Improvement of Biogas Production from Animal Manure via Bacteria Enrichment using CASCADE Technology

C. C. Zhou[* **], Y. F. Zhi[***], D. K. Li[***], X. L. Zheng[***], S. Wei[*] ,J . Her[*]

[*]Cascade Clean Energy, Inc. Cupertino, CA, USA, charles.zhou@ccleanenergy.com
[**] Nanjing University of Technology, Nanjing, China.
[***]Cascade Clean Energy Co. Ltd of Nanjing, Nanjing, China, yfzhi2003@126.com

ABSTRACT

A machine learning system – Computer Assisted Strain Construction and Development Engineering, or CASCADE for short, was applied to identifying microorganisms and finding correlations based on predictive patterns between the organisms' genetic information and their metabolic behavior for selecting baterial strains of more effective phenotypes. Further demonstrated was a bioaugmentation method showing improvement of anaerobic digestion of cow manure for biogas production by adding bacteria selected usiing the CASCADE algorithm. The experiments showed that when compared to the traditional anaerobic digestion, biogas production of the groups with addition of selected bacterial strains increased by 50.3% and 36.2% respectively while the methane contents of biogas were nearly equal to the control. DGGE analysis showed addition of selected bacteria caused the difference of dominant bateria during the digestion process which was useful to the anaerobic digestion. The results showed that adding bacteria selected by CASCADE in the anaerobic digestion process is an effective way to increase biogas production from cow manure.

Keywords: CASCADE, microorganism, animal manure, anaerobic digestion, biogas

1 INTRODUCTION OF CASCADE

We have applied a machine learning system – Computer Assisted Strain Construction and Development Engineering, or CASCADE for short, to a collection of microorganisms and found correlations and predictive patterns between the organisms' genetic information and their metabolic behavior. Five types of genetic information from a total of 366 organisms were analysied: 1) Gene similarity scores along the metabolic pathways (Gene Similarity), 2) Genes in the metabolic pathways (Gene Pathway), 3) Percentage of gene usages in 23 general function categories (Gene Generic Functions), 4) 64 Codon usages (Codon Usage), and 5) Unique number of genes in 137 metabolic function categories (Gene Metabolic Functions). We compiled metabolic pathways from KEGG and BioCyc databases and investigated Metabolic Networks Reconstructions (MNR). Sketch map of the CASCADE core is shown in Figure 1.
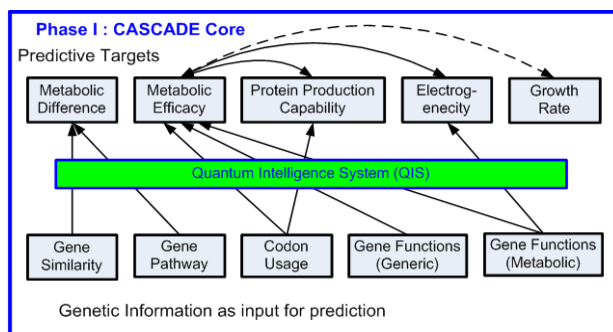

Figure 1: Sketch map of the CASCADE core

We defined a measure "average metabolic efficiency" (AME) which computes, on average, how many times a gene appears in a metabolic pathway. We found that the AME measure is highly correlated with the metabolic capabilities in real life. A higher AME indicates a higher maintenance required for an organism; therefore it might result in a higher cost to express proteins. A lower AME might correlate with a higher yield for expressing a gene product.

We successfully built five predictive models using the five genetic information sources listed above initially as input to learn across multiple organisms and then extract the correlation patterns from these genetic make-ups and the desirable metabolic behavior and properties. Compared to the traditional biological approach, this approach is able to learn the behavior from relatively genetically distant organisms, not just from their direct mutants. This might lead to a better understanding of the actual genetic mechanism underlying the metabolic process.

We first built two predictive models of AME using the gene similarities and gene pathways data of the first two data sets. As a proof-of-concept, the models successfully verified the clusters of bacteria and their scientific classifications. A couple of genes known to have significant contribution to the bacteria metabolism are validated using this model. We then built a predictive model of AME using percentage of gene usages in 23 general function categories. This leads to the discovery of the clusters of organisms that might have better metabolic behavior than commonly used E coli. We also built a predictive model of AME using the codon usages information, which leads to the discovery of

the specific codon usages that are sensitive to certain protein expression.

Our cluster-based predictive model is important and significant in that it is able to handle high number of input dimensions (e.g. ~600 genes) and low number of samples (~150 proteobacteria), which posts a difficult pattern recognition problem in many cases. Also, the system is helpful to biologists to go beyond E. coli, which is considered a golden standard,, when looking for other organisms with more efficient production capability by learning a desired behavior and its contributing factors across a collection of organisms. The correlations and predictive patterns are validated through statistical ANOVA tests and shows that our predictive patterns are indeed statistically significant.

For a given organism, data and text that describes a biological system as a whole can be collected along with information on gene similarity among organisms, gene functions, metabolic functions, biological pathways and pathway substrates/products involving special function (e.g., for energy-generation). This is done for the organisms in various public and private databases. We then apply a network of Knowledge Pattern Search to group the organism population into characteristic groups based on the profiles compiled previously. One or a group of microorganisms are selected based on a profile match score[1~5].

By applying CASCADE, our product is able to link massive genetic and chemical fingerprints in the metabolic and biological pathways to assess an organism's metabolic capability. This makes it possible to customize and find efficient microbes (or even to discover novel microorganisms) and to select the optimal bacteria consortia for maximizing the yields of the desired products.

2 APPLICATIONS OF CASCADE TO BIOGAS PRODUCTION

Biogas production from anaerobic digestion is a traditional method to gain clean energy. Anaerobic digestion of animal manure has been extensively researched and demonstrated. However, this has been hampered by lower gas production due to low biodegradability and unsteady bioprocess during manure digestion. Many methods, for example, pre-treating dairy manure by chemical or physical ways, are used to increase biogas production without much success[6].

Applying the CASCADE method has made it possible to customize and find efficient microbes for higher gas production during anaerobic digestion. For example, we had compiled number of substrates consumed and products produced in the reactions involved in a fermentation process using a biowaste input as the feeding substrates. By applying the CASCADE selection process with combined information of substrates (In), products (Out), metabolic pathways (Pathway) and interested properties (Target) such as electrogenic results in a list of microorganisms that are

likely taking cellulose or actetate as a substrate present in the various wastewater contents. This provides a fast path to find the microorganisms that are able to digest diversified waste contents and recover clean electricity. Main microorganisms involving anaerobic digestion when using CASCADE are shown in Figure 2.
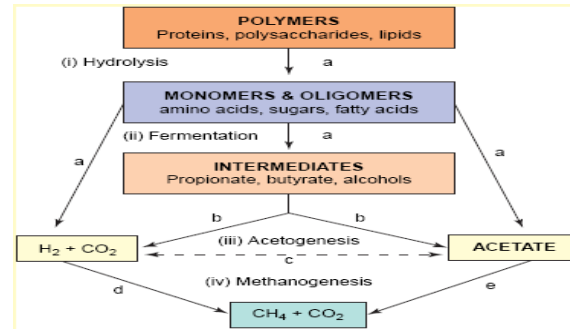
a: Fermentative Bacteria (Acidogen); b: Obligatory Hydrogen-Producing Acetogenic Bacteria; c: Acetogens and Homoacetogens; d: Hydrogenotrophic Methanogens;e: Acetoclastic Methanogens

Figure 2: Main microorganisms involving anaerobic digestion when using CASCADE[7]

A bioaugmentation method by adding bacteria selected by CASCADE to animal manure for biogas production is discussed in this paper. As a demonstration, we use biological data that link the microorganisms to their capabilities of metabolism, e.g. digesting or producing substances of interest, and to perform an initial screening and selection of organisms with these desirable capabilities. Bacteria selected by CASCADE aimed to anaerobic digestion of cow manure are then added to a continuous digestion process. Biogas production and methane contents were measured and related DGGE analysis was studied. The results showed that adding bacteria selected by CASCADE to the anaerobic process is an effective way to increase biogas production from cow manure.

2.1 Materials and Methods

We compiled a list of possilble animal manure contents from various documents reported by animal manure treatment authority or researchers, and used them as the CASCADE input for selecting organisms. We scientifically selected some bacteria, which are distributed in all four stages of Anaerobic Digestion processes, compared with naturally grown bacteria in a series of anaerobic digestion experiments. These serve as baselines to compare with the CASCADE selected microorganisms. Two groups named MAG1 and MAG2 of different CASCADE organisms combination were used in the experiments.

The fresh cow manure was collected form Mengniu Austasia International Dairy Farm in China and this power plant can process 500 tons of dairy manure and produce

12,000 cubic meters of biogas every day by tradtional anaerobic digestion.

The anaerobic digestion process was conducted at 38.5±0.5°C in three manure-containing flasks which were tightly sealed and connected to air bags. We used naturally grown and two groups named MAG1 and MAG2 scientifically selected bacteria to test 600 mL of cow dung. A continuous digestion process was setup. The experiment was first performed in batch fermentation for two days and continuous fermentation began from Day 3 to Day 21 with 25mL manure in/out everyday, and batch fermentation again until the gas production stopped at Day 39.

The gas volume of the gas bags were measured at different time points. The gas contents were analyzed at selected time points. At Day 10, the samples were obtained from each flask and related strains were collected. And DNA extraction and amplification of 16S rRNA gene V3 zone were conducted, then DGGE(denaturing gradient gel electrophoresis) profiles of PCR products were gained to analysis micrroganisms diversity[8].

2.2   Results and Discussions

Biogas production of different group:

Figure 3 shows the biogas production of three different groups. We can see that all of biogas productions increases with time and those of added-baterial groups all exhibit higher production than one of the control with the natural traditional microorganisms. 5570 mL biogas was produced in the control group and 7840 mL biogas for MAG1 group which is 50.3% higher than the control, and 6830 mL biogas for MAG2 group which is 36.2% higher than the control. So from biogas prodution, added-bacterial groups is more effective than the original microorganisms, especially MAG1 group.
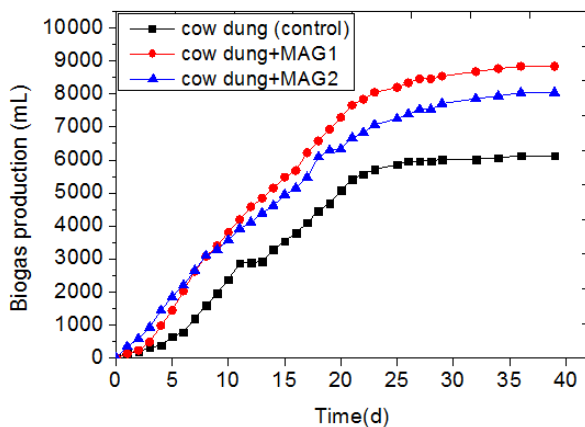

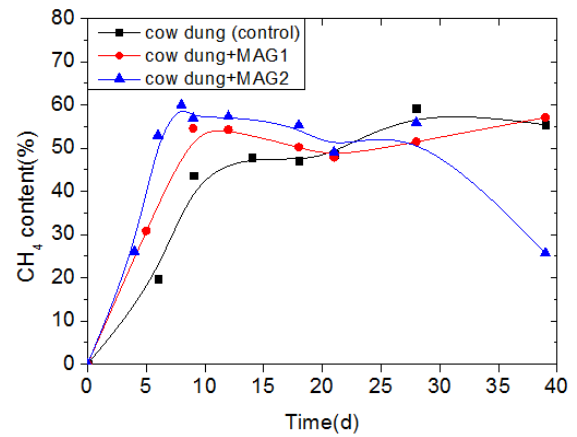Figure 3: Biogas production of different groups

Figure 4: CH₄ contents of different groups

CH₄ contents of different group:

CH₄ contents of the biogas were shown in Figure 4. We can see that at nearly Day 8 all of biogas had almost steady CH₄ contents and after that mean CH₄ contents of the control, MAG1 and MAG2 groups showed respectally 45%-50%, 50%-55% and 55%-60% which means reduction of CH4 contents were not occurred with improvement of biogas production. And with the progress of digestion, contents of the two groups with added bacteria were keeping higher CH₄ contents in most of time which indicated that added bacteria is useful to improve the ability of natural anaerobic digestion microorganisms.

It also can be observed that bacterial MAG1 is more suitable for cow dung fermentation, for it can make the highest gas production and higher CH₄ content.

DGGE analysis:

Figure 5 showed the DGGE profiles of the control group(CA-1) and the group added MAG1(CA-2). We can see that they presented different bands in which CA-1 contained 7 bands from a-g and CA-2 contained 6 bands from h-m. Band a,b, d,e and Band h,i,j are more clear in respective sequence. As shown in Table 1, sequence identification obtained for bands showed that there were different dominant bateria. *Pseudomonas putida*, *Pseudomonas*, and *Rikenellaceae* contained in CA-1 sample(the control) were not the most priority microorganisms contained in CA-2 sample(the group added MAG1) in which *Bacteroidales* has the obvious dominant role. Besides that we noted that *Tissierella creatinophila* of Band m in CA-2 sample were not nearly occurred in CA-1 sample. The results showed that addition of selected bacteria by CASCADE caused the difference of dominant microoragnisms during the digestion process. Meanwhile biogas production and contents results indicated that the difference between natural bateria and selected bacteria was useful to the anaerobic digestion process. So added bacteria selected by CASCADE is an effective way to increase biogas production from cow manure.
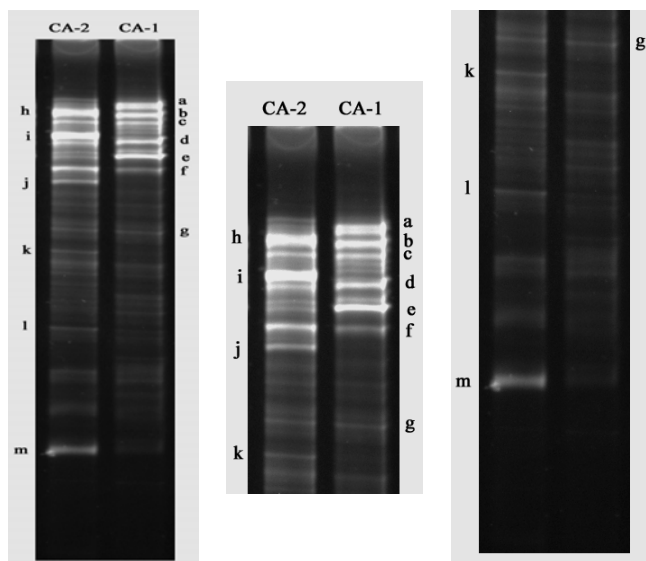
(a)Full DGGE profile   (b)Zoomed top of the DGGE profile   (c)Zoomed bottom of the DGGE profile

Figure 5: DGGE profiles of the control group(CA-1) and the group added MAG1(CA-2)

Sample	Band	V3 Zone	Microorganism
CA-1	a	150bp	*Pseudomonas putida*
	b	148bp	*Pseudomonas*
	c	147bp	*Bacteroidales*
	d	163bp	*Bacteroidales*
	e	155bp	*Rikenellaceae*
	f	133bp	*Alkalibacter*
	g	144bp	*Bacteroidales*
CA-2	h	147bp	*Bacteroidales*
	i	154bp	*Bacteroidales*
	j	147bp	*Bacteroidales*
	k	121bp	*Anaerotruncus*
	l	128bp	*Caldilinea*
	m	125bp	*Tissierella creatinophila*

Table 1: Sequence identification obtained for bands

3 CONCLUSION

"Computer-Assisted Strain Construction and Development Engineering (CASCADE)" is an *In Silico* screening platform to systematically select organisms. By applying CASCADE, our product is able to link massive genetic and chemical fingerprints in the metabolic and biological pathways to assess an organism's metabolic capability. In this paper, CASCADE was used to select bateria which are useful for producing biogas during anaerobic digestion process of cow manure.

The experiments showed that, compared to the traditional anaerobic digestion, biogas production of added group increased by 50.3% and 36.2% respectively vis different groups and methane contents of biogas were nearly equal to the control experiments. Related DGGE analysis showed that addition of selected bacteria caused the difference of dominant microoragnisms during the digestion process which was useful to the anaerobic digestion process of cow manure. The results showed that added bacteria selected by CASCADE is an effective way to increase biogas production from cow manure.

REFERENCES

[1] Y. Zhao, S. Wei, I. Oglesby, and C. C. Zhou, "Utilizing the Quantum Intelligence System for Drug Discovery (QIS D2) for Anti-HIV and Anti-Cancer Cocktail Detection," Journal of Medical Chemical, Biological, & Radiological Defense, 7, 2009.

[2] D. G. Ahren and C. A. Ouzounis, "Robustness of metabolic map reconstruction", Journal of Bioinform Comput Biol. 2, 589-93, 2004

[3] V. J. Nikiforova, C. O. Daub, H. Hesse, L. Willmitzer, and R. Hoefgen, " Integrative gene-metabolite network with implemented causality deciphers informational fluxes of sulphur stress response", J. Exp. Bot. 56:1887-1896, 2005

[4] D. B. Kell, "Metabolomics and systems biology: making sense of the soup", Curr Opin Microbiol. 7, 296-307, 2004.

[5] A. Li, and C. Chan, "Integrating gene expression and metabolic profiles", J. Biol Chem. 279, 27124-27137, 2004.

[6] Y. Chen, J. J. Cheng, and K. S. Creamer, "Inhibition of anaerobic digestion process: A review," Bioresource Technology, 99, 4044-4064, 2008.

[7] L. T. Angenent, K. Karim, M. H. Al-Dahhan, B. A. Wrenn and R Domiguez-Espinosa, "Production of bioenergy and biochemicals from industrial and agricultural wastewater", Trends in Biotechnology, 22, 477-485, 2004.

[8] K.Rantsiou, G. Comi, and L. Cocloin, "The *rpoB* gene as a target for PCR-DGGE analysis to follow lactic acid baterial population dynamics during food fermentations," Food Microbiology, 21, 481-487, 2004.