3-1-2015

Relating Methanogen Community Structure and Anaerobic Digester Function

Benjamin T.W. Bocher
Marquette University

Keerthi Cherukuri
Marquette University

James Maki
Marquette University, james.maki@marquette.edu

M. Johnson
Marquette University

Daniel Zitomer
Marquette University, daniel.zitomer@marquette.edu

Accepted version. *Water Research*, Vol 70, (March 1, 2015): 425-435. DOI.

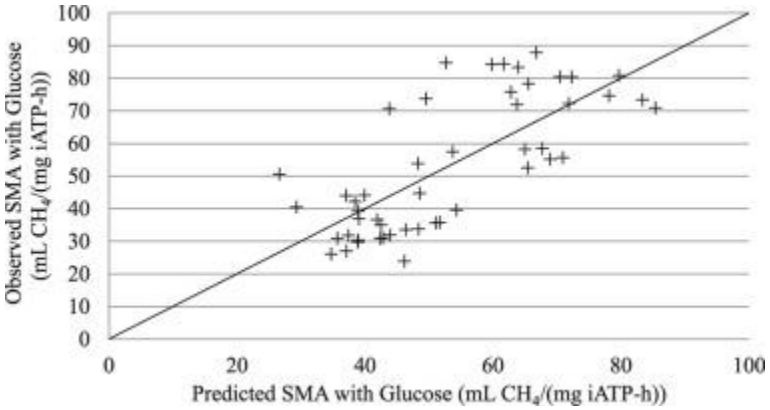
Relating Methanogen Community Structure and Anaerobic Digester Function

B.T.W. Bocher
BP Americas Inc., Petrochemicals Technology: Water Treatment, Naperville, IL
K. Cherukuri
Department of Biological Sciences, Marquette University, Milwaukee, WI
J.S. Maki
Department of Biological Sciences, Marquette University, Milwaukee, WI
M. Johnson
Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI
D.H. Zitomer
Department of Civil, Construction and Environmental Engineering, Marquette University, Milwaukee, WI

Water Research, Vol 70, (March 1, 2015): pg. 425-435. DOI.

Abstract: Much remains unknown about the relationships between microbial community structure and anaerobic digester function. However, knowledge of links between community structure and function, such as specific methanogenic activity (SMA) and COD removal rate, are valuable to improve anaerobic bioprocesses. In this work, quantitative structure–activity relationships (QSARs) were developed using multiple linear regression (MLR) to predict SMA using methanogen community structure descriptors for 49 cultures. Community descriptors were DGGE demeaned standardized band intensities for amplicons of a methanogen functional gene (*mcrA*). First, predictive accuracy of MLR QSARs was assessed using cross validation with training (n = 30) and test sets (n = 19) for glucose and propionate SMA data. MLR equations correlating band intensities and SMA demonstrated good predictability for glucose (q^2 = 0.54) and propionate (q^2 = 0.53). Subsequently, data from all 49 cultures were used to develop QSARs to predict SMA values. Higher intensities of two bands were correlated with higher SMA values; high abundance of methanogens associated with these two bands should be encouraged to attain high SMA values. QSARs are helpful tools to identify key microorganisms or to study and improve many bioprocesses. Development of new, more robust QSARs is encouraged for anaerobic digestion or other bioprocesses, including nitrification, nitritation, denitrification, anaerobic ammonium oxidation, and enhanced biological phosphorus removal.

Graphical abstract



Keywords: Anaerobic digestion, Anaerobic digestion modeling, *mcrA* gene, Microbial community structure, Quantitative structure activity relationship (QSAR), Specific methanogenic activity

1. Introduction

New sustainable waste management approaches including energy and resource recovery are now favored over energy-intensive treatment methods of the past (Angenent et al., 2004 and Lettinga,

2010). Within more sustainable approaches, anaerobic biotechnology plays a central role for low-energy oxygen demand removal, lower biosolids production than aerobic systems, and renewable energy production from biomethane (Speece, 2008, Holm-Nielsen et al., 2009 and Novotny et al., 2010). This, combined with recent advances that include the ability of anaerobic membrane bioreactors treating dilute, municipal wastewater to achieve effluent five-day biochemical oxygen demand (BOD_5) concentrations of <10 mg/L at approximately 10 °C, is making anaerobic biotechnology increasingly attractive (Shin et al., 2014), thereby encouraging the development of more robust anaerobic systems.

One challenge to anaerobic bioprocess improvement is that much remains unknown about the relationships between microbial community structure and digester function, such as biogas production and chemical oxygen demand (COD) removal rates. Although it is known that complex, interacting microbial populations accomplish the overall anaerobic degradation process, the microbial community is almost always unaccounted for in standardized testing, operation and design (Batstone et al., 2002 and Curtis et al., 2003).

1.1. Influence of environmental parameters on digester microbial community

Numerous reports describe the influence of environmental parameters on methanogenic microbial community structure, but reverse approaches describing the influence of microbial community structure on digester functional stability or methanogenic activity are less numerous. Influences of various environmental parameters on community structure include the following: (1) higher influent SO_4^{-2} concentration leads to higher sulfate-reducing bacteria levels (Raskin et al., 1996 and Pender et al., 2004), (2) higher digester acetate concentration leads to higher *Methanosarcina* and lower *Methanosaeta* abundance (Griffin et al., 1997 and McMahon et al., 2001), (3) NH_3–N concentrations greater than approximately 3 g/L leads to lower *Methanosarcina* levels, higher *Methanomicrobiales* levels (Angenent et al., 2002) and a shift from acetoclastic methanogenesis to syntrophic acetate oxidation with hydrogenotrophic methanogenesis (Fotidis et al., 2013), (4) lower temperature leads to higher diversity

at 37 °C versus 55 °C (Karakashev et al., 2005) and sometimes leads to a shift from acetoclastic methanogenesis to hydrogenotrophic methanogenesis at psychrophilic temperatures (Enright et al., 2009 and Zhang et al., 2012), (5) different substrates lead to different community structures, including manure versus wastewater sludge (Karakashev et al., 2005) and glucose versus whey permeate and sewage sludge (Lee et al., 2009), and (6) trace nutrient deprivation causes a shift in community structure, with low cobalt or nickel concentrations causing decreased *Methanosarcina* abundance and decreased COD removal rate in methanol-fed bioreactors (Fermoso et al., 2008a and Fermoso et al., 2008b).

1.2. Influence of community structure on digester functional stability

Approaches describing the influence of microbial community structure on anaerobic functional stability during and after perturbation have been reviewed (Briones and Raskin, 2003 and Allison and Martiny, 2008). These approaches include the work of Hashshan et al. (2000) who concluded that anaerobic digester communities with multiple microorganisms within the same trophic group (i.e., more parallel processing) exhibited greater functional stability after organic overload. The functional stability ostensibly resulted because one or more microorganisms were present and able to function in each critical group during and after upset. In addition, less stable community structure (i.e., greater community flexibility) may increase functional stability upon perturbation since less stable communities are more able to adapt to stress (Fernandez et al., 2000). Communities with higher evenness have been found to be more functionally resistant to selective stress than uneven communities. In this regard, Wittebolle et al. (2009) reported that denitrifying communities with higher evenness exhibited higher denitrification rates when exposed to salt toxicity compared to communities with low evenness. Although methanogenic systems were not their focus, the results may be applicable to anaerobic digesters; the theory being that in highly even communities there is a higher probability that one or more organisms resistant to the stress is present in significant enough numbers to proliferate and maintain functionality.

Water Research, Vol 70, (March 1, 2015): pg. 425-435. DOI.

1.3. Community structure, methanogenic activity and linear relationships

The influence of microbial community structure on methanogenic activity during non-perturbed operation has been investigated. In a multi-year survey of nine full-scale digesters treating brewery wastewater, communities with greater evenness and redundancy exhibited higher specific methanogenic activity (SMA) values and higher COD removal (Werner et al., 2011). In addition, higher *Bacteroidetes* and *Archaea* abundances have been shown to correlate to higher hydrolytic and methanogenic specific activities, respectively (Regueiro et al., 2012).

In our laboratory, Tale et al. (2011) measured SMA against propionate for 14 different biomass samples from full-scale anaerobic digesters and the microbial communities were also compared. Principal components analysis (PCA) depicted a linear relationship between SMA with propionate and microbial community structure defined by denaturing gradient gel electrophoresis (DGGE) banding pattern of methyl coenzyme M reductase (*mcrA*), a gene ubiquitous in methanogens. Biomass with high SMA values clustered together on a PCA plot developed using only DGGE banding patterns, whereas biomass with low SMA values clustered in a different location. In addition, the presence of hydrogenotrophic methanogens closely related to *Methanospirillum hungatei* and *Methanobacterium beijingense* was associated with high propionate SMA values. Building on this work and the work of Freitag and Prosser, 2009 and Morris et al., 2014 found that *mcrA* gene copy numbers in a group of four different enrichment cultures were linearly correlated to SMA with H_2/CO_2 ($r^2 = 0.98$).

Multiple linear regression (MLR) has been used to develop quantitative structure activity relationships (QSARs) between chemical structure descriptors and biological or physiochemical activities, such as toxicity or Henry's Constant values (Nirmalakhandan and Speece, 1988). When Tale (2010) applied MLR to activity data, SMA values were linearly related to community structure as defined by standardized, demeaned DGGE band intensities of *mcrA* amplicons. However, the resulting MLR equation was overfitted and not predictive

Water Research, Vol 70, (March 1, 2015): pg. 425-435. DOI.

because too many independent variables (i.e., 10 bands) were used to predict the SMA values of too few samples (i.e., 14 biomass samples). Therefore, an insufficient number of different biomass samples was employed, preventing development of a predictive relationship (Tale, 2010). However, the reports described above encouraged research using MLR and a larger group of anaerobic microbial communities to develop QSARs.

In this study, a large set of 49 different methanogenic microbial communities was employed to develop predictive QSARs between microbial community structure descriptors and methanogenic activity with glucose and propionate. The objectives were to provide a tool for future research regarding community structure–activity relationships and advance development of models relating microbial community structure and anaerobic digester design and operation.

2. Materials and methods

Initially, an investigation was performed to determine if the activity of a blend of biomass from two different laboratory anaerobic digesters could be predicted from activity of each biomass alone, and if a linear correlation was appropriate. Subsequently, 49 microbial communities were assayed to establish QSARs.

2.1. Blends of two different laboratory biomass samples

For initial investigation of blended biomass, SMA tests were performed using mixtures of biomass from two different mesophilic, continuously mixed digesters as well as each biomass alone. Biomass 1 was from a laboratory anaerobic digester previously operated for over five years and fed non-fat dry milk, 5 g/L $NaHCO_3$ and basal nutrient medium. The Biomass 1 digester was operated at a solids retention time (SRT) of 15 days and organic loading rate (OLR) of approximately 2 g COD/L-day. Biomass 2 was from a laboratory anaerobic co-digester (SRT = 14 days, OLR = 3 g COD/L-day) previously operated for over two months and fed a synthetic primary sludge (Natural Choice Dog Food, NutroProducts, Inc., Industry, CA, USA) and five industrial co-digestates, including soft drink, corn ethanol and cheese production wastes as described by Navaneethan

et al. (2011). Daily, both digesters received 5 g/L $NaHCO_3$ and the basal nutrient medium. Blends of these two digesters were composed of different Biomass 1: Biomass 2 volatile solids (VS) ratios ranging from 0:1 to 1:0 (Navaneethan et al., 2011).

2.1.1. Specific methanogenic activity (SMA) for blended biomass samples

Concentrations of Biomass 1 and 2 ranged from 3.6 to 10.1 g VS/L. Mixtures were prepared in 15 mL of basal nutrient medium, which was placed in a 160 mL serum bottle, sparged with O_2-free gas (7:3 v/v N_2:CO_2), and sealed with solid Balch-type butyl rubber stoppers (Geo-Microbial Technologies, Inc., Ochelata, OK) and aluminum-crimped seals. Biomass was allowed to produce biogas endogenously for three days. Endogenous biogas was then removed on day three to depressurize the headspace and 12 g/L calcium acetate was added.

Produced biogas volume was measured at ambient pressure and 35 °C for 20 days using a glass syringe with a wetted glass barrel. Syringe content was re-injected into the serum bottle after volume measurement. Headspace methane content was measured after 20 days. Cumulative methane volume produced versus time was plotted and MPR (mL CH_4/h) was calculated using linear regression on the portion of the curve within the first 60 h of biogas production. SMA (mL CH_4/g VS-h) was calculated by dividing MPR by VS mass.

2.2. Microbial communities to establish quantitative structure–activity relationships

Microbial communities used to establish QSARs included anaerobic biomass from nine full-scale digesters, four laboratory propionate enrichment cultures and 36 different mixtures of the digester biomasses with the propionate enrichment cultures.

2.2.1. Biomass from full-scale digesters

Digester biomass samples were obtained from full-scale, mesophilic anaerobic digesters with different configurations, including

complete mix stirred tank reactors, plug flow digesters, and upflow anaerobic sludge blanket reactors, treating diverse substrates from brewing, cheese processing, food flavorings production, milk production and soft drink bottling industries as well as municipal wastewater sludge. Biomass samples were also geographically diverse, coming from locations in the U.S., including California, Colorado, Iowa, New Mexico, Pennsylvania and Wisconsin (see Supplementary Data, Table S-1). A variety of configurations, substrates and locations was employed in an effort to obtain a range of microbial community structures.

2.2.2. Propionate enrichment cultures

Propionate enrichment cultures were maintained at 35 ± 1 °C in 2 L polyethylene terephthalate (PETE) vessels containing 1 L of culture on a shaker table in an incubator (Model 4350, Thermo Scientific, Marietta, OH). All enrichment cultures were initially seeded with a mix of eight biomass samples from operating mesophilic anaerobic bioreactors treating agricultural, brewery, beverage industry, food industry, municipal wastewater sludge, and nonfat dry milk wastes. Culture effluent (100 mL) was removed once per day to maintain a SRT and hydraulic residence time (HRT) of ten days. Effluent was replaced with 100 mL of substrate containing calcium propionate (0.77 ± 0.03 g COD/L-day), 5 g/L $NaHCO_3$ and basal nutrient medium. The theoretical maximum Ca^{+2} and Na^{+} concentrations were 1.5 and 1.4 g/L, respectively, which were well below typical inhibitory concentrations (Speece, 2008). Biogas volume was measured at ambient pressure using a wet test gas meter (Precision Scientific Petroleum Instruments, San Antonio, TX). Four enrichment conditions that differed based upon oxygen (O_2) addition rate were employed in triplicate in an effort to obtain different microbial communities. The four O_2 addition rates were equivalent to 0, 1.3, 6.7, and 12.5% of the COD loading rate. Oxygen was added by injecting O_2 (aviator breathing grade; 99.9% purity; approximately 1 atm and 22 °C) once a day into the enrichment vessel headspace with a wetted-barrel, glass syringe immediately after substrate was added. Propionate enrichment cultures receiving low doses of O_2 were anticipated to produce significant amounts of methane under steady-state conditions, as described elsewhere (Zitomer and Shrout, 1998).

2.2.3. Mixtures of digester and propionate enrichment culture

In an effort to generate a larger number of microbial communities for statistical analyses, digester biomass and propionate enrichment cultures were mixed to produce synthetic communities. This technique was inspired by the work of Wittebolle et al. (2009) who mixed 18 pure cultures of denitrifying species in different combinations to generate a large number of synthetic communities to relate community evenness to denitrification activity after a perturbation (Wittebolle et al., 2009). The 36 mixtures employed herein consisted of 4:1 intracellular adenosine-5′-triphosphate (iATP) mass ratio mixtures of each digester biomass with each propionate enrichment culture, respectively. Before mixing, granular biomass samples from upflow anaerobic sludge blanket (UASB) reactors denoted in Supplementary Data Table S-1 were disrupted to make them flocculant.

2.3. Methane production rate (MPR) and specific methanogenic activity (SMA) assays

Methane production rate (MPR) and SMA were determined in triplicate at 35 ± 1 °C and 150 rpm using a gyratory shaker-incubator (model C25KC, New Brunswick Scientific, Edison, NJ) to quantify activity against glucose and propionate (Angelidaki et al., 2007). Biomass was thickened at 6000 rpm (g = 7000) for five minutes in a centrifuge (Clinical 200, VWR International, Radnor, Pennsylvania Germany). Granular samples were disrupted to make flocculant biomass before centrifuging.

VS concentration was used to estimate active biomass fraction in initial studies involving blends of two laboratory biomass samples. However, VS values are inaccurate quantifiers of active biomass (Roe and Bhagat, 1982, Cairns et al., 2005 and Whalen et al., 2006) because the inactive and soluble portions of VS do not metabolize substrate or intermediates, but are included in the VS measurement (Perle et al., 1995). Biomasses fed biologically inert VS, having longer sludge age, or recently subjected to stress may have a significant portion of inactive VS (Quirk and Eckenfelder, 1986). Therefore,

Water Research, Vol 70, (March 1, 2015): pg. 425-435. DOI.

intracellular adenosine triphosphate (iATP) was used to estimate active biomass for microbial communities used to establish QSARs.

For QSAR activity testing an active biomass concentration of 0.74 mg iATP/L was prepared by re-suspending thickened biomass in basal nutrient medium. The suspension (25 mL) was placed in a 160 mL serum bottle, sparged with O_2-free gas (7:3 v/v N_2:CO_2) and sealed with solid Balch-type butyl rubber stoppers (Geo-Microbial Technologies, Inc., Ochelata, OK) and aluminum-crimped seals. Biomass was allowed to produce biogas endogenously for three days. Endogenous biogas was then removed to depressurize the headspace and either glucose (5.6 g COD/L) or calcium propionate (4.6 g COD/L) was added. Triplicate bottles that contained no exogenous substrate were prepared similarly and used as blanks to account for endogenous methane production. Blanks without additional substrate provided endogenous decay MPR and SMA values, which were subtracted from gross MPR and SMA value, respectively, to calculate net SMA values that were employed to establish QSARs.

Produced biogas volume was measured at ambient pressure and 35 °C for 20 days using a glass syringe with a wetted glass barrel. Syringe content was re-injected into the serum bottle after volume measurement. Headspace methane content was measured after 20 days. Cumulative methane volume produced versus time was plotted and MPR (mL CH_4/h) was calculated using linear regression on the portion of the curve within the first 60 h of biogas production when the biogas production rate reached a constant value. SMA (mL CH_4/mg iATP-h) was calculated by dividing MPR by active biomass.

Propionate was chosen as a substrate to quantify activity because its degradation, due partially to the subsequent H_2 transfer, is often the rate-limiting step in the overall anaerobic degradation process; thus, the rate of methane production from numerous substrates, including those that are more complex, may be related to the SMA values with propionate (McCarty and Smith, 1986 and Ito et al., 2012). Glucose was selected to utilize a substrate that involved an additional microbial guild of fermentative organisms. Furthermore, degradation of glucose has previously been employed to investigate methanogenic activity and stability (e.g., Hashsham et al., 2000, Fernandez et al., 2000 and Dearman et al., 2006).

2.4. Microbial community analysis

Biomass (50 mL) was centrifuged for 10 min at 2500× g (AccuSpin Micro 17, Thermo Fischer Scientific, Waltham, MA), and DNA was extracted from 0.75 mL of thickened biomass using a kit (DNA Powersoil™ Total DNA Isolation Kit, MoBio Laboratories, Carlsbad, CA) according to the manufacturer's instructions. The presence of extracted DNA was confirmed using gel electrophoresis (1% agarose in 1× TrisAcetate–EDTA, TAE) with ethidium bromide (0.8 μL/mL) stain (Sambrook and Russell, 2001). DNA samples were stored at −80 °C until further analysis.

DNA was amplified using PCR, a thermocycler (PTC-200 DNA Engine Cycler, Bio-Rad, Foster City, CA) and EconoTaq® PLUS 2X Master Mix, which included the Taq polymerase (Lucigen Corporation, Middleton, WI). Forward and reverse primers were added to the PCR tube with nuclease-free H_2O to make a 100 μL reaction including approximately 100 ng of template. Nested PCR was performed on the extracted DNA according to the program outlined by Tale et al. (2011). PCR was performed using primers for the *mcrA* gene of methanogenic *Archaea* designed by Luton et al. (2002). These primers are: mcrA1f: 5′ – GGTGGTGTMGGATTCACACARTAYGCWACAGC – 3′, (mcrF); GCmcrA1f: 5′ – *GC-clamp-GGTGGTGTMGGATTCACACARTAYGCWACAGC – 3′, (GCmcrA1f), where GC-clamp = 5′–CGCCCGCCGCGCCCCGCGCCCGTGCCGCCGCCCCCGCCCG – 3′, (GC-clamp); mcrA500r: 5′ – TTCATTGCRTAGTTWGGRTAGTT – 3′, (mcrR). The primer product was an approximately 460-bp-long segment of *mcrA*, which codes for the alpha subunit of methyl coenzyme M reductase (Luton et al., 2002). Since the *mcrA* gene is specific to and ubiquitous in methanogens (Thauer, 1998), it has been used to compare methanogen community structure and identify the taxonomically distinct methanogens in many environments (Morris et al., 2014). Because of the degenerate nature of the primers, the recommendations of Luton et al. (2002) were followed, which included a slow increase in temperature ($0.1°Cs^{-1}$) between the annealing and extension steps of the first five cycles to aid in the initial product formation.

Each microbial community was fingerprinted using DGGE, which separated amplified genes into bands on a polyacrylamide gel. DGGE has been used with *mcrA* as a target gene (e.g., Tale et al., 2011). The denaturant concentration used for DGGE varied linearly over 75 mm and ranged from 40% at the top of the gel to 70% at the bottom of the gel (expressed as v/v of the total gel volume). A detection system (BioRad Universal DCode Mutation Detection System, Richmond, CA) was used to run the DGGE gels using 300 ng of DNA product in each lane of the gel with 2× blue loading dye. DGGE was performed on 1-mm-thick 8% polyacrylamide gel following the manufacturer's protocol. An electric potential of 100 V was maintained across the gel for 12 h. A 1% SYBR® gold dye solution (Invitrogen, CA) was used to stain the gel. After immersing the gel in the staining solution and rotating it for 30 min on a shaker table at a speed sufficient to mix the dye solution, it was viewed under ultraviolet light using an imaging system (GDS-8000 Bioimaging System, UVP Inc., Upland, CA).

DGGE band intensities were determined using gel viewing software (Lab Works v. 4.6.00.0 Lablogics, Inc., Mission Viejo, CA) with a minimum band height of 0.050, allowed error of ±5%, and the following options activated: dark bands and bright background, rows of equal molecular weight, maximum OD level for the image, and center peak. Multiple DGGE gels were prepared to accommodate all 49 biomass samples. A ladder was run on every gel to adjust for variations among DGGE gels. The total ladder band intensity on each gel was divided by the average total ladder band intensity among all gels to calculate a standardized ladder band intensity for each gel. The intensity of each band "n" on a gel was then divided by the standardized total ladder band intensity for that gel to yield a normalized band intensity value for each band (I_n). The demeaned, normalized band intensity for band "n" (X_n) was then calculated as follows:

equation(1)

$$X_n = I_n - \left(\sum_{j=1}^{m} I_{n,j}\right) \Big/ m$$

Water Research, Vol 70, (March 1, 2015): pg. 425-435. DOI.

where $I_{n,j}$ is the normalized band intensity value of band "n" in lane "j" and "m" is the number of lanes. The X_n values were used to create MLR equations.

2.5. Multiple linear regression (MLR)

MLR equations expressing SMA (dependent variable) as a function of DGGE demeaned, normalized band intensity values (independent variables) were in the following form:

equation(2)

$$\mathrm{SMA} = \sum_{n=1}^{m} \beta_n(X_n) + \varepsilon$$

where SMA is the specific methanogenic activity with propionate or glucose (mL CH_4/mg iATP-h), β_n is the correlation coefficient for band "n" (mL CH_4/mg iATP-h) and ε is a constant (mL CH_4/mg iATP-h). Initially, stepwise regression using forward selection was performed with MATLAB (v. R2010bSP1, MathWorks®, Natick, MA) using the "regress" function to determine the number of independent variables (bands) to include in Equation (2). For this, MLR equations were generated using all possible combinations of the 15 bands. The maximum R^2 value was observed using a minimum of five bands. A small number of independent variables was desirable since a higher number could result in overfitted MLR equations causing poor predictability. Therefore, the small group of five bands was selected for subsequent regression analysis.

Cross-validation was performed using the leave-group-out (LGO) procedure (Konovalov et al., 2008). Others have jackknifed data in this way to assess the predictability of QSARs (Nirmalakhandan and Speece, 1988a, Eriksson et al., 2003 and Konovalov et al., 2008). A random-number generator was used to select a subset of 30 samples from which to fit MLR equations (i.e., training set). The remaining 19 samples were then used to estimate the predictability of the MLR equations, as indicated by the value of the calculated validation parameter, q^2 (Konovalov et al., 2008). With a LGO approach, the q^2 value is generally accepted as a means to determine if a model is predictive (Eriksson et al., 2003, Benigni and Bossa, 2008 and Konovalov et al., 2008). After confirming that equations

demonstrated good predictability, training and test data were combined and used to determine final MLR equations.

2.6. Basal nutrients and other analysis

The basal nutrient medium included macro- and micro-nutrients required for methanogenic culture growth and was described by Schauer-Gimenez et al. (2010). The iATP concentration was measured using a luciferase-based test kit (QuenchGone21 Wastewater Test Kit, LuminUltra, Fredericton, New Brunswick, Canada). Briefly, extracellular ATP was measured using a proprietary stabilizing reagent and total ATP was measured after cell lysis with NaOH reagent. The iATP concentration was calculated as the difference between total and extracellular ATP concentrations.

Temperature and pH of propionate enrichments were measured daily using a glass electrode and meter (Orion 4 Star pH-DO Benchtop electrode – 9206BN, Thermo Scientific, Marietta, OH). Total solids (TS), VS, volatile suspended solids (VSS), COD, volatile fatty acid (VFA) concentrations (acetic, propionic, butyric, iso-butyric, valeric, and iso-valeric acids), and biogas methane content were determined by gas chromatography using standard methods (APHA et al., 1998). For soluble COD (SCOD) analysis, samples were centrifuged at 13,000 rpm for ten minutes (Clinical 200 VWR International LLC Radnor, Pennsylvania) and supernatant was filtered through a 0.45 μm filter (Whatman International Ltd., Maidstone, England). Filtrate was then tested for SCOD.

3. Results and discussion

3.1. Activity of blended laboratory digester biomass samples

Biomass 1 and 2 samples were blended in various proportions to produce synthetic microbial communities. The methane production rates of the mixtures ($MPRp_{mix}$) as well as the methane production rate of each biomass sample alone were measured with propionate as the substrate. $MPRp_{mix}$ values were linearly correlated to the MPR values of the individual cultures (Fig. 1a):

Water Research, Vol 70, (March 1, 2015): pg. 425-435. DOI.

equation(3)

$$MPRp_{mix}=f_1MPRp_1+f_2MPRp_1$$

where $MPRp_1$ and $MRPp_2$ are the measured MPR values of the individual biomass samples with propionate (mL CH_4/h) and f_1 and f_2 are the decimal volatile solids mass factions of Biomass 1 and 2, respectively (g VS_i/g VS_{total}), with $f_1 + f_2 = 1$. The coefficient of determination (R^2) between measured and predicted $MPRp_{mix}$ values was 0.95, indicating that the linear correlation (Equation (3)) quantified most of the $MPRp_{mix}$ variability (Fig. 1a).

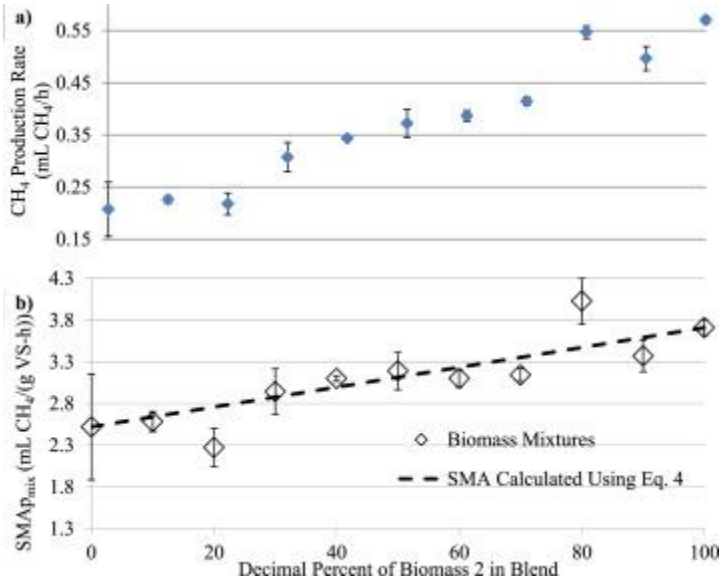


Fig. 1. Relationship between blend ratio of two methanogenic cultures and the methanogenic activity of the culture blend. Methane production rate (a) and specific methanogenic activity (b) of a blend of Biomass 1 and 2. Diamonds represent experimental values, the dotted line represents the values calculated using Equation (4).

MPR data are not normalized by dividing by microbial mass; therefore, SMA values that were normalized to the amount of biomass were used for all subsequent activity testing. The propionate SMA values of biomass mixtures ($SMAp_{mix}$) were linearly correlated to the SMA values of the individual biomass samples (Fig. 1b):

equation(4)

$$SMAp_{mix}=f_1SMAp_1+f_2SMAp_2$$

Water Research, Vol 70, (March 1, 2015): pg. 425-435. DOI.

where $SMAp_1$ and $SMAp_2$ are the SMA values for Biomass 1 and 2, respectively (mL CH_4/gVS h) with propionate. The coefficient of determination (R^2) between measured and predicted $SMAp_{mix}$ values was >0.75, indicating that a linear correlation (Equation (4)) quantified most of the $SMAp_{mix}$ variability (Fig. 1b). The lower coefficient of determination for predicted and calculated SMA in comparison to MPR may be attributed to inaccuracies using VS concentration to quantify active biomass. Therefore, iATP concentration was used in subsequent testing in an attempt to gain a more accurate active biomass estimate.

Additionally, Equation (4) was used to predict SMA values of the 36 blended cultures subsequently used to develop QSARs. The propionate SMA values of each of the two individual cultures ($SMAp_1$ for full-scale digester biomass and $SMAp_2$ for propionate enrichment culture) were multiplied by their respective iATP mass percent in the blend (f_1 = 0.8 for full-scale digester biomass and f_2 = 0.2 for propionate enrichment culture). Observed (experimental) $SMAp_{mix}$ values were graphed against predicted $SMAp_{mix}$ values (Fig. 2). The observed $SMAp_{mix}$ was linearly correlated to $SMAp_1$ and $SMAp_2$, R^2 = 0.905 (Fig. 2), implying that methanogenic activity may be linearly correlated to the abundance of certain methanogens.

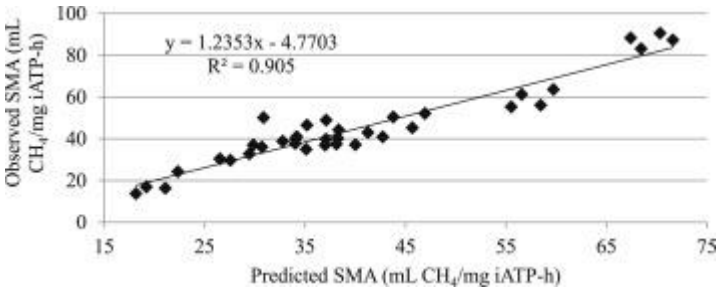


Fig. 2. Observed (experimental) SMA for blends of digester biomass and enrichment cultures versus predicted SMA values using Equation (4).

3.2. Activity of 49 digester, propionate enrichment cultures and their mixtures

Quasi steady state data for the enrichment cultures are presented in Table 1. All enrichment cultures produced methane and achieved typical methanogenic operation. Biomass from the nine full-scale digesters and four enrichments alone were assayed as well as 36

biomass and enrichment mixtures, yielding a total of 49 potentially different methanogenic communities. Methanogen communities of the nine biomass samples and four propionate enrichments were different from each other as depicted using principal component analysis for the key bands that were correlated to SMA values as described below (see Fig. S-1).

Table 1. Quasi steady state propionate enrichment culture parameters.

Enrichment	0%	1.3%	6.7%	12.5%
Oxygen addition (mg O_2/d)	0	11	53	99
pH	7.09 ± 0.07	7.11 ± 0.04	7.04 ± 0.03	6.99 ± 0.05
$SCOD_{effluent}$ (mg/L)	160 ± 10	170 ± 10	210 ± 30	160 ± 30
$SCOD_{removal}$ (%)	98 ± 1.7	97 ± 1.3	97 ± 1.3	98 ± 1.9
CH_4 (L/d)	0.21 ± 0.02	0.22 ± 0.02	0.21 ± 0.03	0.19 ± 0.03
CH_4 Content (%)	78 ± 0.52	79 ± 0.52	73 ± 0.50	73 ± 2.5
$VSS_{effluent}$ (g/L)	0.76 ± 0.14	0.73 ± 0.06	0.84 ± 0.02	0.95 ± 0.03
$CH_3CH_2COO^-_{effluent}$ (mg/L)	49 ± 1.3	86 ± 28	52 ± 15	47 ± 14
$CH_3COO^-_{effluent}$ (mg/L)	<20[a]	<20[a]	<20[a]	40 ± 30

[a]Below 20-mg/L detection limit.

A significant range of activity was observed among the 49 biomass samples; SMA with glucose (SMAg) values varied more than 250% and ranged from 24 ± 5.1 to 88 ± 1.4 mL CH_4/mg iATP-h with an average of 53 ± 20 mL CH_4/mg iATP-h (Fig. 3). The SMA with propionate (SMAp) values varied more than 600% and ranged from 15 ± 0.30 to 91 ± 4.2 mL CH_4/mg iATP-h with an average of 44 ± 18 mL CH_4/mg iATP-h (Fig. 4). The variation was in agreement with previous observations that average SMAp values vary greatly among biomass samples from different anaerobic digesters when VS was used as an active biomass measure (Tale et al., 2011). Even when an ostensibly better metric of active biomass, iATP, was employed rather than VS, a large specific activity range was still observed. The large range of SMA values was desirable for developing a relationship between activity and microbial community structure.

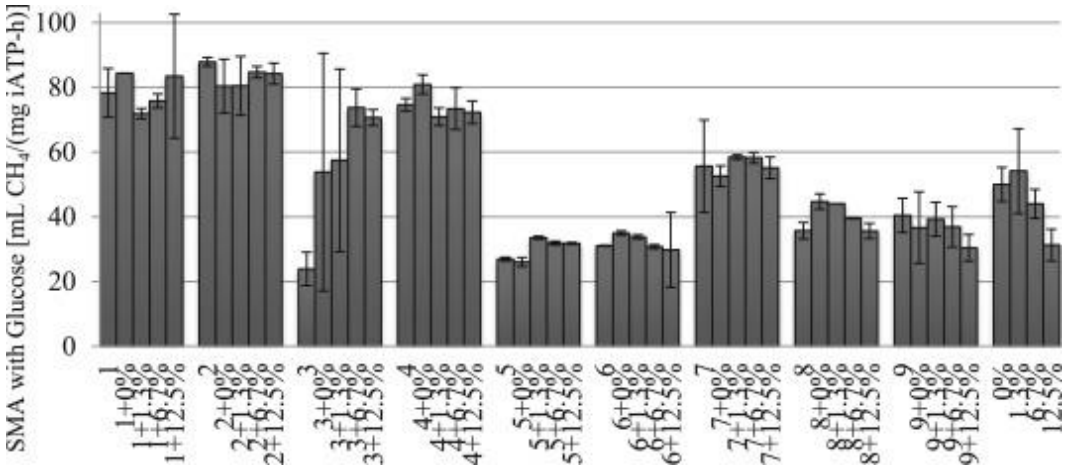


Fig. 3. Observed SMAg values for 49 anaerobic microbial communities. Each group of five bars includes a digester biomass sample and, from left to right, the biomass sample mixed with enrichments that received 0, 1.3, 6.7, and 12.5% of the COD organic loading rate, respectively. The last group on the right represents the enrichment cultures alone.

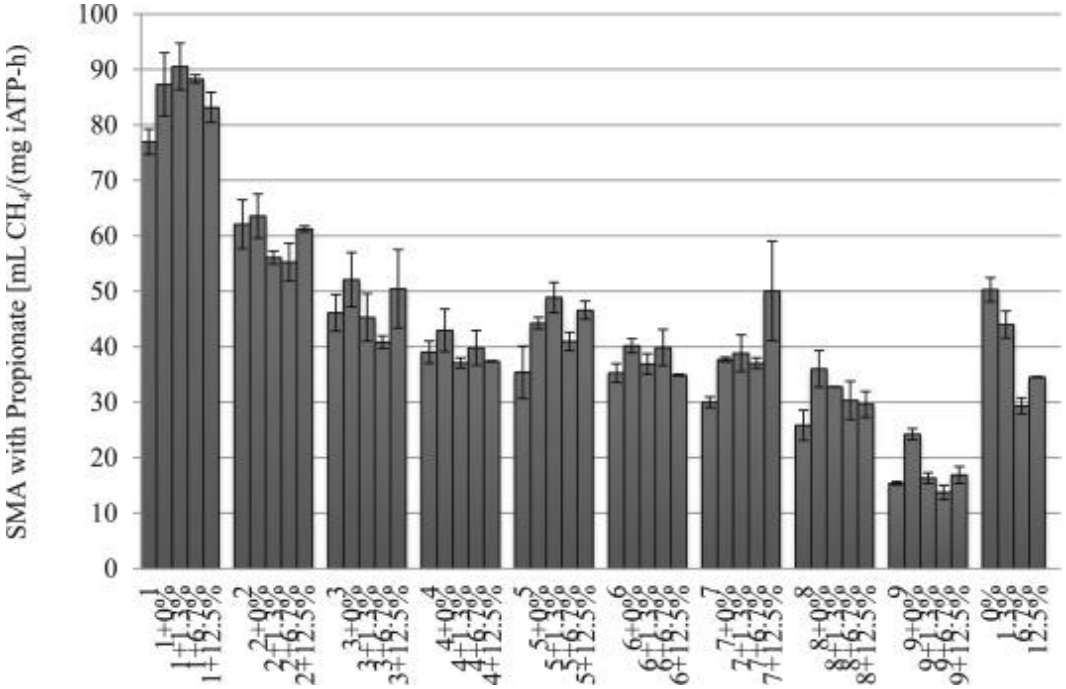


Fig. 4. Observed SMAp values for 49 anaerobic microbial communities. Each group of five bars includes a digester biomass sample and, from left to right, the biomass sample mixed with enrichments that received 0, 1.3, 6.7, and 12.5% of the COD organic loading rate, respectively. The last group on the right represents the enrichment cultures alone.

3.3. Spearman's Rank correlation coefficient

Based on DGGE banding patterns of *mcrA* gene amplicons (i.e., methanogen community structure), SMA values with propionate and with glucose correlated (r_s = 0.561) above the 99.9% level for a two-tailed test using Spearman's Rank Correlation Coefficient (Zar, 1972). Therefore, a greater presence of specific bands indicated a greater methane production rate from both glucose and propionate. These results suggested that (1) the same rate-limiting step (e.g., H_2 transfer) was present for SMA tests with both substrates and (2) the methanogen community could be used to predict SMA values even if acidogenesis and acetogenesis are also occurring in addition to methanogenesis.

3.4. Quantitative structure–activity relationship (QSAR)

Two quantitative, predictive relationships between methanogen community structure and biomass activity (i.e., QSARs) were developed using MLR, demeaned normalized DGGE band intensities and SMA values normalized using iATP mass. No quantitative and predictive relationships resulted when SMA values were normalized by VS mass. In addition, no quantitative and predictive relationship was found using the Shannon Index, Dice Similarity Coefficient, Jaccard Similarity Coefficient, and Pearson's Correlation Coefficient as community descriptors.

Results of stepwise regression using forward selection showed that employing seven of the 15 DGGE bands (Bands 4, 7, 8, 10, 11, 14 and 15) was sufficient to yield MLR equations with $R^2 > 0.50$ (Fig. 5). A small number of independent variables (i.e., bands) was desirable since a higher number could result in overfitted MLR equations with poor predictivity. Therefore, a small group of five bands was selected for subsequent regression analysis for each set of SMA data. This is in line with reports that methanogenesis from propionate and glucose involved four to six key microbes in trophic groups (Ito et al., 2012).

Water Research, Vol 70, (March 1, 2015): pg. 425-435. DOI.

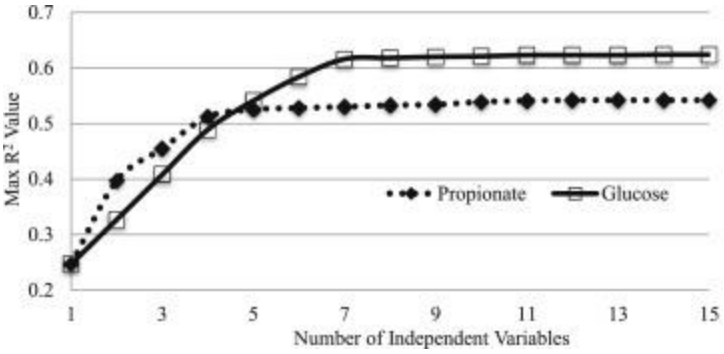


Fig. 5. MLR maximum coefficient of determination (R^2) versus number of independent variables (i.e., bands) employed.

A sample size of 30 has been shown to accurately test five independent variables at the recommended minimum $r^2 = 0.4$ for a predictive model (Topliss and Costello, 1972). Therefore, 30 biomass samples were randomly selected and used as a training set to develop MLR equations for SMAg and SMAp. The remaining 19 biomass samples were used as a test set via the LGO approach. When the MLR developed using the training set was challenged to predict SMA using test set data, the average q^2 value, used to assess the predictability of each MLR equation, was 0.53 for SMAg and 0.52 for SMAp correlations. Both q^2 values were above the 0.5 threshold reported to describe a regression with good quantitative predictability (Eriksson et al., 2003).

Because the LGO approach yielded regressions with good SMA predictability, all 49 biomass samples were then used to develop the following MLR equations that quantitatively predicted methanogenic activities as a function of microbial community structure (Fig. 6 and Fig. 7):

equation(5)

$$\mathrm{SMAg} = -430(\mathrm{X_4}) - 470(\mathrm{X_7}) - 76(\mathrm{X_{11}}) - 79(\mathrm{X_{14}}) + 170(\mathrm{X_{15}}) + 89,\ \mathrm{n} = 49;\ \mathrm{R}^2 = 0.54$$

equation(6)

$$\mathrm{SMAp} = -220(\mathrm{X_4}) - 82(\mathrm{X_8}) + 340(\mathrm{X_{10}}) - 52(\mathrm{X_{14}}) + 180(\mathrm{X_{15}}) + 50,\ \mathrm{n} = 49;\ \mathrm{R}^2 = 0.53$$

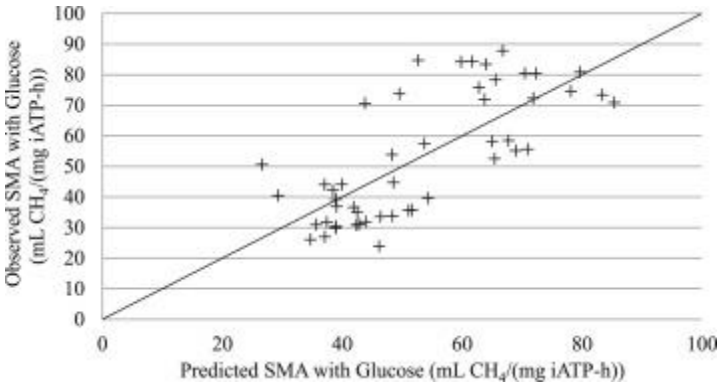


Fig. 6. Comparison of observed and predicted SMAg values using QSAR equation (Equation (5)).

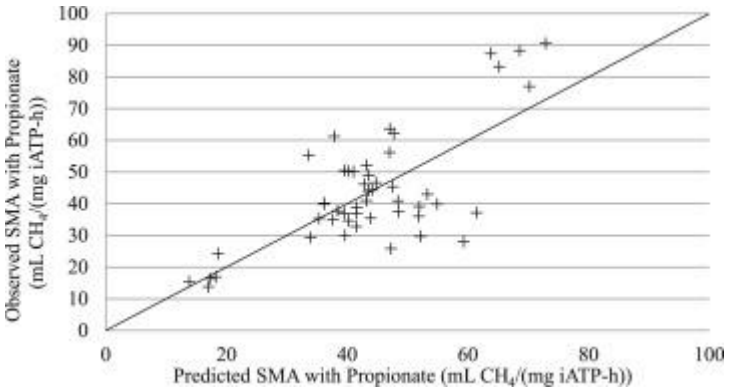


Fig. 7. Comparison of observed and estimated SMAp values using QSAR equation (Equation (6)).

Previous research indicated a quantitative link between SMA value and methanogen community structure existed, but the small number of microbial communities employed was insufficient to develop statistically significant correlations and to estimate predictability (Freitag and Prosser, 2009, Tale, 2010, Tale et al., 2011, Regueiro et al., 2012 and Morris et al., 2014). Equations (5) and (6)), however, were developed using a larger number of microbial communities and were both good correlations and exhibited good predictability.

Even with the inherent variability of current activity test methods and potential DGGE bias and limitations, these models were predictive (von Wintzingerode et al., 1997, Vallaeys et al., 1997, Muyzer and Smalla, 1998 and Muyzer, 1999). Better correlations with greater predictability may result from the use of quantitative polymerase chain reaction (qPCR), next generation sequencing

methods, or other approaches to more accurately characterize microbial communities and activities.

The positive correlation coefficient values (β values) for X_{15} in both glucose and propionate MLR equations (Equations (5) and (6))) as well as for X_{10} in the propionate MLR equation (Equation (6)) suggested that methanogens associated with these bands contribute most significantly to high SMA values. It is possible that digesters engineered specifically to foster growth of these methanogens or their direct addition to anaerobic digesters (i.e., bioaugmentation) may increase maximum methane production rate under stable operating conditions, but additional research is warranted to explore these possibilities. Future research should include excision and sequencing of bands to identify, at a minimum, the key methanogens (i.e., X_{15} and X_{10}) and investigate the relationship between their abundance and methanogenic activity as well as digester function.

Methanogens represented by X_n values with negative correlation coefficients (i.e., negative coefficient methanogens) are associated with lower maximum methane production rates; higher demeaned standardized band intensities associated with these methanogens correlated to lower SMA values.

The model developed in this research correlated SMA and community structure only, not overall digester performance or stability and microbial community structure. Therefore, additional research to develop different QSAR models for function descriptors other than SMA (e.g., resistance or stressed conditions) is needed.

In a broad context, QSARs may be helpful tools to study and improve many bioprocesses, especially when multiple microorganisms exist in the same trophic group (i.e., trophic redundancy), functional genes or other molecular data can be used to characterize critical trophic groups, and linear or other relationships are predictive. Therefore, it is hoped that this research will serve as a template to encourage development of new, more robust QSARs for anaerobic digestion using other functional responses and other microbial community descriptors. In addition, new QSARs could be developed for other bioprocesses, including nitrification, nitritation, denitrification,

anaerobic ammonium oxidation and enhanced biological phosphorus removal.

4. Conclusions

The specific methanogenic activity (SMA) of a blend of two different methanogenic biomasses from digesters fed complex substrates can be predicted from a linear combination of the individual biomass SMA values. Therefore, a linear correlation is appropriate to consider when developing empirical relationships between community structure and SMA.

SMA values for different biomass samples from well-operating methanogenic processes vary greatly (i.e., as much as 600%), even when a putatively better metric, iATP, is employed rather than VS to quantify active biomass. The large range of SMA values indicates that the biomass activity in many low-SMA anaerobic digesters could be increased. In addition, the large range of activity is desirable for developing empirical models relating activity to microbial community structure.

Two quantitative structure activity relationships (QSARs) between methanogen community structure (DGGE band intensities for a methanogen functional gene) and biomass activity (SMA with glucose and propionate) were developed using multiple linear regression. The relationships were observed using a larger number of microbial communities (i.e., 49) and were both good correlations ($R^2 > 0.5$) that exhibited good predictability ($q^2 > 0.5$). Even with the inherent variability of activity testing and potential DGGE bias and limitations, these models were predictive. In the future, better correlations with greater predictability may result from the use of quantitative polymerase chain reaction (qPCR), next generation sequencing methods, or other approaches to more accurately characterize microbial communities and activities.

Positive correlation coefficient values for two of the 15 DGGE bands employed in QSAR equations reveal that higher relative abundance of methanogens associated with these bands contribute to higher SMA values. It is possible that anaerobic digesters engineered specifically to foster growth of these two methanogens or their direct

Water Research, Vol 70, (March 1, 2015): pg. 425-435. DOI.

addition to anaerobic digesters (i.e., bioaugmentation) may increase maximum methane production rate under stable operating conditions.

QSARs are helpful tools to identify key microorganisms or to study and improve many bioprocesses. Therefore, it is hoped that this research will encourage development of new, more robust QSARs for anaerobic digestion using other functional responses and other microbial community descriptors. In addition, new QSARs could be developed for other bioprocesses, including nitrification, nitritation, denitrification, anaerobic ammonium oxidation and enhanced biological phosphorus removal.

Acknowledgments

The authors thank WE Energies, Marquette University and the Seeds of Faith Fellowship Program for funding, LuminUltra Technologies, Ltd. (Fredericton, Canada) for providing iATP testing reagents and luminometer, as well as the following laboratory contributors: foremost, Mike Dollhopf, as well as Caitlin Collins, Kevin Glauber, Kyle Hill, Mike Leipus, John Ross, Vaibhav Tale, and Sarah Walsh.

References

Allison and Martiny, 2008. S.D. Allison, J.B.H. Martiny. Resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. U. S. A.,* 105 (2008), pp. 11415–11519

Angelidaki et al., 2007. I. Angelidaki, M. Alves, D. Bolzonella, L. Borzacconi, L. Campos, A. Guwy, P. Jenicek, S. Kalyuzhnui, J. Van Leir. *Anaerobic Biodegradation, Activity and Inhibition Task Group Meeting Report. Prague Czech Republic, October, 9–10, 2006.* (2007)

Angenent et al., 2002. L. Angenent, S. Sung, L. Raskin. Methanogenic population dynamics during startup of a full-scale anaerobic sequencing batch reactor treating swine waste. *Water Res.,* 36 (18) (2002), pp. 4648–4654

Angenent et al., 2004. L.T. Angenent, K. Karim, M.H. Al-Dahhan, B.A. Wrenn, R. Domíguez-Espinosa. Production of bioenergy and biochemicals from industrial and agricultural wastewater. *Trends Biotechnol.,* 22 (9) (2004), pp. 477–485

APHA et al., 1998. APHA, WEF, AWWA. *Standard Methods for the Examination of Water and Wastewater.* American Public Health Association, Washington, DC (1998)

Batstone et al., 2002. D.J. Batstone, J. Keller, I. Angelidaki, S.V. Kalyuzhnyi, S.G. Pavlostathis, A. Rozzi, W.T.M. Sanders, *et al.* The IWA anaerobic digestion model no. 1 (ADM1). *Water Sci. Technol.,* 45 (10) (2002), pp. 65–73

Benigni and Bossa, 2008. R. Benigni, C. Bossa. Predictivity of QSAR. *J. Chem. Inf. Model.,* 48 (5) (2008), pp. 971–980

Briones and Raskin, 2003. A. Briones, L. Raskin. Diversity and dynamics of microbial communities in engineered environments and their implications for process stability. *Curr. Opin. Biotechnol.,* 14 (3) (2003), pp. 270–276

Cairns et al., 2005. J.E. Cairns, P.A. Whalen, P.J. Whalen, D.R. Tracey, R.E. Palo, K. Street, K. Tower. Dissolved ATP – a new process control parameter for biological wastewater treatment. *Proceedings of Water Environment Federation Technical Exposition and Conference (WEFTEC), October 29 – November 2, Washington D.C., U.S.A (2005)*

Curtis et al., 2003. T.P. Curtis, I.M. Head, D.W. Graham. Theoretical ecology for engineering biology. *Environ. Sci. Technol.,* 37 (3) (2003), pp. 64A–70A

Dearman et al., 2006. B. Dearman, P. Marschner, R.H. Bentham. Methane production and microbial community structure in single-stage batch and sequential batch systems anaerobically co-digesting food waste and biosolids. *Appl. Microbiol. Biotechnol.,* 69 (2006), pp. 589–596

Enright et al., 2009. A. Enright, V. McGrath, D. Gill, G. Collins, V. O'Flaherty. Effect of seed sludge and operation conditions on performance and archaeal community structure of low-temperature anaerobic solvent-degrading bioreactors. *Syst. Appl. Microbiol.,* 32 (1) (2009), pp. 65–79

Eriksson et al., 2003. L. Eriksson, J. Jaworska, A.P. Worth, M.T.D. Cronin, R.M. McDowell, P. Gramatica. Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. *Environ. Health Perspect.,* 111 (10) (2003), pp. 1361–1375

Fermoso et al., 2008a. F.G. Fermoso, G. Collins, J. Bartacek, V. O'Flaherty, P. Lens. Acidification of methanol-fed anaerobic granular sludge bioreactors by cobalt deprivation: induction and microbial community dynamics. *Biotechnol. Bioeng.,* 99 (1) (2008), pp. 49–58

Fermoso et al., 2008b. F.G. Fermoso, G. Collins, J. Bartacek, V. O'Flaherty, P. Lens. Role of nickel in high rate methanol degradation in anaerobic granular sludge bioreactors. *Biodegradation,* 19 (5) (2008), pp. 725–737

Fernandez et al., 2000. A.S. Fernandez, S.A. Hashsham, S.L. Dollhopf, L. Raskin, O. Glagoleva, F.B. Dazzo, R.F. Hickey, *et al.* Flexible community structure correlates with stable community function in methanogenic bioreactor communities perturbed by glucose. *Appl. Environ. Microbiol.,* 66 (9) (2000), pp. 4058–4067

Fotidis et al., 2013. I. Fotidis, D.B. Karakashev, I. Angelidaki. The dominant acetate degradation pathway/methanogenic composition in full-scale anaerobic digesters operating under different ammonia levels. *Int. J. Environ. Sci. Technol.* (2013) http://dx.doi.org/10.1007/s13762-013-0407-9

Freitag and Prosser, 2009. T.E. Freitag, J.I. Prosser. Correlation of methane production and functional gene transcriptional activity in a peat soil. *Appl. Environ. Microbiol.,* 75 (21) (2009), pp. 6679–6687

Griffin et al., 1997. M.E. Griffin, K.D. McMahon, R.I. Mackie, L. Raskin. Methanogenic population dynamics during start-up of anaerobic digesters treating municipal solid waste and biosolids. *Biotechnol. Bioeng.,* 57 (3) (1997), pp. 342–355

Hashsham et al., 2000. S.A. Hashsham, A.S. Fernandez, S.L. Dollhopf, F.B. Dazzo, R.F. Hickey, J.M. Tiedje, C.S. Criddle. Parallel processing of substrate correlates with greater functional stability in methanogenic

bioreactor communities perturbed by glucose. *Appl. Environ. Microbiol.,* 66 (9) (2000), pp. 4050–4057

Holm-Nielsen et al., 2009. J.B. Holm-Nielsen, T. Al Seadi, P. Oleskowicz-Popiel. The future of anaerobic digestion and biogas utilization. *Bioresour. Technol.,* 100 (22) (2009), pp. 5478–5484

Ito et al., 2012. T. Ito, K. Yoshiguchi, H.D. Ariesyady, S. Okabe. Identification and quantification of key microbial trophic groups of methanogenic glucose degradation in an anaerobic digester sludge. *Bioresour. Technol.,* 123 (0) (2012), pp. 599–607

Karakashev et al., 2005. D. Karakashev, D.J. Batstone, I. Angelidaki. Influence of environmental conditions on methanogenic compositions in anaerobic biogas reactors. *Appl. Environ. Microbiol.,* 71 (1) (2005), pp. 331–338

Konovalov et al., 2008. D.A. Konovalov, L.E. Llewellyn, Y. Vander Heyden, D. Coomans. Robust cross-validation of linear regression QSAR models. *J. Chem. Inf. Model.,* 48 (10) (2008), pp. 2081–2094

Lee et al., 2009. C. Lee, J. Kim, K. Hwang, V. O'Flaherty, S. Hwang. Quantitative analysis of methanogenic community dynamics in three anaerobic batch digesters treating different wastewaters. *Water Res.,* 43 (1) (2009), pp. 157–165

Lettinga, 2010. G. Lettinga. Challenges of a feasible route towards sustainability in environmental protection. *Front. Environ. Sci. Eng. China,* 4 (2) (2010), pp. 123–134

Luton et al., 2002. P.E. Luton, J.M. Wayne, R.J. Sharp, P.W. Riley. The mcrA gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. *Microbiology,* 148 (2002), pp. 3521–3530

McCarty and Smith, 1986. P. McCarty, D. Smith. Anaerobic wastewater treatment. *Environ. Sci. Technol.,* 20 (12) (1986), pp. 1200–1206

McMahon et al., 2001. K. McMahon, P.G. Stoot, Mackie, L. Raskin. Anaerobic codigestion of municipal solid waste and biosolids under

various mixing conditions II: microbial population dynamics. *Water Res.,* 35 (7) (2001), pp. 1817–1827

Morris et al., 2014. R. Morris, A. Schauer-Gimenez, U. Bhattad, C. Kearney, C. Struble, D. Zitomer, J. Maki. Methyl conenzyme M reductase (mcrA) gene correlates with activity measurements of methanogenic H2/CO2-enriched anaerobic biomass. *Microb. Biotechnol.,* 7 (1) (2014), pp. 77–84

Muyzer and Smalla, 1998. G. Muyzer, K. Smalla. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonier Leeuwenhoek,* 73 (1) (1998), pp. 127–141

Muyzer, 1999. G. Muyzer. DGGE/TGGE a method for identifying genes from natural ecosystems. *Curr. Opin. Biotechnol.,* 2 (1999), pp. 317–322

Navaneethan et al., 2011. N. Navaneethan, P. Topczewski, S. Royer, D. Zitomer. Blending anaerobic co-digestates: synergism and economics. *Water Sci. Technol.,* 63 (12) (2011), pp. 2916–2922

Novotny et al., 2010. V. Novotny, J. Ahern, P. Brown. *Water Centric Sustainable Communities.* John Wiley & Sons, Inc, Hoboken, NJ (2010)

Nirmalakhandan and Speece, 1988. N. Nirmalakhandan, R.E. Speece. Structure activity relationships: quantitative techniques for predicting the behavior of chemicals in the ecosystem. *Environ. Sci. Technol.,* 22 (6) (1988), pp. 606–615

Nirmalakhandan and Speece, 1988a. N.N. Nirmalakhandan, R.E. Speece. Prediction of aqueous solubility of organic chemicals based on molecular structure. *Environ. Sci. Technol.,* 22 (3) (1988), pp. 328–338

Pender et al., 2004. S. Pender, M. Toomey, M. Carton, D. Eardly, J.W. Patching, E. Colleran, V. O'Flaherty. Long-term effects of operating temperature and sulphate addition on the methanogenic community structure of anaerobic hybrid reactors. *Water Res.,* 38 (3) (2004), pp. 619–630

Perle et al., 1995. M. Perle, S. Kimchie, G. Shelef. Some biochemical aspects of the anaerobic degradation of dairy wastewater. *Water Res.,* 29 (6) (1995), pp. 1549–1554

Quirk and Eckenfelder, 1986. T. Quirk, W.W. Eckenfelder. Active biomass in activated sludge analysis and design. *J. Water Pollut. Control Fed.,* 58 (9) (1986), p. 932

Raskin et al., 1996. L. Raskin, B.E. Rittmann, D.A. Stahl. Competition and coexistence of sulfate-reducing and methanogenic populations in anaerobic biofilms. *Appl. Environ. Microbiol.,* 62 (10) (1996), pp. 3847–3857

Regueiro et al., 2012. L. Regueiro, P. Veiga, M. Figueroa, J. Alonso-Gutierrez, A. Stams, J.M. Lema, M. Carballa. Relationship between microbial activity and microbial community structure in six full-scale anaerobic digesters. *Microbiol. Res.,* 167 (10) (2012), pp. 581–589

Roe and Bhagat, 1982. P.C. Roe, S.K. Bhagat. Adenosine triphosphate as a control parameter for activated sludge processes. *J. Water Pollut. Control Fed.,* 54 (3) (1982), pp. 244–254
http://www.jstor.org/stable/25041280 Water Environment Federation

Sambrook and Russell, 2001. J. Sambrook, D.W. Russell. *Molecular Cloning: a Laboratory Manual.* (third ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (2001)

Schauer-Gimenez et al., 2010. A.E. Schauer-Gimenez, D.H. Zitomer, J.S. Maki, C.A. Struble. Bioaugmentation for improved recovery of anaerobic digesters after toxicant exposure. *Water Res.,* 44 (2010), pp. 3555–3564

Shin et al., 2014. C. Shin, P. McCarty, J. Kim, J. Bae. Pilot-scale temperate-climate treatment of domestic wastewater with a staged anaerobic fluidized bed membrane bioreactor (SAF-MBR). *Bioresour. Technol.,* 159 (2014), pp. 95–103

Speece, 2008. R. Speece. *Anaerobic Biotechnology and Odor/Corrosion Control for Municipalities and Industries.* Archae Press, Nashville, TN (2008)

Tale, 2010. V.P. Tale. *Bioaugmentation for Recovery of Anaerobic Digesters Subjected to Organic Overload.* (Ph.D. dissertation) Marquette University, Milwaukee, WI (2010)

Tale et al., 2011. V.P. Tale, J.S. Maki, C.A. Struble, D.H. Zitomer. Methanogen community structure-activity relationship and bioaugmentation of overloaded anaerobic digesters. *Water Res.,* 45 (16) (2011), pp. 5249–5256

Thauer, 1998. R.K. Thauer. Biochemistry of methanogenesis: a tribute to Marjory Stephenson. *Microbiology,* 144 (1998), pp. 2377–2406 35

Topliss and Costello, 1972. J.G. Topliss, R.J. Costello. Chance correlations in structure-activity studies using multiple regression analysis. *J. Med. Chem.,* 15 (10) (1972), pp. 1066–1068

Vallaeys et al., 1997. T. Vallaeys, E. Topp, G. Muyzer, V. Macheret, G. Laguerre, A. Riguad, G. Soulas. Evaluation of denaturing gradient gel electrophoresis in the detection of 16S rDNA sequence variaton in rhizobia and methanotrophs. *FEMS Microbiol. Ecol.,* 24 (1997), pp. 279–285

von Wintzingerode et al., 1997. F. von Wintzingerode, U.B. Göbel, E. Stackebrandt. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.,* 21 (3) (1997), pp. 213–229 Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/9451814

Werner et al., 2011. J.J. Werner, D. Knights, M.L. Garcia, N.B. Scalfone, S. Smith, K. Yarasheski, T.A. Cummings, A.R. Beers, R. Knight, L.T. Angenent. Bacterial community structures are unique and resilient in full-scale bioenergy systems. *Proc. Natl. Acad. Sci. U. S. A.,* 108 (10) (2011), pp. 4158–4163

Whalen et al., 2006. P.A. Whalen, P.J. Whalen, D.R. Tracey. Cellular ATP – a superior measure of active biomass for biological wastewater treatment processes. *Proceedings of Water Environment Federation Technical Exposition and Conference (WEFTEC), Dallas, TX, U.S.A, vol. 4,* Water Environment Federation (2006), pp. 3025–3037

Wittebolle et al., 2009. L. Wittebolle, M. Marzorati, L. Clement, A. Balloi, D. Daffonchio, K. Heylen, P. De Vos, *et al.* Initial community evenness favours functionality under selective stress. *Nature,* 458 (7238) (2009), pp. 623–626

Zar, 1972. J.H. Zar. Significance testing of the spearman rank correlation coefficient. *J. Am. Stat. Assoc.,* 67 (339) (1972), pp. 578–580

Zhang et al., 2012. D. Zhang, W. Zhu, C. Tang, Y. Suo, L. Gao, X. Yuan, X. Wang. Bioreactor performance and methanogenic population dynamics in a low-temperature (5–18 °C) anaerobic fixed-bed reactor. *Bioresour. Technol.,* 104 (1) (2012), pp. 136–143

Zitomer and Shrout, 1998. D.H. Zitomer, J.D. Shrout. Feasibility and benefits of methanogenesis under oxygen limited conditions. *Waste Manag.,* 18 (1998), pp. 107–116

Appendix A. Supplementary data

7. Supplementary Information

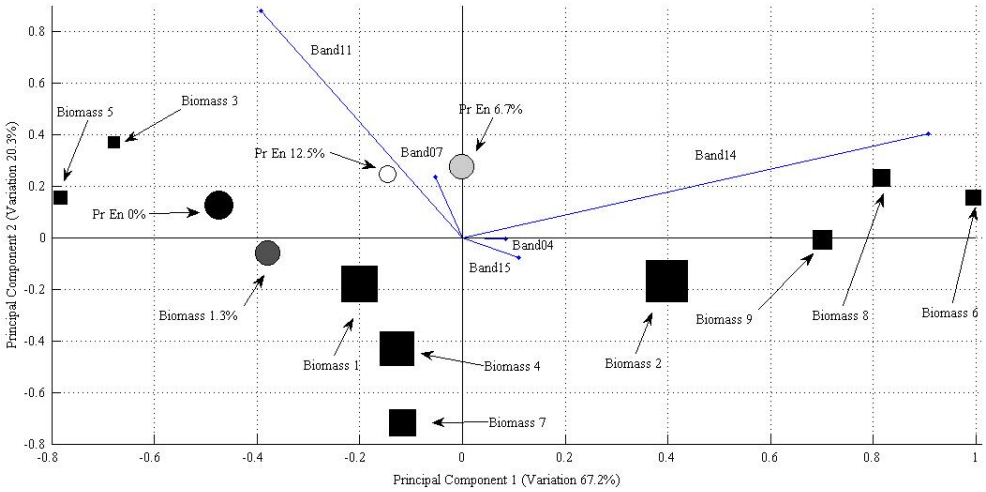
Table S-1. Anaerobic Biomass Samples

Sample Number	Digester Type*	Substrate	Location
1	CSTR	Food Flavorings	Jackson, WI
2	Anaerobic Plug Flow	Brewery Waste	Fort Collins, CO
3	CSTR	Municipal Wastewater Solids	Philadelphia, PA
4	UASB	Brewery Waste	Chico, CA
5	CSTR	Non-Fat Dry Milk	Milwaukee, WI
6	UASB	Soft Drink Bottling Waste	Watertown, WI
7	CSTR	Municipal Wastewater Solids	Des Moines, IA
8	UASB	Brewery Waste	LaCrosse, WI
9	CSTR	Cheese Processing Waste	Las Cruces, NM

* CSTR - Completely mixed stirred tank reactor, UASB - Upflow anaerobic sludge sludge blanket

**WWTP – Wastewater treatment plant

a)



b)

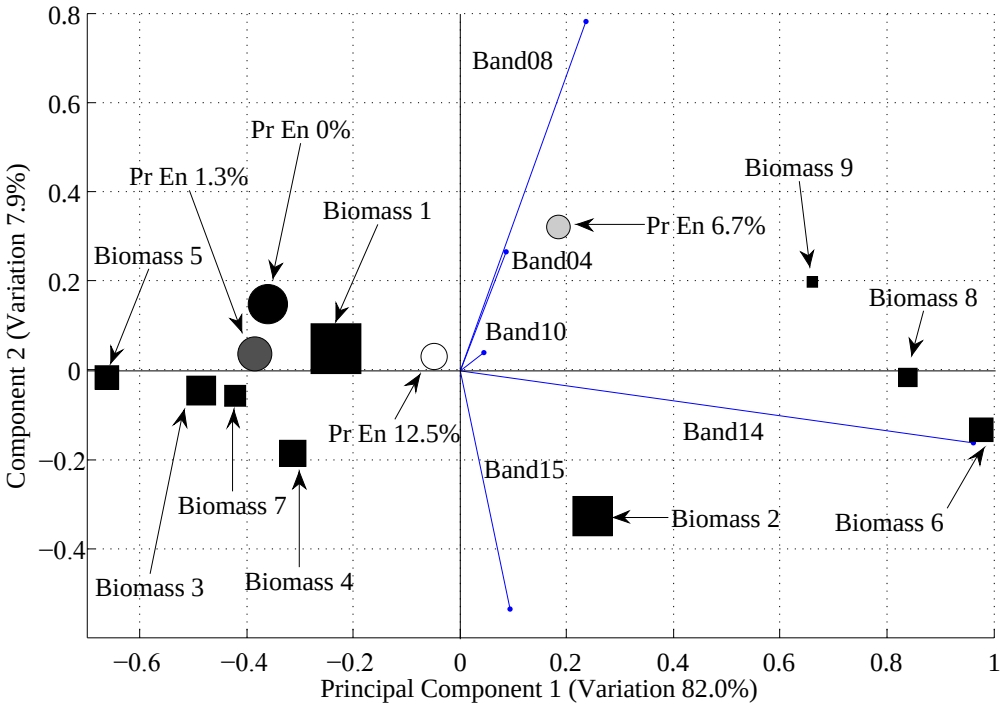


Figure S-1. Principal component analysis plot using all the 13 source cultures (9 biomass and 4 enrichment cultures).
Bands 4, 8, 10, 14, and 15 were used to establish the MLR correlation with propionate SMA (a), and bands 4, 7, 11, 14, and 15 were used to establish MLR correlation with glucose SMA. Propionate enrichment (Pr En) and full-scale digester biomass (Biomass) are depicted.

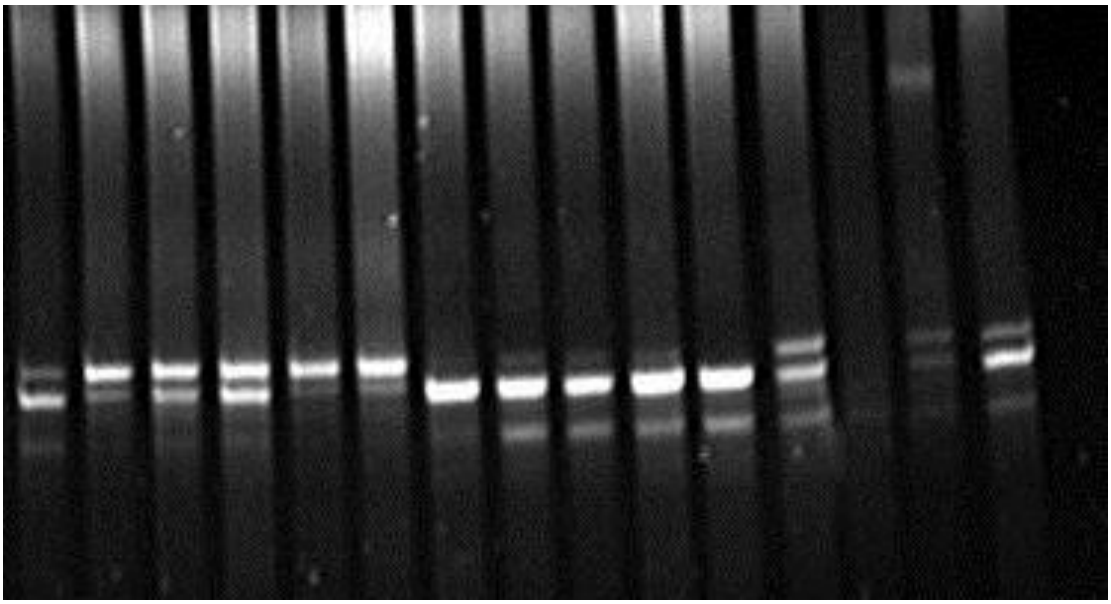


Figure S-2. DGGE gel for 13 source cultures cultures (9 biomass and 4 enrichment cultures).
Lanes (from left to right) include ladder, biomass samples 1 to 9, propionate enrichments receiving 0, 1.3, 6.7, and 12.5% of the oxygen demand and duplicate ladder.