https://doi.org/10.7717/peerj.8102
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 A Villalobos-Lopez 7, Xianwu Guo Corresp. 1

1 Laboratorio de Biotecnología Genómica, Centro de Biotecnología Genómica, 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, Osogbo, 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.
*Stenotrophomonas* sp. Pemsol isolated from crude oil contaminated soil in Mexico that can degrade polycyclic aromatic hydrocarbons and its whole genome sequence analyzed

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

\(^1\) Centro de Biotecnología Genómica, Instituto Politécnico Nacional, 88710, Reynosa, México

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

\(^3\) Centro de Investigación en Dinámica Celular, Instituto de Investigación en Ciencias Básicas y Aplicadas, Universidad Autónoma del Estado de Morelos (UAEM), 62209, Cuernavaca, Morelos, México

\(^4\) Laboratorio de Estudios Ecogenómicos, Centro de Investigación en Biotecnología, Universidad Autónoma del Estado de Morelos (UAEM), 62209, Cuernavaca, Morelos, México

\(^5\) Centro de Investigación en Biotecnología Aplicada, Instituto Politécnico Nacional, Ex-Hacienda San Juan Molino, carr. est. Tecuexcomac-Tepetitla Km. 1.5, 90700, Tepetitla de Lardizábal, Tlaxcala, México

*Corresponding author: xguo@ipn.mx*
Abstract

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 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 bioemulsification 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.

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

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

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

Bacterial remediation of PAHs involved the activities of some genes encoding for oxygenases or peroxidases. Some of the genes associated with the degradation of PAHs are alkane
monooxygenases, such as alkB from *Pseudomonas*; alkm from *Acinetobacter* sp. Strain, ADP-1; alkB1 and alkB2 from *Rhodococcus* sp. Other genes were xylE, catechol-2, 3 dioxygenases from *Pseudomonas putida*; ndoB, naphthalene monooxygenase from *P. putida*; and nidA, pyrene dioxygenase large subunit from *Mycobacterium* sp. strain PYR-1, as well as various dehydrogenases and protocatechuate dioxygenases in *Stenotrophomonas* spp. (Gunsalus, 1951; Seo, Keum & Li, 2009; Urszula et al., 2009; Das & Chandran, 2011). Some PAHs degradation needs the help of biosurfactant produced by bacteria (Van Beilen & Funhoff, 2007; Fritsche & Hofrichter, 2008). The biosurfactants or surface-active substances decreases the surface tension on the surface of water molecules, thereby making entrapped PAH on surfaces available for the use of bacteria (Boonchan, Britz & Stanley, 1998).

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

In this study, we isolated a strain of *Stenotrophomonas* sp. Pemsol from crude oil contaminated soil in the state of Tabasco, Mexico, which showed good potential to degrade several PAHs. The sequence analysis of *Stenotrophomonas* sp. Pemsol revealed that it is rich in genes required for the
 degradation of PAHs and other hydrocarbons. The aim of the study is to elucidate and understand
the genetic basis involved in the uptake and degradation of PAHs in *Stenotrophomonas* sp. Pemsol.

2. Material & Methods

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

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

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

Genomic DNA was extracted from 5 mL Bacterial culture grown in Luria broth using Promega
wizard genomic DNA purification kit (Promega, Madison, USA) as per the manufacturer’s
instruction. The 16S rRNA genes were amplified by PCR using steno1 (5’ AGG GAA ACT TAC
GCT AAT ACC- 3’) and steno2 (5’ CTC TGT CCC TAC CAT TGT AG-3’). The PCR mix contains
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
µL (0.5 µM) of each primer, 0.75 µL (50 mM) MgCl₂, 16.75 µL double distilled water and 2 µL
DNA (10 ng/µL). PCR products were purified and sequenced at the Centro de Biotecnologia
Genomica, Instituto Politecnico Nacional (IPN), Mexico using the ABI 3130 sequencing machine
for species identification. 16S rRNA gene sequences were analyzed with Seqman software version
13 and subjected to similarities search against those retrieved from NCBI using Blastn program
(https://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequences were aligned with related sequences retrieved
from NCBI database using Mega 6.0 (Tamura et al., 2013) and phylogenetic tree was constructed using neighbor-joining algorithm. Reliability of tree topologies was confirmed by bootstrap analysis using 1000 repeat alignment. *Stenotrophomonas sp.* Pemsol sequence has been deposited on NCBI with accession number, KX500117.1.

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

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

2.3 *Identifying metabolic intermediates from the degradation of naphthalene*
Extraction of naphthalene and other hydrocarbon from culture media was performed using equal volume of hexane by triplicate. Then hexane was eliminated with vacuum pressure for further the analysis (FTIR, UPLC-MS and GC-MS).

2.3.1 Fourier-transform infrared spectroscopy (FTIR)

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

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

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

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

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

2.4.1 Genome assembly and annotation

The reads quality was checked with Fastqc (Andrews, 2010) and the adaptors from the raw reads were trimmed with trim-galore version 4.10 which also filtered out reads with poor quality. De novo genome assembly was carried out with a standalone Spades 3.11.1 genome assembler (Center for Algorithmic biotechnology, St. Petersburg State University, Russia) (Bankevich et al., 2012). The assembly’s quality was checked with QUAST (Gurevich et al., 2013). The assembled contigs were ordered and reduced into a single scaffold with MedusaCombo, an online genome multidraft scaffold (Bosi et al., 2015). The assembled genome was annotated with Prokka annotating pipeline (Seemann, 2014). Further functional genome annotation was done with online genome analysis server WebMGA (http://weizhong-lab.ucsd.edu/metagenomic-analysis) (Wu et al., 2011). WebMGA was used to predict the KEGG functions and COG categories present in the genome. The presence of transposon and insertion sequences was predicted with an web based analysis tool software ISsaga (http://issaga.biotoul.fr/ISsaga2/issaga_index.php) (Varani et al., 2011). The Pan core genome analysis for Stenotrophomonas sp. Pemsol and 12 other Stenotrophomonas species to identify the unique genes in Stenotrophomonas sp. Pemsol. The Stenotrophomonas genome compared with Stenotrophomonas so. Pemsol were the finished genome from the genus available on NCBI database as of the time of this write up. (These species include; _S. maltophilia_ JV3, _S. maltophilia_ ASS1, _S. pavani_ LMG, _S. rhizophilia_ QLP4, _S. pictorium_ JCM 9942, _S. maltophilia_
K279a, *S. nitrireducen* 2001, *S. panacihumi, S. maltophilia* ATCC 19687, *S. maltophilia* R551-3, and *Stenotrophomonas* sp.). We retrieved the unique genes from *Stenotrophomonas* sp. Pemsol genome and subject them to further analysis. This include manual blast search analysis of the genes on NCBI database to determine what they encode. We also carried out the synteny analysis of the genes on SyntTax synteny and RAST annotation server.

2.4.2 Prediction of genomic island

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

2.4.3 Comparative genome analysis

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

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

3.0 Results and Discussion

3.1 Isolation and identification of the *Stenotrophomonas* strain Pemsol from crude oil contaminated soil
The objective of this work was to isolate a *Stenotrophomonas* strain that could be used for bioremediation of oil-polluted soil. This is why we used the StenoVIA agar medium for selection of *Stenotrophomonas* strains (Kerr et al., 1996). Several uniform colonies with characteristic yellow color appeared on the selective medium after 48 h of incubation. The fragment of 16S RNA gene of one clone called Pemsol was amplified and sequenced with a pair of primers, steno1 and steno2. The BLAST search of this sequence and phylogenetic analysis with other known strains showed that Pemsol belongs to the genus *Stenotrophomonas*. The closest species being *Stenotrophomonas maltophilia M27*, with 99% identity as shown in Figure 1.

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

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

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

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

### 3.4 Analysis of degradation products using FTIR Spectroscopy

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

### 3.5 UPLC-MS and Gas Chromatography-MS analysis of degradation products

The UPLC-MS and GC-MS analysis were performed to detect the metabolites formed from the degradation of naphthalene after 30-day experiment. The absence of a peak corresponding to naphthalene on the spectra obtained from UPLC-MS and GC-MS analysis confirmed the
degradation of naphthalene, after comparing with the control (Figure 4, S1 (a-b) see supplementary file). Meanwhile, a peak occurred with molecular weight of 109.98 as the major metabolite, estimated to be C_6H_5OH correspondent to the molecular weight for catechol (Figure S4-S5, see supplementary material). It could thus be inferred that the degradation of naphthalene by Pemsol is associated with the formation of catechol. Several studies have reported the degradation of naphthalene and other PAHs to involve the formation of catechol or protocatechuate (Smith, 1990; Johnsen, Wick & Harms, 2005). Although *Stenotrophomonas* sp. Pemsol did not emulsify hydrocarbon, the metabolite analysis showed that it successfully degrade naphthalene.

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

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

*Stenotrophomonas* sp. Pemsol has 147 genes associated with the degradation of PAHs in its genome (Table S1). In this category, nine genes encode enzymes belonging to the lactoylglutathione lyase family (PEM_01474; 01733; 01959; 02738; 02855; 02960; 02961; 03007; 03383) in COG0346, which have been reported to be involved in the degradation of aromatic compound (Mesarch, Nakatsu & Nies, 2000). One of the lactoylglutathione (PEM_03383) was predicted to be a catechol 2, 3 dioxygenases. This gene is essential for the conversion of salicylate aldehyde to catechol in the naphthalene’s metabolic degradation pathway (Grund, Denecke &
Eichenlaub, 1992). It also has a gene encoding salicylate hydroxylase (PEM_02405) (COG0654) (EC:1.14.13.1) (nahG), which converts salicylic acid, an intermediate in the degradation of naphthalene, to catechol by removing the carboxyl group at position 1 and introducing a hydroxyl group in the same position as replacement (Goyal & Zylstra, 1997; Bosch et al., 1999). Two genes, homogentisate 1,2-dioxygenase (PEM_03309) that are involved in the catabolism of aromatic rings (Borowski, Georgiev & Siegbahn, 2005), and 2, 4 dihydroxyacetophenonedioxygenase (PEM_00137), which helps in the cleavage of carbon-carbon bond in a substituent aromatic ring, were also detected (Keegan et al., 2014). The above-mentioned information ratifies the capacity of Stenotrophomonas sp. Pemsol to degrade naphthalene as shown in the experiment.

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

Pemsol has a phosphomannomutase (PEM_00211), and two other glycosyl 2 family transferase genes which are needed for biosurfactant production (Pal et al., 2017). The absence of other genes essential for biosurfactant production may be the reason for its inability to emulsify PAHs in our
experimental studies. Thus, in *Stenotrophomonas* sp. Pemsol, the degradation of PAHs without emulsification has a genetic basis. Unlike in *Pseudomonas aeruginosa* N002, this strain has 25 genes for bio surfactant formation and regulation and can bio-emulsify PAHs.

### 3.7 Specific PAH degradation genes in genomic islands (GI)

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

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

A pan-core genome analysis of *Stenotrophomonas* sp. Pemsol with 12 complete genomes of other *Stenotrophomonas* species was performed. The result of pan-genome analysis showed that *Stenotrophomonas* sp. Pemsol possesses 154 unique genes. Most genes identified to be unique were part of the genes in *Stenotrophomonas* sp. Pemsol genomic island. The predicted functions for these genes are shown in Table S6. Some of these genes are involved in the degradation of PAH (Ghosal, et al., 2016). For example, short dehydrogenase reductase (SDR) can catalyze the reduction of C=C bond between aromatic compound (Kavanagh et al., 2008). SDR gene and two
genes, encoding LysR-type transcriptional regulators (LTTRs, PEM_19060) (accession genome number) and glutathione S-transferase (GST), exist in a gene cluster on the genome (Figure 8). Further, the 3 proteins SDR, LTTR and GST showed the same gene order as the closest orthologs with the identity of 91.6%, 94.24% and 85.65% respectively from Lysobacter gummosus after blast analysis, implying that this gene cluster in the two species has a common origin and Pemsol could have obtained this gene cluster by horizontal gene transfer (Figure 8). The KEGG database clearly showed that GST directly involved in the degradation of many hydrocarbon compounds (benzoapyrene, naphthalene, trichloroethylene, bromobenzene, etc.). The LTTR family protein has been reported to have significant function in regulating genes that are important for the catabolism of aromatic compound, cell motility and quorum sensing (Pal et al., 2017). Thus, LysR gene in this gene cluster could be involved in the regulation of SDR and GST for PAH degradation to help the survival of Stenotrophomonas sp. Pemsol in this crude contaminated environment.

3.9 Comparative COG category analysis of Stenotrophomonas sp. Pemsol

The COG categories in Stenotrophomonas species was compared with the COGs in Pemsol. Stenotrophomonas sp. Pemsol has higher number of genes in some COG categories than the other 12 Stenotrophomonas species compared. The Fischer test statistical analysis of the COG categories showed that Stenotrophomonas species Pemsol has in abundance more genes in the categories energy production and conversion (C) (6.01%), amino acid transport and metabolism coenzyme transport and metabolism (H) (6.78%), cell motility (N) (3.74%), secondary metabolite biosynthesis, transport and metabolism (Q) (2.58%), general function prediction (R) (8.4%), function unknown (S) (6.4%), signal transduction (T) (6.52%), defense mechanism (V) (3.19%), extracellular mechanism (W) (1.77%), (Supplementary file 2). The abundance of the genes in these
The COG categories in Stenotrophomonas species Pemsol was compared with the COGs in five other PAH degrading bacteria (Acinetobacter baylyi_ADP1, Acinetobacter lwofii_SH145, Alcanivorax borkumensis_SK2, Franconibacter pulveris_DJ34) previously reported. The Fisher exact test statistics for the comparison of the COG categories in Stenotrophomonas sp. Pemsol and four other hydrocarbon-degrading bacteria (Acinetobacter baylyi_ADP1, Acinetobacter lwofii_SH145, Alcanivorax borkumensis_SK2, Franconibacter pulveris_DJ34) revealed that Pemsol has more genes in the COG category G, N, T and W than the other hydrocarbon-degrading bacteria (figure 5). These categories are associated with carbohydrate transport and metabolism (4.54%), cell movement (3.74%), signal transduction mechanisms (6.52%) and extracellular structure (1.75%), which could help Stenotrophomonas sp. Pemsol to metabolize hydrocarbons, sense the signals from environment and respond appropriately to the environmental stimuli, for example, moving towards the useful or away from the harmful components of the PAHs.

4.0 Conclusion

Stenotrophomonas sp. Pemsol was isolated from crude oil contaminated soil from Tabasco, Mexico. It grew in the presence of six PAHs (biphenyl, anthracene, anthraquinone, phenanthrene, naphthalene and phenanthridine) as unique carbon source. The identification of Pemsol confirmed that it is a member of the genus Stenotrophomonas and that its closest relative in the genus is Stenotrophomonas maltophilia M27. The ability of Pemsol to degrade PAH was confirmed by its degradation activities on naphthalene as revealed by FTIR, UPLC-MS and GC-MS analysis. The complete genome analysis of Pemsol revealed that it possesses many genes that are involved in
the degradation of PAHs but only 3 gene associated with bioemulsification, leading to no biosurfactant production. The presence of some genes associated with the degradation of PAHs in the genomic islands inferred that those genes were horizontally acquired. Comparing with other four sequenced hydrocarbon-degrading bacteria, Pemsol is much rich in genes for the COG category G, N, T and W, which are mainly relevant to hydrocarbon utilization and interaction with environment. These results give insight into the genetic basis involved in the survival of Pemsol in its oil-contaminated site and provide guide on the possible strategies for the bioremediation of an oil-polluted environment with *Stenotrophomonas* sp. Pemsol without biosurfactant production.
References


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
<table>
<thead>
<tr>
<th>Features</th>
<th>Genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA, total number of bases</td>
<td>4,373,402</td>
</tr>
<tr>
<td>DNA coding number of bases</td>
<td>4,370,616</td>
</tr>
<tr>
<td>DNA G + C content (%)</td>
<td>66.59%</td>
</tr>
<tr>
<td>Misc_RNA</td>
<td>39</td>
</tr>
<tr>
<td>Protein coding genes</td>
<td>3905</td>
</tr>
<tr>
<td>rRNA genes</td>
<td>4</td>
</tr>
<tr>
<td>tRNA genes</td>
<td>70</td>
</tr>
<tr>
<td>tmRNA</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 1 (on next page)

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