A peer-reviewed version of this preprint was published in PeerJ on 6 January 2020.

View the peer-reviewed version (peerj.com/articles/8102), which is the preferred citable publication unless you specifically need to cite this preprint.

Elufisan TO, Rodríguez-Luna IC, Oyedara OO, Sánchez-Varela A, Hernández-Mendoza A, Dantán Gonzalez E, Paz-González AD, Muhammad K, Rivera G, Villalobos-Lopez MA, Guo X. 2020. The Polycyclic Aromatic Hydrocarbon (PAH) degradation activities and genome analysis of a novel strain *Stenotrophomonas sp*. Pemsol isolated from Mexico. PeerJ 8:e8102 https://doi.org/10.7717/peerj.8102

PeerJ Preprints NOT PEER-REVIEWED

Stenotrophomonas sp. Pemsol isolated from crude oil contaminated soil in Mexico that can degrade polycyclic aromatic hydrocarbons and its whole genome sequence analyzed

Temidayo O Elufisan Corresp., 1, 2 , **Isabel C Rodríguez-Luna** [1] , **Omotayo O Oyedara** [3] , **Alejandro Sánchez-Varela** [1] , **Armando Hernandez Mendoza** [4] , **Edgar Dantan Gonzalez** [5] , **Alma D Paz-González** [6] , **Kashif Muhammad** [6] , **Gildardo Rivera** [6] , **Miguel Á Villalobos-Lopez** [7] , **Xianwu Guo** Corresp. 1

[1] Laboratorio de Biotecnologia Genomica, Centro de Biotecnologia Genomica, Instituto Politecnico Nacional, Mexico, Reynosa, Tamaulipas, Mexico

[2] Science policy and Innovation studies, National Center for Technology Management Obafemi Awolowo University campus Ile-ife, Ile-Ife, Osun, Nigeria

[3] Microbiology Department, Osun state University, Oshogbo, Osun, Nigeria

[4] Centro de Investigación en Dinámica Celular,, Instituto de Investigación en Ciencias Básicas y Aplicadas, Universidad Autónoma del Estado de Morelos (UAEM), 62209,, Cuernavaca,, Morelos, Mexico

[5] Laboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos, Cuernavaca, Morelos, Mexico

[6] Laboratorio de Biotecnologia Famaceutica, Centro de Biotecnologia Genomica, Instituto Politecnico Nacional, Mexico, Reynosa, Tamaulipas, Mexico

[7] Centro de Investigación en Biotecnología Aplicada, Instituto Politécnico Nacional, Tepetitla, Tlaxcala, Mexico

Corresponding Authors: Temidayo O Elufisan, Xianwu Guo
Email address: telufisan1500@alumno.ipn.mx, xguo@ipn.mx

Background: Stenotrophomonas are ubiquitous gram-negative bacteria which survive in a wide range of environments. They can use many substances for their growth and are known to be intrinsically resistant to many antimicrobial agents. They have been tested for biotechnological applications, bioremediation and antimicrobial agents because of their recalcitrant nature to many toxic compounds. **Method**. *Stenotrophomonas sp*. Pemsol was isolated from a crude oil contaminated soil. The capability of this isolate to tolerate and degrade polycyclic aromatic hydrocarbons (PAHs) (anthracene, anthraquinone, biphenyl, naphthalene, phenanthrene, phenanthridine and xylene) was evaluated on Bush Nell Hass medium containing PAHs as the unique carbon sources. The metabolites formed after 30-day degradation of naphthalene by Pemsol were analyzed using Fourier Transform Infra-red Spectroscopic (FTIR), Ultra-Performance Liquid Chromatography-Mass Spectrometry (UPLC-MS) and Gas Chromatography-Mass Spectrometry (GC-MS). **Results.** Complete degradation of naphthalene at a concentration of 1 mg/mL was obtained and a newly formed catechol peak obtained from the UPLC-MS and GC-MS confirmed the degradation. The strain Pemsol lacked the ability to produce biosurfactant so that it cannot bio-emulsify PAHs. The whole genome analysis of *Stenotrophomonas* sp. Pemsol revealed a wealth of genes for hydrocarbon utilization and interaction with the environment and the presence

of 147 genes associated with the degradation of PAHs, some of which are strain-specific on the genomic islands. Few genes are associated with bio-emulsification indicated that Pemsol without biosurfactant production has a genetic basis. This is the first report of the complete genome analysis sequence of a PAH-degrading *Stenotrophomonas*. *Stenotrophomonas* sp. Pemsol possesses features that makes it a good bacterium for genetic engineering and will therefore be a good tool for the remediation of crude oil or PAH-contaminated soil.

PeerJ Preprints NOT PEER-REVIEWED

1 **_Stenotrophomonas_ sp. Pemsol isolated from crude oil contaminated soil in Mexico that can**
2 **degrade polycyclic aromatic hydrocarbons and its whole genome sequence analyzed**

3 Temidayo Oluyomi Elufisan[1, 2], Isabel Cristina Rodríguez-Luna[1], Omotayo Opemipo Oyedara[1],
4 Alejandro Sánchez-Varela[1], Armando Hernandez Mendoza[3], Edgar Dantan Gonzalez[4], Alma D.
5 Paz-González[1], Kashif Muhammad[1], Gildardo Rivera[1], Miguel Ángel Villalobos-Lopez[5], Xianwu
6 Guo[1*]

7 [1] Centro de Biotecnología Genómica, Instituto Politécnico Nacional, 88710, Reynosa, México

8 [2] Department of Science Policy and Innovation Studies (SPIS), National Center for Technology
9 Management, (An agency of the Federal Ministry of Science and Technology), Obafemi Awolowo
10 University Campus, 220282, Ile-Ife, Nigeria

11 [3] Centro de Investigación en Dinámica Celular, Instituto de Investigación en Ciencias Básicas y
12 Aplicadas, Universidad Autónoma del Estado de Morelos (UAEM), 62209, Cuernavaca, Morelos,
13 México

14 [4] Laboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad
15 Autónoma del Estado de Morelos (UAEM), 62209, Cuernavaca, Morelos, México.

16 [5] Centro de Investigación en Biotecnología Aplicada, Instituto Politécnico Nacional, Ex-Hacienda
17 San Juan Molino, carr. est. Tecuexcomac-Tepetitla Km. 1.5, 90700, Tepetitla de Lardizábal,
18 Tlaxcala, México

19 *Corresponding author: xguo@ipn.mx

20

21

22 **Abstract**

23 **Background**:

24 Stenotrophomonas are ubiquitous gram-negative bacteria which survive in a wide range of
25 environments. They can use many substances for their growth and are known to be intrinsically
26 resistant to many antimicrobial agents. They have been tested for biotechnological applications,
27 bioremediation and antimicrobial agents because of their recalcitrant nature to many toxic
28 compounds.

29 **Method**.

30 *Stenotrophomonas sp*. Pemsol was isolated from a crude oil contaminated soil. The capability of
31 this isolate to tolerate and degrade polycyclic aromatic hydrocarbons (PAHs) (anthracene,
32 anthraquinone, biphenyl, naphthalene, phenanthrene, phenanthridine and xylene) was evaluated
33 on Bush Nell Hass medium containing PAHs as the unique carbon sources. The metabolites
34 formed after 30-day degradation of naphthalene by Pemsol were analyzed using Fourier Transform
35 Infra-red Spectroscopic (FTIR), Ultra-Performance Liquid Chromatography-Mass Spectrometry
36 (UPLC-MS) and Gas Chromatography-Mass Spectrometry (GC-MS).

37 **Results.**

38 Complete degradation of naphthalene at a concentration of 1 mg/mL was obtained and a newly
39 formed catechol peak obtained from the UPLC-MS and GC-MS confirmed the degradation. The
40 strain Pemsol lacked the ability to produce biosurfactant, so that it cannot bio-emulsify PAHs. The
41 whole genome analysis of *Stenotrophomonas* sp. Pemsol revealed a wealth of genes for
42 hydrocarbon utilization and interaction with environment and the presence of 147 genes associated
43 with the degradation of PAHs, some of which are strain-specific on the genomic islands. Few

44 genes are associated with bioemulsification indicated that Pemsol without biosurfactant production
45 has a genetic basis.

46 This is the first report of the complete genome analysis sequence of a PAH-degrading
47 *Stenotrophomonas*. *Stenotrophomonas* sp. Pemsol possesses features that makes it a good
48 bacterium for genetic engineering and will therefore be a good tool for the remediation of crude
49 oil or PAH contaminated soil.

50 **Keywords:** *Stenotrophomonas,* Polycyclic Aromatic Hydrocarbon (PAH), Biphenyl,
51 Naphthalene, degradation, Sequencing.

52

53

54

55

56

57

58

59

60

61

62

63 **1. Introduction**

64 *Stenotrophomonas* species are ubiquitous bacteria, occupying various habitats including harsh
65 environment (Ryan et al., 2009; Hughes et al., 2016). They can use a wide range of substances for
66 their growth (Juhasz, Stanley & Britz, 2000; Pages et al., 2008; Zhang et al., 2009; Urszula et al.,
67 2009; Iyer, Iken & Leon, 2016). The vast metabolic capability of *Stenotrophomonas* species has
68 encouraged the studies to aim at finding new paths for their biotechnological application, such as
69 bioremediation, biodegradation, plant growth promotion, removal of organophosphate and
70 synthesis of antimicrobial agents (Ryan et al., 2009; Rajkumar et al., 2010; Iyer, Iken & Leon,
71 2016; Arulazhagan et al., 2017a). In particular, several studies have focused on the use of
72 *Stenotrophomonas maltophilia* for the remediation of Polycyclic Aromatic Hydrocarbons (PAHs)
73 or of crude oil contaminated sites (Boonchan, Britz & Stanley, 1998; Juhasz, Stanley & Britz,
74 2000; Arulazhagan et al., 2017b)

75 PAHs are the compounds formed from two or more fused aromatic rings. In the environment,
76 PAHs can be produced from either natural or manmade combustion sources. PAHs range from
77 naphthalene (two fused benzene rings) to coronene (seven fused benzene rings). Accidental
78 petroleum spillage is one of the ways through which PAHs are released into the environment.
79 Human exposure to PAH or its analogs is a great risk to health. Cancer resulting from previous
80 exposure to PAHs has been demonstrated in animal models (Kim et al., 2013). Risks associated
81 with PAH exposure validates the importance of adequate cleanup strategy in the environments.
82 Microbes have been described as the best agents for the bioremediation in oil-spilled sites
83 (Haritash & Kaushik, 2009).

84 Bacterial remediation of PAHs involved the activities of some genes encoding for oxygenases or
85 peroxidases. Some of the genes associated with the degradation of PAHs are alkane

86 monooxygenases, such as *alk*B from *Pseudomonas*; *alk*m from *Acinetobacter* sp. Strain, ADP-1;
87 *alk*B1 and *alk*B2 from *Rhodococcus* sp. Other genes were *xyl*E, catechol-2, 3 dioxygenases from
88 *Pseudomonas putida*; *ndo*B, naphthalene monooxygenase from *P. putida*; and *nidA*, pyrene
89 dioxygenase large subunit from *Mycobacterium* sp. strain PYR-1, as well as various
90 dehydrogenases and protocatechuate dioxygenases in *Stenotrophomonas* spp. (Gunsalus, 1951;
91 Seo, Keum & Li, 2009; Urszula et al., 2009; Das & Chandran, 2011). Some PAHs degradation
92 needs the help of biosurfactant produced by bacteria (Van Beilen & Funhoff, 2007; Fritsche &
93 Hofrichter, 2008). The biosurfactants or surface-active substances decreases the surface tension
94 on the surface of water molecules, thereby making entrapped PAH on surfaces available for the
95 use of bacteria (Boonchan, Britz & Stanley, 1998).

96 Genome sequencing of some bacteria with the potentials to degrade hydrocarbons and PAHs has
97 given deeper insight into the genes involved in the degradation, and mineralization of PAHs
98 (GUNSALUS, 1951; Schneiker et al., 2006; Kim et al., 2008; Das & Chandran, 2011; Pal et al.,
99 2017). It also gives information on the other peripheral pathways involved in this process, for
100 example, the annotation of the genes involved in the production of bio-surfactants and biofilm
101 formation in bacteria. Several bacteria with good potentials for hydrocarbon degradation have been
102 sequenced (Kim et al., 2008; Das et al., 2015; Pal et al., 2017). Although there were reports of
103 PAHs-degrading *Stenotrophomonas* species, no *Stenotrophomonas*' genome sequence has been
104 analyzed for its ability to degrade hydrocarbon so far.

105 In this study, we isolated a strain of *Stenotrophomonas* sp. Pemsol from crude oil contaminated
106 soil in the state of Tabasco, Mexico, which showed good potential to degrade several PAHs. The
107 sequence analysis of *Stenotrophomonas* sp. Pemsol revealed that it is rich in genes required for the

108 degradation of PAHs and other hydrocarbons. The aim of the study is to elucidate and understand
109 the genetic basis involved in the uptake and degradation of PAHs in *Stenotrophomonas* sp. Pemsol.

110 2. Material & Methods

111 *2.1 Sampling, isolation and cultivation of Stenotrophomonas sp. Pemsol*

112 *Stenotrophomonas sp.* Pemsol was isolated from crude oil-contaminated soil, Tabasco, Mexico
113 (17°52'26.9"N 92°29'12.4" W). One gram of soil sample was added into 10 mL of Luria-Bertani
114 broth and the mix was incubated at 30 ºC overnight. 1 mL of the bacterial culture was serially
115 diluted from 10^{-1} to 10^{-8} in phosphate buffer (pH= 6.5). One hundred microliters of each dilution
116 were spread on selective medium (StenoVIA agar, Himedia, India) plates. Colonies formed on
117 plates were selected for further identification.

118 *2.2 Amplification and sequencing of 16S RNA gene and phylogenetic analysis*

119 Genomic DNA was extracted from 5 mL Bacterial culture grown in Luria broth using Promega
120 wizard genomic DNA purification kit (Promega, Madison, USA) as per the manufacturer's
121 instruction. The 16S rRNA genes were amplified by PCR using steno1 (5' AGG GAA ACT TAC
122 GCT AAT ACC- 3') and steno2 (5' CTC TGT CCC TAC CAT TGT AG-3'). The PCR mix contains
123 0.5 µL, 2.5 U Taq DNA polymerase, 0.5 µL of 10 mM d -NTP mix, 2.5 µL of 10× PCR buffer, 1
124 µL (0.5 µM) of each primer, 0.75 µL (50 mM) $MgCl_2$, 16.75 µL double distilled water and 2 µL
125 DNA (10 ng/µL). PCR products were purified and sequenced at the Centro de Biotecnologia
126 Genomica, Instituto Politecnico Nacional (IPN), Mexico using the ABI 3130 sequencing machine
127 for species identification. 16S rRNA gene sequences were analyzed with Seqman software version
128 13 and subjected to similarities search against those retrieved from NCBI using Blastn program
129 (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequences were aligned with related sequences retrieved

PeerJ Preprints NOT PEER-REVIEWED

130 from NCBI data base using Mega 6.0 (Tamura et al., 2013) and phylogenetic tree was constructed
131 using neighbor-joining algorithm. Reliability of tree topologies was confirmed by bootstrap
132 analysis using 1000 repeat alignment. *Stenotrophomonas sp.* Pemsol sequence has been deposited
133 on NCBI with ascension number, KX500117.1.

134 *2.3 Cultivation and growth in PAHs-containing media and bio-emulsification*

135 All PAHs (anthracene, 95%; anthraquinone, 97%; biphenyl, 99%; naphthalene, 99%;
136 phenanthrene, 99%; phenanthridine, 98% and xylene, 98.5%) used for this study were purchased
137 from Sigma Aldrich, Mexico. Growth in PAH tests were carried out using the minimum medium
138 Bushnell Hass (BH medium) with one of the following PAH: naphthalene, phenanthridine,
139 anthraquinones, biphenyl, phenanthrene and xylene. All hydrocarbons were dissolved in dimethyl
140 chloride and the solvent was left to evaporate before introducing the hydrocarbons in experimental
141 system. 100 µL of overnight grown culture of bacteria washed in phosphate buffer was inoculated
142 in 100 mL BH medium containing the above-mentioned PAHs at a concentration of 1 mg/mL in
143 250 mL Erlenmeyer flask while non-inoculated BH medium containing hydrocarbons and BH
144 medium devoid of hydrocarbons but inoculated with *Stenotrophomonas* served as controls. All of
145 them were incubated at 30 ^{o}C in a rotatory incubator with revolution of 200 rpm for 8 days.
146 *Stenotrophomonas'* growth was checked every two days using colony counting. All experiments
147 were in triplicates. Spectrophotometric analysis was also carried out on culture from all
148 experimental set-ups to corroborate the observations from colony counting method. Emulsification
149 was tested according to the previous reports (Boonchan, Britz & Stanley, 1998; Panjiar, Sachan &
150 Sachan, 2015)

151 *2.3 Identifying metabolic intermediates from the degradation of naphthalene*

152 Extraction of naphthalene and other hydrocarbon from culture media was performed using equal
153 volume of hexane by triplicate. Then hexane was eliminated with vacuum pressure for further the
154 analysis (FTIR, UPLC-MS and GC-MS).

155 *2.3.1 Fourier-transform infrared spectroscopy (FTIR)*

156 The air-dried samples were analyzed on Bruker Alpha FT-IR spectrometer with Platinum ATR
157 (AXS Inc., Madison, WI, USA) to determine the presence or absence of specific bonds after
158 degradation.

159 *2.3.2* Ultra-Performance Liquid Chromatographic-Mass Spectrometry (*UPLC-MS) and Gas*
160 *Chromatography-Mass Spectrometry (GC-MS) analysis*

161 A total of 1 mg of extract was dissolved in 1 mL in dichloromethane. Then, 0.1 mL was added to
162 0.9 mL of methanol for analysis by Ultra-Performance Liquid Chromatographic (UPLC) with an
163 ACQUITY QDa mass detector from Waters (Milford, MA, USA) under the following conditions:
164 column: ACQUITY UPLC®BEH C_{18} 1.7µm 2.1x100 mm ; mobile phase A (0.1% formic acid in
165 water) mobile phase B (methanol) and C (Acetonitrile) in a time 0.5-5 min 27%A:25%B;48 °C;
166 total run time: 5 min; flow rate: 0.3 mL/min; injection volume: 3.0 µL; temperature column: 40
167 °C.

168 A total of 1 mg of extract was dissolved in 1 mL in dichloromethane. Then, 0.1 mL was added to
169 0.9 mL of methanol for analysis by Gas Chromatographic (7890A GC System) coupled to a Mass
170 detector (5975C inert MSD with Triple-Axis Detector) from Agilent technologies under the
171 following conditions: column: J&W 19091S-433HP-5MS: 30 m x 250 µm x 0.25 µm; Oven
172 Program 70 °C for 2 min. #1 then 10 °C/min to 160 °C for 2 min; #2 then 5 °C/min to 240 °C
173 for 2 min; #3 then 30 °C/min to 290 °C for 2 min. Run Time 34.667 min; injection volume: 2µL.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27617v1 | CC BY 4.0 Open Access | rec: 28 Mar 2019, publ: 28 Mar 2019

174 *2.4 Whole genome sequencing and analysis*

175 The genomic DNA was extracted as described above using the Promega DNA extraction kit (USA)
176 according to the manufacturer's instruction. The extracted bacterial genomic DNA was sequenced
177 at the Unidad Universitaria de Secuenciación Masiva y Bioinformática at the Instituto de
178 Biotecnología, UNAM with the Illumina MiSeq platform.

179 *2.4.1 Genome assembly and annotation*

180 The reads quality was checked with Fastqc (Andrews, 2010) and the adaptors from the raw reads
181 were trimmed with trim-galore version 4.10 which also filtered out reads with poor quality. De
182 novo genome assembly was carried out with a standalone Spades 3.11.1 genome assembler (Center
183 for Algorithmic biotechnology, St. Petersburg State University, Russia) (Bankevich et al., 2012).
184 The assembly's quality was checked with QUAST (Gurevich et al., 2013). The assembled contigs
185 were ordered and reduced into a single scaffold with MedusaCombo, an online genome multidraft
186 scaffolder (Bosi et al., 2015). The assembled genome was annotated with Prokka annotating
187 pipeline (Seemann, 2014). Further functional genome annotation was done with online genome
188 analysis server WebMGA (http://weizhong-lab.ucsd.edu/metagenomic-analysis) (Wu et al., 2011).
189 WebMGA was used to predict the KEGG functions and COG categories present in the genome.
190 The presence of transposon and insertion sequences was predicted with an web based analysis tool
191 software ISsaga (http://issaga.biotoul.fr/ISsaga2/issaga_index.php) (Varani et al., 2011). The Pan
192 core genome analysis for *Stenotrophomonas* sp. Pemsol and 12 other *Stenotrophomonas* species
193 to identify the unique genes in *Stenotrophomonas* sp. Pemsol. The Stenotrophomonas genome
194 compared with Stenotrophomonas so. Pemsol were the finished genome from the genus available
195 on NCBI database as of the time of this write up. (These species include; *S. maltophilia* JV3, *S.*
196 *maltophilia* ASS1, *S. pavani* LMG, *S. rhizophilia* QLP4, *S. pictorium* JCM 9942, *S. maltophilia*

197 K279a, *S. nitrireducen* 2001, *S. panacihumi*, *S. maltophilia* ATCC 19687, *S. maltophilia* R551-3,
198 and *Stenotrophomonas* sp.). We retrieved the unique genes from *Stenotrophomonas* sp. Pemsol
199 genome and subject them to further analysis. This include manual blast search analysis of the
200 genes on NCBI database to determine what they encode. We also carried out the synteny analysis
201 of the genes on SyntTax synteny and RAST annotation server.

202 *2.4.2 Prediction of genomic island*

203 The genomic island in the sequenced genome was predicted by genomic island viewer 4 (Bertelli
204 et al., 2017). The annotation of the functional content of the genes associated with genomic island
205 and unique gene predicted for *Stenotrophomonas* sp. Pemsol was done with BLAST2GO (Conesa
206 et al., 2005)

207 *2.4.3 Comparative genome analysis*

208 Genetic relatedness with other *Stenotrophomonas* species was determined by analyzing the
209 average nucleotide identity on J speciesWS (Richter & Rossello-Mora, 2009) and Genome-
210 Genome distance hybridization (GGDH) (Auch et al., 2010) tools. Further analysis on Pemsol was
211 carried out in the Integrated Microbial Genome (IMG) server (https://img.jgi.doe.gov) and Kbase
212 Platform (https://narrative.kbase.us/narrative/ws.27061.obj.1).

213 The complete genome sequence has been deposited on DDBJ/EMBL/GenBank under the
214 accession number CP025780.

215 **3.0 Results and Discussion**

216 *3.*1 Isolation and identification of the *Stenotrophomonas* strain Pemsol from crude oil
217 contaminated soil

218 The objective of this work was to isolate a *Stenotrophomonas* strain that could be used for
219 bioremediation of oil-polluted soil. This is why we used the StenoVIA agar medium for selection
220 of *Stenotrophomonas* strains (Kerr et al., 1996). Several uniform colonies with characteristic
221 yellow color appeared on the selective medium after 48 h of incubation. The fragment of 16S RNA
222 gene of one clone called Pemsol was amplified and sequenced with a pair of primers, steno1 and
223 steno2. The BLAST search of this sequence and phylogenetic analysis with other known strains
224 showed that Pemsol belongs to the genus *Stenotrophomonas.* The closest species being
225 *Stenotrophomonas maltophilia M*27, with 99% identity as shown in Figure1

226 *3.2 Utilization of PAHs by Stenotrophomonas sp.* Pemsol *as sole carbon source*

227 The experiments on tolerance and growth properties of *Stenotrophomonas* sp. Pemsol to various
228 PAH showed that this strain grew well at a concentration of 1 mg/mL in the presence of biphenyl,
229 phenanthrene, phenanthridine, naphthalene and anthraquinone in BH medium but did not exhibit
230 growth in BH medium supplemented with xylene as carbon source, as shown in Figure 2. Pemsol
231 also displayed the ability to grow in a mix of the five PAHs at the final concentration of 1 mg/mL
232 for the five compounds (Figure 2b). It showed two growth peaks in the mix, indicating that this
233 strain preferred to use some compounds as sole carbon source rather than others. Previous studies
234 have reported that *Stenotrophomonas* possess the ability to degrade xenobiotic, PAHs and
235 organophosphates (Ryan et al., 2009; Iyer, Iken & Leon, 2016). *S. maltophilia* strains with such
236 capabilities have been isolated from different environments ranging from common environment to
237 extreme environment such as highly acidic or basic environment (Felsenstein, 1985; Boonchan,
238 Britz & Stanley, 1998; Juhasz, Stanley & Britz, 2000; Samanta, Singh & Jain, 2002; Gao et al.,
239 2013; Tebyanian, Hassanshahian & Kariminik, 2013; Arulazhagan et al., 2017b). The prsent study

240 showed that the strain *Stenotrophomonas* sp. Pemsol can grow using some of PAHs as sole carbon
241 source, indicating that Pemsol could degrade those PAHs.

242 *3.3 Bio-emulsion and surfactant production* in *Stenotrophomonas* sp. *Pemsol*

243 Emulsion and surfactant production can help in PAHs degradation. Bio-surfactant production
244 usually enhance the dislodging of PAH attached to surfaces in water, thereby making hydrophobic
245 hydrocarbon available for the use of bacteria (Cameotra & Bollag, 2003). Thus, the emulsion and
246 surfactant production was evaluated, as described by Boochan *et al*., 1998; and Panijah *et al*., 2015
247 (Boonchan, Britz & Stanley, 1998; Panjiar, Sachan & Sachan, 2015). The result showed that
248 Pemsol lacked the ability to bio-emulsify PAHs (Figure 3). Since our study showed that Pemsol
249 can grow in tested PAHs as sole carbon source (Figure 2), It is needed to understand if
250 *Stenotrophomonas* sp. Pemsol really degrades PAHs without the emulsion and surfactant
251 production.

252 *3.4 Analysis of degradation products using FTIR Spectroscopy*

253 The ability of *Stenotrophomonas sp*. Pemsol to degrade naphthalene was analyzed using FTIR
254 spectrometry. New peaks at wavelengths -OH (3200-2800 cm^{-1}); $-C{=}O_{(CH2)}$ (1684 cm^{-1}); $-C{=}O_{(OH)}$
255 (1641 cm^{-1}); $-CH_2$ (2911)) after 15th day of degradation study and -OH (3300-3100 cm^{-1}); -C=O
256 (1690 or 1700 cm^{-1}); $-CH_2$ (3001 cm^{-1}) observed after the 30th day provides evidence of degradation
257 of naphthalene by *Stenotrophomonas* sp. (Figure S2a-c, see supplementary material).

258 *3.5 UPLC-MS and Gas Chromatography-MS analysis of degradation products*

259 The UPLC-MS and GC-MS analysis were performed to detect the metabolites formed from the
260 degradation of naphthalene after 30-day experiment. The absence of a peak corresponding to
261 naphthalene on the spectra obtained from UPLC-MS and GC-MS analysis confirmed the

262 degradation of naphthalene, after comparing with the control (Figure 4, S1 (a-b) see
263 supplementary file). Meanwhile, a peak occurred with molecular weight of 109.98 as the major
264 metabolite, estimated to be C_6H_5OH correspondent to the molecular weight for catechol (Figure
265 S4-S5, see supplementary material). It could thus be inferred that the degradation of naphthalene
266 by Pemsol is associated with the formation of catechol. Several studies have reported the
267 degradation of naphthalene and other PAHs to involve the formation of catechol or protocatechuate
268 (Smith, 1990; Johnsen, Wick & Harms, 2005). Although *Stenotrophomonas* sp. Pemsol did not
269 emulsify hydrocarbon, the metabolite analysis showed that it successfully degrade naphthalene.

270 *3.6 Genetic basis for the degradation of PAH by* Stenotrophomonas *sp. Pemsol*

271 A number of genomes of strains belonging to the genus *Stenotrophomonas* have been sequenced,
272 including clinical isolates (Lira et al., 2012; Iyer, Iken & Leon, 2016), but to date no genome with
273 emphasis on the strain with PAH degradation was reported. The genome of *Stenotrophomonas* sp.
274 Pemsol was sequenced with Illumina technology. The complete sequence of strain Pemsol was
275 assembled de novo to 62 contigs. These contigs were then reduced to one contig with Medusa
276 Scaffolder. The genome is composed of a single circular chromosome of 4.37 Mb (Table 2, Figure
277 5).

278 *Stenotrophomonas* sp. Pemsol has 147 genes associated with the degradation of PAHs in its
279 genome (Table S1). In this category, nine genes encode enzymes belonging to the
280 lactoylglutathione lyase family (PEM_01474; 01733; 01959; 02738; 02855; 02960; 02961; 03007;
281 03383) in COG0346, which have been reported to be involved in the degradation of aromatic
282 compound (Mesarch, Nakatsu & Nies, 2000). One of the lactoylglutathione (PEM_03383) was
283 predicted to be a catechol 2, 3 dioxygenases. This gene is essential for the conversion of salicylate
284 aldehyde to catechol in the naphthalene's metabolic degradation pathway (Grund, Denecke &

285 Eichenlaub, 1992). It also has a gene encoding salicylate hydroxylase (PEM_02405) (COG0654)
286 (EC:1.14.13.1) (*nahG*), which converts salicylic acid, an intermediate in the degradation of
287 naphthalene, to catechol by removing the carboxyl group at position 1 and introducing a hydroxyl
288 group in the same position as replacement (Goyal & Zylstra, 1997; Bosch et al., 1999). Two genes,
289 homogentisate 1,2-dioxygenase (PEM_03309) that are involved in the catabolism of aromatic
290 rings (Borowski, Georgiev & Siegbahn, 2005), and 2, 4 dihydroxyacetophenonedioxygenase
291 (PEM_00137), which helps in the cleavage of carbon-carbon bond in a substituent aromatic ring,
292 were also detected (Keegan et al., 2014). The above-mentioned information ratifies the capacity
293 of *Stenotrophomonas* sp. Pemsol to degrade naphthalene as shown in the experiment.

294 There are several other genes in *Stenotrophomonas* sp. Pemsol that can assist in the degradation
295 of PAHs. These genes encode chloromuconate isomerase (PEM_00043, EC:5.5.1.7),
296 carboxymethylenebutenolidase that converts 4-carboxymethyl-4-methylbut-2-en-4-olide formed
297 from methyl catechol during the degradation of toluene to 4-oxohex-2-enedioate in *Burkholderia*
298 (Dobslaw & Engesser, 2015), 4-oxalocrotonate tautomerase (PEM_00595) known to be associated
299 with the degradation of toluene, o-xylene, 3-ethyltoluene and 1,2,4-trimethylbenzene (Chen et al.,
300 1992), and biphenyl-2,3-diol 1,2-dioxygenase (PEM_03239, EC:1.13.11.39) associated with the
301 degradation of biphenyl and gamma-hexachlorocyclohexane (Yam et al., 2009). Several
302 monooxygenases and different types of dehydrogenase present on the genome could catalyze the
303 degradation of aromatic hydrocarbon and other xenobiotics (Versalovic et al., 2016; Pal et al.,
304 2017)

305 Pemsol has a phosphomannomutase (PEM_00211), and two other glycosyl 2 family transferase
306 genes which are needed for biosurfactant production (Pal et al., 2017). The absence of other genes
307 essential for biosurfactant production may be the reason for its inability to emulsify PAHs in our

308 experimental studies. Thus, in *Stenotrophomonas* sp. Pemsol, the degradation of PAHs without
309 emulsification has a genetic basis. Unlike in *Pseudomonas aeruginosa* N002, this strain has 25
310 genes for bio surfactant formation and regulation and can bio-emulsify PAHs.

311 *3.7 Specific PAH degradation genes in genomic islands (GI)*

312 Microbes have been widely known to acquire new properties via horizontal gene transfer. In
313 Pemsol, 35 genomic islands (GI) (Figure.7, Table S2), 336,552bp in length were identified,
314 constituting 7.7% of the genome. Table S3 contains the detailed description of the genes predicted
315 as GI genes in *Stenotrophomonas* sp. Pemsol. Some genes on the genomic island showed
316 similarity to genes found in bacteria of other taxa. Most genes were predicted to be of unknown
317 function. Twelve PAH degradation genes encoding some transporters and several transcriptional
318 regulators were in the GI. A regulatory protein (PEM_01297) known to be important the regulation
319 of xenobiotics' degradation was found in the genomic island. A Cysteine-liking transporter
320 (PEM_00076) and another sulfite transporter (PEM_03784) were required for the transport of
321 sulfite molecules in PAHs were also found (Takumi & Nonaka, 2016).

322 3.8 Analysis of unique genes in adaptation for survival in crude contaminated environment.

323 A pan-core genome analysis of *Stenotrophomonas* sp. Pemsol with 12 complete genomes of other
324 *Stenotrophomonas* species was performed. The result of pan -genome analysis showed that
325 *Stenotrophomonas* sp. Pemsol possesses 154 unique genes. Most genes identified to be unique
326 were part of the genes in *Stenotrophomona*s sp. Pemsol genomic island. The predicted functions
327 for these genes are shown in Table S6. Some of these genes are involved in the degradation of
328 PAH (Ghosal, et al., 2016). For example, short dehydrogenase reductase (SDR) can catalyze the
329 reduction of C=C bond between aromatic compound (Kavanagh et al., 2008). SDR gene and two

330 genes, encoding LysR-type transcriptional regulators (LTTRs, PEM_19060) (accession genome
331 number) and glutathione S-transferase (GST), exist in a gene cluster on the genome (Figure 8).
332 Further, the 3 proteins SDR, LTTR and GST showed the same gene order as the closest orthologs
333 with the identity of 91.6%, 94.24% and 85.65% respectively from *Lysobacter gummosus* after blast
334 analysis, implying that this gene cluster in the two species has a common origin and Pemsol could
335 have obtained this gene cluster by horizontal gene transfer (Figure 8). The KEGG database clearly
336 showed that GST directly involved in the degradation of many hydrocarbon compounds
337 (benzoapyrene, naphthalene, trichloroethylene, bromobenzene, etc.). The LTTR family protein has
338 been reported to have significant function in regulating genes that are important for the catabolism
339 of aromatic compound, cell motility and quorum sensing (Pal et al., 2017). Thus, LysR gene in
340 this gene cluster could be involved in the regulation of SDR and GST for PAH degradation to help
341 the survival of *Stenotrophomonas* sp. Pemsol in this crude contaminated environment.

342 **3.9 Comparative COG category analysis of *Stenotrophomonas* sp. Pemsol**

343 The COG categories in *Stenotrophomonas* species was compared with the COGs in Pemsol.
344 *Stenotrophomonas* sp. Pemsol has higher number of genes in some COG categories than the other
345 12 *Stenotrophomonas* species compared. The Fischer test statistical analysis of the COG categories
346 showed that *Stenotrophomonas* species Pemsol has in abundance more genes in the categories
347 energy production and conversion (C) (6.01%), amino acid transport and metabolism coenzyme
348 transport and metabolism (H) (6.78%), cell motility (N) (3.74%), secondary metabolite
349 biosynthesis, transport and metabolism (Q) (2.58%), general function prediction (R) (8.4%),
350 function unknown (S) (6.4%), signal transduction (T) (6.52%), defense mechanism (V) (3.19%),
351 extracellular mechanism (W) (1.77%), (Supplementary file 2). The abundance of the genes in these

352 categories could be associated with its survival and adaptation in the crude contaminated
353 environment.

354 The COG categories in *Stenotrophomonas* species Pemsol was compared with the COGs in five
355 other PAH degrading bacteria (*Acinetobacter_baylyi*_ADP1, *Acinetobacter_lwoffii*_SH145,
356 *Alcanivorax_borkumensis*_SK2, *Franconibacter_pulveris*_DJ34) previously reported. The Fisher
357 exact test statistics for the comparison of the COG categories in *Stenotrophomonas* sp. Pemsol and
358 four other hydrocarbon-degrading bacteria (*Acinetobacter_baylyi*_ADP1,
359 *Acinetobacter_lwoffii*_SH145, *Alcanivorax_borkumensis*_SK2, *Franconibacter_pulveris*_DJ34)
360 revealed that Pemsol has more genes in the COG category G, N, T and W than the other
361 hydrocarbon-degrading bacteria (figure 5). These categories are associated with carbohydrate
362 transport and metabolism (4.54%), cell movement (3.74%), signal transduction mechanisms
363 (6.52%) and extracellular structure (1.75%), which could help *Stenotrophomonas* sp. Pemsol to
364 metabolize hydrocarbons, sense the signals from environment and respond appropriately to the
365 environmental stimuli, for example, moving towards the useful or away from the harmful
366 components of the PAHs.

367 *4 .0 Conclusion*

368 *Stenotrophomonas* sp. Pemsol was isolated from crude oil contaminated soil from Tabasco,
369 Mexico. It grew in the presence of six PAHs (biphenyl, anthracene, anthraquinone, phenanthrene,
370 naphthalene and phenanthridine) as unique carbon source. The identification of Pemsol confirmed
371 that it is a member of the genus *Stenotrophomonas* and that its closest relative in the genus is
372 *Stenotrophomonas maltophilia* M27. The ability of Pemsol to degrade PAH was confirmed by its
373 degradation activities on naphthalene as revealed by FTIR, UPLC-MS and GC-MS analysis. The
374 complete genome analysis of Pemsol revealed that it possesses many genes that are involved in

375 the degradation of PAHs but only 3 gene associated with bioemulsification, leading to no
376 biosurfactant production. The presence of some genes associated with the degradation of PAHs in
377 the genomic islands inferred that those genes were horizontally acquired. Comparing with other
378 four sequenced hydrocarbon-degrading bacteria, Pemsol is much rich in genes for the COG
379 category G, N, T and W, which are mainly relevant to hydrocarbon utilization and interaction with
380 environment. These results give insight into the genetic basis involved in the survival of Pemsol
381 in its oil-contaminated site and provide guide on the possible strategies for the bioremediation of
382 an oil-polluted environment with *Stenotrophomonas* sp. Pemsol without biosurfactant production.

383

384

385

386

387

388

389

390

391

392

393

394

395 **References**

396 Andrews S. 2010. FastQC: A quality control tool for high throughput sequence data. *626*
397 *Babraham Institute, Cambridge, UK.* DOI: citeulike-article-id:11583827.

398 Arulazhagan P, Al-Shekri K, Huda Q, Godon JJ, Basahi JM, Jeyakumar D. 2017a.
399 Biodegradation of polycyclic aromatic hydrocarbons by an acidophilic Stenotrophomonas
400 maltophilia strain AJH1 isolated from a mineral mining site in Saudi Arabia. *Extremophiles*
401 21:163–174. DOI: 10.1007/s00792-016-0892-0.

402 Arulazhagan P, Al-Shekri K, Huda Q, Godon JJ, Basahi JM, Jeyakumar D. 2017b.
403 Biodegradation of polycyclic aromatic hydrocarbons by an acidophilic Stenotrophomonas
404 maltophilia strain AJH1 isolated from a mineral mining site in Saudi Arabia. *Extremophiles*
405 21:163–174. DOI: 10.1007/s00792-016-0892-0.

406 Auch AF, von Jan M, Klenk HP, Göker M. 2010. Digital DNA-DNA hybridization for microbial
407 species delineation by means of genome-to-genome sequence comparison. *Standards in*
408 *Genomic Sciences* 2:117–134. DOI: 10.4056/sigs.531120.

409 Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko
410 SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev
411 MA, Pevzner PA. 2012. SPAdes: A New Genome Assembly Algorithm and Its
412 Applications to Single-Cell Sequencing. *Journal of Computational Biology*. DOI:
413 10.1089/cmb.2012.0021.

414 Van Beilen JB, Funhoff EG. 2007. Alkane hydroxylases involved in microbial alkane
415 degradation. *Applied Microbiology and Biotechnology* 74:13–21. DOI: 10.1007/s00253-
416 006-0748-0.

417 Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, Brinkman FSL. 2017.
418 IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic*
419 *Acids Research* 45:W30–W35. DOI: 10.1093/nar/gkx343.

420 Boonchan S, Britz ML, Stanley GA. 1998. Surfactant-enhanced biodegradation of high
421 molecular weight polycyclic aromatic hydrocarbons by stenotrophomonas maltophilia.
422 *Biotechnology and Bioengineering* 59:482–494. DOI: 10.1002/(SICI)1097-
423 0290(19980820)59:4<482::AID-BIT11>3.0.CO;2-C.

424 Borowski T, Georgiev V, Siegbahn PEM. 2005. Catalytic reaction mechanism of homogentisate
425 dioxygenase: A hybrid DFT study. *Journal of the American Chemical Society* 127:17303–
426 17314. DOI: 10.1021/ja054433j.

427 Bosch R, Moore ERB, García-Valdés E, Pieper DH. 1999. NahW, a novel, inducible salicylate
428 hydroxylase involved in mineralization of naphthalene by Pseudomonas stutzeri AN10.
429 *Journal of Bacteriology* 181:2315–2322.

430 Bosi E, Donati B, Galardini M, Brunetti S, Sagot MF, Lió P, Crescenzi P, Fani R, Fondi M.
431 2015. MeDuSa: A multi-draft based scaffolder. *Bioinformatics* 31:2443–2451. DOI:
432 10.1093/bioinformatics/btv171.

433 Cameotra SS, Bollag JM. 2003. Biosurfactant-enhanced bioremediation of polycyclic aromatic
434 hydrocarbons. *Critical Reviews in Environmental Science and Technology* 33:111–126.
435 DOI: 10.1080/10643380390814505.

436 Chen LH, Kenyon GL, Curtin F, Harayama S, Bembenek ME, Hajipour G, Whitman CP. 1992.
437 4-Oxalocrotonate tautomerase, an enzyme composed of 62 amino acid residues per
438 monomer. *Journal of Biological Chemistry* 267:17716–17721.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27617v1 | CC BY 4.0 Open Access | rec: 28 Mar 2019, publ: 28 Mar 2019

439 Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: A universal
440 tool for annotation, visualization and analysis in functional genomics research.
441 *Bioinformatics* 21:3674–3676. DOI: 10.1093/bioinformatics/bti610.

442 Das D, Baruah R, Sarma Roy A, Singh AK, Deka Boruah HP, Kalita J, Bora TC. 2015.
443 Complete genome sequence analysis of Pseudomonas aeruginosa N002 reveals its genetic
444 adaptation for crude oil degradation. *Genomics* 105:182–190. DOI:
445 10.1016/j.ygeno.2014.12.006.

446 Das N, Chandran P. 2011. Microbial Degradation of Petroleum Hydrocarbon Contaminants: An
447 Overview. *Biotechnology Research International* 2011:1–13. DOI: 10.4061/2011/941810.

448 Dobslaw D, Engesser KH. 2015. Degradation of toluene by ortho cleavage enzymes in
449 Burkholderia fungorum FLU100. *Microbial Biotechnology* 8:143–154. DOI: 10.1111/1751-
450 7915.12147.

451 Felsenstein J. 1985. Confidence Limits on Phylogenies: An Approach Using the Bootstrap.
452 *Evolution* 39:783. DOI: 10.2307/2408678.

453 Fritsche W, Hofrichter M. 2008. Aerobic Degradation by Microorganisms. In: *Biotechnology:*
454 *Second, Completely Revised Edition*. 144–167. DOI: 10.1002/9783527620999.ch6m.

455 Gao S, Seo JS, Wang J, Keum YS, Li J, Li QX. 2013. Multiple degradation pathways of
456 phenanthrene by Stenotrophomonas maltophilia C6. *International Biodeterioration and*
457 *Biodegradation* 79:98–104. DOI: 10.1016/j.ibiod.2013.01.012.

458 Goyal AK, Zylstra GJ. 1997. Genetics of naphthalene and phenanthrene degradation by
459 Comamonas testosteroni. *Journal of Industrial Microbiology and Biotechnology* 19:401–

460 407. DOI: 10.1038/sj.jim.2900476.

461 Grund E, Denecke B, Eichenlaub R. 1992. Naphthalene degradation via salicylate and gentisate
462 by Rhodococcus sp. strain B4. *Applied and Environmental Microbiology* 58:1874–1877.

463 GUNSALUS IC. 1951. GROWTH OF BACTERIA. In: *Bacterial Physiology*. 101–125. DOI:
464 10.1016/B978-1-4832-3136-5.50010-3.

465 Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for
466 genome assemblies. *Bioinformatics* 29:1072–1075. DOI: 10.1093/bioinformatics/btt086.

467 Haritash AK, Kaushik CP. 2009. Biodegradation aspects of Polycyclic Aromatic Hydrocarbons
468 (PAHs): A review. *Journal of Hazardous Materials* 169:1–15. DOI:
469 10.1016/j.jhazmat.2009.03.137.

470 Hughes GL, Raygoza Garay JA, Koundal V, Rasgon JL, Mwangi MM. 2016. Genome Sequence
471 of Stenotrophomonas maltophilia Strain SmAs1, Isolated From the Asian Malaria Mosquito
472 Anopheles stephensi. *Genome Announcements* 4:e00086-16. DOI:
473 10.1128/genomeA.00086-16.

474 Iyer R, Iken B, Leon A. 2016. Characterization and comparison of putative Stenotrophomonas
475 maltophilia methyl parathion hydrolases. *Bioremediation Journal*. DOI:
476 10.1080/10889868.2015.1114462.

477 Johnsen AR, Wick LY, Harms H. 2005. Principles of microbial PAH-degradation in soil.
478 *Environmental Pollution* 133:71–84. DOI: 10.1016/j.envpol.2004.04.015.

479 Juhasz AL, Stanley GA, Britz ML. 2000. Microbial degradation and detoxification of high
480 molecular weight polycyclic aromatic hydrocarbons by Stenotrophomonas maltophilia

481 strain VUN 10,003. *Letters in Applied Microbiology* 30:396–401. DOI: 10.1046/j.1472-
482 765x.2000.00733.x.

483 Kavanagh KL, Jörnvall H, Persson B, Oppermann U. 2008. Medium- and short-chain
484 dehydrogenase/reductase gene and protein families. *Cellular and Molecular Life Sciences*
485 65:3895–3906. DOI: 10.1007/s00018-008-8588-y.

486 Keegan R, Lebedev A, Erskine P, Guo J, Wood SP, Hopper DJ, Rigby SEJ, Cooper JB. 2014.
487 Structure of the 2,4'-dihydroxyacetophenone dioxygenase from Alcaligenes sp. 4HAP. *Acta*
488 *Crystallographica Section D: Biological Crystallography* 70:2444–2454. DOI:
489 10.1107/S1399004714015053.

490 Kerr KG, Denton M, Todd N, Corps CM, Kumari P, Hawkey PM. 1996. A new selective
491 differential medium for isolation of stenotrophomonas maltophilia. *European Journal of*
492 *Clinical Microbiology and Infectious Diseases*. DOI: 10.1007/BF01709373.

493 Kim KH, Jahan SA, Kabir E, Brown RJC. 2013. A review of airborne polycyclic aromatic
494 hydrocarbons (PAHs) and their human health effects. *Environment International* 60:71–80.
495 DOI: 10.1016/j.envint.2013.07.019.

496 Kim SJ, Kweon O, Jones RC, Edmondson RD, Cerniglia CE. 2008. Genomic analysis of
497 polycyclic aromatic hydrocarbon degradation in Mycobacterium vanbaalenii PYR-1.
498 *Biodegradation* 19:859–881. DOI: 10.1007/s10532-008-9189-z.

499 Lira F, Hernández A, Belda E, Sánchez MB, Moya A, Silva FJ, Martínez JL. 2012. Whole-
500 genome sequence of Stenotrophomonas maltophilia D457, a clinical isolate and a model
501 strain. *Journal of bacteriology* 194:3563–4. DOI: 10.1128/JB.00602-12.

502 Mesarch MB, Nakatsu CH, Nies L. 2000. Development of catechol 2,3-dioxygenase-specific
503 primers for monitoring bioremediation by competitive quantitative PCR. *Applied and*
504 *Environmental Microbiology* 66:678–683. DOI: 10.1128/AEM.66.2.678-683.2000.

505 Pages D, Rose J, Conrod S, Cuine S, Carrier P, Thierry Heulin WA. 2008. Heavy Metal
506 Tolerance in Stenotrophomonas maltophilia. *PLOS* 3:1–6. DOI:
507 http://dx.doi.org/10.1371/journal.pone.0001539.

508 Pal S, Kundu A, Banerjee T Das, Mohapatra B, Roy A, Manna R, Sar P, Kazy SK. 2017.
509 Genome analysis of crude oil degrading Franconibacter pulveris strain DJ34 revealed its
510 genetic basis for hydrocarbon degradation and survival in oil contaminated environment.
511 *Genomics* 109:374–382. DOI: 10.1016/j.ygeno.2017.06.002.

512 Panjiar N, Sachan SG, Sachan A. 2015. Screening of bioemulsifier-producing micro-organisms
513 isolated from oil-contaminated sites. *Annals of Microbiology* 65:753–764. DOI:
514 10.1007/s13213-014-0915-y.

515 Rajkumar M, Ae N, Prasad MNV, Freitas H. 2010. Potential of siderophore-producing bacteria
516 for improving heavy metal phytoextraction. *Trends in Biotechnology* 28:142–149. DOI:
517 10.1016/j.tibtech.2009.12.002.

518 Richter M, Rossello-Mora R. 2009. Shifting the genomic gold standard for the prokaryotic
519 species definition. *Proceedings of the National Academy of Sciences* 106:19126–19131.
520 DOI: 10.1073/pnas.0906412106.

521 Ryan RP, Monchy S, Cardinale M, Taghavi S, Crossman L, Avison MB, Berg G, van der Lelie
522 D, Dow JM. 2009. The versatility and adaptation of bacteria from the genus
523 Stenotrophomonas. *Nature reviews. Microbiology* 7:514–25. DOI: 10.1038/nrmicro2163.

524 Samanta SK, Singh O V., Jain RK. 2002. Polycyclic aromatic hydrocarbons: Environmental
525 pollution and bioremediation. *Trends in Biotechnology* 20:243–248. DOI: 10.1016/S0167-
526 7799(02)01943-1.

527 Schneiker S, dos Santos VAM, Bartels D, Bekel T, Brecht M, Buhrmester J, Chernikova TN,
528 Denaro R, Ferrer M, Gertler C, Goesmann A, Golyshina O V., Kaminski F, Khachane AN,
529 Lang S, Linke B, McHardy AC, Meyer F, Nechitaylo T, Pühler A, Regenhardt D, Rupp O,
530 Sabirova JS, Selbitschka W, Yakimov MM, Timmis KN, Vorhölter F-J, Weidner S, Kaiser
531 O, Golyshin PN. 2006. Genome sequence of the ubiquitous hydrocarbon-degrading marine
532 bacterium Alcanivorax borkumensis. *Nature Biotechnology* 24:997–1004. DOI:
533 10.1038/nbt1232.

534 Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069.
535 DOI: 10.1093/bioinformatics/btu153.

536 Seo JS, Keum YS, Li QX. 2009. Bacterial degradation of aromatic compounds. *International
537 Journal of Environmental Research and Public Health* 6:278–309. DOI:
538 10.3390/ijerph6010278.

539 Smith MR. 1990. The biodegradation of aromatic hydrocarbons by bacteria. *Biodegradation*
540 1:191–206. DOI: 10.1007/BF00058836.

541 Takumi K, Nonaka G. 2016. Bacterial cysteine-inducible cysteine resistance systems. *Journal of
542 Bacteriology*. DOI: 10.1128/JB.01039-15.

543 Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary
544 genetics analysis version 6.0. *Molecular Biology and Evolution* 30:2725–2729. DOI:
545 10.1093/molbev/mst197.

546 Tebyanian H, Hassanshahian M, Kariminik A. 2013. Hexadecane-degradation by Teskumurella
547 and Stenotrophomonas Strains Isolated From Hydrocarbon Contaminated Soils.
548 *Jundishapur Journal of Microbiology* 6. DOI: 10.5812/jjm.9182.

549 Urszula G, Izabela G, Danuta W, Sylwia L. 2009. Isolation and characterization of a novel strain
550 of Stenotrophomonas maltophilia possessing various dioxygenases for monocyclic
551 hydrocarbon degradation. *Brazilian journal of microbiology : [publication of the Brazilian
552 Society for Microbiology]* 40:285–91. DOI: 10.1590/S1517-838220090002000014.

553 Varani AM, Siguier P, Gourbeyre E, Charneau V, Chandler M. 2011. ISsaga is an ensemble of
554 web-based methods for high throughput identification and semi-automatic annotation of
555 insertion sequences in prokaryotic genomes. *Genome Biology* 12:R30. DOI: 10.1186/gb-
556 2011-12-3-r30.

557 Wu S, Zhu Z, Fu L, Niu B, Li W. 2011. WebMGA: A customizable web server for fast
558 metagenomic sequence analysis. *BMC Genomics* 12. DOI: 10.1186/1471-2164-12-444.

559 Yam KC, D'Angelo I, Kalscheuer R, Zhu H, Wang JX, Snieckus V, Ly LH, Converse PJ, Jacobs
560 WR, Strynadka N, Eltis LD. 2009. Studies of a ring-cleaving dioxygenase illuminate the
561 role of cholesterol metabolism in the pathogenesis of Mycobacterium tuberculosis. *PLoS
562 Pathogens* 5. DOI: 10.1371/journal.ppat.1000344.

563 Zhang H, Yang C, Zhao Q, Qiao C. 2009. Development of an autofluorescent organophosphates-
564 degrading Stenotrophomonas sp. with dehalogenase activity for the biodegradation of
565 hexachlorocyclohexane (HCH). *Bioresource Technology* 100:3199–3204. DOI:
566 10.1016/j.biortech.2009.02.008.

567

568

569

570

Table 1(on next page)

Genome feature for Stenotrophomonas species Pemsol

Figure 1 is the phylogenetic 3 drawn from the alignment of the 16s rRNA fragment of the Stenotrophomonas genome, Figure 2 is the growth of Stenotrophomonas sp. Pemsol using different PAHs as unique carbon, Figure 3. Bio-emulsification activity of Stenotrophomonas sp. Pemsol, The GC- MS/MS analysis spectrum of the test experiment after the 30th day of degradation studies. Figure 5. Circular Genome Map for Stenotrophomonas sp. Pemsol, Figure 6. Genomic Island distribution in Stenotrophomonas sp. Pemsol, Figure 7. LysR and SDR cluster in Stenotrophomonas sp. Pemsol and its comparison with Lysobacter gummosus and Figure 8 . COG Distribution comparison of Pemsol and other hydrocarbon degrading bacteria

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27617v1 | CC BY 4.0 Open Access | rec: 28 Mar 2019, publ: 28 Mar 2019

PeerJ Preprints NOT PEER-REVIEWED

1

Table 1 Genome feature

Features	Genome
DNA, total number of bases	4, 373,402
DNA coding number of bases	4, 370061
DNA G + C content (%)	66.59%.
Misc_RNA	39
Protein coding genes	3905
rRNA genes	4
tRNA genes	70
tmRNA	1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21 \

22

23

Figure 1(on next page)

Figures in manuscript

PeerJ Preprints | NOT PEER-REVIEWED

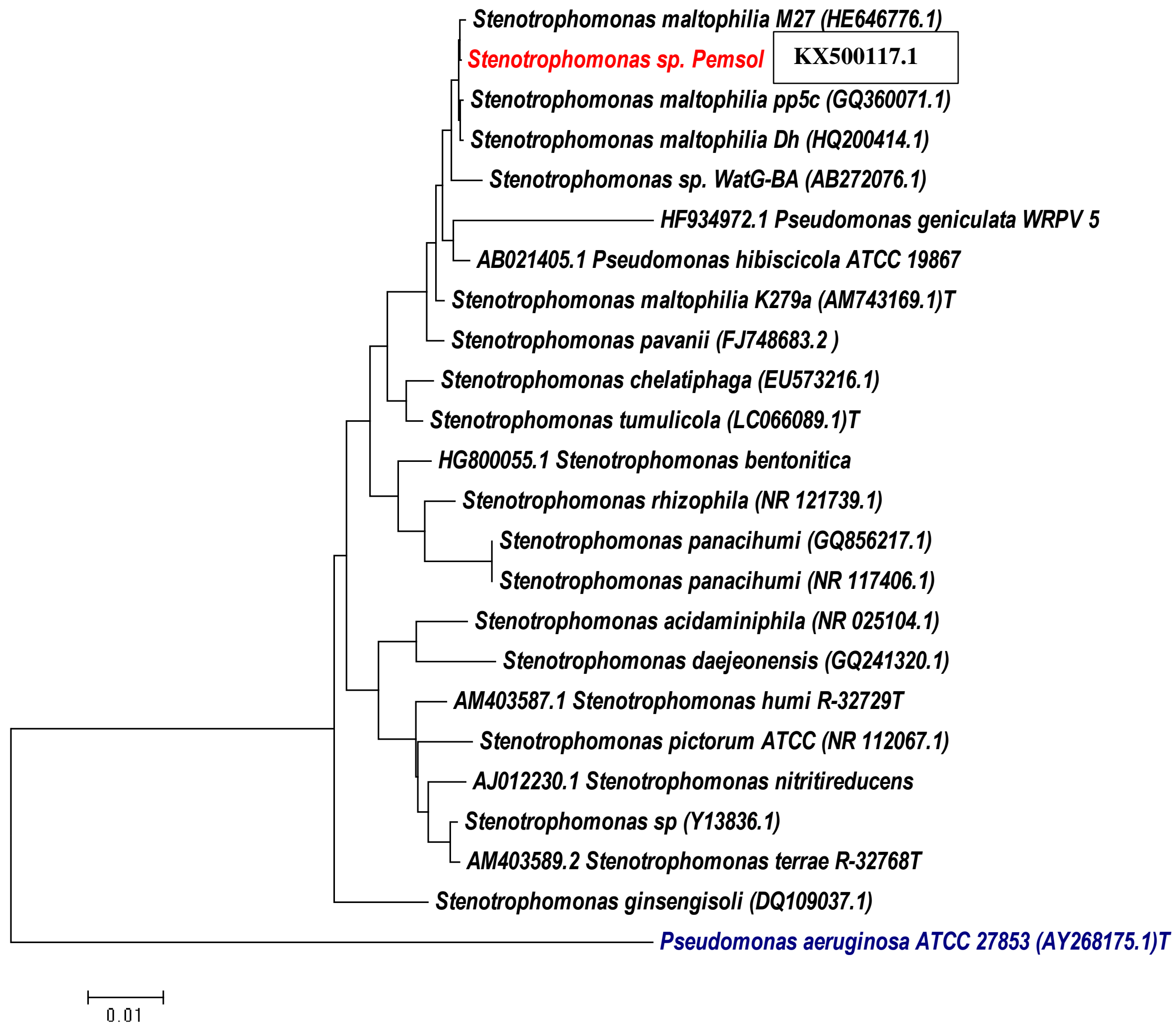


Figure 1. Phylogenetic tree of *Stenotrophomonas sp*. Pemsol with other members of the genus *Stenotrophomonas* based on the sequence of 16S rRNA gene.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27617v1 | CC BY 4.0 Open Access | rec: 28 Mar 2019, publ: 28 Mar 2019

PeerJ Preprints NOT PEER-REVIEWED

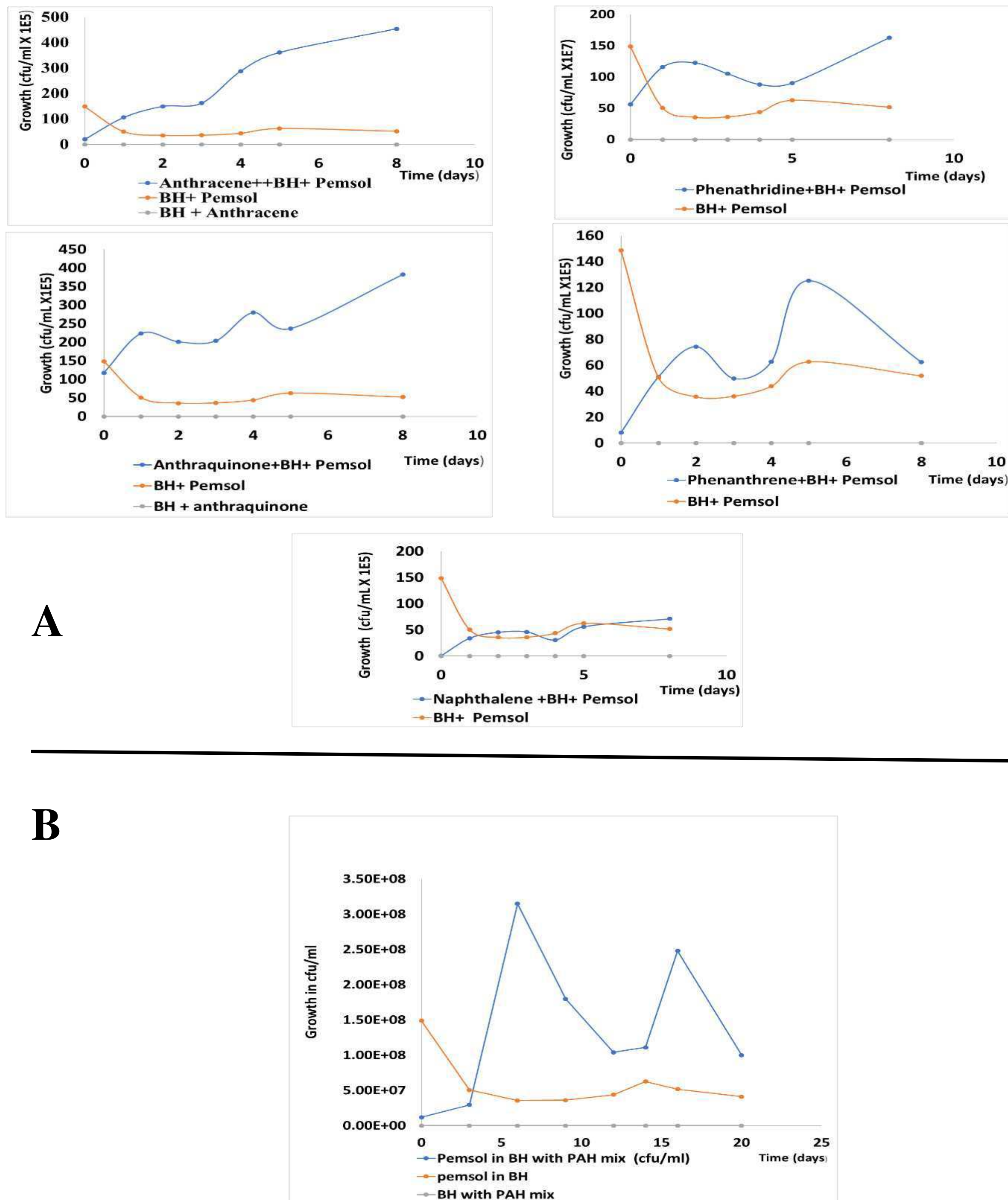


Figure 2. The growth of *Stenotrophomonas* sp. Pemsol using different PAHs as unique carbon source. A shows Pemsol's growth using individual PAH and the controls;

B shows Pemsol's growth using the PAH compound mix.

PeerJ Preprints NOT PEER-REVIEWED

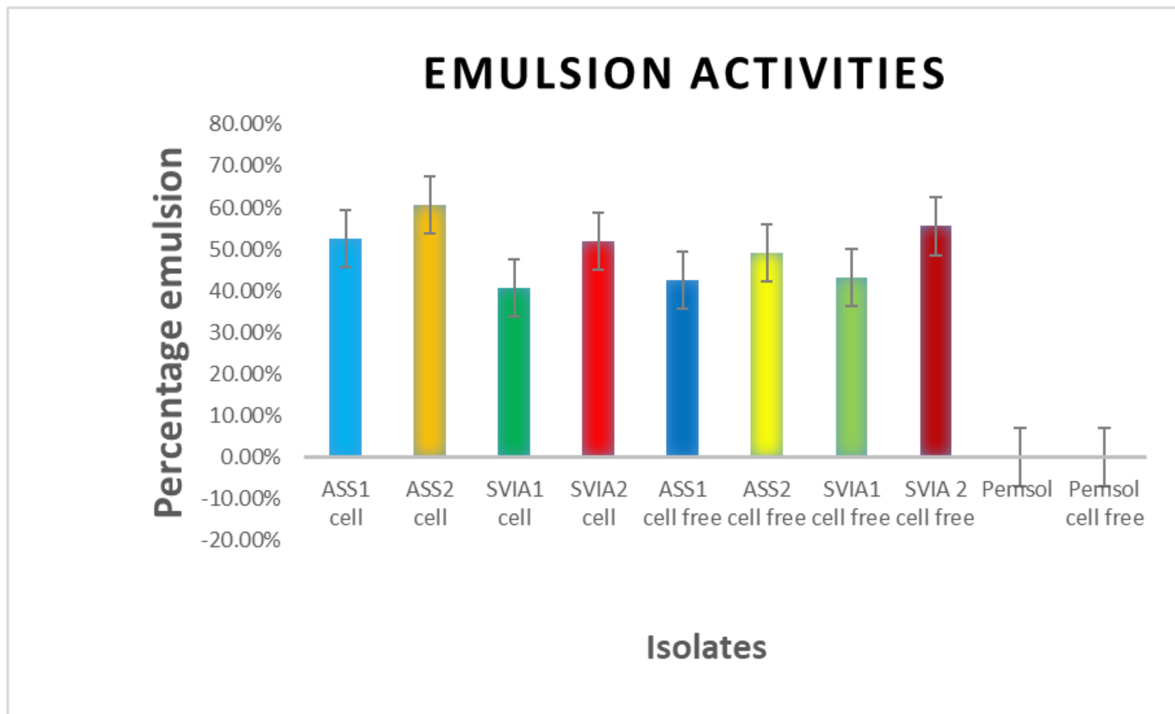
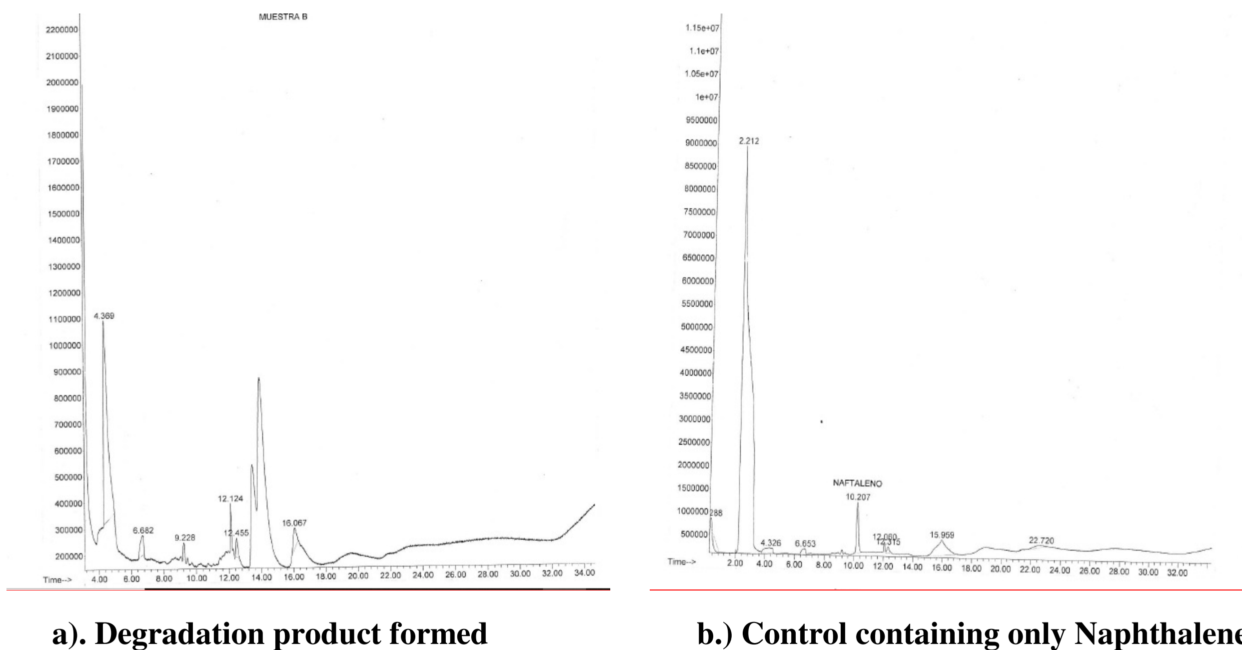


Figure 3. Bio-emulsification activity of *Stenotrophomonas* sp. Pemsol



a). Degradation product formed **b.) Control containing only Naphthalene**

Figure 4. The GC- MS/MS analysis spectrum of the test experiment after the 30th day of degradation studies**.**

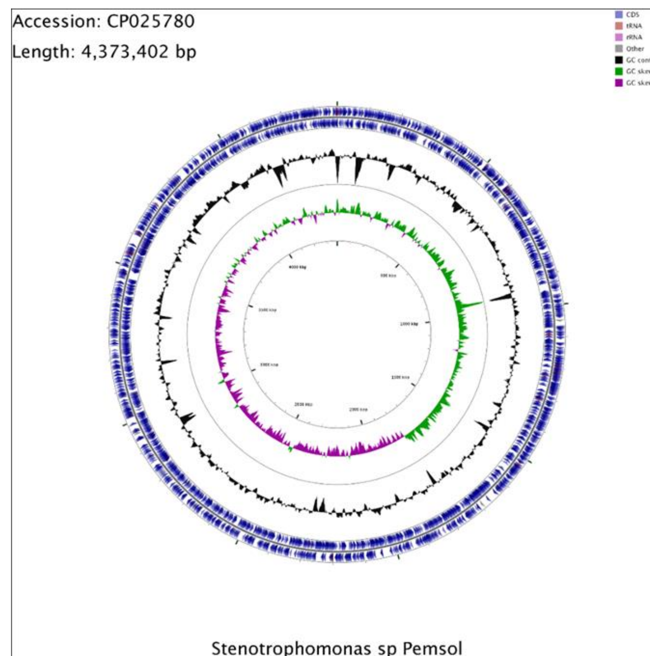


Figure 5. Circular Genome Map for *Stenotrophomonas* sp. Pemsol

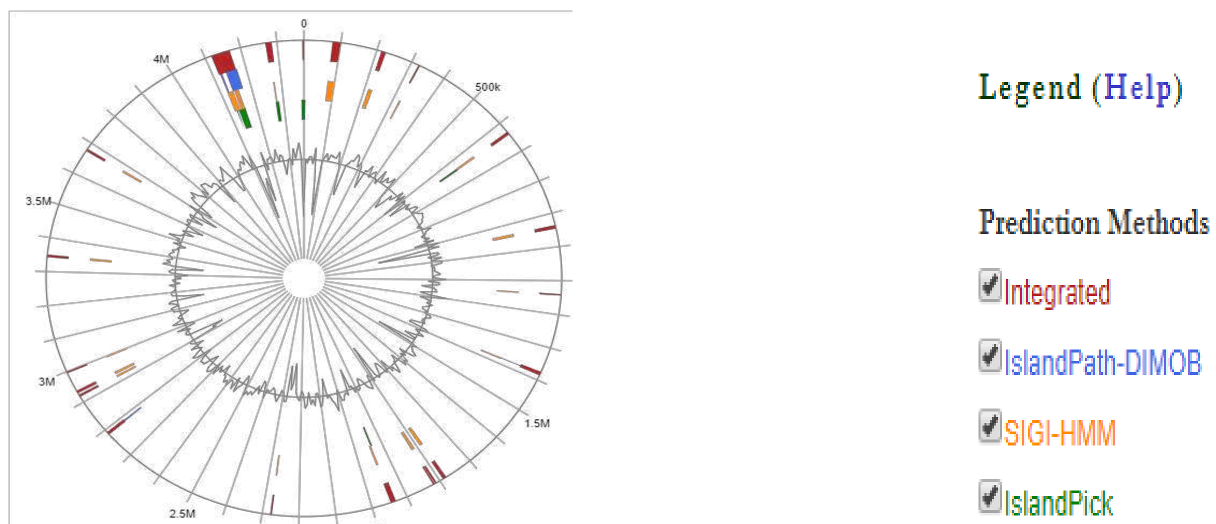


Figure 6. Genomic Island distribution in *Stenotrophomonas* sp. Pemsol

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27617v1 | CC BY 4.0 Open Access | rec: 28 Mar 2019, publ: 28 Mar 2019

PeerJ Preprints NOT PEER-REVIEWED

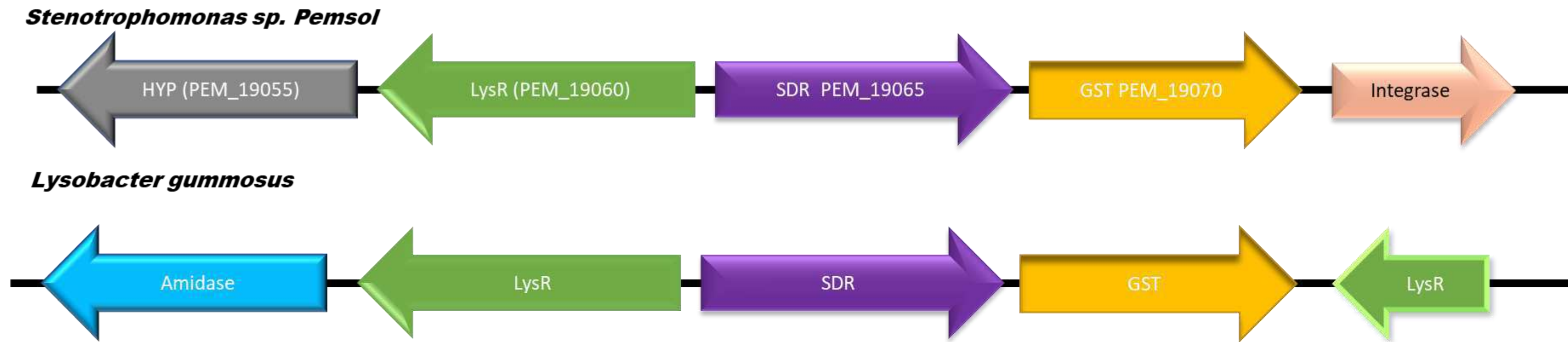


Figure 7. LysR and SDR cluster in *Stenotrophomonas* sp. Pemsol and its comparison with *Lysobacter gummosus*

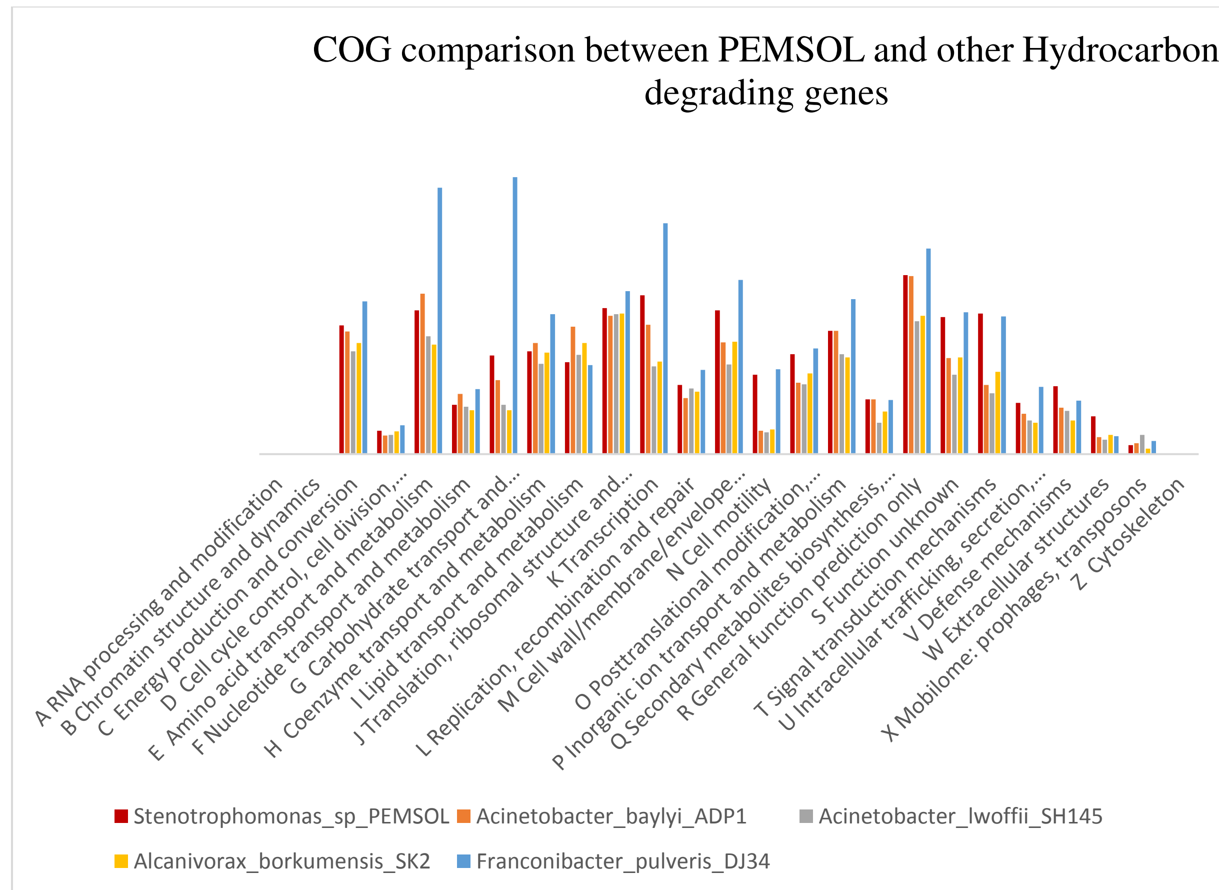


Figure 8 . COG Distribution comparison of Pemsol and other hydrocarbon degrading bacteria

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27617v1 | CC BY 4.0 Open Access | rec: 28 Mar 2019, publ: 28 Mar 2019