UNIVERSITY of York

This is a repository copy of *Biogas production from hydrothermal liquefaction wastewater (HTLWW): Focusing on the microbial communities as revealed by high-throughput sequencing of full-length 16S rRNA genes*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/105843/

Version: Accepted Version

Article:
Chen, Huihui, Wan, Jingjing, Chen, Kaifei et al. (4 more authors) (2016) Biogas production from hydrothermal liquefaction wastewater (HTLWW): Focusing on the microbial communities as revealed by high-throughput sequencing of full-length 16S rRNA genes. Water research. pp. 98-107. ISSN 0043-1354

https://doi.org/10.1016/j.watres.2016.09.052

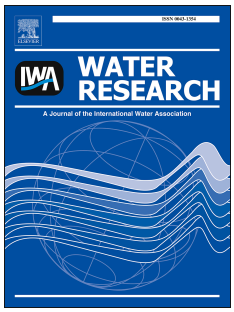
Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Accepted Manuscript



Biogas production from hydrothermal liquefaction wastewater (HTLWW): Focusing on the microbial communities as revealed by high-throughput sequencing of full-length 16S rRNA genes

Huihui Chen, Jingjing Wan, Kaifei Chen, Gang Luo, Jiajun Fan, James Clark, Shicheng Zhang

PII: S0043-1354(16)30729-1

DOI: 10.1016/j.watres.2016.09.052

Reference: WR 12391

To appear in: *Water Research*

Received Date: 17 July 2016

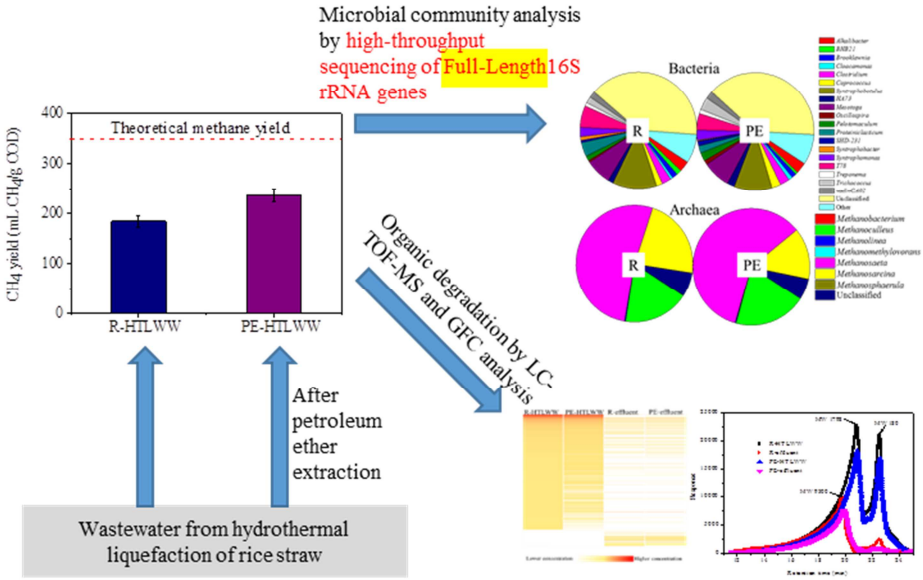
Revised Date: 23 September 2016

Accepted Date: 25 September 2016

Please cite this article as: Chen, H., Wan, J., Chen, K., Luo, G., Fan, J., Clark, J., Zhang, S., Biogas production from hydrothermal liquefaction wastewater (HTLWW): Focusing on the microbial communities as revealed by high-throughput sequencing of full-length 16S rRNA genes, *Water Research* (2016), doi: 10.1016/j.watres.2016.09.052.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

Biogas production from hydrothermal liquefaction wastewater (HTLWW): Focusing on the microbial communities as revealed by high-throughput sequencing of full-length 16S rRNA genes

Huihui Chen[1], Jingjing Wan[1], Kaifei Chen[1], Gang Luo[1]*, Jiajun Fan[2], James Clark[2], Shicheng Zhang[1]*

[1]Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP[3]), Department of Environmental Science and Engineering, Fudan University, Shanghai 200433, China

[2]Green Chemistry Centre of Excellence, Department of Chemistry, University of York, York, YO10 5DD, U.K.

* Corresponding author:

Gang Luo: gangl@fudan.edu.cn, +86 65642297

Shicheng Zhang: zhangsc@fudan.edu.cn, +86 65642297

ACCEPTED MANUSCRIPT

Abstract

Hydrothermal liquefaction (HTL) is an emerging and promising technology for the conversion of wet biomass into bio-crude, however, little attention has been paid to the utilization of hydrothermal liquefaction wastewater (HTLWW) with high concentration of organics. The present study investigated biogas production from wastewater obtained from HTL of straw for bio-crude production, with focuses on the analysis of the microbial communities and characterization of the organics. Batch experiments showed the methane yield of HTLWW (R-HTLWW) was 184 mL/g COD, while HTLWW after petroleum ether extraction (PE-HTLWW), to extract additional bio-crude, had higher methane yield (235 mL/g COD) due to the extraction of recalcitrant organic compounds. Sequential batch experiments further demonstrated the higher methane yield of PE-HTLWW. LC-TOF-MS, HPLC and gel filtration chromatography showed organics with molecular weight (MW)<1000 were well degraded. Results from the high-throughput sequencing of full-length 16S rRNA genes analysis showed similar microbial community compositions were obtained for the reactors fed with either R-HTLWW or PE-HTLWW. The degradation of fatty acids were related with *Mesotoga infera*, *Syntrophomonas wolfei* et al. by species level identification. However, the species related to the degradation of other compounds (e.g. phenols) were not found, which could be due to the presence of uncharacterized microorganisms. It was also found previously proposed criteria (97 % and 98.65 % similarity) for species identification of 16S rRNA genes were not suitable for a fraction of 16S rRNA genes.

Key Words: hydrothermal liquefaction wastewater; biogas production; degradation of

organics; microbial community compositions

1. Introduction

Hydrothermal liquefaction (HTL) is an attractive mean to generate renewable bio-energy from biomass. The organic components of biomass are converted into bio-crude under certain temperatures (200-350 ^{o}C) and pressures (4-22 MPa), and at the same time a large amount of wastewater containing various organic compounds is produced in the process (Fig 1) (Gai et al. 2015). Previous studies mainly focused on the characterization and potential utilization of the bio-crude (Davis et al. 2011, Xu and Lad 2008), and little attention was paid to the utilization of hydrothermal liquefaction wastewater (HTLWW) even though a significant fraction (20-50 %) of the organics in the biomass was converted and entered into HTLWW (Panisko et al. 2015, Tommaso et al. 2015, Xu and Lad 2008). HTLWW may contain cyclopentenones, phenols, acids et al. depending on the feedstocks and reaction conditions (Cheng et al. 2016, Panisko et al. 2015, Villadsen et al. 2012). Inappropriate disposal of HTLWW would result in the environmental pollution considering its high organic contents. The utilization of HTLWW is therefore important in order to achieve overall environmental and economical sustainability of the HTL process(Nelson et al. 2013).

Anaerobic digestion is widely used in the treatment of organic wastes, which can reduce the environmental pollution of organic wastes and at the same time produce energy in the form of biogas. Only one previous study investigated the biogas potential of HTLWW, which was obtained from the HTL of algae (Tommaso et al. 2015). It was reported that around 44 %-61 % of the COD was removed and converted to biogas after anaerobic digestion, which indicated that part of the organics in HTLWW was not bio-degradable or even toxic to the

62 microorganisms. Therefore, the characterization of the organic components in HTLWW and
63 elucidation of their degradation during anaerobic digestion are crucial in order to make full
64 utilization of HTLWW. Lignocellulosic materials, different from algae, are abundant in the
65 world, and their utilization via HTL has been studied before (Gan 2012, Kumagai et al. 2007,
66 Tekin et al. 2014), however, the HTLWW remains to be investigated. Since the organic
67 components of HTLWW strongly depend on the feedstocks, it is necessary to investigate the
68 biogas potential of HTLWW obtained from the HTL of lignocellulosic materials. In addition,
69 several previous studies not only extract the bio-crude from solid phase as shown in Fig 1,
70 but also from HTLWW (Leng et al. 2015, Shuping et al. 2010, Yin et al. 2010). The organic
71 solvents could extract some compounds like ketones, phenols and aloxyphenolic according to
72 Yang's research (Yang et al. 2014), which were recalcitrant or inhibitory molecules for
73 anaerobic digestion. Therefore, the extraction of HTLWW by organic solvents before
74 anaerobic digestion could potentially increase the biogas yield (Cheng et al. 2016, Mottu et al.
75 2000). However, the effects of organic solvents extraction on the subsequent biogas
76 production from HTLWW was still unknown. Organic solvents, including tetrahydrofuran,
77 toluene, ethyl acetate, acetone, ether, methylene chloride, methanol petroleum ether and
78 n-hexane, are organic solvents that can be used to extract bio-oil from HTL mixture products
79 and their extraction properties were main determine by their polarity (Yang et al. 2014).
80 Some organic solvents are highly toxic to human and therefore only four solvents with
81 different polarities and less toxicity to human were chosen (Semenov 1986)."
82 Anaerobic digestion involves various microorganisms for the degradation of organic
83 compounds (Luo et al. 2016b). Considering the complex organics in HTLWW, it is necessary

to reveal the microbial communities responsible for the degradation of organics in HTLWW, what would provide in-depth understanding of anaerobic digestion of HTLWW. The rapid development of next-generation sequencing technologies makes it possible to reveal the diversity and structure of the microbial community, with high sequencing depth (Luo et al. 2013). However, currently 16S rRNA genes analysis were mainly based on the second generation sequencing (e.g. 454 GS Junior (Roche), Miseq (Illumina), and Ion Torrent PGM (Life Technologies)), which could only make sequencing on short sequences (< 600bp) and were not able to provide reliable taxonomic information down to genus and species level (Loman et al. 2012, Mosher et al. 2013). Single molecule, real time sequencing (i.e. third generation sequencing) by PacBio RS SMRT chip can generate longer sequences than the second generation sequencing, and is possible to make high-throughput sequencing of the full-length 16S rRNA genes (Mosher et al. 2013, Mosher et al. 2014). A previous study demonstrated the sequences obtained from high-throughput sequencing of full-length 16S rRNA genes of *Shewanella oneidensis MR1* by Pacific Biosciences RS II sequencer can be accurately assigned to the species level (>99 % accuracy) (Mosher et al. 2014). However, high-throughput sequencing of full-length 16S rRNA genes has not been used for the microbial community analysis in mixed cultures (e.g. anaerobic digestion) until now.

Based on the above considerations, the present study aimed to elucidate the mechanisms involved in biogas production from HTLWW obtained from HTL of rice straw. The biogas production potentials from HTLWW extracted by various commonly used organic solvents were investigated, the organics and their removal during anaerobic digestion were characterized, and the microbial community involved in the anaerobic digestion of HTLWW

were revealed by high-throughput sequencing of full-length 16S rRNA genes using Pacific Biosciences RS II sequencer for the first time.

2. Material and methods

2.1. HTLWW

The HTLWW was obtained from a pilot-scale hydrothermal reactor with a volume of 80 L. 3.0 kg of minced rice straw mixed with 47 kg of water were added into the reactor and then heated to 280 ^{o}C at 12.0 MPa for 30 min (Chen et al. 2015). The mixture was filtered by a 300-mesh screen after HTL, and the filtrate was HTLWW.

HTLWW was then extracted by petroleum ether (PE), cyclohexane (CH), dichloromethane (DM) and ethyl acetate (EA) to separate parts of the organic components (Duan and Savage 2011, Yang et al. 2014), and they were named as PE-HTLWW, CH-HTLWW, DM-HTLWW and EA-HTLWW, respectively. The raw HTLWW was named as R-HTLWW. For the extraction, 125 mL organic solvent was added to a 500 mL bottle, and 250 mL HTLWW was also added. The bottles were then capped tightly and shaken with the speed of 120 rpm for 10 min by a shaker (Duan and Savage 2011). The mixture was then transferred to a funnel for the separation of organic solvents and HTLWW. The above procedure was repeated for the separated HTLWW for the second time extraction. The four samples PE-HTLWW, CH-HTLWW, DM-HTLWW and EA-HTLWW were then obtained. They were all placed in a refrigerator at -20 ^{o}C for further usage. Table 1 presents the COD values of the HTLWW samples and the saturated organic solvents in water.

2.2. Biogas production potentials of HTLWW

Batch experiments were conducted to determine the biogas potentials of HTLWW extracted

by various organic solvents. 118 mL serum bottles were used. 15 mL inoculum and 45 mL BA medium containing a certain amount of HTLWW were added to each bottle. The initial COD value of all the bottles were 0.75 g/L by adding different amounts of HTLWW to the BA medium. The pH value was adjusted to 7.5. All the bottles were flushed with N_2 for 5 min to remove oxygen, and then sealed with butyl rubber stoppers and aluminum screw caps. All the bottles were placed in an incubator with constant temperature 37 ^{o}C. The inoculum was obtained from an anaerobic reactor treating cassava stillage in an ethanol plant (Taicang, Suzhou, China). The bottles with only inoculum were used as control. All the experiments were done in triplicates.

2.3. Semi-continuous experiments

Based on the batch experiments, R-HTLWW and PE-HTLWW were used for the anaerobic sequencing batch reactors (ASBR) to determine the long-term biogas production performances, the degradation of organics, and the microbial community involved in the degradation of organics. ASBR has been widely used in previous studies for the treatment of organic wastewater (Angenent et al. 2002, Timur and Özturk 1999). Two 800 mL ASBR were used with working volume 400 mL. The reactors were fed every two days. The reactors were settled for 2 hours before discharging the supernatant, and new substrates were then fed to the reactors. The hydraulic retention time was controlled at 5 days and sludge retention time was controlled at 40 days by discharging excess sludge periodically for each reactor. Initially, 10 g/L glucose was used as the substrate to ensure both reactors had comparable performances. Then reactor R was fed with R-HTLWW, and reactor PE was fed with PE-HTLWW. For reactor R, R-HTLWW was diluted to the same COD concentration as PE-HTLWW in order

to have the same organic loading rate as reactor PE.

2.4 High-throughput sequencing of full-length 16S rRNA genes and bioinformatic analysis

Samples were obtained during the steady-states of both reactors. Total genomic DNA was extracted from each sample using QIAamp DNA Stool Mini Kit (QIAGEN, 51504). The quantity and purity of the extracted DNA were checked by Nanodrop 2000. PCR was then conducted with the primers 27F (AGAGTTTGATCCTGGCTCAG) and 1492R (GGTTACCTTGTTACGACTT) for bacteria and the primers 20F (TTCCGGTTGATCCYGCCRG) and 1492R for archaea (DeLong 1992). All PCR amplifications were performed using the Taq PCR Core Kit (QIAGEN) with 1 uL template DNA and 20 pmol of each primer. The PCR conditions for bacteria were: 95 °C for 5 min, 28 cycles of three steps: 95 °C for 45 s, 55 °C for 1 min, and 68 °C for 2 min, followed by a final step at 68 °C for 7 min. The PCR conditions for archaea were: 95 °C for 2 min, 27 cycles of three steps: 94 °C for 45 s, 54 °C for 45 s, and 72 °C for 1.5 min, followed by a final step at 72 °C for 7 min. The samples were sent out for sequencing in one cell of the Pacific Biosciences RS II platform combined with the P4/C2 chemistry. The obtained sequences were deposited into the European Nucleotide Archive (ENA) with accession number PRJEB14373. The onboard software provided on the Pacific Biosciences RS II sequencer was used to eliminate CCS (circular consensus sequences) with <99 % predicted accuracy. The low-quality sequences (no exact matches to the forward and reverse primers, and length <1300 bp) and chimeras were removed from the raw sequencing data by MOTHUR program. The numbers of high quality sequences were 7911 (R) and 9099 (PE) for bacteria with

average length of 1390 bp, 1667 (R) and 1905 (PE) for archaea with average length of 1450 bp. The numbers of sequences were normalized to the same sequencing depths (7911 sequences for bacteria and 1667 sequences for archaea) to facilitate the comparison between different samples. The sequences were clustered into operational taxonomic units (OTU) with cutoff 0.03. Rarefaction curves, Shannon diversity index, coverage were also analyzed by MOTHUR program. The sequences were phylogenetically assigned to taxonomic classifications by RDP Classifier with a confidence threshold of 80 %. RDP could only assign the sequences into genus level. In order to get species classification, all the sequences were aligned using BLASTN against NCBI 16S rRNA database with strict criteria (percentage identity at both 97 % and 98.65 %, and alignment length>1300 bp). Both 97 % and 98.65 % of percentage identity were proposed in previous studies for species identification (Kim et al. 2014, Stackebrandt and Goebel 1994, Tindall et al. 2010). MEGAN software was then used to assign the sequences down to species level based on the BLASTN results (Huson et al. 2007). The volumes of gases reported in the present study were at standard temperature and pressure.

2.5. Analytical methods

COD was measured according to APHA (APHA 1995). Gas produced during the anaerobic digestion was detected by GC with thermal conductivity detector. Helium was used as the carrier gas (Liu et al. 2016). GC–MS was used to characterize the chemical compositions of organics extracted from HTLWW by different organic solvents. Gas chromatography was performed on a 30 m HP-INNOWax quartz capillary column with 0.25 mm inner diameter (I.D.) and 0.25 μm film thickness with injection temperature of 250 °C. The column was

initially held at 60 °C for 2 min and heated to 250 °C and held there for 10 min. Helium was used as the carrier gas (1.0 mL/min). A NIST Mass Spectral Database was used for compound identification. HPLC was used to measure the organic acids in the HTLWW samples as described previously (Chen et al. 2015). LC-TOF-MS was used to provide a detailed overview of the organic compounds in the HTLWW samples. It was performed on a Waters ACQUITY UPLC system equipped with a binary solvent delivery manager and a sample manager, coupled with a Waters Micromas Q-TOF Premier Mass Spectrometer equipped with an electrospray interface. Acquity BEH C18 column (100 mm×2.1 mm i.d., 1.7 µm; Waters, Milford, USA) was maintained at 45 °C and eluted with gradient solvent from A:B (99:1) to A:B (0:100) at a flow rate of 0.40 mL/min, where B was acetonitrile (0.1 % (v/v) formic acid) and A was aqueous formic acid (0.1% (v/v) formic acid). The wavelength was 280 nm and the injection volume 5.00 ul, column temperature was 50.0 °C. The source and desolvation temperature were 115 °C and 350 °C respectively. The UV–Vis spectrum was studied using absorptions at 254 wavelength, and the analysis was carried out using a double-beam UV–Vis spectrophotometer from Shimadzu (UV-1800). The molecular weight distributions of HTLWW before and after anaerobic digestion were determined by a GFC analyzer (LC-10ADVP, Shimadzu) according to a previous study (Wen et al. 2012).

3. Results and discussion

3.1. Biogas production potentials of HTLWW extracted by different organic solvents

Fig 2(A) presents the cumulative methane yields of HTLWW extracted by different organic solvents during the biogas potential tests. The methane yields increased fast in the first 10 days for the samples not including DM-HTLWW, which could be related with the

degradation of easy biodegradable organics. Slight increase of methane yields were observed after 10 days for R-HTLWW, PE-HTLWW and CH-HTLWW. Fig 2(B) shows the methane yields of the five HTLWW after 27 days digestion. The methane yield of R-HTLWW (184 mL/g COD) was much lower than the theoretical value (350 mL/g COD), and it indicated there were organics which were difficult to be biodegraded. However, the methane yield of HTLWW was increased after extraction by the organic solvents except DM, which showed that proper organic solvents could improve the anaerobic biodegradability of HTLWW. Further study was conducted to characterize the different HTLWW in order to understand how the organic solvents extraction affected its biodegradability.

3.2 Characterization of HTLWW extracted by different organic solvents

The high methane yield of EA-HTLWW was mainly attributed to the degradation of EA rather than the organics in the HTLWW since EA contributed to more than 98% of the COD in EA-HTLWW (Table 1), and the high methane yield of EA itself was shown in Fig S1. The negligible methane yield of DM-HTLWW was due to the toxicity of DM to the methanogens since no methane was produced when DM alone was used (Fig S1), and the toxicity of DM to methanogens was also reported in previous studies (Kim et al. 1996, McBride and Wolfe 1971). The above results showed that both EA and DM were not suitable as organic solvents since they would increase the difficulty of the subsequent utilization of HTLWW. Both PE and CH had low solubility as demonstrated by their contribution to the total COD in Table 1. Higher methane yield was obtained from PE-HTLWW compared to CH-HTLWW and R-HTLWW, which indicated that PE might have extracted more organics that are difficult to be biodegradated and thereby improved the biodegradability of HTLWW.

GC-MS was conducted to characterize the organics that extracted by different solvents Fig S2, and the relative amounts of major compounds extracted by four different organic solvents were summarized in Table S1. For PE and CH, the two weak polar solvents extracted weak polar components including furans, ketones and phenols. In general, more organics were extracted by PE compared to CH, which might result in the increased methane yield of PE-HTLWW since furans, ketones and phenols were recalcitrant or inhibitory molecules for anaerobic digestion (Speece 1983). Compounds detected from DM and EA organic phases had higher response values than those from PE and CH organic phases in terms of both quantities and types, which was consistent with their higher extracting yields (Table 1), and the results were also agreed with Yang's (Yang et al. 2014) study where DM and EA with higher polarity were found to extract more organic acids, alcohol, ketones and phenols since many polar organic can be produced in HTL process. As GC-MS in our study only detected the extracted compounds by organic solvents, the organic acids in the HTLWW, which was shown to be dominant in the HTLWW in a previous study (Panisko et al. 2015), were further analyzed by HPLC, and the results were show in Table 2. The concentrations of residual organic acids in HTLWW after extraction decreased with the increase of solvent polarity. Lactic acid, acetic acid and propionic acid, which were easy to be converted to methane (Jeris and McCarty 1965, Vandenberg et al. 1976), were not extracted by PE and CH.

3.3 Biogas production from R-HTLWW and PE-HTLWW in ASBR

The two reactors were operated for around 100 days until steady-states were achieved (Fig S3). The methane yield (153 mL/g COD) of R-HTLWW was significantly higher than that (218 mL/g COD) of PE-HTLWW ($P<0.01$, ANOVA). The higher methane yield from

PE-HTLWW compared to R-HTLWW was consistent with the batch experiments. However, the methane yields from both PE-HTLWW and R-HTLWW were relatively lower than that from batch experiments, which could be due to the short HRT since the batch experiments allowed the full conversion of biodegradable organics. The above results further demonstrated that PE extraction improved the biodegradability of HTLWW in a certain extent.

3.4 Degradation of organic compounds in ASBR

The UV-VIS (Fig S4) of R-effluent and PE-effluent spectral absorption decreased compared to R-HTLWW and PE-HTLWW, respectively, which was related with the degradation of organic compounds in the anaerobic reactors. However, the absorption between 210-250 nm and 260-300 nm of R-effluent and PE-effluent suggested that ketones and phenols were not fully degraded in the anaerobic reactors (Cheng et al. 2016), which could resulted in the lower methane yield of both R-HTLWW and PE-HTLWW compared to the theoretical value (350 mL/gCOD).

LC-TOF-MS identified 785 organic compounds from the four samples. As shown in Fig 3, the dominant organic compounds were well degraded, and the detected organic compounds in R-effluent and PE-effluent were less compared to R-HTLWW and PE-HTLWW, further indicating that most of the organic compounds were degraded in anaerobic reactors. NMDS analysis based on LC-TOF-MS results also showed a clear separation of the samples of influent and effluent. The main organic compounds as determined by GC-MS and HPLC were also identified from LC-TOF-MS as shown in Table S2. It is obvious that most of the organic acids, ketones and about half of the phenols were degraded in the anaerobic reactors

and the organics left in the effluent were mainly phenols. The full degradation of organic acids were also demonstrated by HPLC analysis as shown in Table 2. Organic acids are preferable substrates for biogas production and therefore they could be fully degraded. Although phenols were reported to be biodegradable under mesophilic conditions (Agarry et al. 2008, Karlsson et al. 2000, Knoll and Winter 1989), there were various types of phenols detected in HTLWW (Table S1), which might result in the partly degradation of the phenols.

Since LC-TOF-MS only detected compounds with molecular weight (MW) less than 1000 in our study as shown in Table S2, the MW distributions of compounds in the samples were further measured by GFC. As shown in Fig 4, there were two peaks for R-HTLWW and PE-HTLWW, which corresponded to the MW 1798 and 180. The results indicated that a considerable amount of organics in the HTLWW were higher MW compounds, and it might be the polymers of HTL intermediate like carbonhydrates, cellulose, hemicellulose, lignin and repolymerization compounds (Zhu et al. 2015). After anaerobic digestion, most of the compounds with MW less than 1000 were degraded, which was consistent with the LC-TOF-MS and HPLC results. However, one peak corresponding to MW 9300 was still observed for samples R-effluent and PE-effluent, which suggested that the organics with MW higher than 1000 were not well degraded. Therefore, the lower methane yields of both R-HTLWW and PE-HTLWW compared to the theoretical value (350 mL/gCOD) could be mainly attributed to the presence of MW higher than 1000 in the HTLWW. Furthermore, a small peak with MW around 180 was observed for both R-HTLWW and PE-HTLWW, which might relate with the organics which were not fully biodegraded as mentioned before.

3.5 Microbial community compositions as revealed by high-throughput sequencing of full-length 16S rRNA genes

The samples obtained from the continuous reactors were then used for microbial community analysis. The rarefaction curves of all the samples at 0.03 distance is shown in Fig S5. The curves of bacteria and archaea were overlapped for both samples, and it indicated samples R and PE had similar microbial richness, which was also reflected by the similar OTU numbers (Bacteria, around 1500 for both samples; Archaea, around 210 for both samples) (Table S3). The results showed that PE extraction of HTLWW did not have obvious effects on the microbial community richness. It should be noted that the sequencing depths for both bacteria (7911) and archaea (1167) were still not enough to cover the whole microbial diversity since plateaus were not achieved for all the rarefaction curves. However, the coverage values for bacteria (>86%) and archaea (>90%) indicated that most common OTUs were detected. The coverage values were relatively lower compared to previous studies (e.g. coverage value 97.4% with sequencing depth 50000 for bacteria (Luo et al. 2013), coverage value 98.7% with sequencing depth 63699 for bacteria (Pan et al. 2015)), which was mainly due to the sequencing depths was relatively lower in our study. However, it should be noted all the above mentioned studies were based on high-throughput sequencing of partial 16S rRNA genes (less than 500 bp). The Shannon diversity index provides both species richness and the evenness of the species in the microbial community (Lu et al. 2012). Similar with the microbial richness, the microbial diversities were not affected by PE extraction of HTLWW for both bacteria (around 5.44) and archaea (around 3.3). The higher OTU numbers and Shannon diversity of bacteria compared to archaea were consistent with previous studies

(Luo et al. 2013, Zhang et al. 2009), further showing bacteria were more diverse than archaea.

The taxonomic classification of bacterial sequences by RDP classifier is shown in Fig 5(A). The similar taxonomic distribution in phylum, class and genus levels were observed for R and PE, further indicating PE extraction did not affect the bacterial communities. It could be due to that PE might only extract unbiodegradable organic compounds and therefore the degraded organic compounds in both reactors R and PE were similar. *Firmicutes*, *Synergistetes*, *Chloroflexi*, and *Bacteroidetes* were dominant phyla, and their dominance in mesophilic anaerobic reactors were also reported previously (Luo et al. 2016a, Sundberg et al. 2013). Although *Thermotogae* had high relative abundance, its dominance was mainly found in thermophilic anaerobic reactors (Shi et al. 2013). Genus level identification indicated *Thermotogae* were mainly composed of *Mesotoga*, which was recently reported to be the only mesophilic genus (Nesbø et al. 2012). *Mesotoga* was reported to use lactic acid and its dominance might be related with the degradation of lactic acid as seen in Table 2. *Clostridia* and *Synergistia* were the dominant classes in phylum *Firmicutes* and *Synergistetes*, respectively, and they were known as syntrophic partners together with hydrogenotrophic methanogens for the efficient degradation of lactic acid and VFAs (Li et al. 2016). Their dominances were most probably related with the high concentrations of lactic acid and VFAs in HTLWW (Table 2). The relative abundances of *Anaerolineae* and *Bacteroidia* were between 7-9 % in both samples, and they were capable of hydrolysis and fermentation of carbohydrates to VFAs (Narihiro and Sekiguchi 2007, Robert et al. 2007) , however, the carbohydrates were not detected in our study (data not shown), which indicated that their

presence might be related with the degradation of other organics. The genus level classification showed that higher percentages (around 40 %) of sequences were unclassified, which was consistent with previous studies (Lu et al. 2012, Luo et al. 2013), and it could be attributed to that most of biogas reactor's communities are still uncharacterized (Bassani et al. 2015). The dominant genus were *Syntrophobotulus*, *Mesotoga*, and *T78*. *Syntrophobotulus glycolicus* is currently the only known member of the genus *Syntrophobotulus*, however, it can only degrade glyoxylate (Yin et al. 2010), which was not detected in our study. Further species level identification did not detected S*yntrophobotulus glycolicus* (Table 3), and it indicated the genus *Syntrophobotulus* might contain unknown species with different metabolic potentials, which deserves further investigation. The role of *Mesotoga* was mentioned previously for the utilization of lactic acid, while the exact role of *T78* was still unknown (Goux et al. 2015).

Species level identification of full-length 16S rRNA gene sequences would provide more information on the microbial compositions and their metabolic potentials. Table 3 summarized the identified bacterial species. At 97 % similarity, the sequences assigned to species level were 5.6 % and 5.1 % of the total sequences for R and PE, respectively. However, increased sequences (9.9% for R and 8.8% for PE) assigned to species level were obtained at 98.65 % similarity. It would be expected less sequences would be assigned to species level with more critical criteria. The higher sequences assigned at 98.65% similarity was attributed to the algorithm (lowest common ancestor) used by MEGAN (Huson et al. 2007). For instance, one sequence might match two or more species in NCBI 16S rRNA genes database at 97 % similarity, therefore MEGAN could not assign the sequence to

species level. However, the matched species might decrease to one at 98.65% similarity, and therefore it could be assigned to species level. Fig S6 shows that 550 sequences were assigned to the genus *Mesotoga*, however, only 172 sequences were further assigned to species *Mesotoga infera* and *Mesotoga prima* at 97 % similarity, while 488 sequences were assigned to the genus *Mesotoga* at 98.65 % similarity and all of the sequences were further assigned to species level (Fig S8). The above results indicated that 97 % similarity was not enough to make species level identification. Although more sequences were assigned to species level at 98.65 %, still the genus *Trichococcus* was not further assigned to species level (Fig S8 and S12). The sequences belonging to *Trichococcus* (Fig S12) were also extracted, and it was found that all the sequences had more than one match to the species in NCBI 16S rRNA genes database at 98.65 % similarity (Table S4). 98.65 % was previously proposed as the threshold for differentiating two species based on the analysis of 6787 genomes belonging to 1738 species (Kim et al. 2014). However, 98.65 % was not the optimal value in our study since microbial community in anaerobic reactor was more diverse. It should be noted that 98.65 % was still suitable for the species level identification of sequences belonging to most genus except *Trichococcus* (Fig S8 and S12). As shown in Table 2, lactic acid and VFAs were well degraded during anaerobic digestion, and their degradation could be correlated with the several known species as shown in Table 3. *Mesotoga infera*, *Mesotoga prima*, and *Petrimonas sulfuriphila* were reported to use lactic acid as carbon source (Ben Hania et al. 2015, Grabowski et al. 2005). *Syntrophobacter sulfatireducens* were known as propionate-oxidizing bacteria (Chen et al. 2005). *Syntrophomonas wolfei*, *Syntrophus aciditrophicus* and *Syntrophus buswellii* were demonstrated to be able to degrade

saturated four to eight fatty acids (Jackson et al. 1999, McInerney et al. 1981, Wallrabenstein and Schink 1994). Both *Syntrophus aciditrophicus* and *Syntrophus buswellii* could also degrade benzoate, which is the intermediate during phenol degradation (Na et al. 2016). However, the species for the degradation of phenols, ketones and alkenes were not detected, which were major organic compounds in HTLWW and were degraded in different extents during anaerobic digestion (Table S2). There were two reasons. First and most important, only a fraction of the bacterial species were recognized and characterized until now (Bassani et al. 2015, Schloss and Handelsman 2005), and therefore many new species remained to be explored, which was reflected by the large numbers of "not assigned" and "no hits" sequences as seen in Fig S6-S13. Second, the sequences had high similarity to several known species, and therefore they were not assigned to the species as discussed before.

Fig 5(B) shows the taxonomic classification of archaea sequences by RDP classifier, and the similar taxonomic distribution in order and genus levels for R and PE also suggested PE extraction did not affect the archaea communities. The order *Methanosarcinales* was dominant in both samples, and it was composed by the genus *Methanosaeta* and *Methanosarcina*. The microorganisms belonging to *Methanosaeta* were strict aceticlastic methanogens, and the higher percentage of *Methanosaeta* compared to *Methanosarcina* was due to the low acetic acid concentration in biogas reactors as seen in Table 2 (Karakashev et al. 2005). All the rest sequences were assigned to the orders *Methanomicrobiales* and *Methanobacteriales*, mediating hydrogenotrophic methanogenesis, which was consistent with the syntrophic degradation of fatty acids and the detected syntrophic species as described before. The genus *Methanoculleus* (Order *Methanomicrobiales*) was the main

hydrogenotrophic methanogenesis genus, which was also found to be dominant in other biogas reactors (Jaenicke et al. 2011, Krause et al. 2008). The species level identification by MEGAN showed that 40.9 % and 47.9 % of the sequences were assigned to species level at 98.65 % similarity, which was much higher than that (<10 %) for bacteria. It could be due to the higher diversity of bacteria compared to archaea as seen in Table S3 and as reported in previous studies (Luo et al. 2013, Zhang et al. 2009), which resulted in the more uncharacterized species in bacteria than that in archaea. The overwhelming majority of the sequences were assigned to *Methanosaeta concilii*. Although more than 200 sequences were assigned to the genus *Methanosarcina* (Fig S9 and S13), only very few sequences were further classified down to species level. The results further indicated that the 98.65 % threshold for differentiating two species was not fully appropriate for all the archaea genus.

3.6 Outlook

The present study showed that HTLWW contains relatively higher amount of unbiodegradable organic compounds (e.g. phenols and other high MW (>1000) organic compounds), which were still left in HTLWW after anaerobic digestion. Therefore, further studies via aerobic biodegradation or chemical oxidation should be conducted to remove the residual organic compounds before discharging to the environment (Jang et al. 2015, Moreira et al. 2015). In addition, the usage of catalysis and changes of the HTL conditions also deserves further investigation in order to decrease the formation of unbiodegradable organic compounds without affecting the bio-crude production (Anastasakis and Ross 2011, Tekin and Karagöz 2013). For the first time, the third generation sequencing by PacBio RS SMRT was applied for the high-throughput sequencing of full-length 16S rRNA genes of mixed

436 cultures. The present study showed that the previously proposed thresholds (97 % and 98.65 %
437 similarity) for species identification of 16S rRNA genes were not suitable for a fraction of
438 16S rRNA genes since different species might have high similarity (>98.65%) (Table S4).
439 Therefore, the species level identification of 16S rRNA genes based on similarity is still
440 challenging and remains further investigation. In addition, high percentages of "not assigned"
441 and "no hits" sequences for bacteria sequences were observed, which could be related with
442 the uncharacterized bacteria, and it could be solved with the gradually increased numbers of
443 characterized species in 16S rRNA gene database. Recently, there were studies focusing on
444 the identification of the genomes of microorganisms from mixed cultures by metagenomic
445 analysis, which is independent of traditional cultivation methods, and thereby it might expand
446 the sequences in 16S rRNA gene database (Bassani et al. 2015, Campanaro et al. 2016).

447 ## 4.Conclusions

448 The present study showed that the methane yield of HTLWW (R-HTLWW) was 184 mL/g
449 COD, while HTLWW after petroleum ether extraction had higher methane yield (235 mL/g
450 COD) due to the extraction of recalcitrant organic compounds. The higher methane yields of
451 PE-HTLWW (225 mL/gCOD) compared to R-HTLWW (160 mL/gCOD) was also
452 demonstrated in the continuous experiments. Further study showed that organics with
453 molecular weight (MW)<1000 were well degraded by LC-TOF-MS, HPLC and gel filtration
454 chromatography analysis. The results from high-throughput sequencing of full-length 16S
455 rRNA genes showed that similar microbial community compositions were obtained for the
456 reactors fed with either R-HTLWW or PE-HTLWW, and the degradation of fatty acids were
457 related with *Mesotoga infera*, *Syntrophomonas wolfei* et al. by species level identification.

458 However, the species related to the degradation of other compounds (e.g. phenols) were not
459 found, and it could be due to the presence of uncharacterized microorganisms. The study also
460 showed that previously proposed criteria (97 % and 98.65 % similarity) for species
461 identification of 16S rRNA genes were not suitable for a fraction of 16S rRNA genes.

462 ## Acknowledgements

463 This study was funded by National Natural Science Foundation of China (51408133,
464 21577025), National Key Technology Support Program (BAD15B06), Yangfan project from
465 Science and Technology Commission of Shanghai Municipality (14YF1400400), and
466 Shanghai Talent Development Fund (2014, 14).

467 ## Supporting Information

468 Supporting Information includes Tables and Figures as noted in the text.

469 ## References

470 Agarry, S.E., Durojaiye, A.O. and Solomon, B.O. (2008) Microbial degradation of phenols: a
471 review. International Journal of Environment and Pollution 32(1), 12-28.

472 Anastasakis, K. and Ross, A.B. (2011) Hydrothermal liquefaction of the brown macro-alga
473 Laminaria Saccharina: Effect of reaction conditions on product distribution and composition.
474 Bioresource Technology 102(7), 4876-4883.

475 Angenent, L.T., Sung, S. and Raskin, L. (2002) Methanogenic population dynamics during
476 startup of a full-scale anaerobic sequencing batch reactor treating swine waste. Water
477 Research 36(18), 4648-4654.

478 APHA (1995) Standard methods for the examination of water and wastewater, 19 th ed. New
479 York, USA: American Public Health Association.

480 Bassani, I., Kougias, P.G., Treu, L. and Angelidaki, I. (2015) Biogas upgrading via
481 hydrogenotrophic methanogenesis in two-stage continuous stirred tank reactors at mesophilic
482 and thermophilic conditions. Environmental Science & Technology 49(20), 12585-12593.
483 Ben Hania, W., Fadhlaoui, K., Brochier-Armanet, C., Persillon, C., Postec, A., Hamdi, M.,
484 Dolla, A., Ollivier, B., Fardeau, M.-L., Le Mer, J. and Erauso, G. (2015) Draft genome
485 sequence of *Mesotoga strain PhosAC3*, a mesophilic member of the bacterial order
486 *Thermotogales*, isolated from a digestor treating phosphogypsum in Tunisia. Standards in
487 Genomic Sciences 10.
488 Campanaro, S., Treu, L., Kougias, P.G., De Francisci, D., Valle, G. and Angelidaki, I. (2016)
489 Metagenomic analysis and functional characterization of the biogas microbiome using high
490 throughput shotgun sequencing and a novel binning strategy. Biotechnology for Biofuels 9(1),
491 1-17.
492 Chen, K., Lyu, H., Hao, S., Luo, G., Zhang, S. and Chen, J. (2015) Separation of phenolic
493 compounds with modified adsorption resin from aqueous phase products of hydrothermal
494 liquefaction of rice straw. Bioresource Technology 182, 160-168.
495 Chen, S., Liu, X. and Dong, X. (2005) *Syntrophobacter sulfatireducens sp. nov.*, a novel
496 syntrophic, propionate-oxidizing bacterium isolated from UASB reactors. International
497 Journal of Systematic and Evolutionary Microbiology 55(3), 1319-1324.
498 Cheng, J.R., Liu, X.M., Chen, Z.Y., Zhang, Y.S. and Zhang, Y.H. (2016) A novel mesophilic
499 anaerobic digestion system for biogas production and in situ methane enrichment from
500 coconut shell pyroligneous. Applied Biochemistry and Biotechnology 178(7), 1303-1314.
501 Davis, R., Aden, A. and Pienkos, P.T. (2011) Techno-economic analysis of autotrophic

502 microalgae for fuel production. Applied Energy 88(10), 3524-3531.

503 DeLong, E.F. (1992) Archaea in coastal marine environments. Proceedings of the National
504 Academy of Sciences 89(12), 5685-5689.

505 Duan, P. and Savage, P.E. (2011) Hydrothermal liquefaction of a microalga with
506 heterogeneous catalysts. Industrial & Engineering Chemistry Research 50(1), 52-61.

507 Gai, C., Zhang, Y., Chen, W.T., Zhou, Y., Schideman, L., Zhang, P., Tommaso, G., Kuo, C.-T.
508 and Dong, Y. (2015) Characterization of aqueous phase from the hydrothermal liquefaction of
509 Chlorella pyrenoidosa. Bioresource Technology 184, 328-335.

510 Gan, J. (2012) Hydrothermal conversion of lignocellulosic biomass to bio-oils. Hydrothermal
511 Conversion.

512 Goux, X., Calusinska, M., Lemaigre, S., Marynowska, M., Klocke, M. and Udelhoven, T.
513 (2015) Microbial community dynamics in replicate anaerobic digesters exposed sequentially
514 to increasing organic loading rate, acidosis, and process recovery. Biotechnology for Biofuels
515 8.

516 Grabowski, A., Tindall, B.J., Bardin, V., Blanchet, D. and Jeanthon, C. (2005) *Petrimonas*
517 *sulfuriphila gen. nov., sp. nov.*, a mesophilic fermentative bacterium isolated from a
518 biodegraded oil reservoir. International Journal of Systematic and Evolutionary Microbiology
519 55(3), 1113-1121.

520 Huson, D.H., Auch, A.F., Qi, J. and Schuster, S.C. (2007) MEGAN analysis of metagenomic
521 data. Genome Research 17(3), 377-386.

522 Jackson, E.B., Bhupathiraju, K.V., Tanner, S.R., Woese, R.C. and McInerney, J.M. (1999)
523 *Syntrophus aciditrophicus sp. nov.*, a new anaerobic bacterium that degrades fatty acids and

benzoate in syntrophic association with hydrogen-using microorganisms. Archives of Microbiology 171(2), 107-114.

Jaenicke, S., Ander, C., Bekel, T., Bisdorf, R., Droge, M., Gartemann, K.-H., Junemann, S., Kaiser, O., Krause, L., Tille, F., Zakrzewski, F., Puhler, A., Schluter, A. and Goesmann, A. (2011) Comparative and joint analysis of two metagenomic datasets from a biogas fermenter obtained by 454-pyrosequencing. Plos One 6, 1.

Jang, H.M., Ha, J.H., Park, J.M., Kim, M.-S. and Sommer, S.G. (2015) Comprehensive microbial analysis of combined mesophilic anaerobic–thermophilic aerobic process treating high-strength food wastewater. Water Research 73, 291-303.

Jeris, J.S. and McCarty, P.L. (1965) The Biochemistry of Methane Fermentation Using \rm C^14 Tracers. Journal of Water Pollution Control Federation 37(2), 178-192.

Karakashev, D., Batstone, D.J. and Angelidaki, I. (2005) Influence of environmental conditions on methanogenic compositions in anaerobic biogas reactors. Applied and Environmental Microbiology 71(1), 331-338.

Karlsson, A., Ejlertsson, J. and Svensson, B.H. (2000) CO_2-dependent fermentation of phenol to acetate, butyrate and benzoate by an anaerobic, pasteurised culture. Archives of Microbiology 173(5-6), 398-402.

Kim, I.S., Young, J.C. and Tabak, H.H. (1996) Impact of chloroanilines on hydrogenotrophic methanogenesis in ethanol-enriched cultures. Water Research 30(3), 601-612.

Kim, M., Oh, H.-S., Park, S.-C. and Chun, J. (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. International Journal of Systematic and Evolutionary Microbiology 64,

346-351.

Knoll, G. and Winter, J. (1989) Degradation of phenol via carboxylation to benzoate by a defined, obligate syntrophilic consortium of anaerobic-bacteria. Applied Microbiology and Biotechnology 30(3), 318-324.

Krause, L., Diaz, N.N., Edwards, R.A., Gartemann, K.-H., Krömeke, H., Neuweger, H., Pühler, A., Runte, K.J., Schlüter, A., Stoye, J., Szczepanowski, R., Tauch, A. and Goesmann, A. (2008) Taxonomic composition and gene content of a methane-producing microbial community isolated from a biogas reactor. Journal of Biotechnology 136(1–2), 91-101.

Kumagai, S., Yamada, N., Sakaki, T. and Hayashi, N. (2007) Biomass: Characteristics of hydrothermal decomposition and saccharification of various lignocellulosic biomass and enzymatic saccharification of the obtained hydrothermal-residue. Journal of the Japan Institute of Energy 86(86), 712-717.

Leng, L., Yuan, X., Chen, X., Huang, H., Wang, H., Li, H., Zhu, R., Li, S. and Zeng, G. (2015) Characterization of liquefaction bio-oil from sewage sludge and its solubilization in diesel microemulsion. Energy 82, 218-228.

Li, L., He, Q., Ma, Y., Wang, X. and Peng, X. (2016) A mesophilic anaerobic digester for treating food waste: process stability and microbial community analysis using pyrosequencing. Microbial Cell Factories 15(1), 1-11.

Liu, Y., Wan, J., Han, S., Zhang, S. and Luo, G. (2016) Selective conversion of carbon monoxide to hydrogen by anaerobic mixed culture. Bioresource Technology 202, 1-7.

Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J. and Pallen, M.J. (2012) Performance comparison of benchtop high-throughput sequencing

568 platforms. Nat Biotech 30(5), 434-439.

569 Lu, L., Xing, D. and Ren, N. (2012) Pyrosequencing reveals highly diverse microbial
570 communities in microbial electrolysis cells involved in enhanced H_2 production from waste
571 activated sludge. Water Research 46, 2425 - 2434.

572 Luo, G., Fotidis, I.A. and Angelidaki, I. (2016a) Comparative analysis of taxonomic,
573 functional, and metabolic patterns of microbiomes from 14 full-scale biogas reactors by
574 metagenomic sequencing and radioisotopic analysis. Biotechnology for Biofuels 9(1), 1-12.

575 Luo, G., Fotidis, I.A. and Angelidaki, I. (2016b) Comparative analysis of taxonomic,
576 functional, and metabolic patterns of microbiomes from 14 full-scale biogas reactors by
577 metagenomic sequencing and radioisotopic analysis. Biotechnology for Biofuels 9.

578 Luo, G., Wang, W. and Angelidaki, I. (2013) Anaerobic digestion for simultaneous sewage
579 sludge treatment and CO biomethanation: Process performance and microbial ecology.
580 Environmental Science & Technology 47(18), 10685-10693.

581 McBride, B.C. and Wolfe, R.S. (1971) Inhibition of methanogenesis by DDT. Nature
582 234(5331), 551-&.

583 McInerney, M.J., Bryant, M.P., Hespell, R.B. and Costerton, J.W. (1981) *Syntrophomonas*
584 *wolfei gen. nov. sp. nov.*, an anaerobic, syntrophic, fatty acid-oxidizing bacterium. Applied
585 and Environmental Microbiology 41(4), 1029-1039.

586 Moreira, F.C., Boaventura, R.A.R., Brillas, E. and Vilar, V.J.P. (2015) Remediation of a
587 winery wastewater combining aerobic biological oxidation and electrochemical advanced
588 oxidation processes. Water Research 75, 95-108.

589 Mosher, J.J., Bernberg, E.L., Shevchenko, O., Kan, J. and Kaplan, L.A. (2013) Efficacy of a

590 3rd generation high-throughput sequencing platform for analyses of 16S rRNA genes from
591 environmental samples. Journal of Microbiological Methods 95(2), 175-181.
592 Mosher, J.J., Bowman, B., Bernberg, E.L., Shevchenko, O., Kan, J., Korlach, J. and Kaplan,
593 L.A. (2014) Improved performance of the PacBio SMRT technology for 16S rDNA
594 sequencing. Journal of Microbiological Methods 104, 59-60.
595 Mottu, F., Laurent, A., Rufenacht, D.A. and Doelker, E. (2000) Organic solvents for
596 pharmaceutical parenterals and embolic liquids: a review of toxicity data. Pda Journal of
597 Pharmaceutical Science & Technology 54(6), 456-469.
598 Na, J.-G., Lee, M.-K., Yun, Y.-M., Moon, C., Kim, M.-S. and Kim, D.-H. (2016) Microbial
599 community analysis of anaerobic granules in phenol-degrading UASB by next generation
600 sequencing. Biochemical Engineering Journal 112, 241-248.
601 Narihiro, T. and Sekiguchi, Y. (2007) Microbial communities in anaerobic digestion processes
602 for waste and wastewater treatment: a microbiological update. Current Opinion in
603 Biotechnology 18(3), 273-278.
604 Nelson, M., Zhu, L., Thiel, A., Wu, Y., Guan, M., Minty, J., Wang, H.Y. and Lin, X.N. (2013)
605 Microbial utilization of aqueous co-products from hydrothermal liquefaction of microalgae
606 Nannochloropsis oculata. Bioresource Technology 136, 522-528.
607 Nesbø, C.L., Bradnan, D.M., Adebusuyi, A., Dlutek, M., Petrus, A.K., Foght, J., Doolittle,
608 W.F. and Noll, K.M. (2012) *Mesotoga prima gen. nov., sp. nov.*, the first described
609 mesophilic species of the Thermotogales. Extremophiles 16(3), 387-393.
610 Pan, F., Xu, A., Xia, D., Yu, Y., Chen, G., Meyer, M., Zhao, D., Huang, C.H., Wu, Q. and Fu,
611 J. (2015) Effects of octahedral molecular sieve on treatment performance, microbial

612 metabolism, and microbial community in expanded granular sludge bed reactor. Water
613 Research 87, 127-136.
614 Panisko, E., Wietsma, T., Lemmon, T., Albrecht, K. and Howe, D. (2015) Characterization of
615 the aqueous fractions from hydrotreatment and hydrothermal liquefaction of lignocellulosic
616 feedstocks. Biomass & Bioenergy 74, 162-171.
617 Robert, C., Chassard, C., Lawson, P.A. and Bernalier-Donadille, A. (2007) *Bacteroides*
618 *cellulosilyticus sp. nov.*, a cellulolytic bacterium from the human gut microbial community.
619 International Journal of Systematic and Evolutionary Microbiology 57(7), 1516-1520.
620 Schloss, P.D. and Handelsman, J. (2005) Metagenomics for studying unculturable
621 microorganisms: cutting the Gordian knot. Genome Biology 6(8), 1-4.
622 Semenov, A.V. (1986) Toxicity of some highly polar organic solvents [Anthelmintics].
623 [Russian].
624 Shi, J., Wang, Z., Stiverson, J.A., Yu, Z. and Li, Y. (2013) Reactor performance and microbial
625 community dynamics during solid-state anaerobic digestion of corn stover at mesophilic and
626 thermophilic conditions. Bioresource Technology 136, 574-581.
627 Shuping, Z., Yulong, W., Mingde, Y., Kaleem, I., Chun, L. and Tong, J. (2010) Production
628 and characterization of bio-oil from hydrothermal liquefaction of microalgae Dunaliella
629 tertiolecta cake. Energy 35(12), 5406-5411.
630 Speece, R.E. (1983) Anaerobic biotechnology for industrial wastewater treatment.
631 Environmental Science & Technology 17(9), A416-A427.
632 Stackebrandt, E. and Goebel, B.M. (1994) Taxonomic note: A place for DNA-DNA
633 reassociation and 16S rRNA sequence analysis in the present species definition in

634 bacteriology. International Journal of Systematic and Evolutionary Microbiology 44(4),
635 846-849.
636 Sundberg, C., Al-Soud, W., Larsson, M., Alm, E., Yekta, S., Svensson, B., Sorensen, S. and
637 Karlsson, A. (2013) 454 pyrosequencing analyses of bacterial and archaeal richness in 21
638 full-scale biogas digesters. Fems Microbiology Ecology 85, 612 - 626.
639 Tekin, K. and Karagöz, S. (2013) Non-catalytic and catalytic hydrothermal liquefaction of
640 biomass. Research on Chemical Intermediates 39(2), 485-498.
641 Tekin, K., Karagoz, S. and Bektas, S. (2014) A review of hydrothermal biomass processing.
642 Renewable & Sustainable Energy Reviews 40, 673-687.
643 Timur, H. and Özturk, I. (1999) Anaerobic sequencing batch reactor treatment of landfill
644 leachate. Water Research 33(15), 3225-3230.
645 Tindall, B.J., Rosselló-Móra, R., Busse, H.-J., Ludwig, W. and Kämpfer, P. (2010) Notes on
646 the characterization of prokaryote strains for taxonomic purposes. International Journal of
647 Systematic and Evolutionary Microbiology 60(1), 249-266.
648 Tommaso, G., Chen, W.T., Li, P., Schideman, L. and Zhang, Y. (2015) Chemical
649 characterization and anaerobic biodegradability of hydrothermal liquefaction aqueous
650 products from mixed-culture wastewater algae. Bioresource Technology 178, 139-146.
651 Vandenberg, L., Patel, G.B., Clark, D.S. and Lentz, C.P. (1976) Factors affecting rate of
652 methane formation from acetate acid by enriched methanogenic cultures. Canadian Journal of
653 Microbiology 22(9), 1312-1319.
654 Villadsen, S.R., Dithmer, L., Forsberg, R., Becker, J., Rudolf, A., Iversen, S.B., Iversen, B.B.
655 and Glasius, M. (2012) Development and application of chemical analysis methods for

656 investigation of bio-Oils and aqueous phase from hydrothermal liquefaction of biomass.
657 Energy & Fuels 26(11), 6988-6998.

658 Wallrabenstein, C. and Schink, B. (1994) Evidence of reversed electron transport in
659 syntrophic butyrate or benzoate oxidation by *Syntrophomonas wolfei* and *Syntrophus*
660 *buswellii*. Archives of Microbiology 162(1), 136-142.

661 Wen, W., Li, X., Gang, L., Qi, Z. and Qin, L. (2012) Optimization of biohydrogen and
662 methane recovery within a cassava ethanol wastewater/waste integrated management system.
663 Bioresource Technology 120(17), 165-172.

664 Xu, C. and Lad, N. (2008) Production of heavy oils with high caloric values by direct
665 liquefaction of woody biomass in sub/near-critical water. Energy & Fuels 22(1), 635-642.

666 Yang, X., Lyu, H., Chen, K., Zhu, X., Zhang, S. and Chen, J. (2014) Selective extraction of
667 bio-oil from hydrothermal liquefaction of salix psammophila by organic solvents with
668 different polarities through multistep extraction separation. Bioresources 9(3), 5219-5233.

669 Yin, S., Dolan, R., Harris, M. and Tan, Z. (2010) Subcritical hydrothermal liquefaction of
670 cattle manure to bio-oil: Effects of conversion parameters on bio-oil yield and
671 characterization of bio-oil. Bioresource Technology 101(10), 3657-3664.

672 Zhang, H., Banaszak, J.E., Parameswaran, P., Alder, J., Krajmalnik-Brown, R. and Rittmann,
673 B.E. (2009) Focused-Pulsed sludge pre-treatment increases the bacterial diversity and relative
674 abundance of acetoclastic methanogens in a full-scale anaerobic digester. Water Research
675 43(18), 4517-4526.

676 Zhu, Z., Rosendahl, L., Toor, S.S., Yu, D. and Chen, G. (2015) Hydrothermal liquefaction of
677 barley straw to bio-crude oil: Effects of reaction temperature and aqueous phase recirculation.

ACCEPTED MANUSCRIPT

678 Applied Energy 137, 183-192.

679

680

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 1 COD values of HTLWW and organic solvents

Organic solvents	Agent polarity	Saturated solvent (gCOD/L)	HTLWW (gCOD/L)	COD contributed by organic solvents (%)	COD extracting percent (%)
-	-	-	20.74	-	-
PE	0.01	0.09	15.99	0.56	23.33
CH	0.1	0.22	17.63	1.25	16.06
DM	3.4	7.95	13.28	59.86	74.30
EA	4.3	146.35	148.90	98.29	87.70

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 2 The concentrations of organic acids (mg/L)

Name	R-HTLWW	PE-HTLWW	CH-HTLWW	DM-HTLWW	EA-HTLWW	R-effluent	PE-effluent
Lactic acid	3722	3708	3698	1628	592	–	–
Acetic acid	1802	1792	1782	740	–	–	–
Propionic acid	680	657	657	399	–	–	–
N-butyric acid	281	289	260	59	–	–	–
Isovaleric acid	146	142	133	–	–	–	–

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 3 Species level identification of the full-length 16S rRNA sequences

	Number of sequences (97% Similarity)		Number of sequences (98.65% Similarity)	
	R	PE	R	PE
Bacteria				
Acinetobacter seohaensis	5	0	3	0
Advenella faeciporci	9	2	0	0
Alkalibacter saccharofermentans	9	5	0	0
Aminivibrio pyruvatiphilus	10	12	0	2
Halothiobacillus neapolitanus	15	6	14	5
Mesotoga infera	169	185	483	519
Mesotoga prima	3	8	5	9
Ornatilinea apprima	14	8	10	4
Parasporobacterium paucivorans	100	93	25	21
Petrimonas sulfuriphila	5	5	5	5
Pseudomonas caeni	4	7	1	3
Pseudomonas stutzeri	2	1	6	5
Syntrophobacter sulfatireducens	25	2	23	2
Syntrophomonas wolfei	28	37	3	3
Syntrophus aciditrophicus	3	1	14	15
Youngiibacter fragilis	9	6	186	89
Others* (28 species)	30	26	9	13
Total	**440 (5.6%)**	**404 (5.1%)**	**787 (9.9%)**	**695 (8.8%)**
Archaea				
Methanosaeta concilii	746	896	666	792
Methanoculleus palmolei	90	109	0	0
Methanomassiliicoccus luminyensis	30	32	0	0
Others* (11 species)	16	13	17	7
Total	**882 (52.9%)**	**1050 (62.9%)**	**683 (40.9%)**	**799 (47.9%)**

*"Others" are the species with numbers of sequences less than 5

ACCEPTED MANUSCRIPT

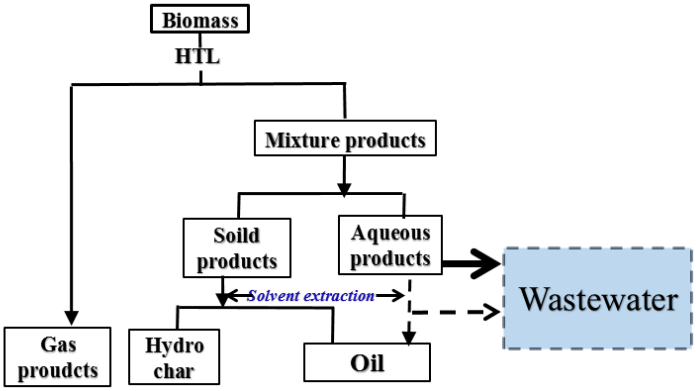
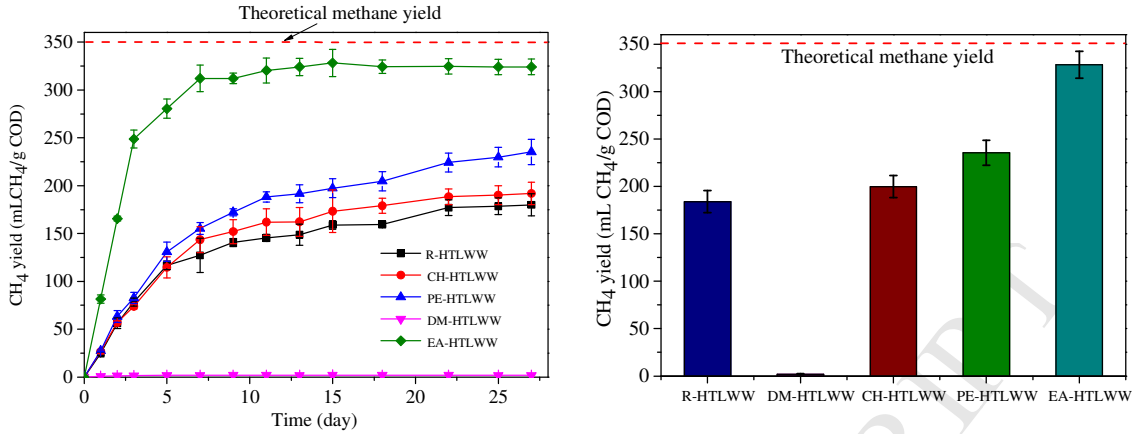


Fig 1 Hydrothermal liquefaction process

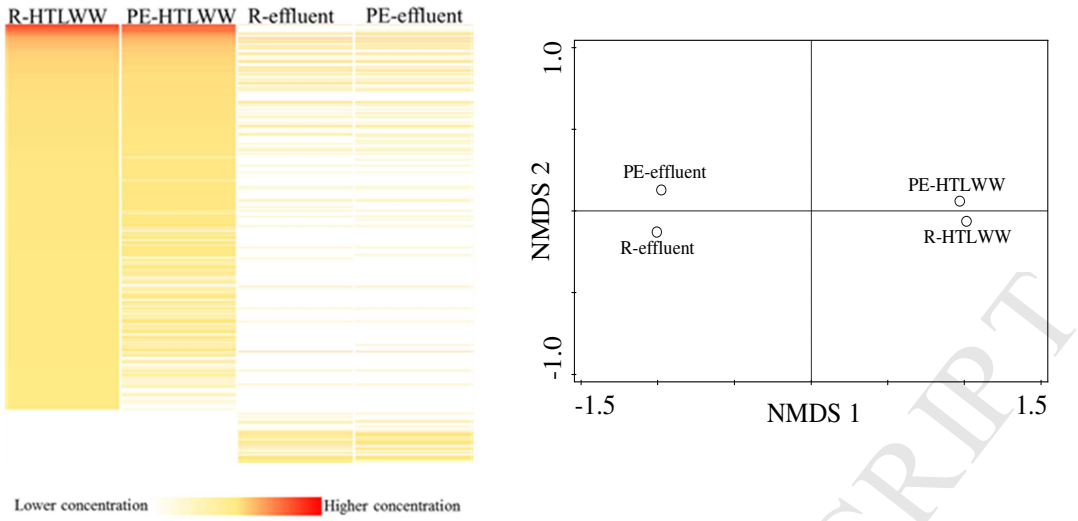
ACCEPTED MANUSCRIPT



(A) (B)

Fig 2 Biogas production potentials of HTLWW (A) time courses of methane production (B) methane yields

ACCEPTED MANUSCRIPT



(A) (B)

Fig 3 Heatmap (A) and NMDS (B) analysis of the samples based on LC-TOF-MS results

ACCEPTED MANUSCRIPT

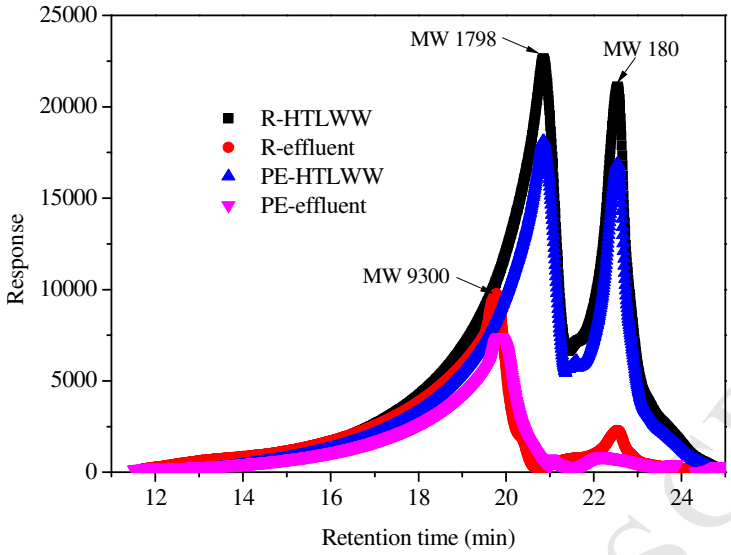


Fig 4 GFC analysis of the samples

ACCEPTED MANUSCRIPT

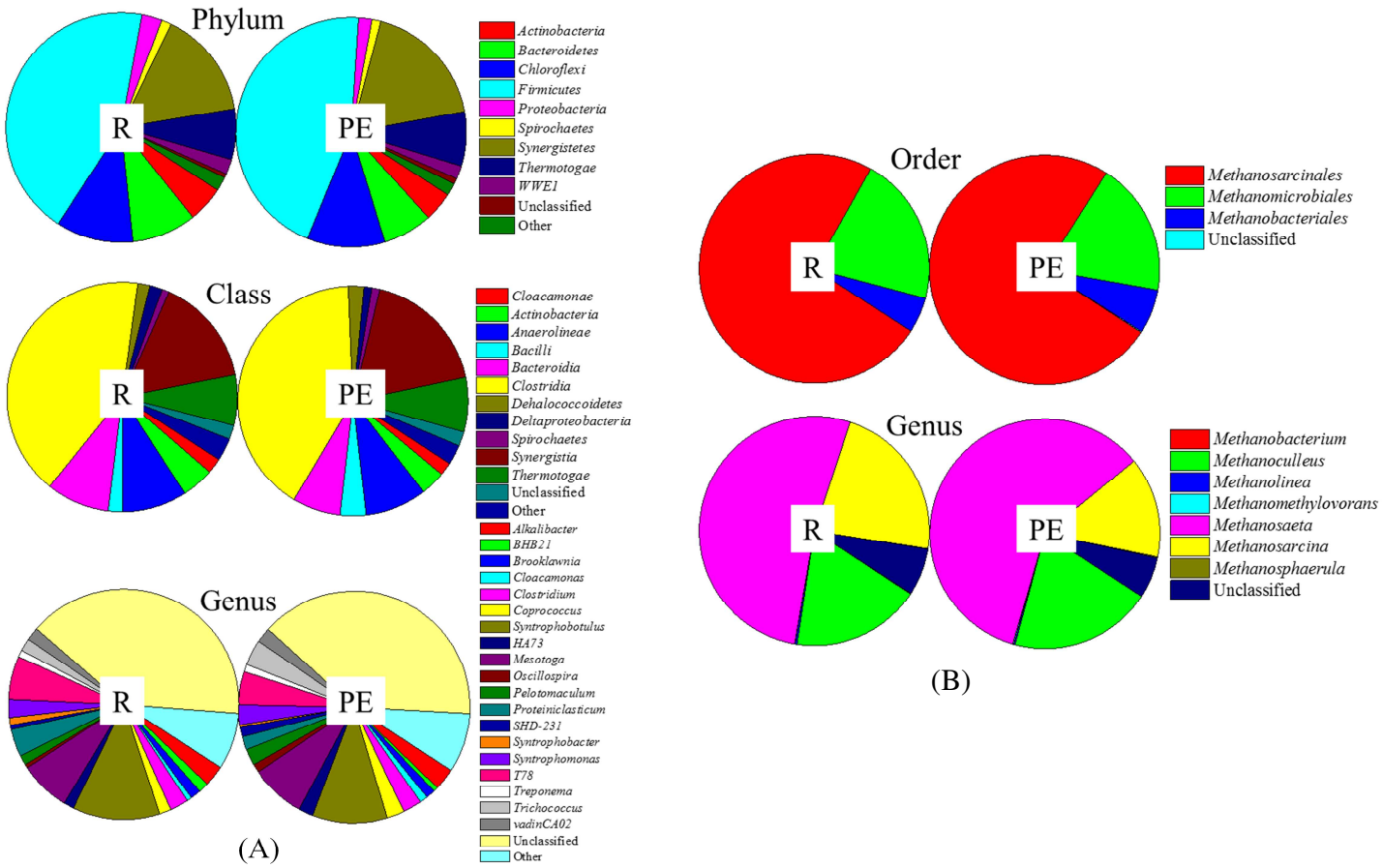


Fig 5 Taxonomic classification of bacteria (A) and archaea (B) sequences based on the high-throughput sequencing of full-length 16S rRNA genes

ACCEPTED MANUSCRIPT

Highlights:

- The methane yield of HTLWW was increased after petroleum ether extraction
- Organics in HTLWW with molecular weight (MW)<1000 were well degraded
- High-throughput sequencing of full-length 16S rRNA genes was applied
- Microbial community compositions were analyzed down to the species level

ACCEPTED MANUSCRIPT