<table>
<thead>
<tr>
<th>Title</th>
<th>Towards a metagenomic understanding on enhanced biomethane production from waste activated sludge after pH 10 pretreatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Wong, MT; Zhang, D; Li, J; Hui, RKH; Tun, HM; Brar, MS; Park, TJ; Chen, Y; Leung, FCC</td>
</tr>
<tr>
<td>Citation</td>
<td>Biotechnology for biofuels, 2013, v. 6 n. 1, article no. 38, p. 1-14</td>
</tr>
<tr>
<td>Issued Date</td>
<td>2013</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10722/190371">http://hdl.handle.net/10722/190371</a></td>
</tr>
<tr>
<td>Rights</td>
<td>Biotechnology for Biofuels. Copyright © BioMed Central Ltd.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
Towards a metagenomic understanding on enhanced biomethane production from waste activated sludge after pH 10 pretreatment

Mabel Ting Wong¹, Dong Zhang², Jun Li¹, Raymond Kin Hi Hui¹, Hein Min Tun¹, Manreetpal Singh Brar¹, Tae-Jin Park¹, Yinguang Chen²* and Frederick C Leung¹,3*

Abstract

Background: Understanding the effects of pretreatment on anaerobic digestion of sludge waste from wastewater treatment plants is becoming increasingly important, as impetus moves towards the utilization of sludge for renewable energy production. Although the field of sludge pretreatment has progressed significantly over the past decade, critical questions concerning the underlying microbial interactions remain unanswered. In this study, a metagenomic approach was adopted to investigate the microbial composition and gene content contributing to enhanced biogas production from sludge subjected to a novel pretreatment method (maintaining pH at 10 for 8 days) compared to other documented methods (ultrasonic, thermal and thermal-alkaline).

Results: Our results showed that pretreated sludge attained a maximum methane yield approximately 4-fold higher than that of the blank un-pretreated sludge set-up at day 17. Both the microbial and metabolic consortium shifted extensively towards enhanced biodegradation subsequent to pretreatment, providing insight for the enhanced methane yield. The prevalence of Methanosaeta thermophila and Methanothermobacter thermautotrophicus, together with the functional affiliation of enzymes-encoding genes suggested an acetoclastic and hydrogenotrophic methanogenesis pathway. Additionally, an alternative enzymology in Methanosaeta was observed.

Conclusions: This study is the first to provide a microbiological understanding of improved biogas production subsequent to a novel waste sludge pretreatment method. The knowledge garnered will assist the design of more efficient pretreatment methods for biogas production in the future.

Keywords: Wastewater treatment plant, Sludge, Pretreatment, Renewable energy, Biomethane, Pyrosequencing, Metagenomic
structure [12]. Methane production from WAS is a complex, multi-step process which involves multiple syntrophic interactions within the microbial consortium [13]. Complex compounds (polysaccharides, proteins, nucleic acids, and lipids) are first converted to oligomers and monomers through the action of extracellular hydrolytic enzymes produced by the primary fermenting bacteria [14]. Subsequently, the intermediate products are further transformed into acetate, carbon dioxide, hydrogen and formate by secondary fermenters [14]. The final methanogenesis step is then conducted by methanogenic archaea, whose energy substrates are highly restricted to acetate, H2, CO2, formate or certain C1 compounds [15]. To improve methane yield from sludge, enormous research efforts have been devoted to the development of pretreatment methods to accelerate sludge hydrolysis, including thermal [16], thermal-alkaline [17], ultrasonic [18], mechanical and thermo-chemical methods [19]. Although the field of pretreatment research has progressed significantly in the past decade, many significant questions related to their effects on the underlying microbial interactions remain unanswered.

As the field of pretreatment method research is nearing a threshold, the accomplishments of the past are pushing the door of microbiology to provide new insights [20,21]. In this study, we aimed to investigate the microbial composition and gene content of sludge subjected to our novel pretreatment method (maintaining pH 10 for 8 days) which leads to significantly enhanced methane generation compared to other documented methods (ultrasonic, thermal and thermal-alkaline) [4]. It was reported in our previous study that both sludge hydrolysis and short-chain fatty acids (SCFAs; eg. acetic, butyric and propionic acid) accumulation were significantly enhanced when WAS was anaerobically fermented under the condition of pH 10 for 8 days [22]. This phenomenon was suspected to be due to biotic factors rather than abiotic ones (e.g. alkaline hydrolysis) since much higher SCFAs accumulation and enzyme activities were observed in un-autoclaved sludge compared to autoclaved sludge [22]. Further investigation by our group suggested that the solubilization of the sludge matrix, usually a hydrolysis event by the embedded extracellular enzymes, may contribute to the significant SCFAs improvement [23].

In this study, a shotgun metagenomic approach was chosen to study potential shifts in microbial communities and/or gene contents that could help explain elevated productions of methane under our novel pretreatment method [24,25]. The latest advances in pyrosequencing technology afford new opportunities to undertake such metagenomic studies to explore the dynamics of microbial communities in time, space or under fluctuating environmental conditions with un-precedented levels of microbial diversity coverage and depth [26-28]. In addition to elucidating the microbiology underpinning the sludge pretreatment process [29], our study sought to improve knowledge of the diversity and physiology of participating syntrophs and methanogens, as well as the mechanism behind the establishment and maintenance of mutualistic cooperation. This knowledge will help to establish a better control over the hydrolysis and methanogenic processes, and promote pretreatment as part of a pertinent strategy for sludge management in WWTP. To the best of our knowledge, this study is the first to adopt a whole genome shotgun approach to link the knowledge between microbiology and engineering of sludge pretreatment methods. Additional statistical soundness was conferred to the surmised conclusion, as the comparative analysis was conducted between meta-datasets generated by the same sequencing method, from well-characterized experimental designs which only differ in known parameters by manipulation.

**Results and discussion**

**Enhanced methane production by using a newly devised pretreatment strategy**

Collected WAS was subjected to a novel pretreatment method (maintaining pH 10 for 8 days) [4], and its effects on methane production enhancement was characterized in this study. Figure 1 illustrates the experimental workflow of obtaining end samples from the biogas-producing anaerobic digesters. Pretreated sludge for 30 days (P30), un-pretreated sludge for 30 days (UP30), or 40 days (UP40) were used as substrate. Pretreated sludge was significantly more effective in producing biogas, as total gas volume produced on day 8 (604.32 ml/g VSS-added) is 7.71 and 1.67 times of that in the un-pretreated sludge and/or gene contents that could help explain elevated productions of methane under our novel pretreatment method [24,25]. The latest advances in pyrosequencing technology afford new opportunities to undertake such metagenomic studies to explore the dynamics of microbial communities in time, space or under fluctuating environmental conditions with un-precedented levels of microbial diversity coverage and depth [26-28]. In addition to elucidating the microbiology underpinning the sludge pretreatment method (maintaining pH 10 for 8 days) [4], and its effects on methane production enhancement was characterized in this study. Figure 1 illustrates the experimental workflow of obtaining end samples from the biogas-producing anaerobic digesters. Pretreated sludge for 30 days (P30), un-pretreated sludge for 30 days (UP30), or 40 days (UP40) were used as substrate. Pretreated sludge was significantly more effective in producing biogas, as total gas volume produced on day 8 (604.32 ml/g VSS-added) is 7.71 and 1.67 times of that in the un-pretreated sludge and/or gene contents that could help explain elevated productions of methane under our novel pretreatment method [24,25]. The latest advances in pyrosequencing technology afford new opportunities to undertake such metagenomic studies to explore the dynamics of microbial communities in time, space or under fluctuating environmental conditions with un-precedented levels of microbial diversity coverage and depth [26-28]. In addition to elucidating the microbiology underpinning the sludge pretreatment

**Biogas producing microbial community residing in anaerobic digester inferred by metagenome sequencing**

To analyze biogas-producing microbes residing in our studied bioreactors in terms of community structure,
gene content, metabolic capabilities and the role of specific organisms in biogas formation, a metagenomic approach using the 454 pyrosequencing was used. Statistical data summarizing the output sequencing quantity of the three independent runs of P30, UP30 and UP40 is given in Table 1. To detect differentially abundant features between the microbial communities, we used the Meta Genome Rapid Annotation using Subsystem Technology (MG-RAST) analysis pipeline, which has shown to be applicable to various metagenomic data performing taxonomic classification as well as functional annotation [30].

Composition of the biogas-producing microbial community was obtained by taxonomic classification of reads based on the M5NR database on the MG-RAST platform. In all three reactor samples, Bacteria were the dominant superkingdom (Table 1). Domain-based allocation was primarily assigned to Bacteria (82.13% on average) and Archaea (11.86% on average), whilst the assigned reads to Eukaryota and Viruses altogether accounted for less than half a percent (Table 1). The bacterial proportion of the microbial community in P30 was the highest (85.46%) followed by UP30 (81.82%) and UP40 (79.10%). As for the Archaea domain, it ranged from 8.32% in P30, to 12.13% in UP30, and to 15.13% in UP40. The clustering pattern shown in Additional file 1 provides a graphical representation of the overall taxonomic similarity—the microbiomes of the unpretreated sludge bioreactors (UP30 and UP40) displayed high resemblance, in which the mature unpretreated sludge bioreactor (UP40, operated for longer time) is comparatively closer to the significantly diverse pretreated sludge bioreactor. The result provided the first insight into both the temporal dynamic and the enormous impact of pretreatment on microbiology. Downstream analyses focused on the Bacteria and Archaea domains (unless otherwise specified, the percentages below are representative of the identified reads within each domain per independent run).

Within the Bacteria domain, the top abundant phyla were Proteobacteria (UP30, 32.17%; UP40, 30.45%; P30, 42.28%), Bacteroidetes (UP30, 22.91%; UP40, 22.25%; P30, 23.38%), Firmicutes (UP30, 16.54%; UP40, 17.65%; P30, 10.97%), Actinobacteria (UP30, 7.13%; UP40, 6.94%; P30, 6.83%) and Chloroflexi (UP30, 5.00%; UP40, 5.21%; P30, 3.37%), they collectively account for over 0.7 of the bacterial reads (normalized between 0 and 1) for each of the three bioreactors (Figure 3A). Proteobacteria, Bacteroidetes and Firmicutes have been reported to be dominant phyla as well in similar analysis of anaerobic digestion of sludge [9]. Further resolution at the class level revealed that the microbial compositions overlapped between the bioreactor samples. To be more precise, the dominant bacterial lineages presented in this study were related to anaerobic digestion, including Bacteroidia (UP30, 13.50%; UP40, 12.69%; P30, 8.90%), δ-proteobacteria (UP30, 13.08%; UP40, 12.71%; P30, 8.93%), Clostridia (UP30, 11.42%; UP40, 12.21%; P30, 7.31%), γ-proteobacteria (UP30, 8.12%; UP40, 7.90%; P30, 11.18%), Actinobacteria (class) (UP30, 7.37%; UP40, 7.20%; P30, 7.06%), α-proteobacteria (UP30, 5.78%; UP40, 5.28%; P30, 9.33%), β-proteobacteria (UP30, 5.73%; UP40, 4.86%; P30, 13.58%), Bacilli (UP30, 5.16%; UP40, 5.25%; P30, 3.53%), Flavobacteria (UP30, 4.70%; UP40, 4.76%; P30, 6.98%) and Chloroflexi (class) (UP30, 2.65%; UP40, 2.56%; P30, 1.84%) (Figure 3B).

With time, there was moderate temporal variation in the relative abundances of these bacterial members in the anaerobic digester fed with un-pretreated sludge (UP30 and UP40). While the relative proportion of these communities...
remained comparable between UP30 and UP40, a larger extent of differences could be observed when P30 was taken into comparison for certain lineages. The over-represented bacterial members in P30 included α-proteobacteria (9.33%), β-proteobacteria (13.58%) and γ-proteobacteria (11.18%); which are respectively 69.06%, 158.21% and 39.60% more than UP30 and UP40 on average. Overall, these sequences were originated from the WAS substrate, AGS inoculum and the biogas-producing symbiont within the system, hence the differences observed in their relative abundance were attributable to the sole variable of either using pretreated or un-pretreated sludge in the bioreactors. These results suggested that the bacterial communities that underlie the anaerobic digesters were dynamic, and they responded rapidly to the pretreated sludge substrate and change substantially over time.

Figure 2 Biogas production profile of pretreated and un-pretreated sludge bioreactors. Efficiency of biogas production in (A) the pretreated sludge bioreactor is significantly higher than (B) the un-pretreated sludge bioreactor, giving rise to a 7.71-fold difference in the produced volume at day 8. (C) The cumulative methane production was 389.8 mL-CH₄/g VSS-added using pretreated sludge as substrate, 3.78 times of that with un-pretreated sludge (103.2 mL-CH₄/g VSS-added) as substrate. Figure 2 (D-E) display methane and carbon dioxide percentages produced by (D) pretreated and (E) un-pretreated sludge bioreactor, respectively; while methane is the major biogas component in both reactors, the methane content was increased by 19.3% with a decrease in carbon dioxide content by 60.9% after pretreatment.
Based on the phylogenetic affiliation of the metagenomic sequences, it is possible to form hypotheses regarding the metabolic functions of the groups [9]. Proteobacteria are microorganisms involved in the initial steps of degradation, studies have shown that they are the main consumers of propionate, butyrate and acetate [31]. Bacteroidetes are known proteolytic bacteria, responsible for the degradation of protein and subsequent fermentation of amino acids known as volatile fatty acids such as butyrate [31]. The H₂ acids into acetate for acetoclastic methanogens [31,32]. The H₂ fermentation was linked to the intra-divergence of each family. The predominance of Methanosaeta — the top dominant methanogens (13.51% on average — UP30, 15.10%; UP40, 13.32%; P30, 12.10%), while Methanosaeta thermophilic (10.67% on average — UP30, 13.13%; UP40, 10.31%; P30, 8.56%) and Methanothermobacter (10.35% on average — UP30, 10.73%; UP40, 9.41%; P30, 10.91%) contributed to slightly lesser portions. Anaerobic digesters are typical habitats for these three genera of methanogens [43], and the enrichment of Methanosarcina species was in congruence with other evaluation studies on primary sludge and WAS anaerobic digesters [44,45]. The observation of a lesser portion of Methanothermobacter was expected, as the operating condition of the bioreactor (approximate 35°C) was not in favour of this thermophilic Archeon’s proliferation (55-65°C) [43].

Methanosaeta species were observed to be more abundant in samples of bioreactor digesting pretreated sludge (UP30; 13.13%) than that with pretreated sludge (P30; 8.56%), but the population decreases with time (UP40; 8.56%). A similar observation was found in an earlier surveillance of the methanogenic population dynamics in anaerobic digesters, where the hitherto abundant Methanosaeta population decreased rapidly as the acetate concentration increased [12]. As acetoclastic methanogens, both Methanosarcina and Methanosaeta are able to split acetate, oxidize the carboxyl-group to CO₂ and reduce the methyl group to CH₄ [43]. At the same time, while Methanosarcina thrive in environments with high acetate concentration; lower acetate concentrations benefit the dominance of Methanosaeta owing to their high affinity for acetate [12,46-48]. As aforementioned, an enhanced bioproduction of acetic acid was reported in the pretreated sludge [4], the revealed shift in the methanogenic community could therefore be interpreted as an ecological consequence [30,49,50] of the sludge pretreatment.

**Methane production via acetoclastic and hydrogenotrophic pathways**

To identify the methane-producing organisms in the bioreactor samples, taxonomic and functional affiliation of the metagenomic reads were evaluated in parallel (represented as percentage of total identified methanogens). On the species level, it is common to all three bioreactor samples that Methanosarcina (12.36% on average — UP30, 14.48%; UP40, 12.36%; P30, 10.25%) and Methanothermobacter thermotrophicus (9.34% on average — UP30, 9.50%; UP40, 8.59%; P30, 9.94%) are the over-represented methanogens (Table 2). Here, Methanosarcina species were not represented as one of the top dominant methanogens (13.51% on average — UP30, 15.10%; UP40, 13.32%; P30, 12.10%), this observation was linked to the intra-divergence of each family

<table>
<thead>
<tr>
<th>Sample</th>
<th>P30</th>
<th>UP30</th>
<th>UP40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretreatment</td>
<td>pH10</td>
<td>Un-pretreated</td>
<td>Un-pretreated</td>
</tr>
<tr>
<td>Semi-continuous</td>
<td>30 d</td>
<td>30 d</td>
<td>40 d</td>
</tr>
<tr>
<td>Sequences count</td>
<td>151,676</td>
<td>114,683</td>
<td>134,522</td>
</tr>
<tr>
<td>Mean sequence length</td>
<td>392±120 bp</td>
<td>289±99 bp</td>
<td>425±115 bp</td>
</tr>
<tr>
<td>Archaea</td>
<td>8.32%</td>
<td>12.13%</td>
<td>15.13%</td>
</tr>
<tr>
<td>Bacteria</td>
<td>85.46%</td>
<td>81.82%</td>
<td>79.10%</td>
</tr>
<tr>
<td>Eukaryota</td>
<td>0.45%</td>
<td>0.30%</td>
<td>0.48%</td>
</tr>
<tr>
<td>Viruses</td>
<td>0.01%</td>
<td>0.02%</td>
<td>0.02%</td>
</tr>
<tr>
<td>other sequences</td>
<td>0.01%</td>
<td>0.01%</td>
<td>0.00%</td>
</tr>
<tr>
<td>unassigned</td>
<td>5.75%</td>
<td>5.72%</td>
<td>5.28%</td>
</tr>
</tbody>
</table>

**Table 1 Statistics of 454 GS Junior pyrosequence datasets presented in this study**
M. thermophila is an obligate methanogen which consumes acetate only [51], whereas M. thermautotrophicus conserves energy by using H₂ to reduce CO₂ to CH₄ [33,52]. While the pH was adjusted to be within the optimum range (7.0) for the two dominate methanogens after pretreatment (Figure 1), it was interesting to detect the prevalence of the thermophilic M. thermautotrophicus in the bioreactors that operated at mesophilic temperature [53]. At the same time, functional enzyme-encoding genes for the two methanogenesis pathways were identified with reference to KEGG and Metacyc pathway database entries (Figure 4, Additional file 2) [54,55]. Nonetheless, based on these results, it was evinced that the production of methane in the studied bioreactors was performed by methanogenic Archaea via acetoclastic and hydrogenotrophic pathways.

Acetoclastic and hydrogenotrophic pathways are indeed common methanogenesis pathways reported in

Figure 3 Bacterial community in pretreated and un-pretreated sludge bioreactor at day 30 and 40. (A) Proteobacteria, Bacteroidetes, Firmicutes, Actinobacteria and Chloroflexi represent top abundant phyla, all account for over 0.7 of the bacterial reads (normalized between 0 and 1) in the bioreactors. Abundance is displayed by both colour scheme and clustering dendrogram. (B) Bacterial consortiums which represent an anaerobic digestion community were manifested with overlapped dominant lineages in the bioreactors, with over-represented α-, β- and γ-proteobacteria in the sludge bioreactor consequent to the pretreatment (compare inner and outermost rings).
various reports [42,56]; the affirmation of these pathways permitted subsequent in-depth investigation of the encompassed enzymology. The first step in acetoclastic methanogenesis is the formation of acetyl-CoA from acetate [51]. Analysis of the recently completed genome of *Methanosaeta thermophilia* confirmed that the majority of the acetoclastic pathway are similar for *Methanosaeta* and *Methanosarcina*, except the enzymes employed for the first step of catalyzation [51]. It was proposed that acetoclastic methanogenesis in *Methanosaeta* proceeded with a modified version of the pathway compared with *Methanosarcina*, which utilizes the acetate kinase/phosphotransacetylase pathway to convert acetate to acetyl-CoA [57,58]. Taking *M. thermophilia* as an example, its genome does not include a readily identifiable acetate kinase, and it has been postulated that this archeon utilizes an acetate transporter coupled with acetyl-CoA synthetases to convert acetate to acetyl-CoA, and the hydrolysis of pyrophosphate by inorganic pyrophosphatase (PPase) drives this reaction forward [51]. In this study, the analysis of metagenomic datasets using SEED annotation indicated the presence of acetyl-CoA synthetases in *Methanosaeta* bin. While the sequencing depth might account for the absence of genes for the acetate kinase/phosphotransacetylase pathway, the discovery of these two enzymes advocated an alternative pathway in acetotrophic methanogenesis for *Methanosaeta* [41]. Results from a recent study on a terephthalate-degrading bioreactor supported this hypothesis [41].

Functional affiliation related to higher methane production after sludge pretreatment

Degradation of the highly organic polymers represents the first and overall rate-limiting step for the mineralization of organic matter in activated sludge and anaerobic digested sludge treatment systems [59-61]. In our work, the studied pretreatment involved alteration of pH, which is an important parameter affecting both bacterial activity and metabolite pathways. KO-based annotations were used to understand how these phylogenetic trends could be used to predict the metabolic potential of these microbes. Figure 5 shows the subsystems that are related to higher methane production, including metabolism of amino acids, energy, carbohydrates, nucleotides, lipids, cofactors and vitamins, xenobiotics, as well as the fermentation of different substrates. These results revealed a general elevated expression of these faculties in the bioreactors fed with un-pretreated sludge over time (compare UP30 and UP40), whereas these levels were highest in the bioreactor digesting pretreated sludge (P30). Distribution of the functional systems was most divergent in P30, which showed predominance in metabolism consistent with a community shifted towards an enhanced biomass degradation metabolism. Herewith, the downstream analysis focused on the degradation of carbohydrates for a number of reasons. Firstly, the matrix of extra-cellular polymeric substances (EPS) that crosslink cells together remains the primary solid part of biological sludge, of which the polysaccharide is the predominant component [62-64]. Secondly, complex carbohydrate is the commonplace recalcitrant in the hydrolysis process, and bioprospecting

---

**Table 2 Relative abundance of Archaea in pretreated and un-pretreated sludge bioreactors at day 30 and 40**

<table>
<thead>
<tr>
<th>Archaeal taxonomic affiliation</th>
<th>UP30</th>
<th>UP40</th>
<th>P30</th>
</tr>
</thead>
<tbody>
<tr>
<td>(phylum/ family/ genus/ species)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euryarchaeota</td>
<td>97.60%</td>
<td>95.60%</td>
<td>94.86%</td>
</tr>
<tr>
<td>Methanosarcinaceae</td>
<td>23.11%</td>
<td>20.56%</td>
<td>19.14%</td>
</tr>
<tr>
<td>Methanosarcina</td>
<td>15.10%</td>
<td>13.32%</td>
<td>12.10%</td>
</tr>
<tr>
<td>Methanobacteriaceae</td>
<td>20.81%</td>
<td>18.61%</td>
<td>21.44%</td>
</tr>
<tr>
<td>Methanothermobacter</td>
<td>10.73%</td>
<td>9.41%</td>
<td>10.91%</td>
</tr>
<tr>
<td>Methanothermobacter thermotrophicus</td>
<td>9.50%</td>
<td>8.59%</td>
<td>9.94%</td>
</tr>
<tr>
<td>Methanosetae</td>
<td>14.00%</td>
<td>11.03%</td>
<td>9.19%</td>
</tr>
<tr>
<td>Methanosaeta</td>
<td>13.13%</td>
<td>10.31%</td>
<td>8.56%</td>
</tr>
<tr>
<td>Methanosaeta thermophila</td>
<td>14.48%</td>
<td>12.36%</td>
<td>10.25%</td>
</tr>
<tr>
<td>Crenarchaeota</td>
<td>2.33%</td>
<td>3.61%</td>
<td>4.40%</td>
</tr>
<tr>
<td>Korarchaeota</td>
<td>0.00%</td>
<td>0.45%</td>
<td>0.37%</td>
</tr>
<tr>
<td>Thaumarchaeota</td>
<td>0.05%</td>
<td>0.27%</td>
<td>0.37%</td>
</tr>
<tr>
<td>Nanoarchaeota</td>
<td>0.02%</td>
<td>0.07%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Legends: Archaeal domain consists three major families belonging to Methanosarcinaceae, Methanobacteriaceae and Methanosetaeace, while *Methanosarcina*, *Methanoseta* and *Methanothermobacter* represent top abundant methanogens on the genus level. The decrease in abundance of *Methanosarcina*-related species in P30 was correlated with an enhanced bioproduction of acetate in the pretreated sludge. Except for phylum level, only candidates of abundance >8% are presented. Species are represented as percentage of all methanogens.

---

Wong et al. Biotechnology for Biofuels 2013, 6:38 Page 7 of 14
http://www.biotechnologyforbiofuels.com/content/6/1/38
for carbohydrate active enzymes is a topic of interest in metagenomics [65]. As a result, the focus of our study was hence placed on the degradation of this dominant recalcitrant component. (Please see Additional file 3 for further insights towards protein and lipid degradation.)

As sequencing technology provides access to a remarkable array of microbial functional capacity, sequence-based data mining is an important prospect in metagenomic projects [66,67]. To understand how the microbial community mediates the solubilization of the sludge matrix, a carbohydrate-active enzyme (CAZy) characterization of the metagenomic datasets was performed after ab initio gene prediction [68,69]. Conventional sequence homology-based enzyme discovery introduces a bias towards the identification of candidates similar to known enzymes, rather than enzymes with low sequence identity and potentially divergent biochemical properties [66]. The entries in the CAZy database contains both experimentally verified and putative carbohydrate-active enzyme domains, hence this search strategy would be able to provide a better insight into the catalysis of biochemical reactions [68,70].

In this study, a total of 1917 and 107 gene modules were recognized across 52 glycosyl hydrolases (GH) and 9 carbohydrate binding modules (CBM) respectively (Tables 3 and 4). GH is defined as a widespread group of enzymes which hydrolyzes the glycosidic bond between carbohydrates or between that and a non-carbohydrate moiety [68]. On the other hand, CBM are contiguous amino acids within a carbohydrate-active enzyme with a discreet fold which bears the carbohydrate-binding activity [68]. In contrast to the small number of enzymes devoted to the hydrolysis of the main chain of cellulose, hemicelluloses and pectins (GH 5, 6, 7, 8, 9, 12, 44, 45, 48 and 74), all three metagenomes displayed a variety of enzymes that digest the side chains of these polymers and oligosaccharides. Families of GH 2 and 3, which contain a large range of glycosidases cleaving non-reducing carbohydrates in oligosaccharides and the side chains of hemicelluloses and pectins, were particularly abundant in the bioreactor samples (29.10% of all GH—UP30, 27.56%; UP40, 29.55%; P30, 24.20%). CBM 48 was the most dominant in the three metagenomes, it binds...
glycogen and is commonly found in a range of enzymes that act on branched substrates, such as the hydrolysis of glycogen, amylpectin and dextrin by isoamylase, pulluanae and branching enzyme. Overall, the number of GH hit returns appeared to increase with time in the bioreactor fed with un-pretreated sludge (468 counts in UP30; 792 counts in UP40); while the bioreactor digesting the pretreated sludge harboured 1.40 times the GH count in the control counterpart at the same time. As for CBM, P30 (51 counts) had a higher abundance than UP40 (38 counts), and made up to approximately 2.83 times that of UP30. Generally, there appeared to be a dynamic process in enzymology within the anaerobic digesters with time and substrate nature. The heightened GH and CBM abundance in the bioreactor with pretreated sludge may possibly correlate with the higher sludge solubilization proposed in previous studies [23].

Conclusions

The treatment and disposal of excess sludge accumulation represents a bottleneck in wastewater treatment plants worldwide due to environmental, economic, social and legal factors. Reduction of excess sludge is becoming one of the biggest challenges in biological wastewater treatment [71,72]. The anaerobic digestion process is a promising technique, along with which waste is eliminated and methane is produced as a valuable renewable energy source. Understanding the impact of pretreatment on the microbiology at different stages of anaerobic digestion is important; it ultimately impacts the performance of a bioreactor. This study represents the first attempt to gain an in-depth metagenomic perspective, with regard to taxonomic and functional aspects, on increased biogas (methane) production when a newly devised pretreatment method is used. Dual taxonomic and functional analysis indicated the microbial and metabolic consortium shifted extensively towards enhanced biodegradation, also seen in the methanogenesis pathways. Altogether, these results presented here help further microbiological understanding in sludge pretreatment research for anaerobic digestion. As microbiology is the ultimate driver for the anaerobic digestion process, this new perspective encourages a closer engagement between the engineering and microbiology knowledge pools. Further studies involving deeper sequencing of the metagenomes to characterize the decomposing and methanogenic microbiome at higher frequency and replicate number parallel to the physical-chemical characterization are required for each WAS pretreatment intervention.

Methods

Collection of waste activated sludge and anaerobic granular sludge

The waste activated sludge (WAS) used as substrate for methane production was collected from the secondary sedimentation tank of a municipal wastewater treatment plant in Shanghai, China. The anaerobic granular sludge (AGS) used as the inoculum for methane production was obtained from the upward-flow anaerobic sludge blanket (USAB) reactor of a food wastewater treatment...
plant in Yixing, China. The main features of WAS and AGS can be found in Additional file 4. The AGS was cultured in a laboratory up-flow anaerobic sludge bed (UASB) by synthetic wastewater (Additional file 5) prior to its use as the inoculum for methane production, where the use of synthetic wastewater for AGS culturing is a common practice in anaerobic digestion studies investigating the microbiology and biodegradability [4,73,74]. The hydrolytic retention time of UASB was 6 h, and the AGS concentration in UASB was approximately 29165 mg/L.

Pretreatment of waste activated sludge

Biogas-producing, anaerobic digester samples were obtained from a reproduction of the experiment described in our previous work [4], which introduced a new sludge pretreatment strategy (pH 10 for 8 d) for enhanced methane production from WAS. During the pretreatment, the pH of the collected WAS was adjusted to 10.0 by addition of 5 M NaOH or 4 M HCl with an automatic titrator, in a tightly-sealed reactor which was mechanically stirred (80 rpm) at 35 ± 1°C for 8 d. A blank test was also conducted, in which the sludge was not pretreated but mechanically stirred for 8 d. Then the pH in all reactors was adjusted to pH 7.0, and 30 mL of AGS was added to each reactor for methane production. All reactors were sealed with rubber stoppers and mechanically stirred at 80 rpm without further pH control. The retention time of all reactors was 17 d and the organic loading rate was 1.88 kg/COD/m³/d. The gas generation was detected by water displacement and the methane concentration was measured by a gas

### Table 3 Identification of glycoside hydrolase (GH) family classified by CAZy database

<table>
<thead>
<tr>
<th>CAZy GH family</th>
<th>Pfam model</th>
<th>UP30</th>
<th>UP40</th>
<th>P30</th>
<th>CAZy GH family</th>
<th>Pfam model</th>
<th>UP30</th>
<th>UP40</th>
<th>P30</th>
</tr>
</thead>
<tbody>
<tr>
<td>GH 1</td>
<td>PF00232.12</td>
<td>11</td>
<td>14</td>
<td>15</td>
<td>GH 38C</td>
<td>PF07748.7</td>
<td>20</td>
<td>42</td>
<td>15</td>
</tr>
<tr>
<td>GH 2</td>
<td>PF00703.15</td>
<td>7</td>
<td>27</td>
<td>28</td>
<td>GH 39</td>
<td>PF01229.11</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>GH 2_C</td>
<td>PF02836.11</td>
<td>18</td>
<td>38</td>
<td>31</td>
<td>GH 42</td>
<td>PF02449.9</td>
<td>8</td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td>GH 2_N</td>
<td>PF02837.12</td>
<td>26</td>
<td>61</td>
<td>44</td>
<td>GH 42C</td>
<td>PF08333.4</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>GH 3</td>
<td>PF00933.15</td>
<td>49</td>
<td>69</td>
<td>59</td>
<td>GH 42 M</td>
<td>PF08324.7</td>
<td>7</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>GH 3_C</td>
<td>PF01915.16</td>
<td>29</td>
<td>39</td>
<td>33</td>
<td>GH 43</td>
<td>PF04616.8</td>
<td>12</td>
<td>27</td>
<td>33</td>
</tr>
<tr>
<td>GH 4</td>
<td>PF02056.10</td>
<td>9</td>
<td>20</td>
<td>10</td>
<td>GH 44</td>
<td>PF12891.1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>GH 8</td>
<td>PF01270.1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>GH 4C</td>
<td>PF11975.1</td>
<td>8</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>GH 9</td>
<td>PF00759.1</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>GH 53</td>
<td>PF07745.7</td>
<td>1</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>GH 10</td>
<td>PF00331.14</td>
<td>4</td>
<td>10</td>
<td>11</td>
<td>GH 57</td>
<td>PF03651.2</td>
<td>22</td>
<td>38</td>
<td>12</td>
</tr>
<tr>
<td>GH 15</td>
<td>PF00723.1</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>GH 62</td>
<td>PF03664.7</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH 16</td>
<td>PF00722.15</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>GH 63</td>
<td>PF03200.1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH 17</td>
<td>PF00332.12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>GH 65C</td>
<td>PF03633.9</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>GH 18</td>
<td>PF00704.22</td>
<td>7</td>
<td>11</td>
<td>7</td>
<td>GH 65 m</td>
<td>PF03629.9</td>
<td>19</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>GH 19</td>
<td>PF0182.13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>GH 65 N</td>
<td>PF03636.9</td>
<td>3</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>GH 20</td>
<td>PF00728.16</td>
<td>25</td>
<td>40</td>
<td>22</td>
<td>GH 67C</td>
<td>PF07477.6</td>
<td>2</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>GH 20b</td>
<td>PF02838.9</td>
<td>12</td>
<td>6</td>
<td>7</td>
<td>GH 67 M</td>
<td>PF07488.6</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>GH 25</td>
<td>PF01183.14</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>GH 67 N</td>
<td>PF03648.8</td>
<td>4</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>GH 26</td>
<td>PF02156.9</td>
<td>5</td>
<td>10</td>
<td>6</td>
<td>GH 76</td>
<td>PF03663.8</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GH 28</td>
<td>PF00295.11</td>
<td>2</td>
<td>16</td>
<td>11</td>
<td>GH 77</td>
<td>PF02446.11</td>
<td>9</td>
<td>26</td>
<td>17</td>
</tr>
<tr>
<td>GH 30</td>
<td>PF02055.10</td>
<td>3</td>
<td>8</td>
<td>11</td>
<td>GH 81</td>
<td>PF03639.7</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GH 31</td>
<td>PF01055.20</td>
<td>20</td>
<td>29</td>
<td>30</td>
<td>GH 88</td>
<td>PF07470.1</td>
<td>9</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>GH 32C</td>
<td>PF08244.6</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>GH 92</td>
<td>PF07971.6</td>
<td>34</td>
<td>49</td>
<td>46</td>
</tr>
<tr>
<td>GH 32 N</td>
<td>PF00251.14</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>GH 97</td>
<td>PF10566.3</td>
<td>15</td>
<td>49</td>
<td>26</td>
</tr>
<tr>
<td>GH 35</td>
<td>PF01301.13</td>
<td>12</td>
<td>4</td>
<td>18</td>
<td>GH 101</td>
<td>PF12905.1</td>
<td>7</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>GH 38</td>
<td>PF01074.16</td>
<td>8</td>
<td>26</td>
<td>21</td>
<td>GH cc</td>
<td>PF11790.2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Total number of counts: 468, 792, 657

Legend: Dynamic process in enzymology with time and substrate nature was observed in the bioreactors, the heightened GH and abundance (in bold) in the pretreated-sludge bioreactor were correlated with a higher sludge solubilization, and in turns a higher SCFA production.
chromatograph (GC-14B, Shimadzu, Japan). The analyses of VSS were conducted in accordance with APHA Standard Methods [75]. In this study, the methane yield was reported as the amount of methane generated per gram of volatile suspended solid added (mL-CH₄/gVSS-added) unless otherwise stated. The full unit of methane production is mL-CH₄/gVSS-added, in which the VSS-added refers to the initial VSS to facilitate comparison of methane production. Moreover, there was no SCFA in the original sludge before anaerobic fermentation, meaning that the initial VSS did not contain volatile fatty acids—hence the risk of overestimation of specific methane yield is improbable [4].

Preparation of anaerobic digester samples
The granular samples cultured with pretreated (pH 10 for 8 d) and un-pretreated WAS were collected when the methane production rate reached the maximum on day 17, and then further cultured by a semi-continuous method to maintain the methane production rate. In replicates of three, the semi-continuous flow reactors with working volume of 1.0 L received 470 mL of pH 10 pretreated sludge and un-pretreated sludge, respectively. Everyday, 52 mL sludge mixture was manually wasted from each reactor, and 52 mL pretreated and un-pretreated sludge was added to the respective reactors; AGS was fed back to reactors by sifting the wasted sludge through a sifter with aperture of 0.2 mm. Further procedure details were described in [4].

As a stable methane generation continued to be maintained by the semi-continuous culturing method, 50 mL anaerobic digester samples were collected from the pretreatment set-ups (as pooled sample) at day 30 (P30), control set-up at day 30 (UP30) and control set-up at day 40 (UP40) for microbial studies. The two samples collected at day 30 (P30, UP30) served to characterize taxonomic and functional difference between pretreated and unpretreated samples. Both sludge bioreactor were seeded with the same synthetic wastewater cultured AGS substrate, and later gradually replaced by the experimental sludge (pretreated or unpretreated WAS) completely as the batch experiment continued. Since both bioreactors received the same inoculums, therefore the only existing variable was the pretreatment, ascribing to the observed difference between the microbiomes. As for UP40, this additional un-pretreated sample at day 40 was prepared to provide further insight of the microbial dynamic in response to time. Figure 1 shows the experimental scheme for the preparation of the anaerobic digester samples.

DNA extraction and 454 sequencing
Samples (P30, UP30 and UP40) were kept at −20°C and thawed once only for nucleic acid extraction using PowerSoil® DNA Isolation Kit (Mobio, Carlsbad, USA), conducted within 24 hours of sample collection. DNA quantification was performed using Nanodrop 2000 spectrophotometer (Nanodrop Technologies, Wilmington, USA). Sequencing of total DNA derived from each was done by the shotgun sequencing approach on the 454 GS Junior system, and the sequencing quality was maintained by 454 propriety software. The raw sequence reads are submitted under the Sequence Read Archive (SRA) accession number SAMN01924662 (P30), SAMN01924663 (UP30), SAMN01924664 (UP40).

Metagenome analysis
To obtain a quantitative picture of the taxonomic composition, sequence datasets were characterized without a prior assembly step [56]. The unassembled sequence reads were submitted to the online metagenome analysis tool Metagenome Rapid Annotation using Subsystem Technology (MG-RAST analysis pipeline 3.0) [76]. Following the quality trimming with default setting, taxonomic and functional profiling was performed using the M5 non-redundant protein database (M5NR) and KEGG Orthology (KO) reference database [54] Parameters used were e-value cutoff of 1 × 10⁻⁵, minimum alignment length of 50 base pairs, and a minimum percentage identity at 50%. Relative abundance of the Archaeal and Bacterial groups were determined by the percentages of respective reads over total assigned reads. Parallel processing was performed similarly on Megan4 to confirm the clustering pattern of taxonomic structure [77]. Only the taxonomic and functional annotation provided by MG-RAST was relied upon for all subsequent analysis.

To identify important carbohydrate-metabolism related gene candidates, pyrosequence reads were first subjected to gene prediction by a heuristic approach.
using MetaGeneMark 1.0 [69]. Next, the amino acid sequences were screened against catalytic domains (CD) characteristic for particular glycoside hydrolase (GH) families or for a particular carbohydrate binding module (CBM) as classified by the CAZY database [68]. HMM models were downloaded from Pfam database version 25.0 and used as a database for pfam_scan [78] (e-value cutoff of 1x10^-4). Methanogenesis related gene candidates were identified with reference to KEGG and Metacyc pathway database [54,55].

Additional files

Additional file 1: Taxonomic clustering of microbiome in pretreated and un-pretreated sludge bioreactor at day 30 and 40.

Additional file 2: Methanogenesis gene identifications in pyrosequencing meta-datasets UP30, UP40, P30 using KO and SEED reference database. Genes related to methanogenesis were identified by comparing the pyrosequence meta-datasets UP30, UP40 and P30 with KO and SEED database. Negative identifications were indicated as 

Additional file 3: Detailed profile related to (a) amino acid and (b) fat metabolism in the meta-datasets UP30, UP40 and P30 using KO reference database. Expression of these metabolism faculties in the un-pretreated sludge bioreactor gradually elevated over time (compare UP30 and UP40), and were highest in pretreated-sludge bioreactor (P30). Label: [KO Map number] [metabolism detail] [path number].

Additional file 4: Properties of WAS and AGS after settling. Description of data: All values are expressed in mg/L except pH. The data are the averages and their standard deviations in duplicate tests.

Additional file 5: Synthetic wastewater composition. Description of data: Chemical characterization of synthetic wastewater.

Abbreviations

AGS: Anaerobic granular sludge; CBM: Carbohydrate binding modules; CAZY: Carbohydrate-active enzyme; CD: Catalytic domains; GH: Glycosyl hydrolases; PPase: Inorganic pyrophosphatase; SRA: Sequence Read Archive; SCFAs: Short-chain fatty acids; USAID: Upward-flow anaerobic sludge blanket; VSS: Volatile suspended solid; WAS: Waste activated sludge; WWTPs: Wastewater treatment plants.

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

MTW performed the sample preparation, sequence analyses, data interpretation and compiled the manuscript. DZ carried out the anaerobic digestion and physical/chemical characterization. JL compiled the computational analyses tools. RRHH performed the next-generation sequencing. HMT, MSB and TJP contributed to data interpretation and manuscript's final revision. FCL and YC conceived and coordinated the study. All authors read and approved the final manuscript.

Authors’ information

FCL is a Full Professor at School of Biological Sciences, former Dean of Faculty of Science, Board member of Initiative on Clean Energy and Environment, The University of Hong Kong, and Fellow of American Association for the Advancement of Science (AAAS). RRHH is a post-doc, JL, HMT, MSB and TJP are PhD students, MTW is an MPhil student in the same institution. YC is a Full Professor and DZ is a post-doc at State Key Laboratory of Pollution Control and Resources Reuse, School of Environmental Science and Engineering, Tongji University.

Acknowledgements

This study was partially supported by Initiative on Clean Energy and Environment, The University of Hong Kong (HKU-ICEE) and a Bioinformatics Center fund from Nanjing Agricultural University. We are grateful to P. Benzé, C. Bush and I. Tasovski for editing the manuscript.

Author details

1 State Key Laboratory of Pollution Control and Resources Reuse, School of Environmental Science and Engineering, Tongji University, 1239 Siping Road, Shanghai 200092, China. 2 Bioinformatics Center, Nanjing Agricultural University, Nanjing, China.

Received: 28 July 2012 Accepted: 6 March 2013
Published: 19 March 2013

References


doi:10.1186/1754-6834-6-38

Cite this article as: Wong et al.: Towards a metagenomic understanding on enhanced biomethane production from waste activated sludge after pH 10 pretreatment. Biotechnology for Biofuels 2013 6:38.