

This is a repository copy of *Insights into microbial community structure and diversity in oil palm waste compost*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/151556/>

Version: Accepted Version

Article:

Khalid, Nurshafika Abd, Rajandas, Heera, Parimannan, Sivachandran et al. (5 more authors) (2019) Insights into microbial community structure and diversity in oil palm waste compost. 3 Biotech. 364. ISSN 2190-572X

<https://doi.org/10.1007/s13205-019-1892-4>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

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.

1 **Insights into microbial community structure and diversity in oil palm waste**
2 **compost**

3 Nurshafika Abd Khalid¹, Heera Rajandas^{2**}, Sivachandran Parimannan², Laurence J. Croft⁴,
4 Stella Loke³, Chun Shiong Chong¹, Neil C. Bruce⁵, Adibah Yahya^{1*}

5 ^[1] *Biorefinery Technology Laboratory, Department of Biosciences, Faculty of Science, 81310*
6 *Universiti Teknologi Malaysia, Malaysia*

7 ^[2] *Centre of Excellence for Omics-Driven Computational Biodiscovery (COMBio), Faculty of*
8 *Applied Sciences, AIMST University 08100 Bedong, Kedah, Malaysia*

9 ^[3] *School of Life and Environmental Sciences, Building M, 221 Burwood Hwy, Burwood VIC*
10 *3125, Deakin University, Burwood Campus, Australia*

11 ^[4] *Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin*
12 *University, Geelong, Victoria 3220, Australia*

13 ^[5] *Centre for Novel Agricultural Products, Department of Biology, University of York,*
14 *Wentworth Way, York, YO10 5DD, United Kingdom*

15 *Corresponding author: Tel: +6 07-5557540

16 Email address: adibah@fbb.utm.my (Adibah Y.)

17 **Second corresponding author: Tel: +6 04-4298159

18 Email address: heraadaas@gmail.com (Heraa R.)

19

20 **Acknowledgements** The authors would like to thank Jamal Al Deen Lawton for compost
21 advice. We gratefully acknowledge oil palm mill Felda Maokil, Labis, Johor for providing us
22 with compost and POME samples. This work was financially supported Ministry of Education
23 Malaysia and Biotechnology and Biological Sciences Research Council (BBSRC) United
24 Kingdom under program of United Kingdom-Southeast Asia Newton Ungku Omar Fund (UK-
25 SEA-NUOF) with project number 4B297 and BB/P027717/1.

26 **Abstract**

27 Empty fruit bunch (EFB) and palm oil mill effluent (POME) are the major wastes generated by the oil
28 palm industry in Malaysia. The practice of EFB and POME digester sludge co-composting has shown
29 positive results, both in mitigating otherwise environmentally damaging waste streams and in producing
30 a useful product (compost) from these streams. In this study, the bacterial ecosystems of 12 week-old
31 EFB-POME co-compost and POME biogas sludge from Felda Maokil, Johor were analysed using 16S
32 metagenome sequencing. Over 10 phyla were detected with Chloroflexi being the predominant phylum,
33 representing approximately 53% of compost and 23% of the POME microbiome reads. The main
34 bacterial lineage found in compost and POME was Anaerolinaceae (Chloroflexi) with 30% and 18% of
35 the total gene fragments, respectively. The significant differences between compost and POME
36 communities were abundances of *Syntrophobacter*, *Sulfuricurvum*, and *Coprococcus*. No methanogens
37 were identified due to the bias of general 16S primers to eubacteria. The preponderance of anaerobic
38 species in the compost, and high abundance of secondary metabolite fermenting bacteria is due to an
39 extended composting time, with anaerobic collapse of the pile in the tropical heat. Predictive functional
40 profiles of the metagenomes using 16S rRNA marker genes suggest the presence of enzymes involved
41 in polysaccharide degradation such as glucoamylase, endoglucanase, arabinofuranosidase, all of which
42 were strongly active in POME. Eubacterial species associated with cellulytic methanogenesis were
43 present in both samples.

44

45 **Keyword** Oil palm empty fruit bunch. Palm oil mill effluent. Compost. Metagenomics. Microbial
46 diversity.

47

48 **1.0 Introduction**

49 **The Malaysian** oil palm industry **is** growing rapidly and Malaysia has become the second largest
50 producer of palm oil after Indonesia (MPOB 2017). Approximately 5.74 Mha of Malaysia's land area
51 was covered with oil palm plantations in 2016 (MPOB 2017) which produced 17,320,000 tonnes of
52 palm oil. This in turn generated a large amount of oil palm derived waste. The oil palm industry
53 produces millions of tonnes of oil palm biomass, especially empty fruit bunch (EFB). There is 1 kg of

54 biomass, such as empty fruit bunch (EFB), palm kernel shell (PKS) and mesocarp fibre (MF), generated
55 for each kg of oil palm extracted (Sulaiman et al. 2011). Conversion of organic waste such as EFB into
56 usable horticultural by-products has been found to be the most efficient way to reuse this raw waste
57 material (Siddiquee et al. 2017).

58 EFB and palm oil mill effluent (POME) are the most abundant waste produced in oil palm
59 mills. At present, EFB and POME have been used as raw materials for co-composting and the resulting
60 co-compost has been used in the oil palm plantations at Felda Maokil. In this case, POME was used to
61 provide moisture to the compost. Composting is one of the most efficient solutions for sustainable
62 management of organic waste, it is an aerobic process that effectively converts cellulosic organic waste
63 into a nutrient-rich organic amendment for agricultural application (Neher et al. 2013). The conversion
64 of organic waste to compost is carried out by a successive microbial community combining both
65 mesophilic and thermophilic activities (Krishnan et al. 2017). However, the microbial community of
66 end product compost in the tropics has not been well characterized.

67 Understanding the microbial diversity of compost systems is important in order to produce high
68 quality compost and determine its effectiveness (Krishnan et al. 2017). Most studies that have explored
69 this rich ecosystem have utilized culture-based methods (Ryckeboer et al. 2003; Ahmad et al. 2007;
70 Vishan et al. 2017). But, culture-based methods are only useful for identifying less than 1% of the total
71 microbial diversity, as the majority of microorganisms are unculturable under standard media and
72 aerobic growth conditions (Handelsman 2004; Ito et al. 2018). The advent of Next-Generation
73 Sequencing (NGS) and metagenomics has opened an avenue to perform comprehensive studies to
74 characterize the total microbial diversity using a culture-independent method. Metagenomics is an
75 alternative that has been widely applied over the last few years (Fernández-Arrojo et al. 2010).

76 Metagenomics refers to the direct isolation of DNA from an environmental sample
77 (Handelsman 2004). There are two different approaches, amplicon sequencing and shotgun
78 metagenomics (Escobar-Zepeda et al. 2015). Amplicon sequencing targets specific regions of DNA
79 from communities by amplifying specific regions using taxonomical informative primer targets such as
80 intergenic transcribed spacers (ITS) and the large ribosomal subunit (LSU) for eukaryotes and the 16S
81 rRNA gene for prokaryotes (Sharpton 2014). Shotgun metagenomics randomly sequences all DNA

82 from a community, which produces a less biased assessment of species abundance but at greater cost.
83 Metagenomics has been commonly used in large and complex samples containing organisms from
84 different life domains or where less bias is required. 16S amplicon sequencing, or metaprofiling
85 (Escobar-Zepeda et al. 2015), is currently the most cost-effective method for DNA library preparation
86 in conjunction with sequencing by platforms such as the Illumina MiSeq. This approach has been widely
87 utilized, not only for studying resident microbiota in wastewater and compost (Krishnan et al. 2017;
88 Wang et al. 2016), but also for studying soil samples (Yan et al. 2016), hot springs (Chan et al. 2015),
89 termite gut (Chew et al. 2018), faecal samples (Costea et al. 2017) and many others.

90 In the present study, the microbial community of the EFB-POME co-compost and POME
91 biogas sludge has been studied using the culture independent 16S amplicon sequencing approach.
92 Metagenomes from the EFB-POME co-compost and POME were directly isolated from the samples
93 without any microorganism cultivation. The V3-V4 regions of prokaryotic 16S rRNA genes were
94 amplified from the metagenome and directly sequenced using Illumina's MiSeq platform. The detailed
95 information on the microbial residents will support further research to improve the duration of the
96 composting process and the quality of final compost by addition of specified microbial species.

97 **2.0 Materials and Methods**

98 **2.1 Collection of samples**

99 The 12 weeks-old EFB-POME co-compost and POME were obtained from Felda Maokil,
100 Labis, Johor (2°17'09.6"N 102°59'37.7"E). A 1 m height compost pile was made with the ratio of 40
101 ton EFB: 120 ton POME, turned over every 3 days to provide aeration. The composting was done by
102 first laying down the EFB on the ground, followed by the addition of POME to the EFB pile at three
103 day intervals to maintain a final moisture content of 65-75% as it is a solid state aerobic fermentation
104 process. A total of 1kg of 12 weeks-old compost was randomly sampled at a depth of 0.5 m inside the
105 compost pile. For POME, 2 litres was collected directly from the anaerobic digester (AD) effluent. The
106 POME and compost samples were collected in sterile containers and stored at 4 °C for further studies.

107 **2.2 Total DNA extraction**

108 Extraction of total DNA from EFB-POME co-compost was done with a modified Griffiths
109 protocol using NucleoSpin® Soil kit (Griffiths et al. 2000; Alessi et al. 2017), while the total DNA from

110 POME was extracted using NucleoSpin® Soil kit (Macherey-Nagel, Germany) based on the
111 manufacturer's protocol (Verma and Satyanarayana 2011). The detailed methods are outlined below.

112 2.2.1 Method 1: Modified Griffiths protocol

113 Half a gram of EFB-POME co-compost was transferred into a microcentrifuge tube containing
114 1 g of garnet beads (OMNI International, USA), 500 µL SL2 lysis buffer (Macherey-Nagel, Germany),
115 500 µL 10% CTAB, 100 µL 1mg/mL lysozyme and 214.3 µL enhancer SX (Macherey-Nagel,
116 Germany). After mixing, the samples were disrupted using Bead Ruptor 4 (OMNI International, USA)
117 for 3 minutes at level 3. The aqueous phase was mixed with inhibitor removal solution SL3 buffer
118 (Macherey-Nagel, Germany) and incubated at 4 °C for 5 min before centrifuging using a NucleoSpin®
119 Inhibitor Removal column (Macherey-Nagel, Germany) to remove any impurities like humic acid and
120 other PCR inhibitors. Following this, equal volumes of phenol: chloroform (1:1) was added to the eluted
121 aqueous phase and separated by centrifugation for 5 minutes at $13,300 \times g$. One-tenth volume of ice-
122 cold sodium acetate and 3 volumes of ice-cold absolute ethanol was added to the aqueous layer before
123 incubating at -80°C for 2 hours to precipitate the DNA. The resulting pellet was washed twice with
124 ice-cold 75% ethanol (Alessi et al. 2017) and the pellet was resuspended in 50 µL TE buffer. The DNA
125 was stored at -20 °C for further use.

126 2.2.2 Method 2: NucleoSpin® Soil kit

127 Two mL of POME was centrifuged at $4,500 \times g$ for 10 minutes and the resulting pellets were
128 transferred into a microcentrifuge containing 1 g of garnet beads (OMNI International). 300 µL SL2
129 buffer (Macherey-Nagel, Germany), 150 µL enhancer SX (Macherey-Nagel, Germany) and 100 µL 1
130 mg/mL lysozyme was added to the pellet. The pellet was homogenized for 3 minutes at level 3 using
131 Bead Ruptor 4 (OMNI International). 100 µL 1 mg/mL lysozyme was added to the homogenate and
132 incubated at 37 °C for 30 minutes. The inhibitor removal solution and column were used to remove any
133 impurities like humic acids or other PCR inhibitors. The binding solution was added to the supernatant
134 before loading onto the spin column. The column was centrifuged to bind the DNA to the column and
135 the column was washed twice with wash buffer provided in the kit. The DNA was finally eluted with
136 the elution buffer and stored at -20 °C.

137

138 2.3 DNA yield and purity determination

139 DNA concentration and quality of the total DNA extracts were determined using NanoDrop™
140 Lite spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and 1% w/v agarose gel
141 electrophoresis, respectively.

142 2.4 16S metagenomics library preparation, sequencing and data analysis

143 The 16S rRNA metagenome libraries were generated using purified total DNA as the template
144 in the polymerase chain reactions (PCR). The V3–V4 region of the 16S rRNA genes were amplified
145 using S-D-Bact-0341-b-S-17, 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGG
146 GNGGCWGCAG-3' and S-D-Bact-0785-a-A-21, 5'-GTCTCGTGGGCTCGGAGATGTGTATAAG
147 AGACAGGACTACHVGGGTATCTAATCC-3'). The underlined oligonucleotide are the Illumina
148 adapter overhang sequences, while the non-underlined sequences are locus-specific sequences which
149 were designed according to a reported primer pair (Klindworth et al. 2013) targeting conserved regions
150 within the V3 and V4 domains of prokaryotic 16S rRNA genes. The metagenome library was then pair-
151 end sequenced on the Illumina MiSeq platform (San Diego, CA, USA) using MiSeq Reagent Kit (v3)
152 for the longest length set to 2 x 300 base pairs (bp).

153 2.5 Bioinformatics analysis

154 The resulting sequencing data were analysed using **the** Mothur software package version 1.41.1
155 (Schloss et al. 2009). Firstly, the read pairs were merged to assemble them into contigs and contigs
156 exhibiting any ambiguous positions were subsequently removed. The sequences were then aligned to
157 the SILVA reference database (Release 132). Upon removal of unaligned sequences, the remaining
158 sequences were further filtered, dereplicated and de-noised before removal of chimeras. Sequences were
159 then classified based on **the** Greengenes database using naïve Bayesian classifier with bootstrap cut-off
160 of 80% before removal of sequences classified **as** unrelated lineages. Finally, the sequences were split
161 into bins based on taxonomy and clustered into OTUs using the *cluster.split* command.

162 The functional composition of EFB-POME compost and POME digester sludge was analysed
163 using **the** Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
164 (PICRUSt) bioinformatics software package (Langille et al. 2013). Firstly, PICRUSt analysis was done
165 by picking OTUs against the August 2013 Greengenes database release of gg_13_8_99 that contained

166 202,421 bacterial and archaeal sequences (McDonald et al. 2012). The OTU counts were normalized
167 and used for metagenome functional predictions with KEGG orthologs (KO). The output was further
168 analysed using the Statistical Analysis of Metagenome Profiles (STAMP) software package (Parks et
169 al. 2014).

170 **2.6 Data accessibility**

171 The raw sequencing data was deposited in the Sequence Read Archive (SRA) of the National
172 Centre for Biotechnology Information (NCBI) database under accession numbers SRR8181848 and
173 SRR8186815 for EFB-POME co-compost and POME, respectively.

174 **3.0 Results and Discussion**

175 **3.1 Microbial diversity analysis of EFB-POME co-compost and POME**

176 EFB-POME co-compost and POME were analysed for their microbial diversity using Illumina
177 MiSeq sequencing of V3-V4 region of the 16S rRNA genes. A total of 72,657 and 92,677 sequence
178 reads were generated from total DNA extracts of EFB-POME co-compost and POME, respectively.
179 After quality filtering and read merging, EFB-POME compost and POME generated 1,272 and 10,705
180 contigs, respectively. POME showed a higher number of OTUs, which indicates that POME has a
181 higher bacterial diversity compared to EFB-POME co-compost. Diversity coverage for each sample
182 was analysed using rarefaction analysis. The rarefaction curve illustrated in Fig.1 was calculated at 3%
183 dissimilarity. Rarefaction analysis shows the samples had reached saturation for genus level and higher
184 taxonomic level. The coverage for mature EFB-POME co-compost was 100%, while for POME was
185 99.4% with Simpson indices of 113.77 and 69, respectively.

186 The most abundant phylum present in both samples was Chloroflexi, which constituted 23% of
187 the total gene fragment abundance in POME and 53% in EFB-POME co-compost (Fig.2A). The
188 remaining phyla present in POME were Firmicutes (19%), Bacteroidetes (16%), Proteobacteria (16%)
189 and Synergistetes (3%), while in EFB-POME co-compost the phyla were Bacteroidetes (15%),
190 Firmicutes (12%), Proteobacteria (7%) and Actinobacteria (3%). According to Chandna et al. (2013),
191 the number of microbial species in early stages of composting depends on the initial substrate used and
192 the prevailing environmental conditions during the composting process, especially the temperature.
193 Composting can be divided into several phases based on the temperature. These include mesophilic,

194 thermophilic and maturing phases, during which different bacterial populations thrive. Chandna et al.
195 (2013) found that Firmicutes are abundant during mesophilic phase and decrease in the maturation
196 phase, while Actinobacteria are stable during mesophilic and thermophilic phases. Neher et al. (2013)
197 found Bacteroidetes dominated at the end of the thermophilic phase, whereas Proteobacteria was
198 dominant after the thermophilic phase. The predominance of anaerobes in the compost sample are
199 strongly suggestive of a secondary fermentation of the pile after collapse of all air spaces and exhaustion
200 of oxygen throughout the mass. The compost method was similar to that developed by Raabe at
201 Berkeley, which takes 18 days in a Mediterranean climate (Raabe 1981). After 84 days in the tropics,
202 the compost has further fermented. The 16S data for the EFB-POME compost is a glimpse of the
203 microbial structure of this mature compost post aerobic composting.

204 The major family that was observed in both samples was Anaerolinaceae, which represents
205 18% abundance in POME and 30% abundance in EFB-POME co-compost as shown in Fig.2B. These
206 OTUs represent a diversity of species, rather than a single dominant species. The other five major
207 families observed in POME were Syntrophaceae (9%), Syntrophomonadaceae (5%),
208 Porphyromonadaceae (5%), Tissierellaceae (3%), and Synergistaceae (3%), while in EFB-POME co-
209 compost, the other major families were Porphyromonadaceae (7%), Lachnospiraceae (3%),
210 Helicobacteraceae (2%), Ruminococcaceae (2%), and Tissierellaceae (2%).

211 The most abundant genus observed in POME was *Syntrophus*, which accounted for 9% of the
212 total gene fragments (Fig.2C). The remaining genus observed in POME include *Syntrophomonas* (5%),
213 *Sedimentibacter* (3%), *Gracilibacter* (3%), *Solibacillus* (3%). Most of the bacteria found in POME
214 digester sludge were anaerobic as methanogenesis is an anaerobic process. In a previous study by
215 Krishnan et al. (2017), *Parabacteroides*, *Levilinea*, *Smithella*, *Prolixibacter* and *Bellilinea* were
216 identified as the common genera found in POME. *Bellilinea* was also found in DNA extracts from
217 POME in the present study. However, this genus represented a small majority, which only accounted
218 for 1% of the community. In the EFB-POME co-compost, on the other hand, *Coprococcus* was
219 identified as the most common genera accounting for 3% of the gene fragments, followed by
220 *Sulfuricurvum* (2%), *Sedimentibacter* (2%) and *Proteiniphilium* (2%). *Coprococcus* are anaerobic
221 bacteria and a major bacterial taxa in the rumen microbiota of some ruminants (Jia et al. 2016). The

222 bacteria from **the** *Coprococcus* family are commonly involved in the degradation of cellulosic materials
223 (Moore et al. 2011) and can be found enriched in xylan based cultures (Jia et al. 2016) which are often
224 found in anaerobic cellulose digestion. Apart from these genera, *Steroidobacter*, *Nitriliruptor*,
225 *Anaeomyxobacter*, *Filomicrobium* and *Truepera* were also found inside lignocellulose biomass
226 compost by Krishnan et al. (2017).

227 The overall population in compost is illustrated in Fig.3A. The most common phyla found in
228 the EFB-POME co-compost was Chloroflexi (53% of the total gene fragments) and this phyla was
229 represented 100% of the total Chloroflexi by Anaerolinaceae family. The second most abundant phyla
230 was Bacteroidetes, which accounted for 15% of the total 16S rRNA gene fragments. The major family
231 in Bacteroidetes was Porphyromonadaceae comprising 47.5% of the total Bacteroidetes. *Petrimonas*
232 and *Proteiniphilum* were identified which represent 9.2% and 32.2% of the total Porphyromonadaceae
233 gene fragments. The remaining family in the Bacteroidetes was unclassified with 52.5% of the total
234 Bacteroidetes gene fragments. The third major phyla in compost was Firmicutes accounting for 12% of
235 the total gene fragments. Fig.3B shows that the major order was Clostridiales, which amounts to 89%
236 of the total Firmicutes gene fragments. The remaining were Bacillales and unclassified Firmicutes with
237 2.7% and 8.3%, respectively. Clostridiales were represented commonly by the family of
238 Lachnospiraceae (24.6% of the total Clostridiales gene fragments) followed by Tissierellaceae (23.8%),
239 Ruminococcaceae (23%), Syntrophomonadaceae (7.7%), Gracilibacteraceae (3.8%), Christenellaceae
240 (2.3%) and Clostridiaceae (1.5%) as shown in Fig.3B.

241 The overall bacterial population of POME **are** shown in Fig.4A. Similar to EFB-POME co-
242 compost, Cloroflexi was also identified as the dominant phyla (23% of the total gene fragments) which
243 was 100% represented (of the total Cloroflexi gene fragments) by **the** Anearolinaceae family.
244 Anearolinaceae **dominates** the population of POME and EFB-POME co-compost. Anearolinaceae is
245 anaerobic and involved in methanogenesis. This family of bacteria is indigenous in many environments
246 rich in oil and hydrocarbon (Liang et al. 2015) and associated with the anaerobic degradation of oil-
247 related compounds (Sutton et al. 2013). Anaerolinaceae has been reported as the predominant species
248 isolated from anaerobic digester systems and has a fermentative metabolism, utilizing carbohydrates
249 and proteinaceous carbon sources under anaerobic conditions (McIlroy et al. 2017; Sun et al. 2016;

250 Yamada et al. 2006). The absence of the archaean methanogens in the data is due to the known poor
251 ability of the standard 16S primers to amplify these organisms (Klindworth et al. 2013). As all known
252 methanogens are archaea, methanogenic archaea ecosystems are only served by specific 16S primers
253 such as S-D-Arch-0349-a-S-17 and S-D-Arch-0786-a-A-20 primer pair (Fischer et al. 2016). The
254 second common phyla identified in the POME population was Firmicutes with 12% of the total gene
255 fragments (Fig.4B). The major order observed from this phylum was Clostridiales with 77.3% of the
256 total Firmicutes gene fragments. The family of Syntrophomonadaceae was the most dominant,
257 accounting for 31.7% of the total Clostridiales gene fragments. The remaining families observed were
258 Gracilibacteraceae (22.5%), Tissierellaceae (21.7%), Ruminococcaceae (5.7%), Clostridiaceae (5.7%)
259 and Lachnospiraceae (3.1%) as shown in Fig.4B. The family of bacteria found in Clostridiales were
260 similar to the community found in co-compost, with slight variations in abundances as many bacteria
261 in this family are thermotolerant and are to survive the composting process.

262 3.2 Comparative analysis of EFB-POME co-compost and POME bacterial communities

263 Sequence data sets retrieved from EFB-POME co-compost and POME digester sludge were
264 compared. Fig.5 compared the relative abundance of the 12 major genus represented in both EFB-
265 POME compost and POME. There were fewer bacterial genera in POME than during the composting
266 process. The dominant bacteria in the compost were *Syntrophobacter*, *Sulfuricurvum* and *Coprococcus*.
267 There is limited evidence that these bacteria are able to produce compost, and in fact are anaerobes.
268 These organisms likely represent the secondary fermentation of the compost, once the pile had collapsed
269 and oxygen and easily metabolisable carbon had been exhausted. Their fermentative abilities are
270 directed to metabolites likely present after thermophilic composting of woody biomass. Bacteria that
271 did not survive aerobic composting but were found in the POME sludge were *Petrimonas*, *Syntrophus*,
272 *Treponema*, *Bellilinea*, *Solibacillus*, *Clostridium*, *Gracilibacter*, *Syntrophomonas*, and *Acholeplasma*.
273 Most of these bacteria are anaerobes and facultative anaerobes as POME is an anaerobic digester
274 effluent.

275 In this study, *Sulfuricurvum* was identified as the predominant bacteria in the mature compost.
276 *Sulfuricurvum* is chemolithoautotrophic and a sulphur-oxidizing bacterium, capable of thriving under
277 microaerobic and anaerobic conditions (Kodama and Watanabe 2004). The condition inside the EFB-

278 POME compost is facultative anaerobic which is therefore, favourable for the growth of *Sulfuricurvum*.
279 *Sulfuricurvum* has previously been identified in contaminated soil (Liu et al. 2015), river sediments (Liu
280 et al. 2018), underground crude-oil storage (Kodama and Watanabe 2004) and wastewater sludge
281 (Hatamoto et al. 2011). The presence of a chemolithoautotroph demonstrates how limited the nutrients
282 were and how mature the compost was. Liu et al. (2015) reported that the abundance of *Sulfuricurvum*
283 increases with higher moisture, since high moisture content is associated with low redox potential and
284 anaerobic environments (Brockett et al. 2012).

285 *Syntrophobacter* was also found in a higher relative abundance in EFB-POME compost
286 compared to POME. *Syntrophobacter* has the ability to degrade propionate, which is usually isolated
287 from methanogenic ecosystems (Boone and Bryant 1980), characteristic of the POME sludge. In
288 anaerobic digestion, the acetogenesis stage is predominantly acetogenic bacteria such as
289 *Syntrophobacter*, which converts fermentation products with more than two carbon atoms, alcohols and
290 aromatics fatty acids into acetate and hydrogen (Kangle et al. 2012). In this stage, the bacteria convert
291 products from the first phase (hydrolysis) to butyric acid, propionic acid, ethanol, acetic acid, carbon
292 dioxide and hydrogen (Nalo et al. 2014).

293 3.3 Predicted functional metagenome profiles

294 Metagenome functional prediction was carried out using Phylogenetic Investigation of
295 Communities by Reconstruction of Unobserved States (PICRUSt) analysis based on the Greengenes
296 16S rRNA database and KO. A bar graph was plotted to compare the abundance of the metabolic
297 features between the two samples as shown in Fig.6. From the result of this study, energy metabolism
298 was found to be highly represented in POME and compost community. The subfunctions in energy
299 metabolism included carbon fixation pathways in prokaryote, oxidative phosphorylation, nitrogen
300 metabolism, sulphur metabolism and methane metabolism. The proportion of sequences for the energy
301 metabolism is higher in POME sample compared to compost, as the secondary fermentation of the
302 compost in anaerobic conditions is relatively energy limited. For the xenobiotics biodegradation and
303 metabolism, the subfunctions benzoates degradation, bisphenol degradation, drug metabolism by
304 cytochrome P450, naphthalene degradation and polycyclic aromatic hydrocarbon degradation were
305 observed to have higher sequence proportion in the compost community compared to the POME

306 community. These functional classes are explained by secondary fermentation of the remaining
307 recalcitrant substrates in the now anoxic pile.

308 Carbohydrate metabolism such as nucleotide sugar, fructose and mannose, starch and sucrose
309 and butanoate were observed to be **slightly** higher in proportion in the POME community compared to
310 compost **due to** active anaerobic processing of the oil palm products **in POME**. The degradation of
311 cellulose and hemicellulose during the composting process can produce carbohydrates (Toledo et al.
312 2017). Those compounds are easily degradable substances, which are preferentially degraded by aerobic
313 eubacteria. Carbohydrate metabolism plays an important role in degradation of hemicellulose and
314 cellulose during the composting process (Wei et al. 2018). Furthermore, amino acids are sources of
315 energy and carbon for bacterial metabolism produced throughout the composting process (López-
316 González et al. 2015). Wu et al. (2017) suggest that a higher abundancy of bacteria with **active** amino
317 acid metabolism increases humic substance synthesis.

318 In order to determine the potential roles of microbial communities in the decomposition of plant
319 polymers, carbon degradation enzymes were identified and their presence is illustrated in Fig.7. The
320 enzymes include genes encoding alpha-amylase, glucoamylase and neopullulanase for starch
321 degradation; beta-glucanase, endoglucanase, and beta-glucosidase for cellulose degradation;
322 arabinofuranosidase and xylanase for hemicellulose degradation; and lastly, beta-hexosaminidase,
323 chitinase, and peptidoglycan hydrolase involved in degradation of chitins derived from fungal
324 decomposition of the plant mass. Apart from that, genes related **to** chemotaxis was also more abundant
325 in the POME sample as this was a liquid culture.

326 Due to the poor amplification of archaean 16S sequences however, methanogenesis is only
327 observed in a limited way in both samples. Limitations of 16S primers targeting the V3-V4 domains
328 are clear in this study, where the primary fermentation **was** not observed. To overcome this limitation
329 and obtain an unbiased view of the archaean diversity, shotgun **metagenomics** sequencing could be
330 employed.

331

332 **Conflict of interest**

333 On behalf of all authors, the corresponding author states that there is no conflict of interest.

334 **References**

- 335 1. Ahmad R, Jilani G, Arshad M, Zahir ZA, Khalid A (2007) Bio-conversion of organic wastes
336 for their recycling in agriculture: an overview of perspectives and prospects. *Annals of*
337 *Microbiology* 57 (4):471-479. doi:10.1007/bf03175343
- 338 2. Alessi AM, Bird SM, Bennett JP, Oates NC, Li Y, Dowle AA, Polikarpov I, Young JPW,
339 McQueen-Mason SJ, Bruce NC (2017) Revealing the insoluble metasecretome of
340 lignocellulose-degrading microbial communities. *Scientific Reports* 7 (1):2356.
341 doi:10.1038/s41598-017-02506-5
- 342 3. Boone DR, Bryant MP (1980) Propionate-Degrading Bacterium, *Syntrophobacter wolinii* sp.
343 nov. gen. nov., from Methanogenic Ecosystems. *Applied and environmental microbiology* 40
344 (3):626-632
- 345 4. Brockett BFT, Prescott CE, Grayston SJ (2012) Soil moisture is the major factor influencing
346 microbial community structure and enzyme activities across seven biogeoclimatic zones in
347 western Canada. *44* (1):9-20. doi:10.1016/j.soilbio.2011.09.003
- 348 5. Chan CS, Chan K-G, Tay Y-L, Chua Y-H, Goh KM (2015) Diversity of thermophiles in a
349 Malaysian hot spring determined using 16S rRNA and shotgun metagenome sequencing.
350 *Frontiers in Microbiology* 6:177. doi:10.3389/fmicb.2015.00177
- 351 6. Chandna P, Nain L, Singh S, Kuhad RC (2013) Assessment of bacterial diversity during
352 composting of agricultural byproducts. *BMC Microbiology* 13 (1):99. doi:10.1186/1471-
353 2180-13-99
- 354 7. Chew YM, Lye S, Md. Salleh M, Yahya A (2018) 16S rRNA metagenomic analysis of the
355 symbiotic community structures of bacteria in foregut, midgut, and hindgut of the wood-
356 feeding termite *Bulbitermes* sp. *Symbiosis* 76 (2):187-197. doi:10.1007/s13199-018-0544-5
- 357 8. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, Tramontano M, Driessen
358 M, Hercog R, Jung F-E, Kultima JR, Hayward MR, Coelho LP, Allen-Vercocoe E, Bertrand L,
359 Blaut M, Brown JRM, Carton T, Cools-Portier S, Daigneault M, Derrien M, Druesne A, de
360 Vos WM, Finlay BB, Flint HJ, Guarner F, Hattori M, Heilig H, Luna RA, van Hylckama
361 Vlieg J, Junick J, Klymiuk I, Langella P, Le Chatelier E, Mai V, Manichanh C, Martin JC,

- 362 Mery C, Morita H, O'Toole PW, Orvain C, Patil KR, Penders J, Persson S, Pons N, Popova
363 M, Salonen A, Saulnier D, Scott KP, Singh B, Slezak K, Veiga P, Versalovic J, Zhao L,
364 Zoetendal EG, Ehrlich SD, Dore J, Bork P (2017) Towards standards for human fecal sample
365 processing in metagenomic studies. *Nature Biotechnology* 35:1069. doi:10.1038/nbt.3960
366 <https://www.nature.com/articles/nbt.3960#supplementary-information>
- 367 9. Escobar-Zepeda A, Vera-Ponce de León A, Sanchez-Flores A (2015) The Road to
368 Metagenomics: From Microbiology to DNA Sequencing Technologies and Bioinformatics.
369 *Frontiers in Genetics* 6 (348). doi:10.3389/fgene.2015.00348
- 370 10. Fernández-Arrojo L, Guazzaroni M-E, López-Cortés N, Belouqui A, Ferrer M (2010)
371 Metagenomic era for biocatalyst identification. *Current Opinion in Biotechnology* 21 (6):725-
372 733. doi:<https://doi.org/10.1016/j.copbio.2010.09.006>
- 373 11. Fischer MA, Güllert S, Neulinger SC, Streit WR, Schmitz RA (2016) Evaluation of 16S
374 rRNA Gene Primer Pairs for Monitoring Microbial Community Structures Showed High
375 Reproducibility within and Low Comparability between Datasets Generated with Multiple
376 Archaeal and Bacterial Primer Pairs. *Frontiers in microbiology* 7:1297-1297.
377 doi:10.3389/fmicb.2016.01297
- 378 12. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ (2000) Rapid Method for Coextraction
379 of DNA and RNA from Natural Environments for Analysis of Ribosomal DNA- and rRNA-
380 Based Microbial Community Composition. *Applied and Environmental Microbiology* 66
381 (12):5488-5491
- 382 13. Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms.
383 *Microbiology and molecular biology reviews* 68 (4):669-685
- 384 14. Hatamoto M, Miyauchi T, Kindaichi T, Ozaki N, Ohashi A (2011) Dissolved methane
385 oxidation and competition for oxygen in down-flow hanging sponge reactor for post-
386 treatment of anaerobic wastewater treatment. *Bioresource technology* 102 (22):10299-10304.
387 doi:<https://doi.org/10.1016/j.biortech.2011.08.099>

- 388 15. Ito T, Sekizuka T, Kishi N, Yamashita A, Kuroda M (2018) Conventional culture methods
389 with commercially available media unveil the presence of novel culturable bacteria. *Gut*
390 *Microbes*:1-15. doi:10.1080/19490976.2018.1491265
- 391 16. Jia Y, Wilkins D, Lu H, Cai M, Lee PKH (2016) Long-Term Enrichment on Cellulose or
392 Xylan Causes Functional and Taxonomic Convergence of Microbial Communities from
393 Anaerobic Digesters. *Applied and Environmental Microbiology* 82 (5):1519-1529.
394 doi:10.1128/aem.03360-15
- 395 17. Kangle K, Kore S, Kore V, Kulkarni G (2012) Recent trends in anaerobic codigestion: a
396 review. *Universal Journal of Environmental Research and Technology* 2 (4):210-219
- 397 18. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2013)
398 Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-
399 generation sequencing-based diversity studies. *Nucleic Acids Research* 41 (1):e1-e1.
400 doi:10.1093/nar/gks808
- 401 19. Kodama Y, Watanabe K (2004) *Sulfuricurvum kujiense* gen. nov., sp. nov., a facultatively
402 anaerobic, chemolithoautotrophic, sulfur-oxidizing bacterium isolated from an underground
403 crude-oil storage cavity. *International Journal of Systematic and Evolutionary Microbiology*
404 54 (6):2297-2300. doi:doi:10.1099/ijs.0.63243-0
- 405 20. Krishnan Y, Bong CPC, Azman NF, Zakaria Z, Othman NA, Abdullah N, Ho CS, Lee CT,
406 Hansen SB, Hara H (2017) Co-composting of palm empty fruit bunch and palm oil mill
407 effluent: Microbial diversity and potential mitigation of greenhouse gas emission. *Journal of*
408 *Cleaner Production* 146:94-100. doi:https://doi.org/10.1016/j.jclepro.2016.08.118
- 409 21. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC,
410 Burkepille DE, Vega Thurber RL, Knight R, Beiko RG, Huttenhower C (2013) Predictive
411 functional profiling of microbial communities using 16S rRNA marker gene sequences.
412 *Nature Biotechnology* 31:814. doi:10.1038/nbt.2676
413 <https://www.nature.com/articles/nbt.2676#supplementary-information>
- 414 22. Liang B, Wang L-Y, Mbadinga SM, Liu J-F, Yang S-Z, Gu J-D, Mu B-Z (2015)
415 *Anaerolineaceae* and *Methanosaeta* turned to be the dominant microorganisms in alkanes-

- 416 dependent methanogenic culture after long-term of incubation. *AMB Express* 5 (1):117-117.
417 doi:10.1186/s13568-015-0117-4
- 418 23. Liu J, Chen X, Shu H-y, Lin X-r, Zhou Q-x, Bramryd T, Shu W-s, Huang L-n (2018)
419 Microbial community structure and function in sediments from e-waste contaminated rivers at
420 Guiyu area of China. *Environmental Pollution* 235:171-179.
421 doi:<https://doi.org/10.1016/j.envpol.2017.12.008>
- 422 24. Liu J, He X-x, Lin X-r, Chen W-c, Zhou Q-x, Shu W-s, Huang L-n (2015) Ecological Effects
423 of Combined Pollution Associated with E-Waste Recycling on the Composition and Diversity
424 of Soil Microbial Communities. *Environmental Science & Technology* 49 (11):6438-6447.
425 doi:10.1021/es5049804
- 426 25. López-González JA, Suárez-Estrella F, Vargas-García MC, López MJ, Jurado MM, Moreno J
427 (2015) Dynamics of bacterial microbiota during lignocellulosic waste composting: Studies
428 upon its structure, functionality and biodiversity. *Bioresource technology* 175:406-416.
429 doi:<https://doi.org/10.1016/j.biortech.2014.10.123>
- 430 26. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL,
431 Knight R, Hugenholtz P (2012) An improved Greengenes taxonomy with explicit ranks for
432 ecological and evolutionary analyses of bacteria and archaea. *The ISME journal* 6 (3):610-
433 618. doi:10.1038/ismej.2011.139
- 434 27. McIlroy SJ, Kirkegaard RH, Dueholm MS, Fernando E, Karst SM, Albertsen M, Nielsen PH
435 (2017) Culture-Independent Analyses Reveal Novel Anaerolineaceae as Abundant Primary
436 Fermenters in Anaerobic Digesters Treating Waste Activated Sludge. *Frontiers in*
437 *Microbiology* 8 (1134). doi:10.3389/fmicb.2017.01134
- 438 28. Moore AM, Munck C, Sommer MOA, Dantas G (2011) Functional metagenomic
439 investigations of the human intestinal microbiota. *Frontiers in microbiology* 2:188-188.
440 doi:10.3389/fmicb.2011.00188
- 441 29. MPOB MPOB (2017) Malaysian Oil Palm Industry Performance 2016 and Prospects for
442 2017. Malaysian Palm Oil Board (MPOB). Ministry of Plantation Industries & Commodities
443 (MPIC).

- 444 30. Nalo T, Tasing K, Kumar S, Bharti A (2014) Anaerobic digestion of municipal solid waste: a
445 critical analysis. *Int J Innov Res Sci Eng Technol* 3 (4):2347-6710
- 446 31. Neher DA, Weicht TR, Bates ST, Leff JW, Fierer N (2013) Changes in Bacterial and Fungal
447 Communities across Compost Recipes, Preparation Methods, and Composting Times. *PLOS*
448 *ONE* 8 (11):e79512. doi:10.1371/journal.pone.0079512
- 449 32. Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of
450 taxonomic and functional profiles. *Bioinformatics (Oxford, England)* 30 (21):3123-3124.
451 doi:10.1093/bioinformatics/btu494
- 452 33. Raabe RD (1981) *The Rapid Composting Method*. Cooperative Extension, University of
453 California,
- 454 34. Ryckeboer J, Mergaert J, Vaes K, Klammer S, De Clercq D, Coosemans J, Insam H, Swings J
455 (2003) A survey of bacteria and fungi occurring during composting and self-heating
456 processes, vol 53.
- 457 35. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,
458 Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber
459 CF (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported
460 Software for Describing and Comparing Microbial Communities. *Applied and Environmental*
461 *Microbiology* 75 (23):7537-7541. doi:10.1128/AEM.01541-09
- 462 36. Sharpton TJ (2014) An introduction to the analysis of shotgun metagenomic data. *Frontiers in*
463 *Plant Science* 5 (209). doi:10.3389/fpls.2014.00209
- 464 37. Siddiquee S, Shafawati SN, Naher L (2017) Effective composting of empty fruit bunches
465 using potential *Trichoderma* strains. *Biotechnology Reports* 13:1-7.
466 doi:<https://doi.org/10.1016/j.btre.2016.11.001>
- 467 38. Sulaiman F, Abdullah N, Gerhauser H, Shariff A (2011) An outlook of Malaysian energy, oil
468 palm industry and its utilization of wastes as useful resources. *Biomass and Bioenergy* 35
469 (9):3775-3786
- 470 39. Sun L, Toyonaga M, Ohashi A, Matsuura N, Tourlousse DM, Meng X-Y, Tamaki H, Hanada
471 S, Cruz R, Yamaguchi T, Sekiguchi Y (2016) Isolation and characterization of *Flexilinea*

- 472 flocculi gen. nov., sp. nov., a filamentous, anaerobic bacterium belonging to the class
473 Anaerolineae in the phylum Chloroflexi. *International Journal of Systematic and Evolutionary*
474 *Microbiology* 66 (2):988-996. doi:doi:10.1099/ijsem.0.000822
- 475 40. Sutton NB, Maphosa F, Morillo JA, Abu Al-Soud W, Langenhoff AAM, Grotenhuis T,
476 Rijnaarts HHM, Smidt H (2013) Impact of long-term diesel contamination on soil microbial
477 community structure. *Applied and environmental microbiology* 79 (2):619-630.
478 doi:10.1128/AEM.02747-12
- 479 41. Toledo M, Gutiérrez MC, Siles JA, García-Olmo J, Martín MA (2017) Chemometric analysis
480 and NIR spectroscopy to evaluate odorous impact during the composting of different raw
481 materials. *Journal of Cleaner Production* 167:154-162.
482 doi:<https://doi.org/10.1016/j.jclepro.2017.08.163>
- 483 42. Verma D, Satyanarayana T (2011) An Improved Protocol for DNA Extraction from Alkaline
484 Soil and Sediment Samples for Constructing Metagenomic Libraries. *Applied Biochemistry*
485 *and Biotechnology* 165 (2):454. doi:10.1007/s12010-011-9264-5
- 486 43. Vishan I, Sivaprakasam S, Kalamdhad A (2017) Isolation and identification of bacteria from
487 rotary drum compost of water hyacinth. *International Journal of Recycling of Organic Waste*
488 *in Agriculture* 6 (3):245-253. doi:10.1007/s40093-017-0172-8
- 489 44. Wang C, Dong D, Wang H, Müller K, Qin Y, Wang H, Wu W (2016) Metagenomic analysis
490 of microbial consortia enriched from compost: new insights into the role of Actinobacteria in
491 lignocellulose decomposition. *Biotechnology for Biofuels* 9:22. doi:10.1186/s13068-016-
492 0440-2
- 493 45. Wei H, Wang L, Hassan M, Xie B (2018) Succession of the functional microbial communities
494 and the metabolic functions in maize straw composting process. *Bioresource technology*
495 256:333-341. doi:<https://doi.org/10.1016/j.biortech.2018.02.050>
- 496 46. Wu J, Zhao Y, Qi H, Zhao X, Yang T, Du Y, Zhang H, Wei Z (2017) Identifying the key
497 factors that affect the formation of humic substance during different materials composting.
498 *Bioresource technology* 244:1193-1196. doi:<https://doi.org/10.1016/j.biortech.2017.08.100>

- 499 47. Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, Kamagata Y (2006)
500 *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and
501 *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the
502 new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the bacterial phylum
503 *Chloroflexi*. *International Journal of Systematic and Evolutionary Microbiology* 56 (6):1331-
504 1340. doi:doi:10.1099/ijs.0.64169-0
- 505 48. Yan X, Luo X, Zhao M (2016) Metagenomic analysis of microbial community in uranium-
506 contaminated soil. *Applied Microbiology and Biotechnology* 100 (1):299-310.
507 doi:10.1007/s00253-015-7003-5
- 508
509
510
511

List of figures captions:

Fig. 1 Rarefaction curve for a dissimilarity of 3% from sample POME and compost

Fig. 2 Relative abundance of prokaryotes 16S rRNA sequences in **(A)** phyla **(B)** family level and **(C)** genera level for POME and compost samples

Fig. 3 Taxonomic affiliation of 16S rRNA metagenomics fragment. The phylogenetic distribution for **(A)** the overall population and **(B)** Firmicutes in compost

Fig. 4 Taxonomic affiliation of 16S rRNA metagenomics fragment. The phylogenetic distribution for **(A)** the overall population and **(B)** Firmicutes in POME

Fig. 5 Heatmap visualizing the z-score distribution of the major genera among POME and compost samples

Fig. 6 Proportion of sequences for the predicted metabolic profile in POME and mature compost

Fig. 7 Number of sequences for the predicted enzymes involved in carbon (polysaccharides) degradation

List of figures:

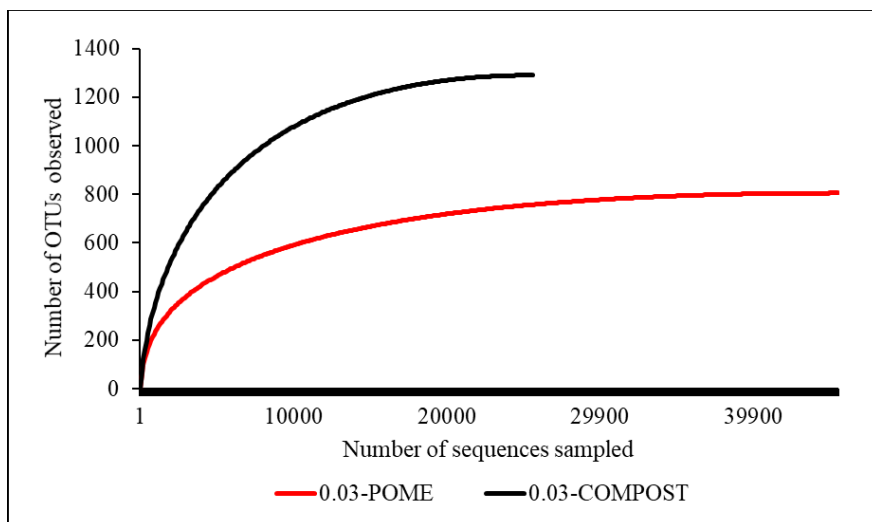


Fig. 1

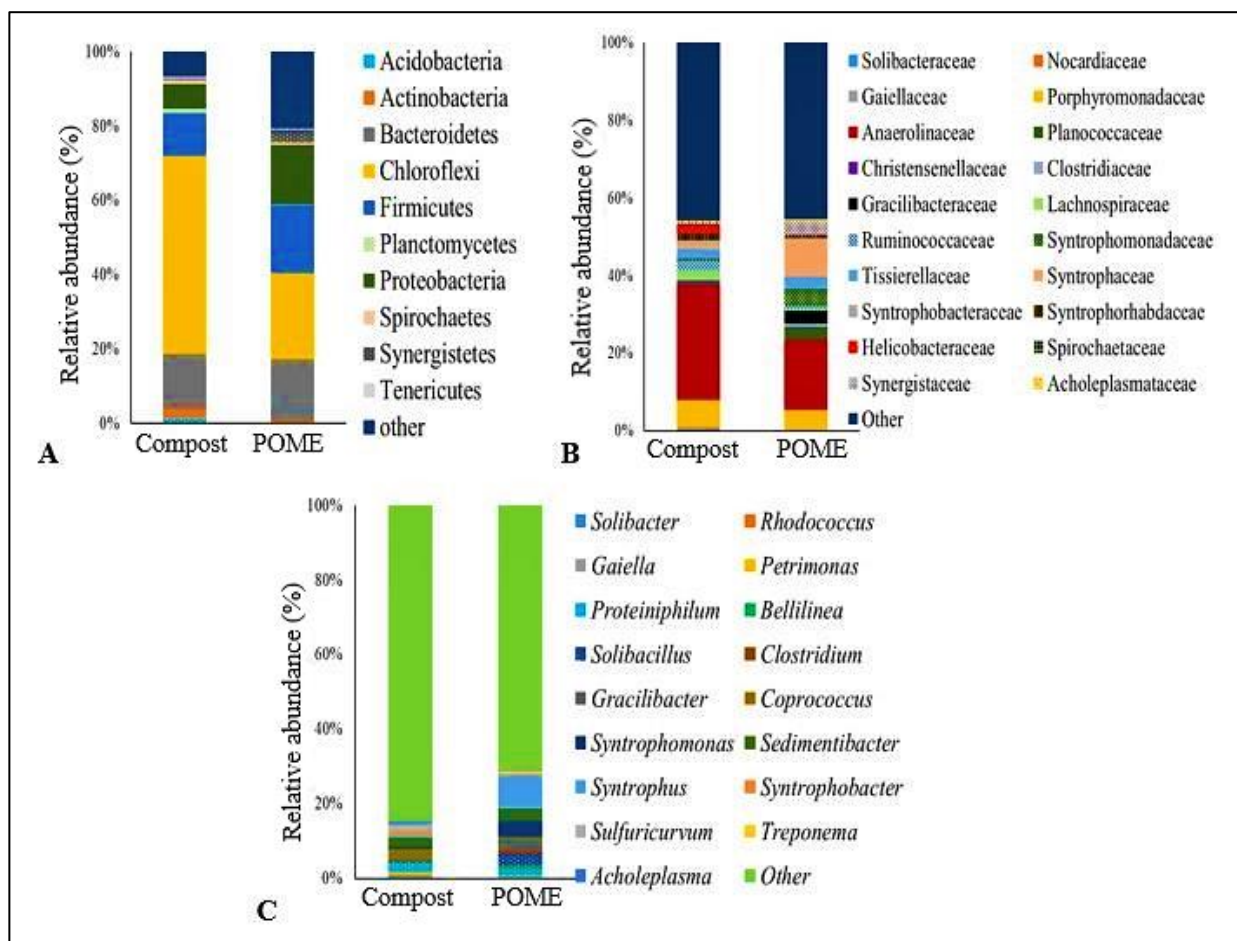


Fig. 2

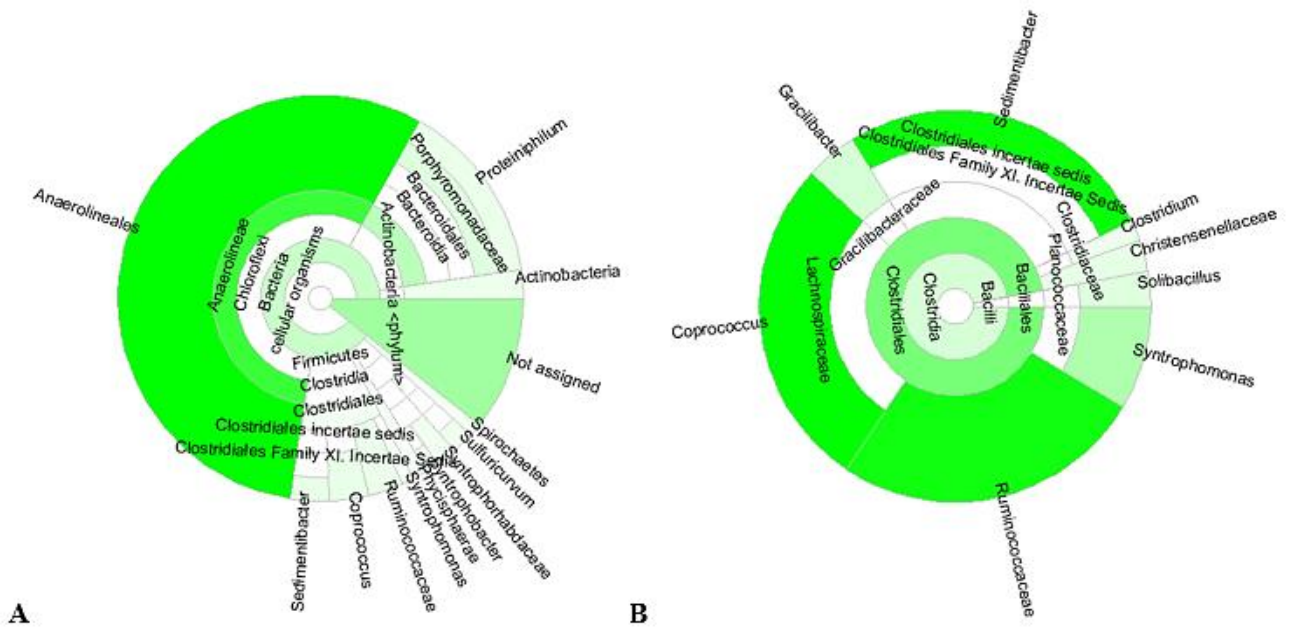


Fig. 3

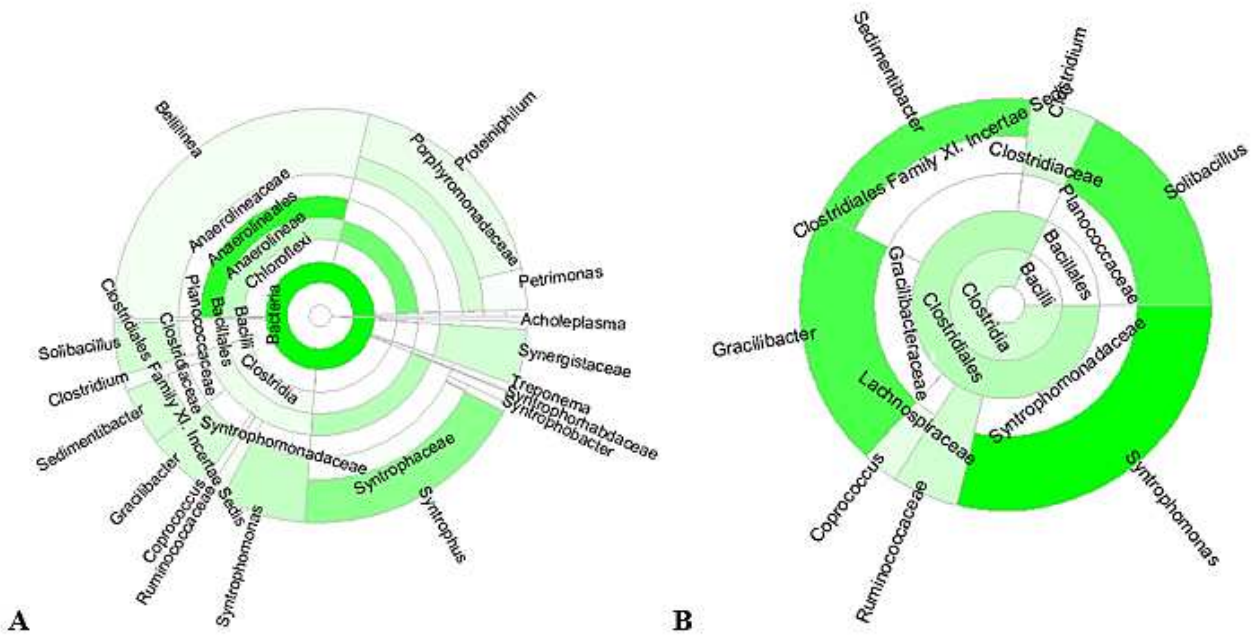


Fig. 4

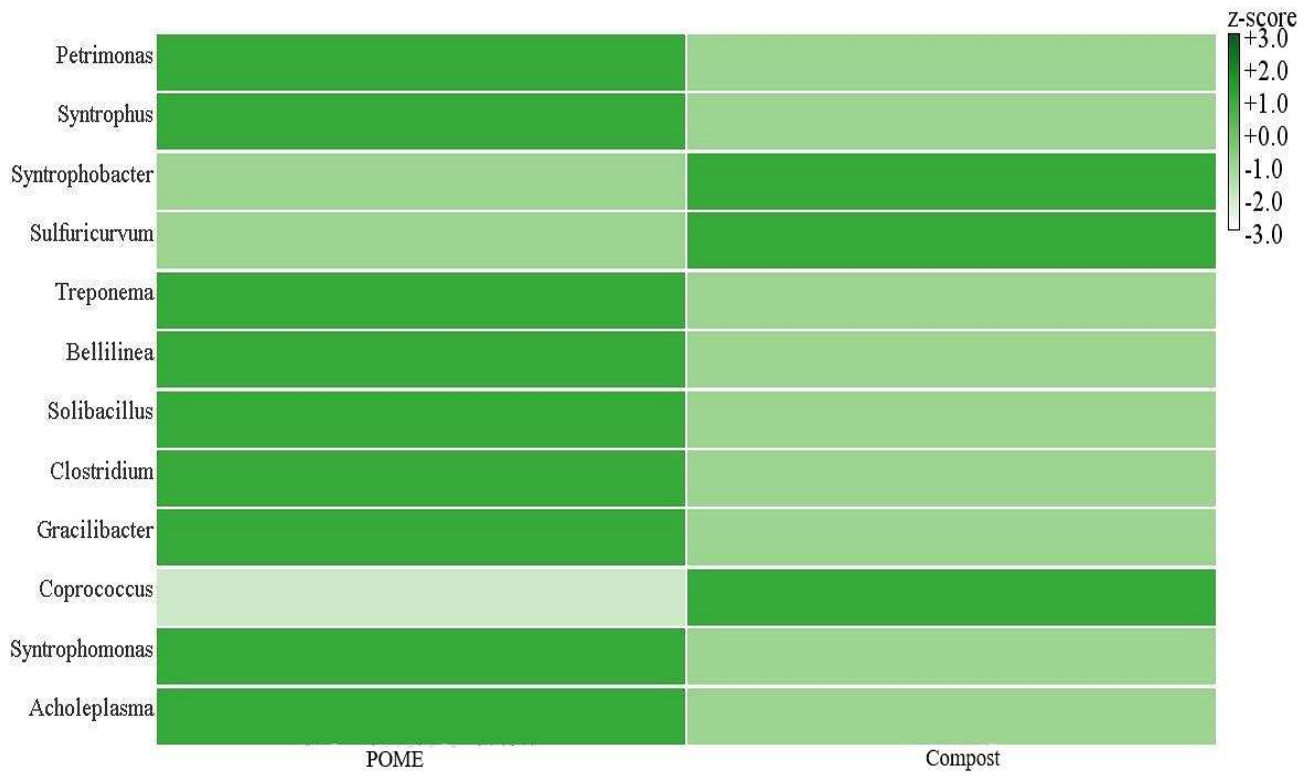


Fig.5

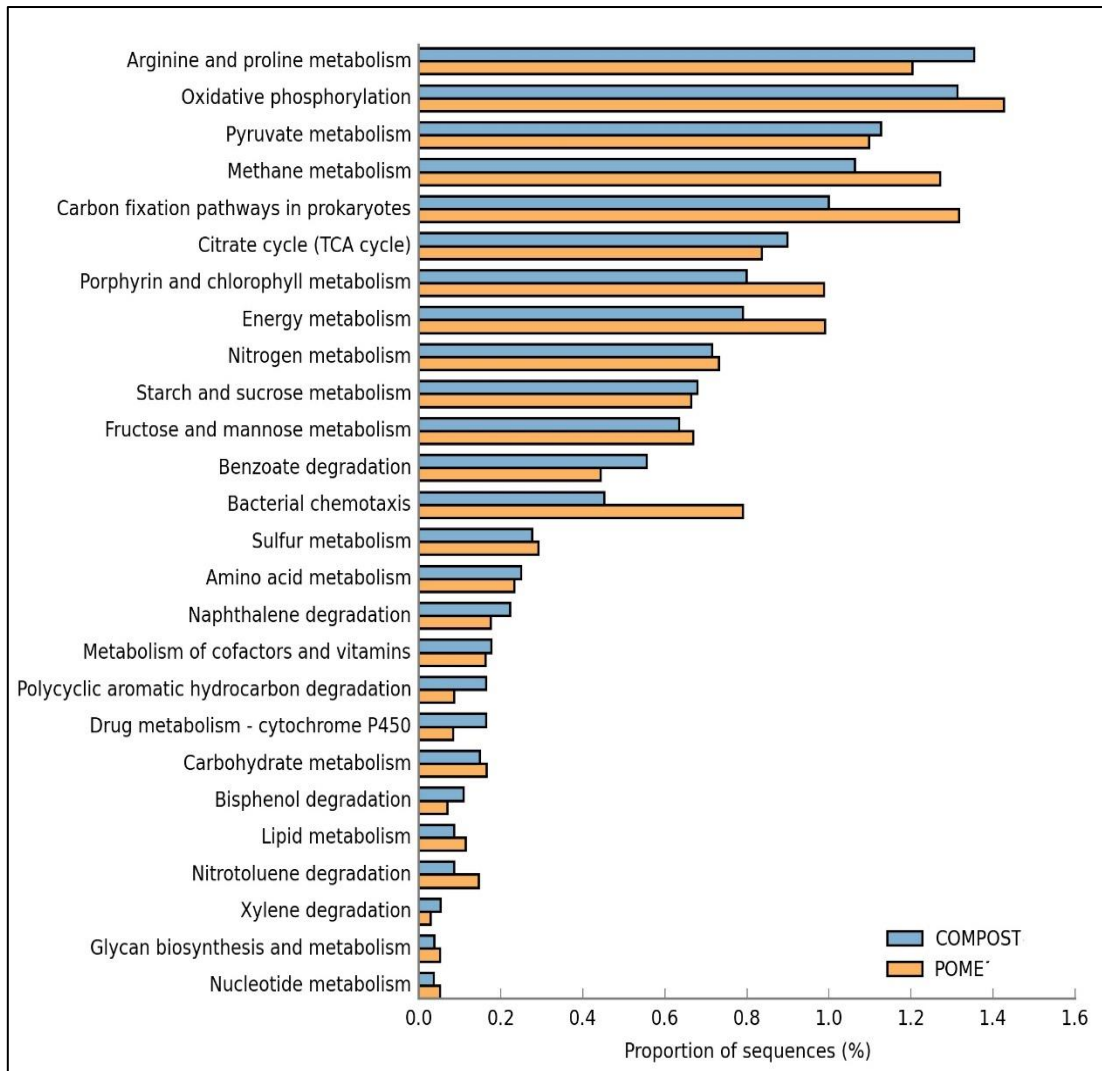


Fig. 6

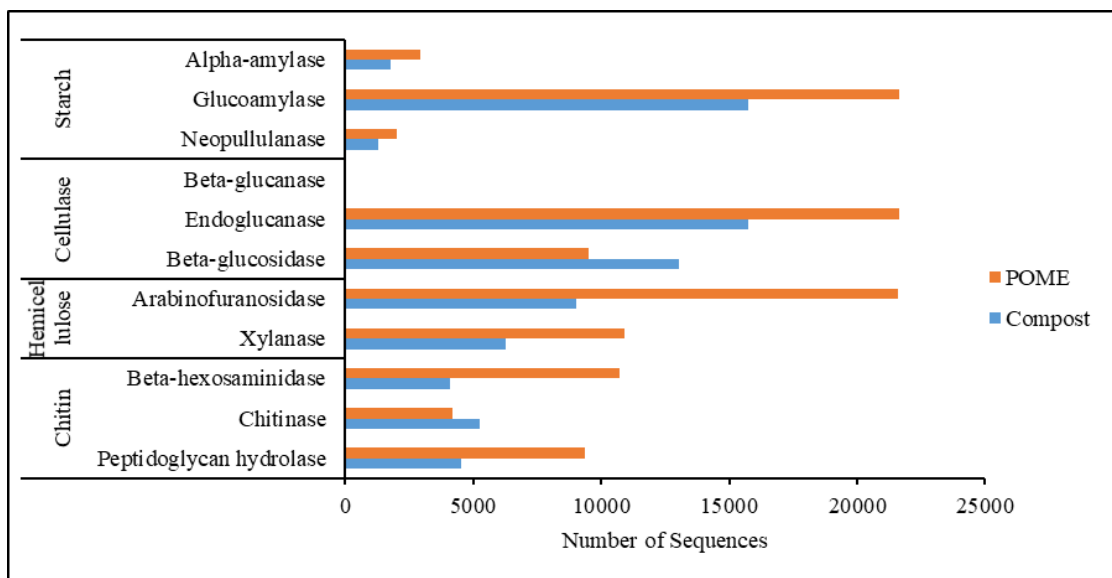


Fig. 7