A Novel High-Throughput Multi-Parameter Flow Cytometry Based Method for Monitoring and Rapid Characterization of Microbiome Dynamics in Anaerobic Systems

Abhishek S. Dhoble¹, Sadia Bekal¹, William Dolatowski³, Connor Yanz⁴, Kris N. Lambert², and
Kaustubh D. Bhalerao¹, *

¹Department of Agricultural and Biological Engineering, ²Department of Crop Sciences,
University of Illinois at Urbana-Champaign, Urbana 61801, United States.

³Tempus, Inc., Wheaton, IL. ⁴SunPower Corporation, Los Angeles, CA.

*Author for correspondence.

E-mail: bhalerao@illinois.edu, Tel: +1-217-244-6569; Fax: +1-217-244-0323

Abstract

A novel multidimensional flow cytometry based method has been demonstrated to monitor and rapidly characterize the dynamics of the complex anaerobic microbiome associated with perturbations in external environmental factors. While community fingerprinting provides an estimate of the meta genomic structure, flow cytometry provides a fingerprint of the community morphology including its autofluorescence spectrum in a high-throughput manner. Using anaerobic microbial consortia perturbed with the controlled addition of various carbon sources, it is possible to quantitatively discriminate between divergent microbiome analogous to community fingerprinting techniques using automated ribosomal intergenic spacer analysis (ARISA). The utility of flow cytometry based method has also been demonstrated in a fully functional industry scale anaerobic digester to distinguish between microbiome composition caused by varying hydraulic retention time (HRT). This approach exploits the rich multidimensional information from flow cytometry for rapid characterization of the dynamics of microbial communities.

Keywords: *Flow Cytometry, Anaerobic Digestion, Microbial Community, Labe Free*

Monitoring, Autofluorescence, Microbial Dynamics, Microbiome Characterization

1. Introduction

In today's era of sustainable energy sources, production of energy from wastes is an essential component (Angenent et al., 2004). Anaerobic bioreactors are a prominent source of bioenergy (Pandit et al., 2015) commonly employed for the treatment of wastewater, stabilization of sludge, and the treatment of hazardous and solid waste streams (Dhoble and Pullammanappallil, 2014). However, there is no simple, in-line, cost effective method available to distinguish between bioreactors that perform well from those that perform inadequately (Leitão et al., 2006). Therefore, there is a general perception that anaerobic bioreactors are unreliable or unstable (Amani et al., 2010).

A healthy and functioning microbiome is considered crucial to the stability and performance of anaerobic bioreactors and their functional stability is governed by dynamic interaction within these microbial communities (Werner et al., 2011). In order to monitor and characterize perturbations in the structure and function of a microbiome, there is a need to develop tools that measure microbiome features beyond genomic or functional diversity (Müller et al., 2012). The strict anaerobic requirements are an impediment to culture based studies of underlying microbial communities (Koch et al., 2014a). It has been demonstrated previously that microbial communities in anaerobic bioreactors can be characterized using molecular probes (Raskin et al., 1994). However the complexity arising from probe design, characterization and availability of rRNA sequences make this approach difficult to implement in an online monitoring system. There is a demand for a quick, simple and **effective method for monitoring** the structure and function of the bioreactor microbial communities (Kinet et al., 2016).

1 Low cost next generation sequencing (NGS) is not high throughput enough to resolve
2 dynamic changes in the structure of the microbiome over time (Yuan et al., 2015). Flow
3 cytometry can provide this information in a high throughput manner since its fast, inline,
4 automated, permits sample labeling, and requires small sample volumes and minimal sample
5 preparation (Müller and Nebe-von-Caron, 2010). It also requires low capital investment, has low
6 per-sample cost, produces rich high dimensional information and is capable of sorting and
7 classifying a sample (Koch et al., 2014b).

8 The central hypothesis of this work is that flow cytometry can classify microbial
9 consortia based on characteristics such as viability, metabolic activity, and morphology. The
10 rationale for the proposed methodology is that the combinations of these characteristics form
11 unique ‘cytometric fingerprints’ (Koch et al., 2014a) that can complement existing technologies
12 (Park et al., 2005) and may facilitate rapid characterization of the dynamics of microbial
13 communities. The goal of this study is to demonstrate the potential usefulness and sensitivity of a
14 novel multi-dimensional flow cytometry based method to allow rapid high-throughput
15 characterization of the dynamics of microbial community evolution and intracommunity
16 dynamics. Using anaerobic microbial consortia perturbed with controlled addition of carbon
17 sources (Owen et al., 1979), the microbial diversity information obtained from established
18 culture-independent technique like ARISA (Fisher and Triplett, 1999) has been demonstrated in
19 this study to be comparable with flow cytometry based analysis. Furthermore, flow cytometry
20 with non-invasive autofluorescence has been demonstrated to potentially detect and classify
21 (type) microbial consortia in an industrial scale anaerobic digester operated at varying hydraulic
22 retention times (HRTs), which can provide high throughput microbiological bioprocess

evaluation and appropriate in-line as well as off-line intervention strategies in bioprocess industries.

Here, a novel methodology has been demonstrated using multiple flow cytometry signals: cell size (FSC or forward scatter), cell granularity / morphology (SSC or side scatter) and autofluorescence (corresponding to an excitation of 405 nm and emission wavelength of 450 nm) towards a high-throughput tool to assess microbial community structure and dynamics. This approach has potential for rapid, label free monitoring and characterizing dynamics of the microbial communities in natural as well as engineered environments.

2. Materials and Methods

2.1 Experimental setup and source of anaerobic digestion culture

Samples of the anaerobic microbial communities were collected from the mesophilic anaerobic digester of Urbana and Champaign Sanitary District (UCSD), Urbana, IL as shown in Fig. S4. The waste activated sludge used in this study was withdrawn from the secondary settling tank of a municipal wastewater treatment plant of UCSD. Laboratory cultures were setup to enrich the community with different carbon sources to introduce biases in the underlying microbial communities. The cultures were incubated in Corning No.1460, 250 ml serum bottles. In each serum bottle, 100 ml of inoculum and the corresponding nutrient solution was added. Each bottle was fed with 2000 mg/L COD of glucose (GLUC), cellulose (CELL), propionate (PROP), butyrate (BUTY), acetate (ACET) and waste activated sludge (SLUD).

The contents of the bottles were sparged with nitrogen gas for about five minutes before sealing. Bottles were sealed with butyl rubber serum caps and crimped with aluminum seals of appropriate size. The sealed bottles were incubated at $35\pm 2^{\circ}\text{C}$. Each experiment was conducted

1 in triplicate accompanied with controls containing only inoculated medium. In all cases, biogas
2 production was measured every 24 h using a water displacement column. After gas measurement
3 the reactors were shaken once a day manually. Moisture content, total and particulate volatile
4 solids, soluble chemical oxygen demand and pH were measured according to Standard Methods
5 for the Examination of Water and Wastewater (APHA, 1998). The peak day samples correspond
6 to the maximum biogas production observed as shown in the supplementary figure S1-S3.
7 Samples with no additional carbon sources were represented as NONE_0 for the zero day sample
8 and NONE_2 and NONE_5 for peak biogas day and fifth day samples respectively.

9 For UCSD experiments, samples were collected on Day 1, 2, 4, 14 and 26 from Digester
10 1 (DIG 1) and Digester 2 (DIG 2) in order to capture the microbiome dynamics beyond the usual
11 residence time. Both the digesters have a same volume but different flow rates making them
12 operational at different Hydraulic Retention Times (HRTs). DIG 1 has the HRT of 13 d while
13 DIG 2 has the HRT of 26 d. In order to explore potential applications of flow cytometry in
14 detecting changes in the microbial community at the industrial scale anaerobic digesters, the
15 focus of the present study was on differences caused by varying HRTs.

16 For autofluorescence control experiments, *E. coli* HB101 strain was obtained from
17 Promega, Madison, WI in which bacterial cells were grown overnight at 37°C then diluted 1:100
18 into 10 ml of fresh LB medium to the optimal optical density (OD) (Rezaeinejad and Ivanov,
19 2013) in a 50-ml tube and grown at 37°C to an OD₆₀₀ of ~1.0. 1:100 diluted subcultures were
20 then grown in 5 mL of 3 different LB concentrations (10mg/L, 20 mg/L and 40 mg/L). Multiple
21 samples were collected during different growth phases and the flow cytometry analyses were
22 performed as described below.

2.2 Flow cytometric analysis

Samples of the anaerobic microbial communities were collected from each serum bottle via a syringe with an 18 gauge needle every 24 h following a biogas measurement. Initial sample (NONE_0) was the same for all the assays, which was the fresh inoculum collected from UCSD. 750 μ L of sample from each serum bottle was strained prior to flow cytometry using BD Falcon 12x75 mm Tube with Cell Strainer Cap having a 35 μ m nylon mesh (Catalog No. 352235). The strained samples were suspended in PBS-1X. Analyses were performed immediately on a Biosciences LSR II Flow Cytometry Analyzer. The excitation laser was tuned for 405 nm. Autofluorescence was measured as light passing a 450/50 PMT B band pass filter with no long pass dichroic mirror. Signals were amplified with a 4-decade log amplifier and collected at a rate of approximately 1000-events/s. Total numbers of events collected were 100,000 for each sample.

2.3 ARISA analysis

Samples for DNA isolation were also collected accompanying the flow cytometry sampling. Total genomic DNA was extracted from 1.5 mL of leachates from each assay using the MoBio PowerFecal DNA Isolation Kit (MoBio Laboratories, Inc., Carlsbad, CA) following manufacturer's instructions. Purity of the extracted DNA was determined by measuring the 260/280 nm absorbance ratios. The extracted DNA samples were stored at -20°C before analysis. ARISA was conducted as described by Fisher and Triplett (1999) with the following modifications: DNA in the range of 15 to 250 ng was used for ARISA analysis. Each PCR reaction contained 10 pmoles of each primer pair; five primer pairs were tested for each DNA sample, and 4 μ L of Taq 5X mix (New England BioLabs) in a total volume of 20 μ L. The reaction was run on a thermocycler using touchdown PCR at 95°C (30 s), [95°C (30 s), 65°C (30 s)-

0.5°C/cycle, 68°C (1:00 min) x 20 cycles], [95°C (30 s), 55°C (30 s), 68°C (1:00 min) x 16 cycles], 68°C (7 m). The resulting PCR products were confirmed by agarose gel electrophoresis, then diluted three fold and sent for fragment analysis at the W. M. Keck Center for Comparative and Functional Genomics at the University of Illinois at Urbana- Champaign. DNA samples were analyzed by denaturing capillary electrophoresis using an ABI 3730XL Genetic Analyzer (Applied Biosystems Inc.). Electrophoresis conditions were 63°C and 15 kV with a run time of 120 m using POP-7 polymer. DNA fragment sizes were calculated using the Peak Scanner 2 software package (Applied Biosystems Inc.) using a threshold of 100 to 500 fluorescence units. The total peak area for each peak was normalized by dividing the area of individual peaks by the total fluorescence detected in each sample run. The primers used are described in Table 1. All oligonucleotides were synthesized by Integrated DNA Technologies.

2.4 Statistical analysis

For the analysis of flow cytometry data, the .fcs files obtained from a Biosciences LSR II Flow Cytometry Analyzer were imported and the various distance matrices were calculated using Root Mean Square Distance (RMSD) analysis. A 3D structure based on the FSC-A, SSC-A and Autofluorescence-A data was created and a bin number of 50 were defined along each dimension. Data visualization and ordination analyses were conducted using R programming language. The inter sample variability in the ARISA data set was obtained by first making a matrix of distances between each observation and then obtaining distribution of distances of points within replicates. The data from six equally spaced time points was utilized to do an autocorrelation analysis on the flow cytometry data. A vector that consists of the distances from the NONE_0 sample was assembled all the way through the 5d samples for every carbon source, for each replicate. For a

given carbon source, the matrix containing the time series data was obtained for that particular replicate.

The primary tools used for ARISA data analysis was the Euclidean Distance Matrix. ARISA profiles were analyzed using PeakScanner Software (Applied Biosystems Inc.). The base pair lengths were quantized to 4 units and the area was normalized. One of the bacterial primers listed in Table 1, archaeal primer and fungal primer were concatenated and the area was normalized for every reading for individual primers. The between and within sample variability was obtained in the same manner as flow cytometry data. To obtain the significance of difference between the between sample variability distribution and the within sample variability distribution, Kolmogorov- Smirnov Test was performed.

3. Results and discussion

3.1 Autocorrelation between divergent microbial communities as a quick tool to monitor perturbations.

Fig.1 (A) shows that the culture continues to evolve over time. The dissimilarity for samples is smallest for a time delay of 1d, and it continues to increase over time. This figure is similar to an autocorrelation function commonly seen in time series data. The increased distance with delay suggests a divergent culture, as seen in GLUC whereas PROP and BUTY are rather stable. ACET and the controls (NONE, SLUD) showed similar behavior (increasing sharply for a delay of 4). Fig.1 (B) looks at day-to-day (delay=1) changes over time plotted as a function of the day with respect to which the day is measured. The point corresponding to minimum change is an indicator of a stable extremum in the colony structure. The dissimilarity is greatest between

days 1 and 2, which can be interpreted as the time at which the greatest perturbation was observed. The extremum in the distance can be interpreted as the time at which the phenotype reaches a stable point under perturbation.

Previous perturbation experiments show that, when anaerobic systems were subjected to substrate shock loadings, the recovery periods could last from a few days to weeks (McCarty and Mosey, 1991). Depending upon the magnitude of the perturbation, an anaerobic community takes varying amount of time to its fully functional condition (Cohen et al., 1982). ACET along with controls NONE and SLUD showed sharp increases for a delay of 4 (Fig.1-B). Certain communities can grow and remain in the digester at subdominant levels and yet be capable of immediate and significant responses to environmental changes. Such subdominant communities take part in the reactive capacity of the digester ecosystem and may help it face perturbations and foster functional stability. In this autocorrelation function based approach, the holistic cytometric fingerprints based on morphology (FSC), granularity/internal complexity (SSC) and metabolic activity (Autofluorescence) has been exploited to capture the microbiome dynamics ranging from active to subdominant communities. This demonstrates the utility of flow cytometry as a reliable, high throughput tool to resolve dynamic changes in the stable vs perturbed colony structure over time.

3.2 Flow cytometry is comparable to traditional culture-independent techniques.

3.2.1 Inter sample variability

Fig. 2(A) shows the inter sample variability for ARISA data and Fig. 2(B) shows the inter sample variability in the flow cytometry data set. All the between sample variabilities were

clearly larger than the within sample variabilities as per two-sample Kolmogorov-Smirnov Test, indicating that there is a significant difference between the two distributions. This test is crucial to determine if the distance between replicates is smaller than the average distance between treatments.

3.2.3 Multidimensional Spacing Analysis

Fig. 3 A & B shows the multidimensional scaling plots from the distance matrices obtained separately from flow cytometry and ARISA data respectively. All the peak value samples cluster separately from the 5 d samples in the ARISA plots with the exception of GLUC_1. However in the flow cytometry plot, NONE_5, NONE_O, ACET_5 and SLUD_5 clustered separately. The ARISA plot indicated a similarity between ACET_5 and NONE_O which is expected because acetate treatment was administered to enrich methanogens since they consume acetic acid to generate methane gas (Berg et al., 1976). Interestingly, as illustrated in Fig. S3, the putative methanogens i.e. samples fed with acetate, sludge and blank samples reached the peak biogas production on day 2. Thus cytometric fingerprints of putative methanogens obtained by flow cytometry in real time **can be** considered as a promising high- throughput tool for routine on-site monitoring of anaerobic digesters (Kinet et al., 2016).

3.2.4 Correlation between Flow Cytometry and ARISA Data

In order to obtain the correlation between ARISA and flow cytometry data, the flow cytometry data was reduced to include only those points present in the ARISA data set. A linear model was constructed **as shown in Fig. 4 (A)** to see if the distance of any sample from the NONE_0 sample in the ARISA space be predicted from the distance of that sample from the NONE_0 sample in

the flow cytometry space. From the Fig. 4 (B), it is apparent that the SLUD_2, GLUC_5 and PROP_5 are outliers.

Flow cytometry looks more at the morphology of the cells (Müller and Nebe-von-Caron, 2010) while ARISA is more of the genetic technique exploiting the highly conserved Internal transcribed spacer (ITS) region (Fisher and Triplett, 1999). The similarities in the clustering of the samples illustrate that one can be used in place of the other to study the complex microbial population dynamics. These results are in agreement with the recently published study (Kinet et al., 2016) where even though there was a lack of significant correlation at genus or species level, same phenotypical profiles of microbiota during assays matched to several 16S rRNA gene sequencing ones. The exception of the SLUD_2, GLUC_5 and PROP_5 can be attributed to the fact that the ecological dynamics of these syntrophic and resilient populations (Werner et al., 2011) may not be evident in the method we are proposing here. Further analysis of ARISA data might resolve this divergence since a single organism may contribute more than one peak to the ARISA community profile, or alternatively, unrelated organisms can have similar length spacers, which leads to underestimation of community diversity (Fisher and Triplett, 1999). The precise sizing information from ARISA profiles could be used to identify the major bands of interest in a separate manual polyacrylamide gel which could then be further characterized through band excision and subsequent sequencing as illustrated by Fisher and Triplett (1999).

3.4 Flow cytometry detects changes in the microbial community at different HRTs

Fig. 5 shows multidimensional spacing (MDS) plot-demonstrating clusters based on the RMSD3D tool (described in Statistical Analysis section above) for the samples obtained UCSD's

DIG 1 and DIG 2 at different time points. All the DIG 2 samples cluster separately from DIG 1 samples with the exception of 3 d sample. In agreement to the previously published work, the difference in the clustering can be attributed to the fact that varying HRTs can impact microbial abundance and composition in anaerobic digesters (Sasaki et al., 2011; Zhang et al., 2006). As shown in the Fig. S8(A), for each of the samples, DIG 2 shows a similar fraction of live cells as DIG 1 and the ratio of live cells between different digesters is constant at $1.04 \pm (1 \text{ Std. Dev.})$. Furthermore, as shown in the Fig. S8 (B), Alexa Fluor 488 mean frequency for DIG1 is always greater than that in DIG 2 (p-value < 0.001 for the paired two-sample test). In order to confirm the validity of this experiment, the control monoculture experiment was performed with *E. coli* as described in the supplementary material. Various molecular markers like DAPI, SYBRgreen, SYTO etc. have been used previously for discriminating microbial cells on the basis of their DNA content (Müller et al., 2012; Koch et al., 2014a; Koch et al., 2014b; Kinet et al., 2016). This study shows that even in the absence of labels, it is possible to discriminate between samples on the basis of from the most commonly observed fluorescing molecules like NADPH and flavins (See supplementary material for more information), along with autofluorescence coming from intracellular proteins (Monici, 2005). Flow cytometry can provide this information in a high throughput manner since it is fast, requires small sample volumes and minimal sample preparation. All the steps can be performed in less than an hour and hence has a potential to be used for controlling biogas reactor based on HRT. The ability of flow cytometry to rapidly detect changes in microbiome composition and dynamics caused by varying HRTs in an operational, industrial scale anaerobic digester is a major step towards its application in monitoring and potentially controlling biodigesters in the future.

4. Conclusion

The flow cytometry based method is sensitive enough to changes within 5 d of perturbation to the complex microbial community as well as community dynamics in operational 1.25 MG municipal anaerobic digester. While ARISA measures diversity at the genomic level and flow cytometry measures diversity at the morphological level, there is an observed correspondence between the two measures at the phylum-level. In addition to the reliable, rapid, high-throughput, routine characterization tool for bioprocess monitoring, the RMSD based tool has a potential to benefit research community in various fields of life sciences and microbial ecology.

Acknowledgements

The authors would like to thank the Roy J. Carver Biotechnology Center (CBC)'s Flow Cytometry Facility at the University of Illinois, Urbana-Champaign for their services, equipment and resources to carry out this work. The authors are also grateful to Dr. Barbara Pilas, Dr. Angela Kent, Dr. Michael J Miller and Dr. Chinmay Soman for their valuable insights in carrying out the proposed work. Mr. Xiaofeng Ban contributed in the E. coli control experiments.

The present work is supported by NSF award 0749028 to KDB and independent grant from the Institute of Sustainability, Energy, and Environment to ASD.

References

- Amani, T., Nosrati, M., Sreekrishnan, T.R., 2010. Anaerobic digestion from the viewpoint of microbiological, chemical, and operational aspects — a review. *Environ. Rev.* 18, 255–278. doi:10.1139/A10-011
- Angenent, L.T., Karim, K., Al-Dahhan, M.H., Wrenn, B.A., Domínguez-Espinosa, R., 2004. Production of bioenergy and biochemicals from industrial and agricultural wastewater. *Trends Biotechnol.* 22, 477–485. doi:10.1016/j.tibtech.2004.07.001
- APHA. 1998. Standard methods for the examination of water and wastewater.**
- Berg, L. van den, Patel, G.B., Clark, D.S., Lentz, C.P., 1976. Factors affecting rate of methane formation from acetic acid by enriched methanogenic cultures. *Can. J. Microbiol.* 22, 1312–1319. doi:10.1139/m76-194
- Cardinale, M., Brusetti, L., Quatrini, P., Borin, S., Puglia, A.M., Rizzi, A., Zanardini, E., Sorlini, C., Corselli, C., Daffonchio, D., 2004. Comparison of different primer sets for use in automated ribosomal intergenic spacer analysis of complex bacterial communities. *Appl. Environ. Microbiol.* 70, 6147–56. doi:10.1128/AEM.70.10.6147-6156.2004
- Cohen, A., Breure, A.M., van Andel, J.G., van Deursen, A., 1982. Influence of phase separation on the anaerobic digestion of glucose—II. *Water Res.* 16, 449–455. doi:10.1016/0043-1354(82)90170-1
- Dhoble, A.S., Pullammanappallil, P.C., 2014. Design and operation of an anaerobic digester for waste management and fuel generation during long term lunar mission. *Adv. Sp. Res.* 54, 1502–1512. doi:10.1016/j.asr.2014.06.029
- Fisher, M.M., Triplett, E.W., 1999. Automated Approach for Ribosomal Intergenic Spacer

Analysis of Microbial Diversity and Its Application to Freshwater Bacterial Communities.
 Appl. Envir. Microbiol. 65, 4630–4636.

Kinet, R., Dzaomuh, P., Baert, J., Taminiau, B., Daube, G., Nezer, C., Brostaux, Y., Nguyen, F., Dumont, G., Thonart, P., Delvigne, F., 2016. Flow cytometry community fingerprinting and amplicon sequencing for the assessment of landfill leachate cellulolytic bioaugmentation. *Bioresour. Technol.* 214, 450–459. doi:10.1016/j.biortech.2016.04.131

Koch, C., Harms, H., Müller, S., 2014a. Dynamics in the microbial cytome-single cell analytics in natural systems. *Curr. Opin. Biotechnol.* 27, 134–41. doi:10.1016/j.copbio.2014.01.011

Koch, C., Müller, S., Harms, H., Harnisch, F., 2014b. Microbiomes in bioenergy production: From analysis to management. *Curr. Opin. Biotechnol.* 27, 65–72. doi:10.1016/j.copbio.2013.11.006

Leitão, R.C., van Haandel, A.C., Zeeman, G., Lettinga, G., 2006. The effects of operational and environmental variations on anaerobic wastewater treatment systems: a review. *Bioresour. Technol.* 97, 1105–18. doi:10.1016/j.biortech.2004.12.007

McCarty, P.L., Mosey, F.E., 1991. Modelling of Anaerobic Digestion Processes (A Discussion of Concepts). *Water Sci. Technol.* 24, 17–33.

Monici M., 2005. Cell and tissue autofluorescence research and diagnostic applications. *Biotechnol Annu. Rev.* 11, 227–56. doi:10.1016/S1387-2656(05)11007-2.

Müller, S., Hübschmann, T., Kleinstuber, S., Vogt, C., 2012. High resolution single cell analytics to follow microbial community dynamics in anaerobic ecosystems. *Methods* 57, 338–349. doi:http://dx.doi.org/10.1016/j.ymeth.2012.04.001

Müller, S., Nebe-von-Caron, G., 2010. Functional single-cell analyses: flow cytometry and cell

1 sorting of microbial populations and communities. *FEMS Microbiol. Rev.* 34, 554–87.
 2 doi:10.1111/j.1574-6976.2010.00214.x
 3 Owen, W.F., Stuckey, D.C., Healy, J.B., Young, L.Y., McCarty, P.L., 1979. Bioassay for
 4 monitoring biochemical methane potential and anaerobic toxicity. *Water Res.* 13, 485–492.
 5 doi:10.1016/0043-1354(79)90043-5
 6 Pandit, P.D., Gulhane, M.K., Khardenavis, A.A., Vaidya, A.N., 2015. Technological Advances
 7 for Treating Municipal Waste, in: *Microbial Factories*. Springer India, New Delhi, pp. 217–
 8 229. doi:10.1007/978-81-322-2598-0_13
 9 Park, H.S., Schumacher, R., Kilbane 2nd, J.J., 2005. New method to characterize microbial
 10 diversity using flow cytometry. *J Ind Microbiol Biotechnol* 32, 94–102.
 11 doi:10.1007/s10295-005-0208-3
 12 Qu, X., Mazéas, L., Vavilin, V.A., Epissard, J., Lemunier, M., Mouchel, J.-M., He, P., Bouchez,
 13 T., 2009. Combined monitoring of changes in delta13CH4 and archaeal community
 14 structure during mesophilic methanization of municipal solid waste. *FEMS Microbiol. Ecol.*
 15 68, 236–45. doi:10.1111/j.1574-6941.2009.00661.x
 16 Ranjard, L., Poly, F., Lata, J.-C., Mougél, C., Thioulouse, J., Nazaret, S., 2001. Characterization
 17 of Bacterial and Fungal Soil Communities by Automated Ribosomal Intergenic Spacer
 18 Analysis Fingerprints: Biological and Methodological Variability. *Appl. Environ.*
 19 *Microbiol.* 67, 4479–4487. doi:10.1128/AEM.67.10.4479-4487.2001
 20 Raskin, L., Stromley, J.M., Rittmann, B.E., Stahl, D.A., 1994. Group-specific 16S rRNA
 21 hybridization probes to describe natural communities of methanogens. *Appl. Envir.*
 22 *Microbiol.* 60, 1232–1240.

- 1 Rezaeinejad, S., Ivanov, V., 2013. Assessment of correlation between physiological states of
2 *Escherichia coli* cells and their susceptibility to chlorine using flow cytometry. *Water Sci.*
3 *Technol. Water Supply* 13, 1056. doi:10.2166/ws.2013.083
- 4 Sasaki, D., Hori, T., Haruta, S., Ueno, Y., Ishii, M., Igarashi, Y., 2011. Methanogenic pathway
5 and community structure in a thermophilic anaerobic digestion process of organic solid
6 waste. *J. Biosci. Bioeng.* 111, 41–6. doi:10.1016/j.jbiosc.2010.08.011
- 7 Werner, J.J., Knights, D., Garcia, M.L., Scalfone, N.B., Smith, S., Yarasheski, K., Cummings,
8 T.A., Beers, A.R., Knight, R., Angenent, L.T., 2011. Bacterial community structures are
9 unique and resilient in full-scale bioenergy systems. *Proc. Natl. Acad. Sci. U. S. A.* 108,
10 4158–63. doi:10.1073/pnas.1015676108
- 11 Yuan, C., Lei, J., Cole, J., Sun, Y., 2015. Reconstructing 16S rRNA genes in metagenomic data.
12 *Bioinformatics* 31, i35–43. doi:10.1093/bioinformatics/btv231
- 13 Zhang, Z.-P., Show, K.-Y., Tay, J.-H., Liang, D.T., Lee, D.-J., Jiang, W.-J., 2006. Effect of
14 hydraulic retention time on biohydrogen production and anaerobic microbial community.
15 *Process Biochem.* 41, 2118–2123. doi:10.1016/j.procbio.2006.05.021

Figure legends

Fig. 1 (A) Autocorrelation analysis on flow cytometry data with six equally spaced time periods.

(B) Autocorrelation analysis on flow cytometry data with day-to-day changes.

Fig. 2 (A) Inter sample variability in the ARISA data set ($D^+ = 0.4246$, p-value = $1.894e-07$). (B)

Inter sample variability in the Flow Cytometry data set (For RMSD 3D Voxel: $D^+ = 0.6208$, p-value = $2.2e-16$).

Fig. 3 Multidimensional spacing (MDS) plots for (A) flow cytometry data (B) ARISA data.

Fig. 4 (A) Flow Cytometry RMSD3D vs ARISA distance model (B) Normal Q-Q Plot

Fig. 5 MDS plot showing clustering of samples obtained from Digester 1 (DIG 1) vs Digester 2 (DIG 2) at different time points.

Fig. 1A

Fig. 1B

Fig. 2 (A)

Fig. 2 (B)

Fig. 3 (A)

Fig. 3 (B)

Fig. 4 (A)

Fig. 4 (B)

Fig. 5

Table 1 Primers used for ARISA analysis.

Primer	Sequence (5'-3')	Reference(s)
<i><u>Bacterial</u></i>		
1406F	FAM-TGYACACACCGCCCGT	(Cardinale et al., 2004)
23Sr	GGGTTBCCCCATTCRG	(Cardinale et al., 2004)
S-D-Bact-1522	FAM-TGCGGCTGGATCCCCTCCTT	(Cardinale et al., 2004)
L-D-Bact-132-a-A-18	CCGGGTTTCCCCATTCGG	(Cardinale et al., 2004)
ITSF	FAM-GTCGTAACAAGGTAGCCGTA	(Cardinale et al., 2004)
ITSReub	GCCAAGGCATCCACC	(Cardinale et al., 2004)
<i><u>Fungal</u></i>		
2234C	GTTTCCGTAGGTGAACCTGC	(Ranjard et al., 2001)
3126T-56	FAM-ATATGCTTAAGTTCAGCGGGT	(Ranjard et al., 2001)
<i><u>Archaeal</u></i>		
1389F	FAM-ACGGGCGGTGTGTGCAAG	(Qu et al., 2009)
71R	TCGGYGCCGAGCCGAGCCATCC	(Qu et al., 2009)