

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](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# 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](https://doi.org/10.1016/j.watres.2016.09.052)

Reference: WR 12391

To appear in: *Water Research*

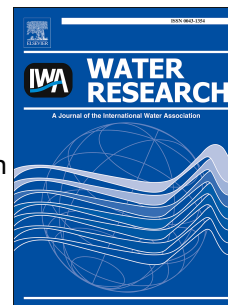
Received Date: 17 July 2016

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.





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

4  
5 Huihui Chen<sup>1</sup>, Jingjing Wan<sup>1</sup>, Kaifei Chen<sup>1</sup>, Gang Luo<sup>1\*</sup>, Jiajun Fan<sup>2</sup>, James Clark<sup>2</sup>, Shicheng  
6 Zhang<sup>1\*</sup>

7  
8 <sup>1</sup>Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP<sup>3</sup>),  
9 Department of Environmental Science and Engineering, Fudan University, Shanghai 200433,  
10 China

11 <sup>2</sup>Green Chemistry Centre of Excellence, Department of Chemistry, University of York, York,  
12 YO10 5DD, U.K.

13  
14 \* Corresponding author:

15 Gang Luo: [gangl@fudan.edu.cn](mailto:gangl@fudan.edu.cn), +86 65642297

16 Shicheng Zhang: [zhangsc@fudan.edu.cn](mailto:zhangsc@fudan.edu.cn), +86 65642297

17

**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

40 organics; microbial community compositions

## 41 **1. Introduction**

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

56 Anaerobic digestion is widely used in the treatment of organic wastes, which can reduce the  
57 environmental pollution of organic wastes and at the same time produce energy in the form of  
58 biogas. Only one previous study investigated the biogas potential of HTLWW, which was  
59 obtained from the HTL of algae (Tommaso et al. 2015). It was reported that around 44 %-61 %  
60 of the COD was removed and converted to biogas after anaerobic digestion, which indicated  
61 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

84 to reveal the microbial communities responsible for the degradation of organics in HTLWW,  
85 what would provide in-depth understanding of anaerobic digestion of HTLWW. The rapid  
86 development of next-generation sequencing technologies makes it possible to reveal the  
87 diversity and structure of the microbial community, with high sequencing depth (Luo et al.  
88 2013). However, currently 16S rRNA genes analysis were mainly based on the second  
89 generation sequencing (e.g. 454 GS Junior (Roche), Miseq (Illumina), and Ion Torrent PGM  
90 (Life Technologies)), which could only make sequencing on short sequences (< 600bp) and  
91 were not able to provide reliable taxonomic information down to genus and species level  
92 (Loman et al. 2012, Mosher et al. 2013). Single molecule, real time sequencing (i.e. third  
93 generation sequencing) by PacBio RS SMRT chip can generate longer sequences than the  
94 second generation sequencing, and is possible to make high-throughput sequencing of the  
95 full-length 16S rRNA genes (Mosher et al. 2013, Mosher et al. 2014). A previous study  
96 demonstrated the sequences obtained from high-throughput sequencing of full-length 16S  
97 rRNA genes of *Shewanella oneidensis* MRI by Pacific Biosciences RS II sequencer can be  
98 accurately assigned to the species level (>99 % accuracy) (Mosher et al. 2014). However,  
99 high-throughput sequencing of full-length 16S rRNA genes has not been used for the  
100 microbial community analysis in mixed cultures (e.g. anaerobic digestion) until now.

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



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

## 108 **2. Material and methods**

### 109 **2.1. HTLWW**

110 The HTLWW was obtained from a pilot-scale hydrothermal reactor with a volume of 80 L.  
111 3.0 kg of minced rice straw mixed with 47 kg of water were added into the reactor and then  
112 heated to 280 °C at 12.0 MPa for 30 min (Chen et al. 2015). The mixture was filtered by a  
113 300-mesh screen after HTL, and the filtrate was HTLWW.

114 HTLWW was then extracted by petroleum ether (PE), cyclohexane (CH), dichloromethane  
115 (DM) and ethyl acetate (EA) to separate parts of the organic components (Duan and Savage  
116 2011, Yang et al. 2014), and they were named as PE-HTLWW, CH-HTLWW, DM-HTLWW  
117 and EA-HTLWW, respectively. The raw HTLWW was named as R-HTLWW. For the  
118 extraction, 125 mL organic solvent was added to a 500 mL bottle, and 250 mL HTLWW was  
119 also added. The bottles were then capped tightly and shaken with the speed of 120 rpm for 10  
120 min by a shaker (Duan and Savage 2011). The mixture was then transferred to a funnel for  
121 the separation of organic solvents and HTLWW. The above procedure was repeated for the  
122 separated HTLWW for the second time extraction. The four samples PE-HTLWW,  
123 CH-HTLWW, DM-HTLWW and EA-HTLWW were then obtained. They were all placed in a  
124 refrigerator at -20 °C for further usage. Table 1 presents the COD values of the HTLWW  
125 samples and the saturated organic solvents in water.

### 126 **2.2. Biogas production potentials of HTLWW**

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

128 by various organic solvents. 118 mL serum bottles were used. 15 mL inoculum and 45 mL  
129 BA medium containing a certain amount of HTLWW were added to each bottle. The initial  
130 COD value of all the bottles were 0.75 g/L by adding different amounts of HTLWW to the  
131 BA medium. The pH value was adjusted to 7.5. All the bottles were flushed with N<sub>2</sub> for 5 min  
132 to remove oxygen, and then sealed with butyl rubber stoppers and aluminum screw caps. All  
133 the bottles were placed in an incubator with constant temperature 37 °C. The inoculum was  
134 obtained from an anaerobic reactor treating cassava stillage in an ethanol plant (Taicang,  
135 Suzhou, China). The bottles with only inoculum were used as control. All the experiments  
136 were done in triplicates.

### 137 **2.3. Semi-continuous experiments**

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

150 to have the same organic loading rate as reactor PE.

151 **2.4 High-throughput sequencing of full-length 16S rRNA genes and bioinformatic**  
152 **analysis**

153 Samples were obtained during the steady-states of both reactors. Total genomic DNA was  
154 extracted from each sample using QIAamp DNA Stool Mini Kit (QIAGEN, 51504). The  
155 quantity and purity of the extracted DNA were checked by Nanodrop 2000. PCR was then  
156 conducted with the primers 27F (AGAGTTTGATCCTGGCTCAG) and 1492R  
157 (GGTTACCTTGTTACGACTT) for bacteria and the primers 20F  
158 (TTCCGGTTGATCCYGCCRG) and 1492R for archaea (DeLong 1992). All PCR  
159 amplifications were performed using the Taq PCR Core Kit (QIAGEN) with 1 uL template  
160 DNA and 20 pmol of each primer. The PCR conditions for bacteria were: 95 °C for 5 min, 28  
161 cycles of three steps: 95 °C for 45 s, 55 °C for 1 min, and 68 °C for 2 min, followed by a final  
162 step at 68 °C for 7 min. The PCR conditions for archaea were: 95 °C for 2 min, 27 cycles of  
163 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  
164 72 °C for 7 min. The samples were sent out for sequencing in one cell of the Pacific  
165 Biosciences RS II platform combined with the P4/C2 chemistry. The obtained sequences  
166 were deposited into the European Nucleotide Archive (ENA) with accession number  
167 PRJEB14373. The onboard software provided on the Pacific Biosciences RS II sequencer  
168 was used to eliminate CCS (circular consensus sequences) with <99 % predicted accuracy.  
169 The low-quality sequences (no exact matches to the forward and reverse primers, and length  
170 <1300 bp) and chimeras were removed from the raw sequencing data by MOTHUR program.  
171 The numbers of high quality sequences were 7911 (R) and 9099 (PE) for bacteria with

172 average length of 1390 bp, 1667 (R) and 1905 (PE) for archaea with average length of 1450  
173 bp. The numbers of sequences were normalized to the same sequencing depths (7911  
174 sequences for bacteria and 1667 sequences for archaea) to facilitate the comparison between  
175 different samples. The sequences were clustered into operational taxonomic units (OTU) with  
176 cutoff 0.03. Rarefaction curves, Shannon diversity index, coverage were also analyzed by  
177 MOTHUR program. The sequences were phylogenetically assigned to taxonomic  
178 classifications by RDP Classifier with a confidence threshold of 80 %. RDP could only assign  
179 the sequences into genus level. In order to get species classification, all the sequences were  
180 aligned using BLASTN against NCBI 16S rRNA database with strict criteria (percentage  
181 identity at both 97 % and 98.65 %, and alignment length>1300 bp). Both 97 % and 98.65 %  
182 of percentage identity were proposed in previous studies for species identification (Kim et al.  
183 2014, Stackebrandt and Goebel 1994, Tindall et al. 2010). MEGAN software was then used  
184 to assign the sequences down to species level based on the BLASTN results (Huson et al.  
185 2007). The volumes of gases reported in the present study were at standard temperature and  
186 pressure.

## 187 **2.5. Analytical methods**

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

194 initially held at 60 °C for 2 min and heated to 250 °C and held there for 10 min. Helium was  
195 used as the carrier gas (1.0 mL/min). A NIST Mass Spectral Database was used for  
196 compound identification. HPLC was used to measure the organic acids in the HTLWW  
197 samples as described previously (Chen et al. 2015). LC-TOF-MS was used to provide a  
198 detailed overview of the organic compounds in the HTLWW samples. It was performed on a  
199 Waters ACQUITY UPLC system equipped with a binary solvent delivery manager and a  
200 sample manager, coupled with a Waters Micromas Q-TOF Premier Mass Spectrometer  
201 equipped with an electrospray interface. Acquity BEH C18 column (100 mm×2.1 mm i.d.,  
202 1.7 µm; Waters, Milford, USA) was maintained at 45 °C and eluted with gradient solvent  
203 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 %  
204 (v/v) formic acid) and A was aqueous formic acid (0.1% (v/v) formic acid). The wavelength  
205 was 280 nm and the injection volume 5.00 µl, column temperature was 50.0 °C. The source  
206 and desolvation temperature were 115 °C and 350 °C respectively. The UV-Vis spectrum  
207 was studied using absorptions at 254 wavelength, and the analysis was carried out using a  
208 double-beam UV-Vis spectrophotometer from Shimadzu (UV-1800). The molecular weight  
209 distributions of HTLWW before and after anaerobic digestion were determined by a GFC  
210 analyzer (LC-10ADVP, Shimadzu) according to a previous study (Wen et al. 2012).

### 211 **3. Results and discussion**

#### 212 **3.1. Biogas production potentials of HTLWW extracted by different organic solvents**

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

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

### 225 **3.2 Characterization of HTLWW extracted by different organic solvents**

226 The high methane yield of EA-HTLWW was mainly attributed to the degradation of EA  
227 rather than the organics in the HTLWW since EA contributed to more than 98% of the COD  
228 in EA-HTLWW (Table 1), and the high methane yield of EA itself was shown in Fig S1. The  
229 negligible methane yield of DM-HTLWW was due to the toxicity of DM to the methanogens  
230 since no methane was produced when DM alone was used (Fig S1), and the toxicity of DM to  
231 methanogens was also reported in previous studies (Kim et al. 1996, McBride and Wolfe  
232 1971). The above results showed that both EA and DM were not suitable as organic solvents  
233 since they would increase the difficulty of the subsequent utilization of HTLWW. Both PE  
234 and CH had low solubility as demonstrated by their contribution to the total COD in Table 1.  
235 Higher methane yield was obtained from PE-HTLWW compared to CH-HTLWW and  
236 R-HTLWW, which indicated that PE might have extracted more organics that are difficult to  
237 be biodegraded and thereby improved the biodegradability of HTLWW.

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

### 256 **3.3 Biogas production from R-HTLWW and PE-HTLWW in ASBR**

257 The two reactors were operated for around 100 days until steady-states were achieved (Fig  
258 S3). The methane yield (153 mL/g COD) of R-HTLWW was significantly higher than that  
259 (218 mL/g COD) of PE-HTLWW ( $P < 0.01$ , ANOVA). The higher methane yield from

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

### 266 **3.4 Degradation of organic compounds in ASBR**

267 The UV-VIS (Fig S4) of R-effluent and PE-effluent spectral absorption decreased compared  
268 to R-HTLWW and PE-HTLWW, respectively, which was related with the degradation of  
269 organic compounds in the anaerobic reactors. However, the absorption between 210-250 nm  
270 and 260-300 nm of R-effluent and PE-effluent suggested that ketones and phenols were not  
271 fully degraded in the anaerobic reactors (Cheng et al. 2016), which could resulted in the  
272 lower methane yield of both R-HTLWW and PE-HTLWW compared to the theoretical value  
273 (350 mL/gCOD).

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



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

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

304 **3.5 Microbial community compositions as revealed by high-throughput sequencing of**  
305 **full-length 16S rRNA genes**

306 The samples obtained from the continuous reactors were then used for microbial community  
307 analysis. The rarefaction curves of all the samples at 0.03 distance is shown in Fig S5. The  
308 curves of bacteria and archaea were overlapped for both samples, and it indicated samples R  
309 and PE had similar microbial richness, which was also reflected by the similar OTU numbers  
310 (Bacteria, around 1500 for both samples; Archaea, around 210 for both samples) (Table S3).  
311 The results showed that PE extraction of HTLWW did not have obvious effects on the  
312 microbial community richness. It should be noted that the sequencing depths for both bacteria  
313 (7911) and archaea (1167) were still not enough to cover the whole microbial diversity since  
314 plateaus were not achieved for all the rarefaction curves. However, the coverage values for  
315 bacteria (>86%) and archaea (>90%) indicated that most common OTUs were detected. The  
316 coverage values were relatively lower compared to previous studies (e.g. coverage value 97.4%  
317 with sequencing depth 50000 for bacteria (Luo et al. 2013), coverage value 98.7% with  
318 sequencing depth 63699 for bacteria (Pan et al. 2015)), which was mainly due to the  
319 sequencing depths was relatively lower in our study. However, it should be noted all the  
320 above mentioned studies were based on high-throughput sequencing of partial 16S rRNA  
321 genes (less than 500 bp). The Shannon diversity index provides both species richness and the  
322 evenness of the species in the microbial community (Lu et al. 2012). Similar with the  
323 microbial richness, the microbial diversities were not affected by PE extraction of HTLWW  
324 for both bacteria (around 5.44) and archaea (around 3.3). The higher OTU numbers and  
325 Shannon diversity of bacteria compared to archaea were consistent with previous studies

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

328 The taxonomic classification of bacterial sequences by RDP classifier is shown in Fig 5(A).  
329 The similar taxonomic distribution in phylum, class and genus levels were observed for R  
330 and PE, further indicating PE extraction did not affect the bacterial communities. It could be  
331 due to that PE might only extract unbiodegradable organic compounds and therefore the  
332 degraded organic compounds in both reactors R and PE were similar. *Firmicutes*,  
333 *Synergistetes*, *Chloroflexi*, and *Bacteroidetes* were dominant phyla, and their dominance in  
334 mesophilic anaerobic reactors were also reported previously (Luo et al. 2016a, Sundberg et al.  
335 2013). Although *Thermotogae* had high relative abundance, its dominance was mainly found  
336 in thermophilic anaerobic reactors (Shi et al. 2013). Genus level identification indicated  
337 *Thermotogae* were mainly composed of *Mesotoga*, which was recently reported to be the  
338 only mesophilic genus (Nesbø et al. 2012). *Mesotoga* was reported to use lactic acid and its  
339 dominance might be related with the degradation of lactic acid as seen in Table 2. *Clostridia*  
340 and *Synergistia* were the dominant classes in phylum *Firmicutes* and *Synergistetes*,  
341 respectively, and they were known as syntrophic partners together with hydrogenotrophic  
342 methanogens for the efficient degradation of lactic acid and VFAs (Li et al. 2016). Their  
343 dominances were most probably related with the high concentrations of lactic acid and VFAs  
344 in HTLWW (Table 2). The relative abundances of *Anaerolineae* and *Bacteroidia* were  
345 between 7-9 % in both samples, and they were capable of hydrolysis and fermentation of  
346 carbohydrates to VFAs (Narihiro and Sekiguchi 2007, Robert et al. 2007) , however, the  
347 carbohydrates were not detected in our study (data not shown), which indicated that their

348 presence might be related with the degradation of other organics. The genus level  
349 classification showed that higher percentages (around 40 %) of sequences were unclassified,  
350 which was consistent with previous studies (Lu et al. 2012, Luo et al. 2013), and it could be  
351 attributed to that most of biogas reactor's communities are still uncharacterized (Bassani et al.  
352 2015). The dominant genus were *Syntrophobotulus*, *Mesotoga*, and *T78*. *Syntrophobotulus*  
353 *glycolicus* is currently the only known member of the genus *Syntrophobotulus*, however, it  
354 can only degrade glyoxylate (Yin et al. 2010), which was not detected in our study. Further  
355 species level identification did not detected *Syntrophobotulus glycolicus* (Table 3), and it  
356 indicated the genus *Syntrophobotulus* might contain unknown species with different  
357 metabolic potentials, which deserves further investigation. The role of *Mesotoga* was  
358 mentioned previously for the utilization of lactic acid, while the exact role of *T78* was still  
359 unknown (Goux et al. 2015).

360 Species level identification of full-length 16S rRNA gene sequences would provide more  
361 information on the microbial compositions and their metabolic potentials. Table 3  
362 summarized the identified bacterial species. At 97 % similarity, the sequences assigned to  
363 species level were 5.6 % and 5.1 % of the total sequences for R and PE, respectively.  
364 However, increased sequences (9.9% for R and 8.8% for PE) assigned to species level were  
365 obtained at 98.65 % similarity. It would be expected less sequences would be assigned to  
366 species level with more critical criteria. The higher sequences assigned at 98.65% similarity  
367 was attributed to the algorithm (lowest common ancestor) used by MEGAN (Huson et al.  
368 2007). For instance, one sequence might match two or more species in NCBI 16S rRNA  
369 genes database at 97 % similarity, therefore MEGAN could not assign the sequence to

370 species level. However, the matched species might decrease to one at 98.65% similarity, and  
371 therefore it could be assigned to species level. Fig S6 shows that 550 sequences were  
372 assigned to the genus *Mesotoga*, however, only 172 sequences were further assigned to  
373 species *Mesotoga infera* and *Mesotoga prima* at 97 % similarity, while 488 sequences were  
374 assigned to the genus *Mesotoga* at 98.65 % similarity and all of the sequences were further  
375 assigned to species level (Fig S8). The above results indicated that 97 % similarity was not  
376 enough to make species level identification. Although more sequences were assigned to  
377 species level at 98.65 %, still the genus *Trichococcus* was not further assigned to species  
378 level (Fig S8 and S12). The sequences belonging to *Trichococcus* (Fig S12) were also  
379 extracted, and it was found that all the sequences had more than one match to the species in  
380 NCBI 16S rRNA genes database at 98.65 % similarity (Table S4). 98.65 % was previously  
381 proposed as the threshold for differentiating two species based on the analysis of 6787  
382 genomes belonging to 1738 species (Kim et al. 2014). However, 98.65 % was not the optimal  
383 value in our study since microbial community in anaerobic reactor was more diverse. It  
384 should be noted that 98.65 % was still suitable for the species level identification of  
385 sequences belonging to most genus except *Trichococcus* (Fig S8 and S12). As shown in Table  
386 2, lactic acid and VFAs were well degraded during anaerobic digestion, and their degradation  
387 could be correlated with the several known species as shown in Table 3. *Mesotoga infera*,  
388 *Mesotoga prima*, and *Petrimonas sulfuriphila* were reported to use lactic acid as carbon  
389 source (Ben Hania et al. 2015, Grabowski et al. 2005). *Syntrophobacter sulfatireducens* were  
390 known as propionate-oxidizing bacteria (Chen et al. 2005). *Syntrophomonas wolfei*,  
391 *Syntrophus aciditrophicus* and *Syntrophus buswellii* were demonstrated to be able to degrade

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

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

414 hydrogenotrophic methanogenesis genus, which was also found to be dominant in other  
415 biogas reactors (Jaenicke et al. 2011, Krause et al. 2008). The species level identification by  
416 MEGAN showed that 40.9 % and 47.9 % of the sequences were assigned to species level at  
417 98.65 % similarity, which was much higher than that (<10 %) for bacteria. It could be due to  
418 the higher diversity of bacteria compared to archaea as seen in Table S3 and as reported in  
419 previous studies (Luo et al. 2013, Zhang et al. 2009), which resulted in the more  
420 uncharacterized species in bacteria than that in archaea. The overwhelming majority of the  
421 sequences were assigned to *Methanosaeta concilii*. Although more than 200 sequences were  
422 assigned to the genus *Methanosarcina* (Fig S9 and S13), only very few sequences were  
423 further classified down to species level. The results further indicated that the 98.65 %  
424 threshold for differentiating two species was not fully appropriate for all the archaea genus.

### 425 **3.6 Outlook**

426 The present study showed that HTLWW contains relatively higher amount of  
427 unbiodegradable organic compounds (e.g. phenols and other high MW (>1000) organic  
428 compounds), which were still left in HTLWW after anaerobic digestion. Therefore, further  
429 studies via aerobic biodegradation or chemical oxidation should be conducted to remove the  
430 residual organic compounds before discharging to the environment (Jang et al. 2015, Moreira  
431 et al. 2015). In addition, the usage of catalysis and changes of the HTL conditions also  
432 deserves further investigation in order to decrease the formation of unbiodegradable organic  
433 compounds without affecting the bio-crude production (Anastasakis and Ross 2011, Tekin  
434 and Karagöz 2013). For the first time, the third generation sequencing by PacBio RS SMRT  
435 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., Kougiyas, 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., Fadhlouli, 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 digester treating phosphogypsum in Tunisia. *Standards in*  
487 *Genomic Sciences* 10.
- 488 Campanaro, S., Treu, L., Kougiyas, 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 pyrolytic. *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

- 524 benzoate in syntrophic association with hydrogen-using microorganisms. Archives of  
525 Microbiology 171(2), 107-114.
- 526 Jaenicke, S., Ander, C., Bekel, T., Bisdorf, R., Droge, M., Gartemann, K.-H., Junemann, S.,  
527 Kaiser, O., Krause, L., Tille, F., Zakrzewski, F., Puhler, A., Schluter, A. and Goesmann, A.  
528 (2011) Comparative and joint analysis of two metagenomic datasets from a biogas fermenter  
529 obtained by 454-pyrosequencing. Plos One 6, 1.
- 530 Jang, H.M., Ha, J.H., Park, J.M., Kim, M.-S. and Sommer, S.G. (2015) Comprehensive  
531 microbial analysis of combined mesophilic anaerobic–thermophilic aerobic process treating  
532 high-strength food wastewater. Water Research 73, 291-303.
- 533 Jeris, J.S. and McCarty, P.L. (1965) The Biochemistry of Methane Fermentation Using  $\text{C}^{14}$   
534 Tracers. Journal of Water Pollution Control Federation 37(2), 178-192.
- 535 Karakashev, D., Batstone, D.J. and Angelidaki, I. (2005) Influence of environmental  
536 conditions on methanogenic compositions in anaerobic biogas reactors. Applied and  
537 Environmental Microbiology 71(1), 331-338.
- 538 Karlsson, A., Ejlertsson, J. and Svensson, B.H. (2000)  $\text{CO}_2$ -dependent fermentation of phenol  
539 to acetate, butyrate and benzoate by an anaerobic, pasteurised culture. Archives of  
540 Microbiology 173(5-6), 398-402.
- 541 Kim, I.S., Young, J.C. and Tabak, H.H. (1996) Impact of chloroanilines on hydrogenotrophic  
542 methanogenesis in ethanol-enriched cultures. Water Research 30(3), 601-612.
- 543 Kim, M., Oh, H.-S., Park, S.-C. and Chun, J. (2014) Towards a taxonomic coherence between  
544 average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation  
545 of prokaryotes. International Journal of Systematic and Evolutionary Microbiology 64,

- 546 346-351.
- 547 Knoll, G. and Winter, J. (1989) Degradation of phenol via carboxylation to benzoate by a  
548 defined, obligate syntrophic consortium of anaerobic-bacteria. *Applied Microbiology and*  
549 *Biotechnology* 30(3), 318-324.
- 550 Krause, L., Diaz, N.N., Edwards, R.A., Gartemann, K.-H., Krömeke, H., Neuweger, H.,  
551 Pühler, A., Runte, K.J., Schlüter, A., Stoye, J., Szczepanowski, R., Tauch, A. and Goesmann,  
552 A. (2008) Taxonomic composition and gene content of a methane-producing microbial  
553 community isolated from a biogas reactor. *Journal of Biotechnology* 136(1–2), 91-101.
- 554 Kumagai, S., Yamada, N., Sakaki, T. and Hayashi, N. (2007) Biomass: Characteristics of  
555 hydrothermal decomposition and saccharification of various lignocellulosic biomass and  
556 enzymatic saccharification of the obtained hydrothermal-residue. *Journal of the Japan*  
557 *Institute of Energy* 86(86), 712-717.
- 558 Leng, L., Yuan, X., Chen, X., Huang, H., Wang, H., Li, H., Zhu, R., Li, S. and Zeng, G. (2015)  
559 Characterization of liquefaction bio-oil from sewage sludge and its solubilization in diesel  
560 microemulsion. *Energy* 82, 218-228.
- 561 Li, L., He, Q., Ma, Y., Wang, X. and Peng, X. (2016) A mesophilic anaerobic digester for  
562 treating food waste: process stability and microbial community analysis using  
563 pyrosequencing. *Microbial Cell Factories* 15(1), 1-11.
- 564 Liu, Y., Wan, J., Han, S., Zhang, S. and Luo, G. (2016) Selective conversion of carbon  
565 monoxide to hydrogen by anaerobic mixed culture. *Bioresource Technology* 202, 1-7.
- 566 Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J. and  
567 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<sub>2</sub> 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., Adebisuyi, 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 Öztürk, 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.

678 Applied Energy 137, 183-192.

679

680

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

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

Name	R-HTLWW	PE-HTLW W	CH-HTL WW	DM-HTLW W	EA-HTL WW	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	–	–	–	–

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

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

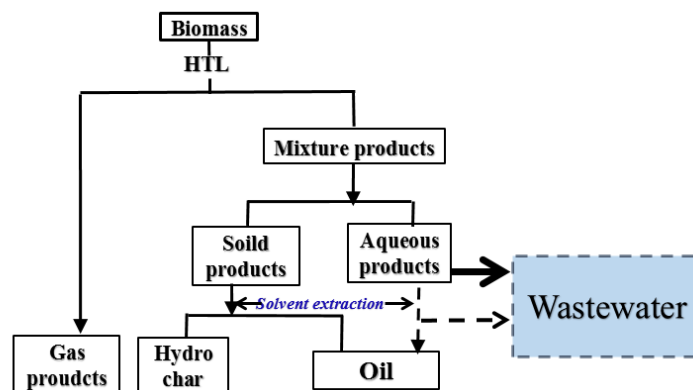


Fig 1 Hydrothermal liquefaction process

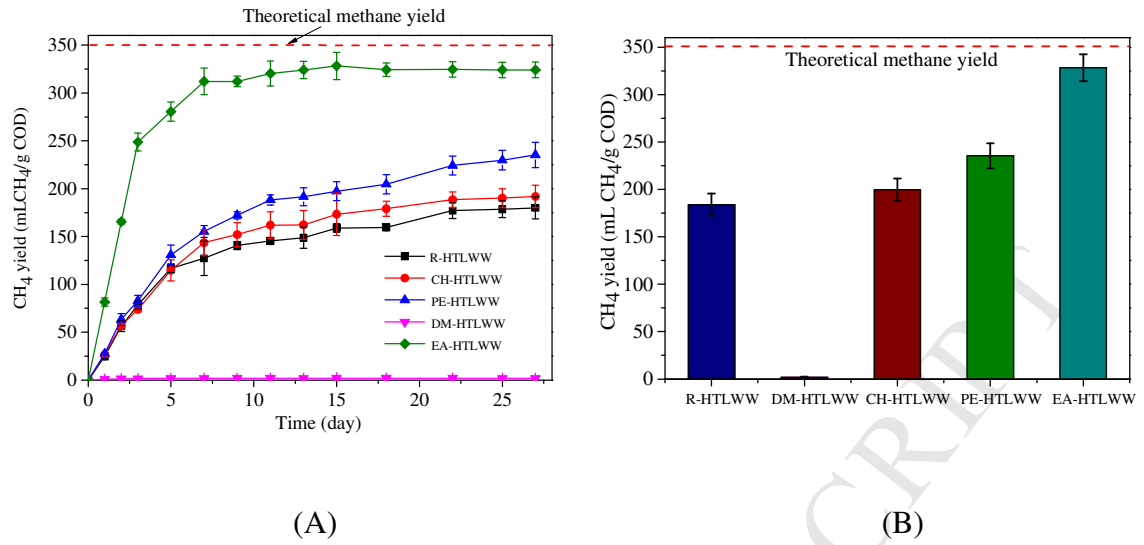


Fig 2 Biogas production potentials of HTLWW (A) time courses of methane production (B) methane yields



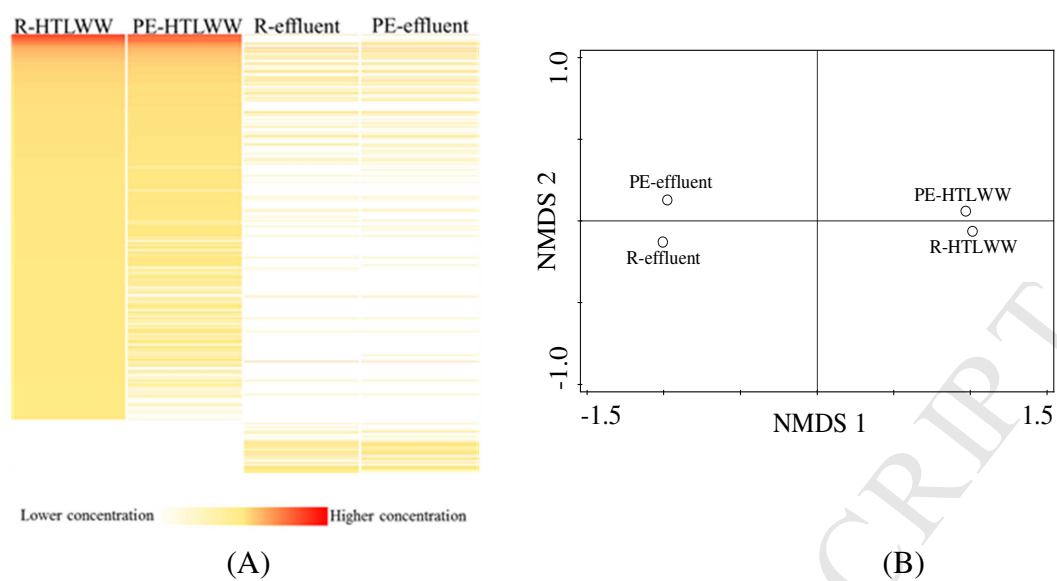


Fig 3 Heatmap (A) and NMDS (B) analysis of the samples based on LC-TOF-MS results

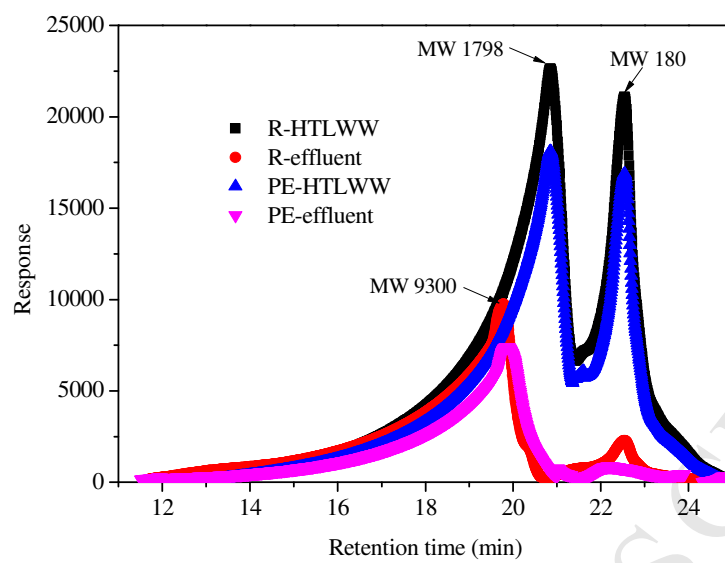


Fig 4 GFC analysis of the samples

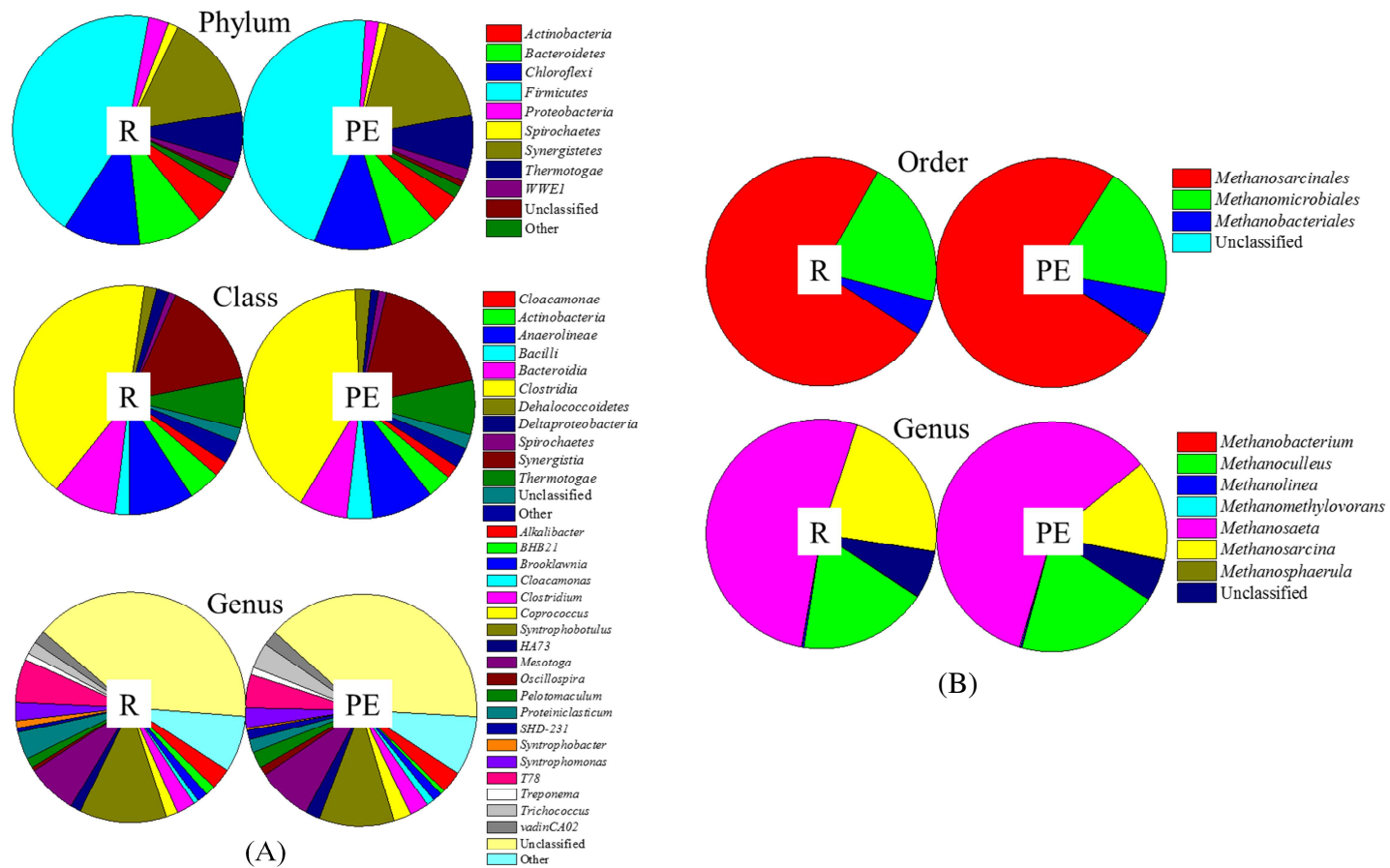


Fig 5 Taxonomic classification of bacteria (A) and archaea (B) sequences based on the high-throughput sequencing of full-length 16S rRNA genes

**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