



Biodiversity and Phylogenetic Relationship of Total Hydrocarbon Degrading Genes in Selected Bacteria Species

R. B. Agbor^{1*}, S. P. Antai² and S. E. Ubi³

¹*Environmental Biotechnology Unit, Department of Genetics and Biotechnology, Faculty of Biological Science, University of Calabar, Calabar, Nigeria.*

²*Department of Microbiology, Faculty of Biological Science, University of Calabar, Calabar, Nigeria.*

³*Department of Civil Engineering, Cross River University of Technology, Calabar, Nigeria.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJOB/2021/v12i330164

Editor(s):

(1) Dr. Jehad m. H. Ighbareyeh, Al-quds Open University, Palestine.

Reviewers:

(1) Patricia Canteri de Souza, State University of Londrina, Brazil.

(2) Nawar Jasim Alsalih, Al Muthanna University, Iraq.

Complete Peer review History: <https://www.sdiarticle4.com/review-history/70907>

Original Research Article

Received 10 May 2021
Accepted 15 July 2021
Published 24 July 2021

ABSTRACT

Hydrocarbons which forms the bulk of soil and water pollutants in the Niger Delta region of Nigeria differs in their susceptibility to microbial attack and degradation. Considering the importance of hydrocarbons in the economy of any nation, a complete knowledge of the hydrocarbons degrading capabilities of associated microbial species becomes increasingly important and indispensable. The research was aimed at assessing the biodiversity and phylogenetic relationship of hydrocarbon degrading genes in selected bacteria species. Sequences of nucleotides and amino acids of hydrocarbons degrading genes in 12 species of bacteria such as *Pseudomonas aeruginosa* I & II, *P. stutzeri*, *Thalassospirillum* spp. I & II, *Alcavorex* spp., *Arthrobacter* spp., *Martellella* spp., *P. taenensis*, *Aneuribacillus* species, *Rhodococcus* spp. and uncultured bacteria, were retrieved from the National Center for Biotechnology Information (NCBI) and analysed for their variability in physicochemical properties, percentage identity and similarity, G-C content, secondary and tertiary structures, their biodiversity and their phylogenetic relationship using MEGA 6 soft-wares, GOR IV, Phyre, Genscan and SIB Protparam. The analysis showed great genetic diversity and polymorphism in the

*Corresponding author: E-mail: agborreagan@yahoo.com;

hydrocarbon degrading genes present in the bacteria species. The phylogenetic tree showed that the genes in most species had undergone mutation. *Martella spp* had the highest probability or chance of mutating with E-value of $7e-129$. Only two of the bacteria species had low instability index indicating that majority of the bacteria species contains or possessed the unstable nature of the hydrocarbon degrading genes. The biodiversity of bacterial species and their genetic distance was ascertained.

Keywords: Polymorphism; genetics; bioinformatics; sequence; environment.

1. INTRODUCTION

It is estimated that the annual global production of petroleum is between 1.7 and 8.8 million metric tons, the majority of which is derived from anthropogenic sources. Biodegradation of hydrocarbons by natural populations of microorganisms represents one of the primary mechanisms by which petroleum and other hydrocarbon pollutants are eliminated from the environment. Hydrocarbon pollution is a major challenge to the global economy, especially in oil producing nations [1]. The sustainability of mankind in the universe is dependent on the quality of the environment. Pollution infested ecosystem tend to affect the overall performance of the microflora and fauna which invariably alter the food chain and thus has a gross effect on the population. The adverse growth of oil industries in Africa, and the distribution process of hydrocarbon products has made oil pollution a major problem of environmental concern. However, during the loading and off-loading process, lots of the products are spill into the tertiary and aquatic environment. Today, some of the plants of economic importance and aquatic species are extinct because of the resultant effect of oil spillage. The pollution of soils by hydrocarbon products stimulate indigenous microbial populations, which are capable of utilizing the hydrocarbons as source of carbon and energy, thereby, degrading the pollutants [1].

The degradability of the hydrocarbons is dependent on the specific abilities of microorganisms to utilize the hydrocarbons as their sole carbon source [2]. According to [3] high levels of heterotrophic and hydrocarbon utilizing microorganisms were found in abundance in Imo river estuary [4]. Reported on the biological efficacy of the leaves of *Tithonia diversifolia* and *Calapogonium mucunoides* in the enhancement of microbial degradation of crude oil polluted soils [5]. Examined the bacterial diversity in tropical crude oil polluted soils and concluded that actinobacteria possesses strong bioremediation influence in polluted soils [6,7-8].

Identified the roles of some Gram positive as strong hydrocarbon degraders. This study was aimed at examining the biodiversity and phylogenetic relationship of total hydrocarbon degrading genes in selected bacteria species.

2. MATERIALS AND METHODS

2.1 Retrieval of Nucleotides and Amino Acid Sequences

The nucleotides and amino acid sequences of total hydrocarbon degrading genes in bacteria. A plastid –coding gene in bacteria which is associated with the degradation of crude oil and total hydrocarbon contents in the microfauna cutting across numerous bacteria species. However, notably amongst the bacteria species are *Pseudomonas aeruginosa*, *Pseudomonas aeruginosa* II, *Pseudomonas stutzeri*, *Thalassospire* spp, *Alcanivox* sp., *Arthrobacter* sp, *Martelella* sp, *Pseudomonas* sp, *Aneurirubacilus* sp, *Thalassospire* sp II, *Rhodococcus* sp. and uncultured bacterium. The nucleotides and amino acid sequence were downloaded from the Gene bank. This was done by obtaining the FASTA format for the nucleotides and amino acids sequences of the legumes form the National Centre for Biotechnology Information (NCBI, USA) databases site.

The basic alignment search tool (BLAST) was used to obtain similar sequences or predicted sequences of uncultured relatives of the bacteria species.

The accession numbers of the sequences retrieved for the various bacteria were recorded, the protein and gene names, sequence length as well as the organism names were retrieved and pasted in a notepad using the FASTA format option.

Pairwise and multiple sequence alignments were carried out to align all retrieved sequences using MEGA 6 software as modified by [9].

2.2 Determination of Percent Identity and Similarity (Homology)

Percentage identity and similarity among the nucleotides and amino acid sequences of the retrieved total hydrocarbon degrading genes in *Pseudomonas aeruginosa*, *Pseudomonas aeruginosa* II, *Pseudomonas stutzeri*, *Thalassospire* sp, *Alcanivox* sp., *Arthrobacter* sp, *Martelella* sp, *Pseudomonas* sp, *Aneurirubacilus* sp, *Thalassospire* sp II, *Rhodococcus* sp. and uncultured bacterium were determined by homology comparison tool for more than two sequences option of the basic alignment search tool.

2.3 Determination of Phylogenetic and Evolutionary History of Hydrocarbon Degrading Genes in Selected Bacteria Species

The phylogenetic analysis and evolutionary history were determined using the Molecular Evolution and Genetic Analysis (MEGA 6) software with maximum likelihood option for the construction of phylogenetic tree for *Pseudomonas aeruginosa*, *Pseudomonas aeruginosa* II, *Pseudomonas stutzeri*, *Thalassospire* spp, *Alcanivox* sp., *Arthrobacter* sp, *Martelella* sp, *Pseudomonas* sp, *Aneurirubacilus* sp, *Thalassospire* sp II, *Rhodococcus* sp. and uncultured bacterium using their MEGA aligned retrieved sequences from the NCBI databases.

“The evolutionary history or pathway was traced using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) option of analysis based on the Jones –Taylor-Thompson (JTT) matrix – based model”.

The reliability of the inferred phylogenetic tree was evaluated using the Bootstrap analysis of 500 replications. The time of divergence or evolutionary history of petroleum degrading genes of the bacteria for each species was estimated based on the nucleotides percent substitution obtained per site.

2.4 Determination of Physico-chemical Properties of the Hydrocarbon Degrading Genes in Selected Bacteria Species

The physico-chemical properties of the petroleum degrading genes for selected twelve (12) bacteria species were determined using the

Expert Protein Analysis System (EXPASY) which is a proteomic server of the Swiss Institute of Bioinformatics (SIB) using the following operations of the EXPASY site.

The selected bacteria species which was informed by the availability of their nucleotides and amino acid sequences in the NCBI databases with reasonable sequence length. The FASTA format of the nucleotides and amino acid sequences of the hydrocarbon degrading genes earlier retrieved and aligned were used for the physico-chemical analysis. Important physico-chemical properties and data for the hydrocarbon degrading genes such as

- i. Theoretical PI,
- ii. Molecular weight of the mat k genes,
- iii. Number of amino acid residues,
- iv. Amino acid composition,
- v. Atomic composition
- vi. Total number of negatively charged residue (arginine + lysine)
- vii. Extinction Co-efficient (m/cm) at 280nm wavelength
- viii. Instability index
- ix. Aliphatic index
- x. Grand average hydrophobicity

Were deduced from the Expasy software information after analysis using the ProtParam (Protein physical and chemical parameters).

2.5 Determination of Predicted Protein Motifs and Structures for Hydrocarbon Degrading Genes for Selected Bacteria Species

The amino acid sequences of petroleum degrading genes of *Pseudomonas aeruginosa*, *Pseudomonas aeruginosa* II, *Pseudomonas stutzeri*, *Thalassospire* sp, *Alcanivox* sp., *Arthrobacter* sp, *Martelella* sp, *Pseudomonas* sp, *Aneurirubacilus* sp, *Thalassospire* sp II, *Rhodococcus* sp. and uncultured bacterium were predicted with the prosite software by using the software for the prediction of motif for secondary structure was achieved using GORIV software as modified by [9].

The motif for the predicted tertiary structure (3D protein structure) for hydrocarbon degrading genes was obtained using the Phyre 2 (Protein Homology Analysis Recognition Engine) software based on Canonical amino acid sequence earlier retrieved from the NCBI databases and modified by [10]. The Rasmol (Raswin) software was used to fine tune the 3D protein structure to ribbons with desired colours and magnitudes.

2.6 Determination of Start and End Codons, and C-G Content of Hydrocarbon Degrading Genes for Selected Bacteria Species

The start and end codons for the mat k genes of the selected bacteria species that is, the putative regions was determined using the GENSCAN software as modified by Burge, 2011. The GENSCAN software was also used to determine the Cytosine – Guanine (C-G) content for each amino acid sequence for each bacteria species [11].

3. RESULTS

3.1 Retrieval of Nucleotide and Amino Acid Sequence

The FASTA format of the nucleotides and the amino acid sequences of total hydrocarbon degrading genes from 12 bacteria species were retrieved from the National Centre for Biotechnology Information (NCBI).

3.2 Determination of Percentage Identity and Similarity

The percentage identity obtained as presented in Table 2 revealed that the highest identity was observed between the hydrocarbon degrading genes as found in *Pseudomonas aeruginosa I and II* with 92% identity while the least identical species with the gene was *Martelella* species sowing 69% identity. On the other hand percent similarity presented in Table 3 showed that highest similarity in the hydrocarbon degrading genes was shared by *Pseudomonas aeruginosa I and II* with each species recording 87 percent similarity. The least percent similarity was however observed in uncultured bacteria which showed only 61 percent similarity. The E-value shows the probability and chances of mutation occurring in the hydrocarbon degrading genes in the bacteria species. As shown in the Table 1 below, the highest mutation rate in the hydrocarbon degrading genes was observed in *Martelella species* with E-value of $7e - 129$ while the least E-value was obtained in *Arthrobacter* species with E-value of $2e-115$.

3.3 Evolutionary Relationships of Taxa

The analysis involved 12 amino acid sequences (Table 2). All positions containing gaps and missing data were eliminated. There were a total of 119 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 [12].

Table 1. Percent identity, percent similarity and E-value for total hydrocarbon degrading genes for selected bacteria species

Bacteria species	E-value	% identity	% Similarity	GAPs (%)
<i>Alcanivorax</i> species	3e – 165	82	81	0.0
<i>Aneurinibacillus</i> species	3e – 114	82	85	0.0
<i>Arthrobacter</i> species	2e – 115	84	78	2.0
<i>Martelella</i> species	7e - 129	69	64	0.0
<i>Pseudomonas aeruginosa</i> II	3e – 141	92	87	2
<i>Pseudomonas stutzeri</i>	4e - 137	84	84	0.0
<i>Pseudomonas taeanensis</i>	5e – 159	90	86	3.0
<i>Thalassospire</i> species	4e – 125	83	86	1.0
<i>Thalassospire</i> species II	2e – 135	76	76	2
<i>Rhodococcus</i> species	4e – 102	92	83	0.0
Uncultured bacterium	3e – 161	84	61	1

Table 2. Hydrocarbon Degrading gene bacteria species, Accession number, number of nucleotides and number of amino acid sequences retrieved

Bacteria species	Accession Number	Amino acid number	Nucleotide sequence
<i>Alcanivorax</i> sp.	EUC70824.1	402	1026
<i>Aneurinibacillus</i> sp.	AMA74041.1	102	306
<i>Arthrobacter</i> sp.	ALQ29439.1	107	321
<i>Martelella</i> sp.	AMM83862.1	351	1053
<i>P. aeruginosa</i> I	KUG316861.1	292	876
<i>P. aeruginosa</i> II	KUG31686.1	241	723
<i>Pseudomonas stutzeri</i>	AKN2795.1	312	936
<i>P. taenensis</i>	KFX71301.1	443	1329
<i>Rhodococcus</i> sp.	AJE25657.1	122	366
<i>Thalassospira</i> spp.II	KJE34172.1	361	1083
Uncultured bacteria	AID55599.1	142	426

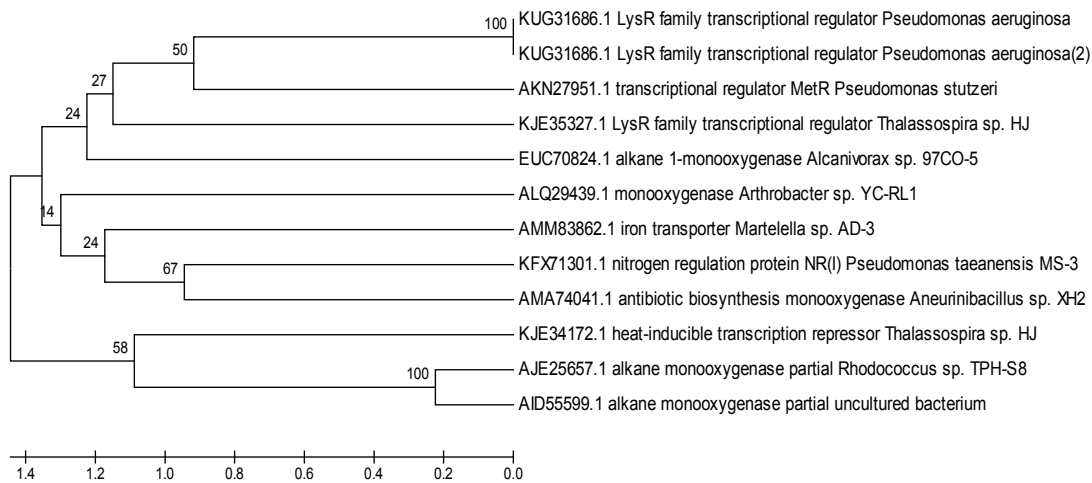


Fig. 1. Phylogram showing evolutionary pathway of twelve species of bacteria with hydrocarbon degrading gene

Table 3. G-C contents, stop codons and predicted peptides and CDS of Nucleotide sequence of hydrocarbon degrading genes in selected bacteria species

Bacteria Species	G-C content (%)	Ploy A+ tail	Ploy A-tail
<i>Alcanivorax</i> species	26.14	231 -236	145 – 149
<i>Aneurinibacillus</i> species	21.12	234-239	261 - 256
<i>Arthrobacter</i> species	32.43	204-209	238 -233
<i>Pseudomonas aeruginosa</i>	28.43	655 - 660	288 - 283
<i>Pseudomonas aeruginosa</i> II	28.25	653 - 658	323 - 318
<i>Pseudomonas stutzeri</i>	34.73	215 - 220	338 - 333
<i>Martelella</i> species	26.78	313 – 318	235 - 230
<i>Pseudomonas taenensis</i>	25.98	255-260	256 – 251
<i>Rhodococcus</i> species	42.22	222 - 227	184 – 179
<i>Thalassospire</i> species	40.75	212 -217	294 - 289
<i>Thalassospire</i> species II	34.37	307 – 312	422 - 417
Uncultured bacterium	32.09	245 - 250	394 – 389

3.4 Inferring Time of Evolutionary Divergence among Hydrocarbon Degrading Genes in Bacteria Species

The divergence time among the different species of polycyclic hydrocarbon degrading bacteria was evaluated using the phylogenetic inferred tree as presented on an evolutionary or phylogenetic tree as shown in Fig. 1 above. The phylogenetic tree of the hydrocarbon degrading gene for the different bacteria species based on the differentiation of the hydrocarbon degrading gene agrees with the taxonomy of the NCBI. The smallest time of divergence of the gene among the bacteria species was zero million years ago (MYA) was observed in most closely related species like *Pseudomonas aeruginosa* I and II species (Table 3). However, from the lowest common ancestor, the average time of divergence observed amongst *Rhodococcus* species and the uncultured bacteria was about 12 MYA. The largest time of divergence observed among the hydrocarbon degrading genes in the bacteria species was 85 MYA as inferred from the phylogram above.

3.5 Physicochemical Properties

The physicochemical properties of the hydrocarbon degrading gene in the selected bacteria species as presented in Table 5 revealed that the highest theoretical pI was observed among *Thalassospiraspp* I with 9.80 pI. While in hydrocarbon degrading gene in *Arthrobacillus spp* recorded the least pI value of 9.01. The molecular weights of *Pseudomonas stutzeri* was greater than the molecular weights of the hydrocarbon degrading genes in *P. aeruginosa* I and others with the lowest molecular weight observed in *Martellella species*. Total number of amino acids was highest in *Pseudomonas taenensis* with 443 Daltons greater than all other pulses evaluated with the least amino acids of 102 Daltons founds in *Aneurinibacillus* species. The GRAVY value of *P. aeruginosa* I and *P. aeruginosa* II were the same (-0.023). The highest GRAVY value of 0.095 was obtained from hydrocarbon degrading gene in *Martellella* species while the least GRAVY value of -0.155 was obtained from *C. arietinum*. Total number of negatively charged protein residue (aspartate + glutamine) was similar in *Martellella* species and *Pseudomonas stutzeri* (Table 4).

The negatively charged parameter was however, highest in *Pseudomonas taenensis* (42) and was least (18) in *Martellella* species hydrocarbon degrading gene. Total number of positively

charged protein residue (arginine + lysine) was highest in *Arthrobacillus* spp. with (72) and was least (21) in *Pseudomonas stutzeri*. Total number of atoms numbering 4210 were obtained from uncultured bacterium while the smallest number of atoms of 3443 were counted from *Thalassospira* species I and II have the same number of atoms.

A high instability index of 55.94 was observed from hydrocarbon degrading gene in while the least instability index of 38.18 was obtained from *Arthrobacter* species. Aliphatic index was highest, 105.10 in *Thalassospira* species I and smallest index of 97.06 obtained from uncultured bacteria pulse species. Extinction coefficient was determined as presented in Table 5. It showed that *Aneurinibacillus* species had the highest coefficient of 96065 with *Rhodococcus* species showing the least coefficient.

3.6 Prediction of Secondary Structures of Hydrocarbon Degrading Genes in Bacteria Species

Prediction of secondary structures of hydrocarbon degrading gene in bacteria species using GOR IV software showed that the hydrocarbon degrading gene contained alpha helix (Hh), the extended strand (Ec) and the random coils (Cc) as presented in Table 5. The entire hydrocarbon degrading protein molecule showed distinct regions of alpha helices and extended strands which were separated by the random coils. The random coils were the highest elements or structures in the secondary structures of all the hydrocarbon degrading genes found in the bacteria species.

The region of secondary structure which made the alpha extended strand are shown in Fig. 2 below varies from one species to another.

3.6 Prediction of Tertiary Structure

The ribbon model of the tertiary structures of the hydrocarbon degrading gene in the cultured and uncultured bacteria species. All the illustrated tertiary structures were in the ribbon model to clearly show the areas of the structures covered and occupied by alpha helix, beta sheets, extended strand and random coil elements. Basic primary and secondary colours were used to depict the areas covered by the elements in the protein tertiary structures. Alpha helices were represented with pink coloured spiral sheet, blue coloured strands represented the random coil elements, green colour represented the extended strand while the yellow coloured region

of the protein represented the beta sheet elements of the proteins.

4. DISCUSSION

Crude oil been a non-renewable resource will diminish and finish when the hydrocarbon degrading genes in the bacteria species found around their deposit are totally exhausted or wipe out through evolutionary mutation and their functionality distorted through evolutionary changes. The implication of this inference to the species is that most of the hydrocarbon degrading gene in the bacteria species especially the closely related once with zero divergence time indicates that the genes are still intact and are yet to undergo serious mutation. This is an indication that their structure and functionality has been minimally distorted or disturbed by mutation throughout their evolutionary history. Sequence length variation observed in the hydrocarbon degrading gene in gene of the selected pulses must have resulted to differentiation and evolution. Sequence lengths variations is also a product of mutation caused by insertions and deletions as reported by [13]. The presence of similar amino acids residues in the amino acid sequence of the selected pulses is an indication that the hydrocarbon degrading gene in bacteria species have undergone sequence of evolutionary changes at these horizontal level. The percentage similarity which was very high amongst some of the pulse species was very high, which implies a high similar pattern of evolution and differentiation as earlier posited by [14]. In their studies on divergence time in legumes, had indicated that high similarity percentage in genes is an indication of high similar pattern of evolution.

The identity of retrieved sequences showing high identity of at least 70% which is an indication that the hydrocarbon degrading gene in the bacteria species has identical homology, function similarity and are the most highly conserved hydrocarbon degrading gene in *Thalassospira* species I in the selected bacteria species. The functionality of highly conserved genes increases with increasing sequence conservation as earlier demonstrated by [15]. Also as posited by [16], sequences found to possess more than 70% identity, would possess about 90% chances of sharing and performing similar functions and other biological processes. Studies by [17], have indicated that individual genes and proteins are generally assigned to families by the degree of their similarity in amino acid sequences which may be up to 30% identity

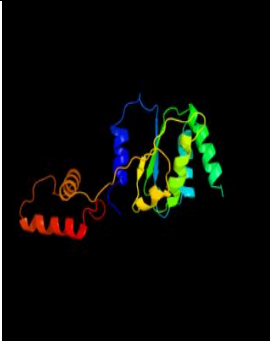

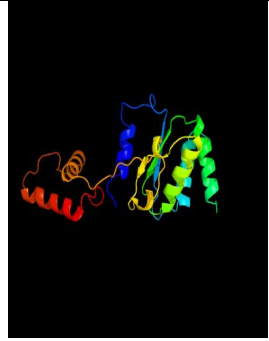
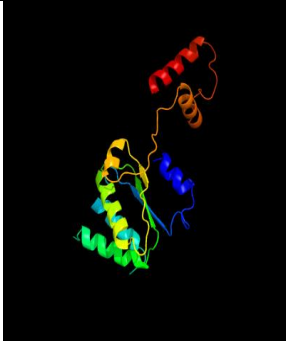
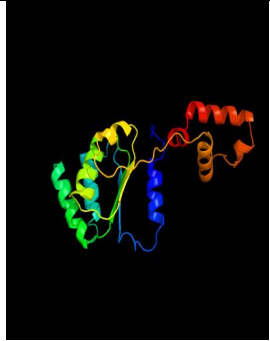
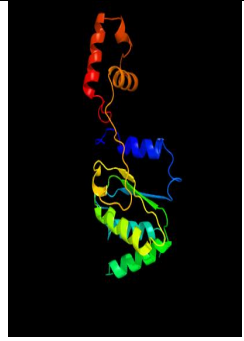

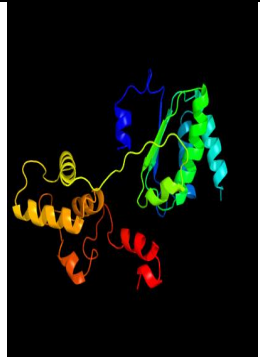
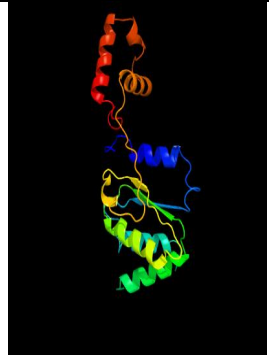

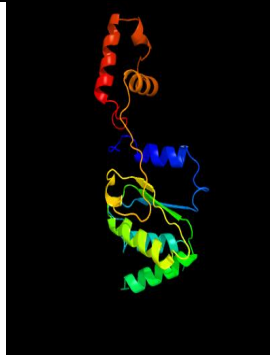
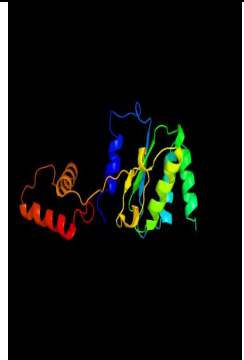
of the sequences or more which confers same or similar characteristics and structural functionality. The E – value is the number of matches that would be expressed by chance from which an E-value of 0.05 is significant which indicates that the similarity and identity among the selected bacteria species were significant for proper detection of characters and functions. Theoretical pI (Isoelectric potential) of all the selected pulses showed net charges since their pH was above the isoelectric point of 5.40 indicating that their protein carries negative charges. This characteristics is of immense importance in during molecular analysis.

Molecular weights of the hydrocarbon degrading gene in bacteria species in some of the species were quite relatively higher than those of others indicating their heaviness and high density of the gene in that species than in others. GRAVY score was high in some pulse species than in others. This also by implication showed that such species with high GRAVY value will be relatively hydrophobic (that is, water intolerance) than others. Higher negative residues observed among some of the bacteria is an indication that their respective intracellular proteins have higher negatively charged residue than the positively charged residues. This finding agrees with the reports of [18]. The instability index value as obtained from the pulse species indicates that species with very high instability index are very and highly unstable *in-vitro* and can undergo thermal dissociation easily. The high instability of some of the pulse species like *Arthrobacillus* species would suggest why it is unstable than *Martelella* species [19]. The results of phylogenetic relationship among the leguminous species has clearly revealed that the hydrocarbon degrading gene have not undergone serious mutation with the resultant zero divergence time among the closely related species. This by implication, further suggest that the hydrocarbon degrading gene still possess structural and functional capabilities of total hydrocarbon degradation of organic matter and fossils in the species different from other non-related species. This low mutation inferences from the phylogenetic relationship of *Pseudomonas aeruginosa* I and II in the species must have accounted for the low instability index obtained from the analysis of the physicochemical properties of the gene among these bacteria species. The biodiversity of bacterial species and their genetic distance was ascertained.

Table 4. Physico-chemical Properties of Nucleotide sequence of hydrocarbon degrading gene in selected bacteria species

Bacteria species	No. Of Aa	Molecular Wt. (KDa)	Theoretical pI	Total No. -ve charges	Total No. +ve charges	Total No. of atom	Extinction coefficient	Instability index (II)	Aliphatic index	GRAVY
<i>Alcanivorax</i> species	478	35874.44	9.41	18	24	3654	40465	48.52	97.46	-0.121
<i>Aneurinibacillus</i> species	183	29684.49	9.55	18	28	4210	47790	54.69	98.67	-0.014
<i>Arthrobacter</i> species	101	20876.	9.50	20	30	4377	49280	48.25	99.76	0.024
<i>Martellella</i> species	98	10302.14	9.23	16	21	3443	32320	48.63	92.31	- 0.021
<i>Pseudomonas aeruginosa</i>	494	31199.14	9.65	31	62	6653	86043	30.17	88.43	0.124
<i>Pseudomonas aeruginosa</i> II	296	19576.34	9.41	19	24	4187	47915	38.18	98.10	-0.095
<i>Pseudomonas stutzeri</i>	308	41390.75	9.60	36	68	8732	99045	43.86	96.32	0.042
<i>Pseudomonas stutzeri</i>	304	51390.77	9.84	36	68	8732	99045	43.86	90.32	0.028
<i>Pseudomonas taeanensis</i>	183	20653.82	9.54	42	31	8618	93210	51.15	97.06	-0.098
<i>Rhodococcus</i> species	352	44390.75	9.14	31	31	2976	41879	36.87	94.03	0.022
<i>Thalassospire</i> species	353	29313.23	9.84	19	29	4160	46300	55.94	99.88	-0.023
<i>Thalassospire</i> species II	304	21390.75	9.24	23	32	3298	35874	43.09	88.90	-0.032
Uncultured bacterium	376	2390.75	9.59	28	27	6312	56312	39.62	9713	0.334

Fig 2. Ribbon Tertiary Structure of Hydrocarbon Degrading Gene in *Bacterial* as displayed using RASMOL view

					
<i>Alcanivorax</i> species	<i>Thalassospira</i> species I	<i>Pseudomonas sturtgeri</i>	<i>Martelella</i> species	<i>Pseudomonas taenensis</i>	<i>Aneurinibacillus</i> species
					
<i>Arthrobacter</i> species I	<i>Rhodococcus</i> species	<i>Pseudomonas aeruginosa</i> II	<i>Thalassospira</i> species II	Uncultured bacterium species	<i>Thalassospira</i> species I

Pink = Alpha helix, Blue = Random Coil, Green = Extended Strand, Yellow = Beta Sheet

Table 5. Secondary Structure of amino acids sequences of hydrocarbon degrading genes in selected bacteria species

Bacteria species	Alpha Helix	Extended Strand (%)	Random coil (%)
<i>Alcanivorax</i> species	34.58	16.93	47.83
<i>Aneurinibacillus</i> species	34.14	19.08	46.91
<i>Arthrobacter</i> species	41.27	18.24	40.58
<i>Martelella</i> species	36.68	13.55	51.75
<i>Pseudomonas aeruginosa</i>	33.26	26.62	42.58
<i>Pseudomonas aeruginosa</i> II	39.42	15.19	46.47
<i>Pseudomonas stutzeri</i>	35.64	21.90	43.34
<i>Pseudomonas taeanensis</i>	37.50	17.53	45.63
<i>Rhodococcus</i> species	36.13	22.99	41.37
<i>Thalassospire</i> species	31.64	18.24	49.82
<i>Thalassospire</i> species II	37.57	15.71	47.72
Uncultured bacterium	35.44	20.89	44.10

5. CONCLUSION

Phylogenetic analysis has been a useful tool in comparing the similarities between different bacterial species. It used in this study has shown how interrelated the bacteria species are.

DISCLAIMER

The products used for this research are commonly and predominantly use products in our area of research and country. There is absolutely no conflict of interest between the authors and producers of the products because we do not intend to use these products as an avenue for any litigation but for the advancement of knowledge. Also, the research was not funded by the producing company rather it was funded by personal efforts of the authors.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Agbor RB, Antai SP. Enhancement of microbial degradation of crude oil polluted

soil using Mexican sunflower (*Tithonia diversifolia*) and wild groundnut (*Calapogonium mucunoides*), Inter. J Eco Eco-sol. 2018;5(1):1-7.

2. Agbor RB, Antai SP, Ekpo IA. Phylogenetic relationship of hydrocarbon degrading fungi species in bioremediation, Glo J. Ear Env Sci. 2018;3(2):8-15.
3. Agbor RB, Antai SP. Phylogenetic relationship of bacterial species involved in bioremediation of hydrocarbon polluted soils. Ann. R. Rev. Bio. 2019;32(6):1-13.
4. Antai SP, Unimke AA, Agbor RB. Assessment of the heterotrophic and crude oil utilizing microorganisms of Imo river estuary of the Niger Delta mangrove ecosystem, Am Int J Bio. 2014;2(1):29-42.
5. Chikere CB, Okpokwasili GC, Chikere BO. Bacterial diversity in a tropical oil-polluted soil undergoing bioremediation. Afri J Biotech. 2009;8(11):2535-2540.
6. Hamamura N, Olson SH, Ward DM, Inskeep WP. Microbial population dynamics associated with crude oil biodegradation in diverse soils. Appl. Environ. Microbiol. 2006;72:6316-6324.

7. Quatrini P, Scaglione G, De Pasquale C, Reila S, Puglia AM. Isolation of gram-positive n-alkane degraders from a hydrocarbon contaminated Mediterranean shoreline. *J. Appl. Microbiol.* 2008;104:251.
8. Okpokwasili GSC. Microbes and the environmental challenge. Inaugural lecture series no. 53. University of Port Harcourt Press. Port Harcourt. 2006;31-56.
9. Thompson DM, Moritz C, Mable BK. *Molecular systematics*, 2nd edition. Sinauer, Sunderland, MA. *Comprehensive Coverage of Techniques for Phylogenetic Tree Reconstruction*; 2014.
10. Kelly E, Stemberg VH. "History, objectivity, and the construction of molecular phylogenies". *Stud. Hist. Phil. Biol. and Biomed. Sci.* 2009;39(4):451-468.
11. Sneath G, Sokai H. Phylogenetic analysis in molecular evolutionary genetics. *Ann. Rev. Genetics. Brief Review of Tree-building Techniques.* 1973;30:371-403.
12. Tamura K, Stecher G, Peterson D, Filipski A, Kumar P. *SMEGA6: Molecular evolutionary genetics analysis version 6.0.* *Mol Bio and Evo.* 2013;30:2725-2729.
13. Kang J, Reghain G, Gabriel D. Universal tree of life. In *encyclopedia of life sciences.* Nature publishing group, Macmillan; 2008.
14. Jones D, Peter R, James H, Jane S. "Phylogenetic". *Merriam-webster dictionary. English pronouncing dictionary,* Cambridge: Cambridge University Press; 2013.
15. Durosaro S, Benny G, Hedges SB. The origin and evolution of model organisms. *Nat Rev.* 2015;3:838-849.
16. Joshi P, Xu D. Molecular evolution: Introduction. In *encyclopedia of life sciences.* Nature publishing group, Macmillan; 2007.
17. Lehninger A, Martin AP, Idrissa T. Molecular clocks. In *encyclopedia of life sciences.* Nature publishing group, Macmillan; 2013.
18. Andrade Y, Atlas RM, Bartha. Biodegradation of petroleum in seawater at low temperatures. *Can. J Microb.* 2009;18:1851-1855.
19. Gurruprassad G, Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. Edited in *evolving genes and proteins* by V. Bryson and H.J. Vogel. Academic press, New York. 2009;97-166.

© 2021 Agbor et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:

The peer review history for this paper can be accessed here:
<https://www.sdiarticle4.com/review-history/70907>