutrient availability between these systems. The anaerobic digesters have substantial energy and nutrient sources, enabling the coexistence of a wide range of microbes. The reactors, fed on acetate medium, were designed to select for a subset of the anaerobic digester community. The lower viral diversity in the reactors reflects the restricted number of host species. The alpha diversity estimates may be artificially high, based on comparison of the results from the MG-RAST calculation to those generated by PHACCS (Angly *et al*., 2005), a tool specifically designed for viral metagenome diversity analysis, for the WWTP viromes of Tamaki *et al*. (2011). However, the relative estimated diversity among the samples should be similar using either method. The rarefaction curves of the assembled reads (Figure A.1) indicate that more sequence data may be beneficial to more fully characterize the viral communities in these systems, particularly for the reactors. Similarly, the coverage depth across all samples was relatively low, ranging from 3.4 to 13.1. Increasing the coverage depth by gathering more data could improve the quality and length of the assemblies by providing stronger evidence for the correct bases where differences in base reads arise, reducing the probability of forming chimaeras, and providing more sequences with which to extend contigs.

The majority of sequences from the anaerobic digesters and laboratory reactors had no significant similarity to M5NR database sequences. This result is consistent with other virome studies, which typically demonstrate ~70% or more of sequences having no significant matches to database sequences, and may reflect uncharacterized viral diversity and demonstrates the paucity of current

knowledge of viruses in the environment. This unknown fraction may predominantly be of viral origin (Edwards and Rohwer, 2005). Virus-specific annotation pipelines have been shown to yield more functional annotations for viromes compared to MG-RAST, which bases its functional classification on bacterial and archaeal sequences (Lorenzi *et al*., 2011). This result supports the idea that a portion of the unknown fractions likely contain viral sequences.

The smaller proportion of sequences in the reactor viromes, compared to the anaerobic digester viromes, that had significant taxonomic matches to database sequences reflects that the reactors have been enriched for a relatively unexplored community. The two anaerobic digesters sampled and that of Tamaki *et al*. (2011) had similar percentages of taxonomic assignment (A 2); and were comparable to many other virome studies (Mokili *et al*., 2012). This similarity suggests that the anaerobic digester systems harbor viral communities that are relatively similar to many other environments. The isolation of the numerically non-dominant component of the broader community may present a useful means of assessing more fully the uniqueness of viruses in different environments. These species are minimally represented in current viral metagenomes, which may mask their actual diversity and relative abundances. Further investigation of systems similar to the reactors would be of interest in assessing the uniqueness of this relatively unknown community.

The taxonomic classification of all viromes was dominated by matches to bacteria. Similarly, all viromes had significant matches to archaea, with the reactor samples having relatively large percentages of archaeal matches. Sequences classified as viral made up only a small percentage of each sample (Table A.2). These results are consistent with prior studies (Tamaki *et al*., 2011; Anderson *et al*., 2011; Berg Miller *et al*., 2012). Analysis of the 16S rRNA gene content indicated

no detectable contamination in the reactor samples and minimal contamination in the anaerobic digester samples. Thus, some of the sequences in the anaerobic digester samples may reflect actual prokaryotic sequences. Though no contamination was detected in either reactor sample, the imperfect nature of the viral purification process may have resulted in undetected prokaryotic DNA in these viromes. However, the abundance of matches to prokaryotes may also reflect the bias of many current databases due to having many more non-viral than viral sequences, unidentified prophages in these databases, and horizontal gene transfer between viruses and their hosts (Lorenzi *et al*., 2011; Krupovic *et al*., 2010; Dinsdale *et al*., 2008). As the contamination of the reactor samples appears to be minimal, the relatively large percentage of matches to archaea may reflect some combination of unidentified prophages and horizontal gene transfer, suggesting that these systems may have enriched for archaeoviruses.

Sequences identified as viral were classified predominantly into families of the order *Caudovirales*. *Siphoviridae* was the most frequently matched family in the daily fed reactor and both WWTP samples, while *Myoviridae* was the most frequently matched family for the hourly fed reactor. These classifications are supported by TEM images. If the classification is representative of the unclassified majority of the viromes, then these systems are dominated by *Caudovirales*. Family-level abundances in the reactors should be interpreted with caution given the small number of sequences classified as viral for the daily fed and hourly fed reactors (8 and 14, respectively). The profiles of both WWTP samples appear similar to the anaerobic digester virome of Tamaki *et al*. (2011). The classification of viruses into families with hosts that are not present in WWTP systems could reflect biases in current databases due to differences in the extents to which virus groups have been studied, or could reflect viruses that have entered and persisted in the

WWTPs. The identification of *Mimiviridae* suggests database bias as a more likely cause, as these viruses should have been filtered out during sample processing.

Further investigation of the reactor samples specifically for phages of methanogens by comparison to the ACLAME database yielded several matches. The only matched gene was the large terminase subunit (TerL) of phage psiM2 of *Methanothermobacterium autotrophicus*, or the largely similar prophage psiM100 of *Methanothermobacter wolfeii*, thermophilic anaerobic digester hydrogenotrophic methanogens (Liu and Whitman, 2008). The temperature of the reactors (31-34°C) is outside of the optimum range for *Methanothermobacter* spp., suggesting that these species are unlikely to be found in the reactors. The TerL is relatively conserved in the order *Caudovirales* and is used as a marker gene for phylogenetic classification (Roux *et al*., 2011). As such, matches to this TerL gene suggest the presence of at least one virus population similar to psiM2 and psiM100. Tailed archaeal proviruses appear to group based on gene content in a manner similar to the hosts' taxonomic grouping, indicating possible coevolution (Krupovic *et al*., 2010). Thus, the presence of these similar viruses suggests the corresponding presence of a host similar to *Methanothermobacter* spp. These matches indicate the possible presence of viruses of a methanogen in the reactor systems.

A level two COG comparison of the viromes in this study to each other and to the anaerobic digester of Tamaki *et al*. (2011) indicated that the distribution of functional abundance is similar across these viromes, though the hourly fed reactor appears to be somewhat divergent (Figure A.3, Figure A.4). This divergence may be an artifact of the relatively small number of sequences assigned to COGs. Metagenomes, both microbial and viral, from different environments have been

shown to have metabolic profiles that reflect the perceived relative importance of various functions in each environment. These profiles have been proposed to be independent of differences in taxonomic profile at different locations within a given environment (Dinsdale *et al*., 2008). The similar functional profiles observed across the anaerobic digesters and daily fed reactor suggests that these systems exert similar demands on the viral communities. This agreement may indicate that the daily fed reactor has successfully maintained conditions that demand functionality that is similar to what was required in the original anaerobic digester system, suggesting that the reactor community may be a realistic representation of the initial community that was isolated. The apparent divergence of the hourly fed reactor, if real, could indicate that this system has become relatively artificial in comparison to its source environment. The larger proportion of COGs related to carbohydrate and nucleotide transport and metabolism and smaller proportion related to replication, recombination, and repair in this reactor suggests that it may present different functional requirements on the microbial community. It has been hypothesized that decreased abundance of metabolic genes and increased abundance of replication, recombination, and repair genes may be reflective of the viral survival strategy in environments with abundant nutrient and carbon supplies (Tamaki *et al*., 2011). This hypothesis is supported in this study by the statistical comparison of the metabolic profiles of the reactor and anaerobic digester viromes to a collection of 42 other viromes, which showed significantly lower abundances of several categories of metabolic genes (Figure A.5) in the reactor and digester viromes. Further support was generated in a study of cow rumen viromes (Berg Miller *et al*., 2012), which showed similar patterns of significantly lowered metabolic gene abundances. The hourly fed reactor receives aliquots of acetate medium that result in relatively small concentration increases. This less concentrated food supply may exert stronger selective pressure related to microbial metabolism, leading to a

corresponding decrease in replication, recombination, are repair genes, compared to the other systems.

The most abundant COG functions were generally similar across samples, and showed an abundance of DNA methylase genes. The five most abundant COGs were identical and in the same order between the South and West Point viromes and coincided closely with those of the Tamaki *et al*. (2011) anaerobic digester, further supporting that there is a high degree of functional similarity among anaerobic digesters. As with the level two COG profiles, the daily fed reactor shared a similar profile of most abundant COGs, while the hourly fed reactor appeared to be relatively divergent, only sharing the observed high proportion of DNA methylases. The abundance of methylase genes agrees with prior studies of anaerobic digestion environments (Tamaki *et al*., 2011; Berg Miller *et al*., 2012). The high proportion of these genes was found in all examined compartments of the WWTP of Tamaki *et al*. (2011). Methylation of phage DNA may act as a defense mechanism against restriction digestion by prokaryotic endonucleases (Kruger and Bickle, 1983). Alternatively, these DNA methylation genes could be acting upon the host DNA, acting to manipulate gene expression to enhance their fitness. Phages in an enhanced biological phosphorus removal system have been found to contain a gene encoding H-NS, a repressor protein. It was proposed that the phages were using this protein to repress host defense mechanisms, such as CRISPRs (Skennerton *et al*., 2011). The DNA methylase genes may be acting in a similar manner.

The presence of multiple CRISPR sequences in many methanogens points to a history of infection by viruses. The majority of spacer matches in the daily fed virome and several in the hourly fed

virome were to spacers from *Methanosaeta concilii* (Table A.3). These matches provide evidence for the presence of one or more viruses targeting this acetoclastic methanogen. The dominance of *Methanosaeta* in these reactors further supports this possibility. The large number of matches for the daily fed virome to spacers from *M. concilii* may reflect that these viral sequences dominate the system along with their potential host. This possibility is supported by the fact that several of the contigs that matched the *M. concilii* were among the longest and most deeply covered contigs (data not shown). Analysis of the species richness and evenness in these samples would be beneficial in further assessing this possibility. However, the associated viruses could have other hosts as well. For instance, the virus phiF1 has been shown to target several *Methanobacterium* species (Nolling and Groffen, 1993). If an alternative host is in this system, then it would likely be similar to *M. concilii* (Krupovic *et al*., 2010). Alternatively, the spacers may have been transferred to another species (Brodt *et al*., 2011), though this event rarely occurs across genera. Analysis of the bacterial 16S rRNA genes isolated from these reactors yielded few matches to identified species, none of which are present in the CRISPRdb (data not shown). The smaller number of matches to CRISPR sequences in the hourly fed reactor could indicate that this viral community targets a larger proportion of these previously undescribed hosts. The presence of a relatively small proportion of sequences matching *M. concilii* spacers, in spite of being dominated by this methanogen like the daily fed reactor, could reflect that this system has greater species evenness, or could be an artifact of community dynamics, reflecting a relatively low point in the abundance of these viral sequences. Alternatively, it could reflect a higher rate of genomic coevolution with host species due to intense selective pressure from CRISPRs (Weinberger *et al*., 2012).

The presence of sequences matching spacers from syntrophic bacteria suggests that the reactors

have maintained this component of the original anaerobic digester community. The presence of syntrophic bacteria is supported by 16S rRNA gene analysis (data not shown). The presence of syntrophic bacteria suggests the possible presence of hydrogenotrophic methanogens or homoacetogens as well. The presence of hydrogenotrophic methanogens may explain the matches to the TerL gene of psiM2 as well as the matches to spacers of *Methanothermococcus* and *Methanocaldococcus* the hourly fed reactor. However, as these species inhabit marine geothermal sediment; these matches may be false positives. The observed diversity of the viral community and the apparent diversity of hosts indicate that multiple functional groups from the original anaerobic digester system have established stable populations in the reactors.

This study demonstrated that phages are an abundant component of the acetate-fed anaerobic community. Evidence was generated for the presence of phages of acetoclastic methanogens in anaerobic digesters. The reason for the apparent disparity in functional and taxonomic profile of the hourly fed reactor is uncertain. Further investigation of anaerobic digesters is needed in order to better explain these differences. The relatively small proportion of reactor virome sequences that matched characterized genes demonstrates the need for further study of phages that target less common species like the archaea. The ability to link unidentified viruses to hosts is an important aspect of describing viral communities that is currently limited in scope. Isolating selected portions of natural microbial communities facilitates the connection of viruses to their respective hosts. As the microbial and viral communities of more environments are explored, patterns in functional profiles and the abundance of specific gene groups, such as DNA methylases, will become more apparent. The samples analyzed in this study represent non-replicated communities and a viral community that for the digesters was a grab sample and for the reactors was averaged

over the course of several days, providing no indication of the possible variability or dynamics of these communities. More work is needed in order to understand these aspects of the relationship between viruses and both the methanogenic community and the overall anaerobic digester. There is still much that remains to be described about the viruses in anaerobic digester systems. Understanding the function of these viruses may lead to future improvements in the stability of anaerobic digesters.

References

Anderson, R.E., Brazelton, W.J., and Baross, J.A. (2011) Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiol Ecol* **77**: 120-133.

Angly, F., Rodriguez-Brito, B., Bangor, D., McNairnie, P., Breitbart, M., Salamon, P., *et al*. (2005) PHACCS, an online tool for estimating the structure and diversity of uncultured viral communities using metagenomic information. *BMC Bioinformatics* **6**: 41.

Berg Miller, M.E., Yeoman, C.J., Chia, N., Tringe, S.G., Angly, F.E., Edwards, R.A., *et al*. (2012) Phage-bacteria relationships and CRISPR elements revealed by a metagenomic survey of the rumen microbiome. *Environ Microbiol* **14**: 207-227.

Brodt, A., Lurie-Weinberger, M.R., and Gophna, U. (2011) CRISPR loci reveal networks of gene exchange in archaea. *Biol Direct* **6**: 65.

Dinsdale, E.A., Edwards, R.A., Hall, D., Angly, F., Breitbart, M., Brulc, J.M., *et al*. (2008) Functional metagenomic profiling of nine biomes. *Nature* **453**: 629-632.

Edwards, R.A., and Rohwer, F. (2005) Viral metagenomics. *Nat Rev Microbiol* **3**: 504-510.

Kruger, D.H., and Bickle, T.A. (1983) Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiol Rev* **47**: 345-360.

Krupovic, M., Forterre, P., and Bamford, D.H. (2010) Comparative analysis of the mosaic genomes of tailed archaeal viruses and proviruses suggests common themes of virion architecture and assembly with tailed viruses of bacteria. *J Mol Biol* **39**: 144-160.

Liu, Y., and Whitman, W.B. (2008) Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Ann NY Acad Sci* **1125**: 171-189.

Lorenzi, H.A., Hoover, J., Inman, J., Safford, T., Murphy, S,. Kagan, L., and Williamson, S.J. (2011) Theviral metagenome annotation pipeline (VMGAP): an automated tool for the functional annotation of viral metagenomic shotgun sequencing data. *Standards Genomic Sci* **4**: 418-429.

Mokili, J.L., Rohwer, F., and Dutilh, B.E. (2012) Metagenomics and future perspectives in virus discovery. *Curr Opin Virol* **2**: 63-77.

Nolling, J., Groffen, A., and de Vos, W.M. (1993) phiF1 and phiF3, two novel virulent, archaeal phages infecting different thermophilic strains of the genus *Methanobacterium*. *J Gen Microbiol* **139**: 2511-2516.

Roux, S., Enault, F., Robin, A., Ravet, V., Personnic, S., Theil, S., *et al*. (2012) Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS ONE* **7**: e33641.

Skennerton, C.T., Angly, F.E., Breitbart, M., Bragg, L., He, S., McMahon, K.D., *et al*. (2011) Phage encoded H-NS: a potential Achilles heel in the bacterial defence system. *PLoS ONE* **6**: e20095.

Tamaki, H., Zhang, R., Angly, F.E., Nakamura, S., Hong, P-Y., Yasunaga, T., *et al*. (2011) Metagenomic analysis of DNA viruses in a wastewater treatment plant in tropical climate. *Environ Microbiol* **14**: 441-452.

Weinberger, A.D., Sun, C.L., Plucinski, M.M., Denef, V.J., Thomas, B.C., Horvath, P., *et al*. (2012) Persisting viral sequences shape microbial CRISPR-based immunity. *PLoS ONE* **8**: e1002475.