OMICS A Journal of Integrative Biology
Volume 24, Number 9, 2020

DOI: 10.1089/omi.2020.0070

Research Article

Comparative Metagenomic Screening of Aromatic Hydrocarbon Degradation and Secondary Metabolite-Producing Genes in the Red Sea, the Suez Canal, and the Mediterranean Sea

Esraa Elsaeed,[1,*] Shymaa Enany,[2,*] Amro Hanora,[2] and Nora Fahmy[2]

Abstract

Marine and ecosystem pollution due to oil spills can be addressed by identifying the aromatic hydrocarbon (HC)-degrading microorganisms and their responsible genes for biodegradation. Moreover, screening for genes coding for secondary metabolites is invaluable for drug discovery. We report here, the first metagenomic study investigating the shotgun metagenome of the Suez Canal water sampled at Ismailia city concerning its aromatic HC degradation potential in comparison to the seawater sampled at Halayeb city at the Red Sea and Sallum city at the Mediterranean Sea. Moreover, for an in-depth understanding of marine biotechnology applications, we screened for the polyketide synthases (PKSs) and nonribosomal peptide synthetase (NRPS) domains in those three metagenomes. By mapping against functional protein databases, we found that 13, 6, and 3 gene classes from the SEED database; 2, 1, and 3 gene classes from the EgGNOG; and 5, 4, and 2 genes from the InterPro2GO database were identified to be differentially abundant among Halayeb, Ismailia, and Sallum metagenomes, respectively. Also, Halayeb metagenome in the Red Sea reported the highest number of PKS domains showing higher potential in secondary metabolite production in addition to the oil degradation potential.

Keywords: metagenomics, aromatic hydrocarbons, ecogenomics, ecology, Suez Canal, Mediterranean Sea, Red Sea

Introduction

Suez Canal is a human-made passage constructed for commercial and navigation targets. It confers a direct way between the Atlantic and the Indian oceans. It is situated in Egypt and passes across three cities, ''Suez'' city at the Red Sea, ''Port Said'' city at the Mediterranean Sea, and ''Ismailia'' city at the midpoint of the canal. It contains several ports while 100 million tons of oil are transferred annually through the canal (Zaki et al., 2014), contributing to serious consequences such as an elevated hydrocarbon (HC) load in the attendant marine ecosystem. Moreover, the real threat of oil tanker crash by virtue of poor naval navigation and chaotic maritime movements further compound the ecological damages (Ismail, 2019).

El Samra et al. (1983) found 13.74 μg/L of petroleum HC in the south of the Suez Canal and 45 μg/g in dry sediments in Port Said. Moreover, Temsah lake, which is a part of the Suez Canal at Ismailia, Egypt, various types of aromatic and aliphatic HCs have contaminated it through cargo tanker discharges of ballast water, local oil production resulting in a worrisome and intensive decline in biodiversity (El-Gendy and Moustafa, 2007; Tundo et al., 2005).

Petroleum products are major sources of marine ecotoxicity (Pacheco and Santos, 2001). The US Environmental Protection Agency (EPA) has noted a list of 16 polycyclic aromatic hydrocarbons (PAHs) that contribute to ecosystem instability (Balachandran et al., 2012; Matsubara et al., 2006). Addressing the problems of oil spills, ecosystem pollution, and instability cannot be solved by chemical dispersants that make matters worse; they only transform oil to another physical phase that cannot be isolated from the water (Schaum et al., 2010; Zheng et al., 2014).

On the other hand, bioremediation is a promising solution in which the microorganisms living under given ecological stress develop adaptive strategies (Bargiela et al., 2015).

[1]Department of Microbiology and Immunology, Faculty of Pharmacy, Delta University, Gamsa, Egypt.
[2]Department of Microbiology and Immunology, Faculty of Pharmacy, Suez Canal University, Ismailia, Egypt.
*Both these authors contributed equally to this work.

Bioremediation allows a biochemical process of oil degradation by enzymatic reactions carried out by microorganisms, such as oxygenases, hydroxylases, and dehydrogenases that transform the crude oil to intermediate compounds (Kumari et al., 2012).

Nonribosomal peptides and polyketides are large groups of secondary metabolites that include many natural products such as antitumor compounds, antimicrobials, surfactants, and immunomodulatory compounds (Kleinkauf and Von Döhren, 1996; Koglin and Walsh, 2009). Both types of secondary metabolites are synthesized by a multifunctional enzyme system called polyketide synthases (PKSs) for polyketides and nonribosomal peptide synthetases (NRPSs) for nonribosomal peptides, respectively (Graça et al., 2015).

Hence, to the best of our knowledge, we present here the first comparative shotgun metagenomic study in the Suez Canal. The study aimed to screen the metagenome of the Ismailia canal water for HC degradation genes available for bioremediation and biotechnological applications, in comparison to the metagenomes in Halayeb at the southmost border of Egypt at the Red Sea, and Sallum at the westmost border of Egypt at the Mediterranean Sea. For an in-depth understanding of marine biotechnological potentials, we screened for PKS and NRPS domains in these three metagenomes.

Materials and Methods

Compliance with ethical standards

Since we collected water samples, ethics approval was not applicable in our study.

Sample collection

Three water samples were collected with a handmade Niskin tube in March of 2019. The first sample was taken from the Mediterranean Sea at Sallum city, which is the farthest Mediterranean Sea point from the Suez Canal in Egypt; the second one was picked from the middle of the Suez Canal at Ismailia city; and the third one was taken from the Red Sea at Halayeb city, which is the farthest Red Sea point from the Suez Canal in Egypt (Fig. 1). For more details about sampling physical parameters, see Table 1. All samples were kept in icebox during transferring to the central laboratory of faculty of pharmacy, Delta University, Egypt.

Water filtration

Vacuum filtration (Rocker) was done on two successive steps, the prefiltration through 10 μm filter papers (Double Ring) and the filtration through 0.22 μm and 45 mm diameter filter membranes (Merck), and all filters were stored at −80°C.

Environmental DNA extraction, quantification, and purification

Extraction from filter membranes directly was done by the Qiagen DNeasy Kits for blood and tissue (Qiagen) (Thomsen et al., 2012) following the protocol of Cowart et al. (2018) with some modifications; 3 h incubation for lysis step, RNA degradation by 40 μL of RNase A, and the addition of 550 μL of buffer AL and 475 μL of AW1. Environmental DNA (eDNA) quantification and qualification were assessed by the NanoDrop (Implen p330).

Library preparation and DNA sequencing

Library preparation and DNA sequencing were carried out at IGA Technology. The Ovation Ultralow Library System V2 Kit (NuGEN, San Carlos, CA) was used for library preparation following the manufacturer's instructions. Both input and final libraries were quantified by Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA) and quality tested by Agilent 2100 Bioanalyzer High Sensitivity DNA assay (Agilent Technologies, Santa Clara, CA). Libraries were sequenced on NovaSeq 6000 in paired-end 150 mode. The number of total reads produced per sample was 10.27, 20.54, and 19.27 M for Sallum, Ismailia, Halayeb, respectively.

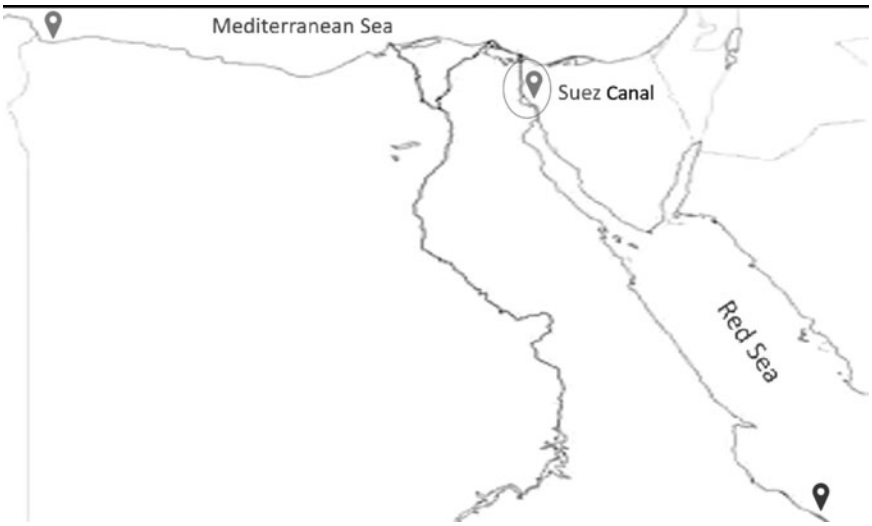

FIG. 1. Sampling locations include ''Sallum'' at the Mediterranean Sea, ''Ismailia'' at the Suez Canal, and ''Halayeb'' at the Red Sea.

Table 1. Samples Metadata

Samples	Depth (m)	Temperature (°C)	Salinity (g/kg)	Location
Sallum (Mediterranean Sea)	10	17.7	38.7	31°34′41.1″N 25°10′49.1″E
Ismailia (Suez Canal)	10	17.7	39.6	30°34′01.1″N 32°18′16.4″E
Halayeb (Red Sea)	10	23.6	41.5	22°13′39.5″N 36°40′00.6″E

Bioinformatic analysis

Reads were processed for format conversions and demultiplexing by Bcl2Fastq 2.0.2 (Illumina, San Diego, CA), adapters were masked with Cutadapt v1.11 (Martin, 2011), reads quality was checked by Multiqc V 1.6 (Ewels et al., 2016), FASTQ reads were denovo assembled by Megahit 1.2.7 (Ewels et al., 2016), the resulted contigs were assessed by Metaquast option of Quast 5.0.2 (Mikheenko et al., 2018) software, sequences were aligned to the manually curated MarRef (Klemetsen et al., 2017), which contain complete reference genomes of marine microbes using Diamond 0.9.29 aligner tool (Buchfink et al., 2015).

Aligned sequences were uploaded to Megan6 (Klemetsen et al., 2017) for taxonomic and functional analysis. In Megan6, National Center for Bioecology Information (NCBI) nonredundant accessions were mapped against functional SEED (Overbeek et al., 2005), EgGNOG (Jensen et al., 2008), and InterPro2GO (Camon et al., 2005) databases and taxonomic classes.

Screening of PKS and NRPS genes

Contigs were aligned to condensation (C) and ketosynthase (KS) domain databases of NaPDoS (Ziemert et al., 2012) by diamond blastx (Buchfink et al., 2015). All obtained hits were below 80% identity. Contigs of the resulting hits were mapped by blastx against the NCBI nonredundant database (NCBI Resource Coordinators, 2018). Sequences that were assigned to PKS and NRPS domains were translated to proteins of the mapped open reading frames by EMBOSS Transeq (www.ebi.ac.uk/Tools/st/emboss_transeq).

The resulting proteins were multiple sequences aligned with muscle plugin in SeeView4 (Gouy et al., 2009) and the phylogenetic tree were constructed by MEGA X (Tamura et al., 2007) after examining the best model for maximum likelihood tree. It was built by using AMG + F model, gamma distribution of 5 and bootstrapping of 500. Also, PKS and NRPS domains were determined by hidden Markov model in SBSPKSv2 webserver (Khater et al., 2017).

Results

After quality control step, the remaining 10,218,258, 10,273,734, and 9,635,108 reads for Halayeb, Ismailia, and Sallum metagenomes were assembled into 332,074 contigs with a total of 98,026 bp, an average length of 769 bp, and N50 equal to 893 bp for Sallum metagenome, 343,311 contigs with a total 255,344,819 bp, an average length of 743 bp, and N50 equal to 836 bp for Ismailia metagenome, and 232,993 contigs with a total 182,036,347 bp, an average length of 781 bp, and N50 equal to 912 bp for Halayeb metagenome. Samples' rarefaction curve (Fig. 2) was plotted relied on the NCBI taxonomic rank at the species level; it demonstrated a reasonable genome coverage from the three metagenomes.

Taxonomic diversity

We found that 166,110, 164,720, and 164,680 reads were assigned to bacteria and 964, 1152, and 1390 to archaea for Halayeb, Ismailia, and Sallum, respectively. Supplementary Figure S1C showed that Sallum metagenome is the most diverse sample as the radial chart of Sallum was mostly eventually divided highlighting the higher degree of evenness. Halayeb is the least diverse sample (Supplementary Fig. S1A) followed by Ismailia metagenome (Supplementary Fig. S1B) because most areas of Halayeb radial chart was occupied by certain species.

Functional annotation to protein databases

The site-specific abundance of aromatic HC-degrading genes according to the SEED database. We recorded 13, 6, and 3 gene classes that were found to be more abundant in Halayeb, Ismailia, and Sallum metagenomes, respectively (Supplementary Fig. S2). For reviewing genes in each class, see Supplementary Tables S1, S4, and S7.

Halayeb metagenome has more red cells (highest z-score) than the other two metagenomes. Discriminative aromatic HC-degrading gene classes in Halayeb metagenome are aromatic amin degradation, 4-hydroxyphenylaceticacid catabolic pathway, anaerobic benzoate catabolism, biphenyl degradation, the central metacleavage pathway of aromatic compound degradation, gallic acid utilization, gentisate degradation, homogentisate pathway of aromatic compound degradation, p-hydroxy benzoate degradation, protocatechuate branch of beta-ketoadipate pathway, salicylate and gentisate catabolism, salicylate ester degradation, and toluene degradation.

The following aromatic HC-degrading gene classes are more common in Ismailia metagenome; 2-aminophenol metabolism, Benzoate transport and degradation cluster, N-heterocyclic aromatic compound degradations, naphthalene and anthracene degradation, phenylacetyl-CoA catabolic pathway (core), and quinate degradation.

Sallum metagenome has more cool cells (lowest z-score) than the others but it is higher in anaerobic aromatic degradation (benzoate), benzoate catabolism, and chloroaromatic degradation pathway.

The site-specific abundance of aromatic HC-degrading genes according to the EgGNOG database. We determined 2, 1, and 3 gene classes that were more abundant in Halayeb, Ismailia, and Sallum metagenomes, respectively (Supplementary Fig. S3). For Halayeb metagenome, hottest cells represent COG3435 gentisate 1,2 dioxygenase and ENOG410YF62 phenol hydroxylase C-terminal dimerization domain. For Ismailia metagenome, the hottest cells represent

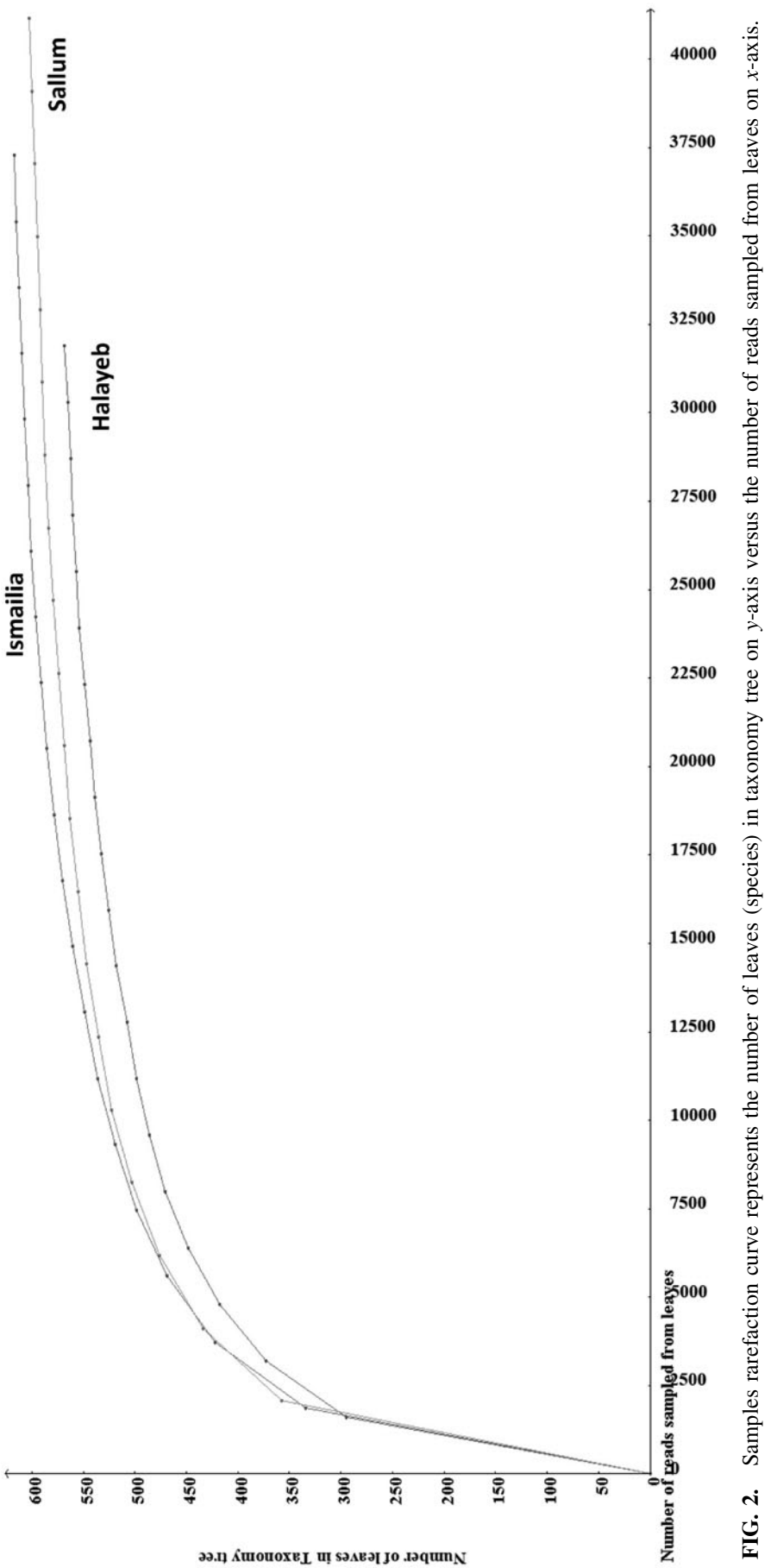

FIG. 2. Samples rarefaction curve represents the number of leaves (species) in taxonomy tree on *y*-axis versus the number of reads sampled from leaves on *x*-axis.

COG2854 toluene tolerance family protein. For Sallum metagenome, COG33396 phenylacetate CoA oxygenase subunit, COG3135 benzoate transporter activity, and COG3460 phenylacetate catabolic process are the predominant genes. For reviewing genes in each class, see Supplementary Tables S3, S6, and S9.

The site-specific abundance of aromatic HC-degrading genes according to the InterPro2GO database. We found 5, 4, and 2 genes that were more abundant in Halayeb, Ismailia, and Sallum metagenomes, respectively (Supplementary Fig. S4). The highest z-score was assigned to IPR023789 2,3 dihydroxyphenyl propionate /2,3-dihydroxycinnamic acid 1,2-dioxygenase, IPR012733 4-hydroxybenzoate 3-monooxygenase, IPR020875 phenylpropionate/cinnamic acid dioxygenase alpha subunit, IPR012083 arylsulfatase, and IPR014436 extradiol aromatic ring-opening dioxygenase DODA-A type in Halayeb metagenome; IPR011982 4-hydroxyphenylacetate 3-monooxygenase reductase component, IPR023786 3-(3-hydroxy-phenyl) propionate/3-hydroxycinnamic acid hydroxylase, IPR022893 shikimate dehydrogenase, and IPR005956 hydroxyphenylpyruvate dioxygenase in Ismailia metagenome; and IPR001273 aromatic amino acid hydroxylase and IPR004674 alkyl-hydroperoxidase AhpD in Sallum metagenome. For more details, see Supplementary Tables S2, S5, and S8.

Screening for genes coded for secondary metabolites producing enzymes

Mapping to NaPDoS databases revealed no hits with percent identity more than 80%. Mapping against NCBI nonredundant database (Table 2) obtained 5, 6, and 2 hits with percent identity more than 85% with expected (e) value lower than 1e-5 for Halayeb, Ismailia, and Sallum, respectively, indicating to be the same or similar to the reference hits (Ziemert et al., 2012). We found 14 hits with percentage identity lower than 80%, and this group represents the possibility of being new compounds. Identification of PKS and NRPS domains by SBSPKSv2 webserver and NCBI blastx (Table 2) explored eight KS domains in Halayeb metagenome, one KS domain in Ismailia metagenome, two KS domains in Sallum metagenome, and six acyl carrier protein (ACP) domains in each of Ismailia and Halayeb metagenomes (Table 2).

Phylogenetic analysis of PKS and NRPS genes

Phylogenetic tree (Fig. 3) of the resulting hits gave more clear investigation to the blastx hits with low percent identity, which are supposed to be new compounds. K141_144235_Halayeb was mapped with percent identity <80% in NCBI and clustered with K141_175987_Sallum, which codes for KS domain in the same clade giving the higher probability for being a KS domain. K141_124819_Halayeb and K141_81012_Halayeb were clustered together with K141_114597_Sallum, which coded for a KS domain in the same clade rising the probability of them to be KS domains.

Discussion

Chronic pollution in the Suez Canal is caused by oil spills and accumulation of PAHs (El-Gendy and Moustafa, 2007; Zaki et al., 2014). Thus tailoring bacterial genomes to stress adaptation by building aromatic HC degradation genes. We screened for these genes in the seawater of Ismailia metagenome as a representative of the Suez Canal in comparison to metagenomes of Halayeb at the Red Sea and Sallum at the Mediterranean Sea. These three locations were selected as Ismailia is in the middle of the canal, while Sallum and Halayeb are the furthest cities from the Suez Canal at the coast of the Mediterranean Sea and the Red Sea; respectively. The further away from the Suez Canal, the less the canal affects both seas. Minimizing the effect of the canal on both seas displays the true capacity of the aromatic HC degradation of both seas, it is somewhat like going back in time before construction of Suez Canal.

Multiple previous studies reported our resulted species as aromatic HC degrader (Bacosa et al., 2015; Devpura et al., 2017; Li et al., 2014; McKew et al., 2007; Radwan et al., 1998; Sorkhoh et al., 1990; Tanaka et al., 2008). We noted that almost all the species in Halayeb are aromatic HC-degrading microorganisms but only half of the species are aromatic HC-degraders in Sallum. Ismailia is in between and that has happened because oiling reduces the diversity of bacteria in combination with a rapid and powerful selection of specialized HC-degrading bacteria that play a critical role in the natural purification of oil-contaminated marine systems (Golyshin et al., 2003).

Biodegradation of aromatic HC requires cooperation network between various bacterial species to degrade a wide array of HCs (Nikhil et al., 2013; Rahman et al., 2002). Pseudomonas sp. DS10-129, Micrococcus sp. GS2-22, Flavobacterium sp. DS5-73, Corynebacterium sp. GS5-66, and Bacillus sp. DS6-86 were utilized in laboratory degradation of crude oil (Rahman et al., 2002) and the combination of Micrococcus sp. and Pseudomonas sp. was utilized for the degradation of diesel engine oil (Nikhil et al., 2013). Our results confirmed this information as we found that multiple species were assigned to each aromatic HC-degrading gene (Supplementary Tables S1–S9).

Halayeb is the least diverse metagenome followed by Ismailia because Halayeb may be exposed to chronic environmental stress (Supplementary Fig. S1A–C). In heat maps, Halayeb metagenome always has the highest number of red-hot cells, but Sallum has the highest number of blue-cool cells, which supports our conclusion that Halayeb metagenome has the highest number of aromatic HC-degrading genes. This may be attributed to the tropical climate of Halayeb. Our explanation agrees with Bargiela et al. (2015) who studied the relationship of oil degradation with the temperature at different locations along the coasts of the Mediterranean Sea and the Red Sea and noted increasing aromatic HC degradation potential by the elevating temperature.

Also, other studies observed increasing in secondary metabolisms by higher temperature due to the enhancement of enzyme kinetics (Bianchi, 2018; Fuhrman and Azam, 1983; López-Urrutia and Morán, 2007; Rivkin et al., 1996; Vázquez-domínguez et al., 2007; White et al., 1991).

Water current does not appear to be an influencing factor because water current through the Suez Canal varies from season to season, the direction to the north is from November to June and the direction to the south is from July to October (Morcos, 1960). Sample collection was in March during the northward current so bringing aromatic HCs to Halayeb through the flow of water is impossible to be the reason behind the high level of aromatic HC-degrading genes.

Table 2. Summary of the Best National Center for Bioecology Information blastx Hits, Enzyme Types, and Available Domains

Contigs	*Name of the best NCBI blastx hits*	*Accession No. of the best NCBI blastx hit*	*% Identity to the best NCBI blastx hit*	*Type of enzyme*	*Available domains*
k141_88193 (Ismailia)	Type I polyketide synthase (Halobacteriovoraceae bacterium)	MBC77776.1	54.47	PKS	Not detected
k141_200902 (Ismailia)	NRPS (Chitinophaga sp. MD30)	WP_095840277.1	33.81	NRPS	Not detected
k141_210540 (Ismailia)	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein	NBQ02056.1	99.597	PKS	ACP (NCBI blastx)
k141_217316 (Ismailia)	Beta-ketoacyl synthase	RZO23865.1	91.637	PKS	ACP (NCBI blastx)
k141_249165 (Ismailia)	Type I polyketide synthase	MBJ01119.1	70.874	PKS 1	Not detected
k141_278182 (Ismailia)	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein	NDH16748.1	96.078	PKS	ACP (NCBI blastx)
k141_296166 (Ismailia)	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein	NCV19295.1	99.078	PKS	ACP (NCBI blastx)
k141_319120 (Ismailia)	Beta-ketoacyl-[acyl-carrier-protein] synthase II	MBS92556.1	100	PKS	ACP (NCBI blastx)
k141_321295 (Ismailia)	Beta-ketoacyl synthase	OUX63826.1	98.605 (to ACP domain)	NRPS-PKS hybrid (ACP (blastx), KS and OX (SBSPKSV2))	ACP, KS, and OX
k141_121487 (Sallum)	NRPS (Mycobacterium colombiense)	WP_082280136.1	56.38	NRPS	Not detected
k141_114597 (Sallum)	Polyketide synthase (Ostreococcus lucimarinus CCE9901)	XP_001416177.1	100	PKS	KS (SBSPKSv2)
k141_175987 (Sallum)	Polyketide synthase (Ostreococcus lucimarinus CCE9901)	XP_001416378.1	100	PKS	KS (SBSPKSv2)
k141_262916 (Sallum)	Type I polyketide synthase (Saccharothrix sp. NRRL B-16314)	WP_081915631.1	61.29	PKS 1	Not detected
k141_283106 (Sallum)	Type I polyketide synthase (Saccharothrix sp. NRRL B-16314)	WP_081915631.1	61.29	PKS 1	Not detected
k141_295832 (Sallum)	Polyketide synthase (Ostreococcus lucimarinus CCE9901)	XP_001416177.1	57.53	PKS	Not detected
k141_6439 (Halayeb)	Modular polyketide synthase (uncultured bacterium)	ABQ50542.1	74.07	PKS (modular)	Not detected
k141_81012 (Halayeb)	Polyketide synthase (Emiliania huxleyi CCMP1516)	XP_005764102.1	66.41	PKS	KS (SBSPKSv2)
k141_81300 (Halayeb)	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein (Rhodobacteraceae bacterium)	NCV67034.1	98.96	PKS	ACP (NCBI blastx), KS (SBSPKSv2)
k141_124819 (Halayeb)	Polyketide synthase (Chrysochromulina tobinii)	KOO31876.1	74.17	PKS	KS (SBSPKSv2)
k141_125204 (Halayeb)	Polyketide synthase (Emiliania huxleyi CCMP1516)	XP_005762094.1	61.36	PKS	KS (SBSPKSv2)
k141_144235 (Halayeb)	Polyketide synthase (Chrysochromulina tobinii)	KOO31005.1	69.44	PKS	KS (SBSPKSv2)
k141_148716 (Halayeb)	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein (Rhodobacterales bacterium)	NCX28560.1	100	PKS	ACP (NCBI blastx), KS (SBSPKSv2)
k141_149434 (Halayeb)	Polyketide synthase (Emiliania huxleyi CCMP1516)	XP_005793027.1	52.17	PKS	Not detected
k141_155029 (Halayeb)	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein (Rhodobacteraceae bacterium)	NDI13436.1	94.34	PKS	ACP (NCBI blastx), KS (SBSPKSv2)
k141_158140 (Halayeb)	Modular polyketide synthase (uncultured bacterium)	ABQ50542.1	71.43	PKS (modular)	Not detected
k141_166915 (Halayeb)	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein (Rhodobacteraceae bacterium)	NDI13436.1	98.92	PKS	ACP (NCBI blastx)
k141_194423 (Halayeb)	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein (Rhodobacteraceae bacterium)	NCV67034.1	99.23	PKS	ACP (NCBI blastx), KS (SBSPKSv2)
k141_212798 (Halayeb)	Beta-ketoacyl-[acyl-carrier-protein] synthase family protein (Rhodobacteraceae bacterium)	NCV67034.1	99.17	PKS	ACP (NCBI blastx)

ACP, acyl carrier protein; KS, ketosynthase; NCBI, National Center for Bioecology Information; NRPS, nonribosomal peptide synthetase; PKS, polyketides synthase.

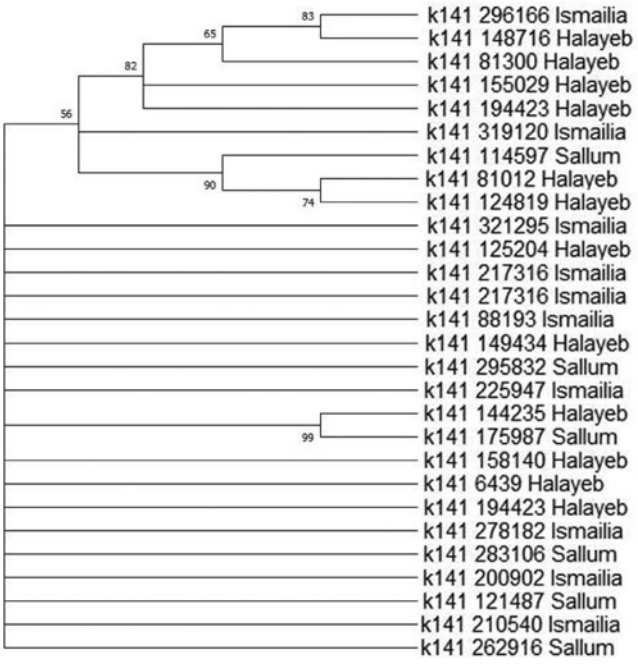

FIG. 3. The phylogenetic tree of the contigs that were mapped to PKS or NRPS domains in NCBI and SBSPKSv2 webserver. Some of unknown contigs (contigs <80% identity) were clustered with known KS domains. KS, ketosynthase; NCBI, National Center for Bioecology Information; NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase.

Although Halayeb has achieved higher levels of aromatic HC degradation genes than Ismailia and Sallum, Ismailia's aromatic HC degradation genes are highly related to petroleum HCs and oil spills, such as naphthalene, anthracene, toluene, and benzene-degrading genes. It can be explained by continuous exposure of the Suez Canal to various petroleum aliphatic and aromatic HCs originating from shipping activities (Tundo, 2005). We collected the Ismailia sample from a Suez Canal region near Temsah Lake, which hosts a small port offering a variety of maritime activities, including the maintenance of the Suez Canal (Tundo, 2005). Naphthalenes, benzenes, and other bicyclic aromatics have previously been detected with high-absorption beaks in chromatographic analysis of Temsah Lake (El-Gendy, 2007).

Naphthalene indicates the presence of a relative fresh oil and unburned petroleum factories (Kennish, 2017; Said and El Agroudy, 2006) encouraging the microbial community to develop naphthalene-degrading genes in Ismailia metagenome (Supplementary Fig. S2). Ismailia metagenome characterized by the COG2854 toluene tolerance genes (Supplementary Fig. S3) offering bacterial compatibility with toluene, which is one of the most toxic petroleum compounds and has anticarcinogenic effects (Chapelle, 1999; Zilli et al., 2000).

We found Benzoate degradation genes (Supplementary Fig. S2) in Ismailia and Sallum metagenomes. Researchers often used Benzoate as a model for studying bacterial degradation of oil spills and aromatic compounds (Carmona et al., 2009; Gibson and Harwood, 2002). IPR022893 shikimate dehydrogenase gene is seen in its highest abundance in Ismailia metagenome (Supplementary Fig. S4). Shikimate dehydrogenase is the main route of plant degradation to aromatic HCs (ElAhwany et al., 2015).

Unlike Sallum and Ismailia, Halayeb metagenome characterized by distinguished aromatic compound degrader genes, such as aromatic amine degradation genes (Supplementary Fig. S2), they merit particular concern as aromatic amines are potent carcinogens (Benigni and Passerini, 2002; Benigni et al., 2007) and widely applied in agricultural chemicals, polymers, rubber, and manufacturing of pharmaceuticals (Benigni et al., 2007), central metacleavage pathway (Supplementary Fig. S2) that come after bacterial oxidation of PAHs to dihydroxy compounds (Gao et al., 2013), salicylate, and gentisate catabolism (Supplementary Fig. S2) that were noticed in sediment metagenomic screening of San Jacinto River (Benigni et al., 2007); IPR012083 arylsulfatase (Supplementary Fig. S4) which was enhanced by the elevated level of PAHs in soil (Lipińska et al., 2014); IPR039091 arylhydrocarbon receptor (Supplementary Fig. S4), which is a transcription factor in various toxicological process of organic pollution (Zhou et al., 2009); and IPR14436 Extradiol aromatic ring opening dioxygenase (Supplementary Fig. S4) that was previously identified in the metatranscriptomic analysis of soil HC degradation (Gonzalez et al., 2018).

There are two major strategies for the degradation of aromatic compounds depending on the presence or absence of oxygen. In aerobic biodegradation, hydroxylation and oxygenolytic aromatic ring cleavage often occurred by oxygenase enzymes (Parales and Resnick, 2006; Vaillancourt et al., 2006). The anaerobic degradation depends on attacking the aromatic ring through reductive reactions (Fuchs, 2008; Gibson and Harwood, 2002; Harwood et al., 2020).

Oxygenases are widely distributed in the three metagenomes (Supplementary Figs. S3 and S4) such as IPR011982 4-hydroxyphenylacetate 3-monooxygenase reductase component and ENOG410XUSH phenylacetate catabolic CoA dioxygenase in Ismailia metagenome and COG33396 phenylacetate CoA oxygenase subunit and ENOG410YZWX0 phytanoyl-CoA dioxygenase in Sallum metagenome. Halayeb metagenome includes more oxygenases such as IPR023789 2,3-dihydroxyphenyl propionate /2,3-dihydroxycinnamic acid 1,2-dioxygenase, IPR020875 phenylpropionate/cinnamic acid dioxygenase alpha subunit, IPR012733 4-hydroxybenzoate 3-monooxygenase, and COG3435 gentisate 1,2 dioxygenase.

Sallum and Halayeb metagenomes marked by their anaerobic degradation of benzoate (Supplementary Fig. S2) that was also found in methanotrophic mates of The Black Sea (Kube et al., 2005). Sallum metagenome is also marked by IPR004674 alkylhydroperoxidase AhpD (Supplementary Fig. S4), which is known as responding mechanisms to the oxidative stress in bacterial isolates from crude oil-contaminated environments. Increasing the level of catalase and peroxidase activities are enzymatic defenses, which are not only critical for the degradation of pollutants but also to respond to reactive oxygen species (Bučková et al., 2010).

Our next step was searching for PKS and NRPS domains, comparing the three metagenomes, and mapping them against NaPDoS database, which could reveal the false-positive results to fatty acid synthase (FAS) that is homologous to PKSs (Jenke-Kodama et al., 2005); those sequences were removed from the downstream analysis. Halayeb metagenome reported the highest number of PKS domains (eight KS and six ACP domains) showing the extra potential of the

Red Sea in secondary metabolite production besides the oil degradation power. Many studies of marine invertebrates agreed with the Red Sea potential for secondary metabolite production (Shreadah et al., 2018; Xi et al., 2012).

Conclusions

Seawater is a valuable source for biotechnological product, searching for bacterial aromatic HC-degrading genes and secondary metabolite-coding genes that enable further bioremediation application and drug discovery. We compared Halayeb, Ismailia, and Sallum seawater metagenomes in terms of aromatic HC-degrading genes and PKS and NRPS domain-coding genes. We found Halayeb as a representative of the Red Sea to be a least diverse and most valuable source for aromatic HCs degradation genes and PKSs coding genes, followed by Ismailia in the middle of the Suez Canal and then Sallum at the Mediterranean Sea that has the highest species diversity. More comparative studies between the Mediterranean Sea and the Suez Canal on other sites with a larger number of samples are required to answer which of them are valuable in aromatic HC-degrading genes and PKS and NRPS-coding genes.

Acknowledgments

The authors thank Mr. Saeed Sobhy and the diver, Hossam El-Fekky, for their effort in making the handmade Niskin-like tube.

Author Disclosure Statement

The authors declare they have no competing financial interests.

Funding Information

The authors received no specific funding for this work.

Supplementary Material

Supplementary Figure S1
Supplementary Figure S2
Supplementary Figure S3
Supplementary Figure S4
Supplementary Table S1
Supplementary Table S2
Supplementary Table S3
Supplementary Table S4
Supplementary Table S5
Supplementary Table S6
Supplementary Table S7
Supplementary Table S8
Supplementary Table S9

References

Bacosa HP, Liu Z, and Erdner DL. (2015). Natural sunlight shapes crude oil-degrading bacterial communities in Northern Gulf of Mexico surface waters. Front Microbiol 6, 1325.

Balachandran C, Duraipandiyan V, Balakrishna K, and Ignacimuthu S. (2012). Petroleum and polycyclic aromatic hydrocarbons (PAHs) degradation and naphthalene metabolism in *Streptomyces* sp. (ERI-CPDA-1) isolated from oil contaminated soil. Bioresour Technol 112, 83–90.

Bargiela R, Mapelli F, Rojo D, et al. (2015). Bacterial population and biodegradation potential in chronically crude oil-contaminated marine sediments are strongly linked to temperature. Sci Rep 5, 11651.

Benigni R, Bossa C, Netzeva T, Rodomonte A, and Tsakovska I. (2007). Mechanistic QSAR of aromatic amines: New models for discriminating between homocyclic mutagens and nonmutagens, and validation of models for carcin ogens. Environ Mol Mutagen 48, 754–771.

Benigni R, and Passerini L. (2002). Carcinogenicity of the aromatic amines: From structure–activity relationships to mechanisms of action and risk assessment. Mutat Res 511, 191–206.

Bianchi P. (2018). Characterization of bacterial hydrocarbon degradation potential in the Red Sea through metagenomic and cultivation methods. Doctoral dissertation. King Abdullah University of Science and Technology Thuwal, Kingdom of Saudi Arabia.

Buchfink B, Xie C, and Huson DH. (2015). Fast and sensitive protein alignment using DIAMOND. Nat Methods 12, 59–60.

Bučková M, Godočíková J, Zámocký M, and Polek B. (2010). Screening of bacterial isolates from polluted soils exhibiting catalase and peroxidase activity and diversity of their responses to oxidative stress. Curr Microbiol 61, 241–247.

Camon EB, Barrell DG, Dimmer EC, et al. (2005). An evaluation of GO annotation retrieval for BioCreAtIvE and GOA. BMC Bioinformatics 6, S17.

Carmona M, Zamarro MT, Blázquez B, et al. (2009). Anaerobic catabolism of aromatic compounds: A genetic and genomic view. Microbiol Mol Biol Rev 73, 71–133.

Chapelle FH. (1999). Bioremediation of petroleum hydrocarbon-contaminated ground water: The perspectives of history and hydrology. Groundwater 37, 122–132.

Cowart DA, Murphy KR, and Cheng C-HC. (2018). Metagenomic sequencing of environmental DNA reveals marine faunal assemblages from the West Antarctic Peninsula. Mar Genomics 37, 148–160.

Devpura N, Jain K, Patel A, Joshi CG, and Madamwar D. (2017). Metabolic potential and taxonomic assessment of bacterial community of an environment to chronic industrial discharge. Int Biodeterior Biodegradation 123, 216–227.

El Samra MI, El Deeb KZ, Askar AI, Wahby SD, and El Shazly MSH. (1983). Preliminary study of petroleum hydrocarbon pollution along the Suez Canal. Bull Nat Inst Oceanogr Fish 9, 97–101.

El-Gendy NS, and Moustafa YM. (2007). Environmental assessment of petroleum hydrocarbons contaminating temsah lake, Suez Canal, Egypt. Orient J Chem 23, 11–26.

ElAhwany AMD, Ghozlan HA, ElSharif HA, and Sabry SA. (2015). Phylogenetic diversity and antimicrobial activity of marine bacteria associated with the soft coral *Sarcophyton glaucum*. J Basic Microbiol 55, 2–10.

Ewels P, Magnusson M, Lundin S, and Käller M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048.

Fuchs G. (2008). Anaerobic metabolism of aromatic compounds. Ann N Y Acad Sci 1125, 82–99.

Fuhrman JA, and Azam F. (1983). Adaptations of bacteria to marine subsurface waters studied by temperature response. Mar Ecol Prog Ser 13, 95–98.

Gao S, Seo JS, Wang J, et al. (2013). Multiple degradation pathways of phenanthrene by Stenotrophomonas maltophilia C6. Int Biodeterior Biodegradation 79, 98–104.

Gibson J, and Harwood CS. (2002). Metabolic diversity in aromatic compound utilization by anaerobic microbes. Annu Rev Microbiol 56, 345–369.

Golyshin PN, Dos Santos VAPM, Kaiser O, et al. (2003). Genome sequence completed of Alcanivorax borkumensis, a hydrocarbon-degrading bacterium that plays a global role in oil removal from marine systems. J Biotechnol 106, 215–220.

Gonzalez E, Pitre FE, Pagé AP, et al. (2018). Trees, fungi and bacteria: Tripartite metatranscriptomics of a root microbiome responding to soil contamination. Microbiome 6, 53.

Gouy M, Guindon S, and Gascuel O. (2009). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27, 221–224.

Graça AP, Viana F, Bondoso J, et al. (2015). The antimicrobial activity of heterotrophic bacteria isolated from the marine sponge Erylus deficiens (Astrophorida, Geodiidae). Front Microbiol 6, 389.

Harwood CS, Burchhardt G, Herrmann H, and Fuchs G. (2020). Anaerobic metabolism of aromatic compounds via the benzoyl-CoA pathway. FEMS Microbiology Reviews 22, 439–458.

Ismail SA (2019). Exxon Valdez oil spill: What we can learn to avoid second mistake in oil transportation. ILKKM Journal of Medical and Health Sciences 1, 1–6.

Jenke-Kodama H, Sandmann A, Müller R, and Dittmann E. (2005). Evolutionary implications of bacterial polyketide synthases. Mol Biol Evol 22, 2027–2039.

Jensen LJ, Julien P, Kuhn M, et al. (2008). eggNOG: Automated construction and annotation of orthologous groups of genes. Nucleic Acids Res 36, D250–D254.

Kennish M. (2017). Practical handbook of estuarine and marine pollution. CRC Press, Boca Roton, Florida, USA.

Khater S, Gupta M, Agrawal P, et al. (2017). SBSPKSv2: Structure-based sequence analysis of polyketide synthases and non-ribosomal peptide synthetases. Nucleic Acids Res 45, W72–W79.

Kleinkauf H, and Von Döhren H. (1996). A nonribosomal system of peptide biosynthesis. Eur J Biochem 236, 335–351.

Klemetsen T, Raknes IA, Fu J, et al. (2017). The MAR databases: Development and implementation of databases specific for marine metagenomics. Nucleic Acids Res 46, D692–D699.

Koglin A, and Walsh CT. (2009). Structural insights into nonribosomal peptide enzymatic assembly lines. Nat Prod Rep 26, 987–1000.

Kube M, Beck A, Meyerdierks A, et al. (2005). A catabolic gene cluster for anaerobic benzoate degradation in methanotrophic microbial Black Sea mats. Syst Appl Microbiol 28, 287–294.

Kumari B, Singh SN, and Singh DP. (2012). Characterization of two biosurfactant producing strains in crude oil degradation. Process Biochem 47, 2463–2471.

Li C, Lai Q, Li G, et al. (2014). Multilocus sequence analysis for the assessment of phylogenetic diversity and biogeography in hyphomonas bacteria from diverse marine environments. PLoS One 9, e101394.

Lipińska A, Kucharski J, and Wyszkowska J. (2014). Activity of arylsulphatase in soil contaminated with polycyclic aromatic hydrocarbons. Water Air Soil Pollut 225, 2097.

López-Urrutia Á, and Morán XAG. (2007). Resource limitation of bacterial production distorts the temperature dependence of oceanic carbon cycling. Ecology 88, 817–822.

Martin M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17, 10–12.

Matsubara M, Lynch JM, and De Leij F. (2006). A simple screening procedure for selecting fungi with potential for use in the bioremediation of contaminated land. Enzyme Microb Technol 39, 1365–1372.

McKew BA, Coulon F, Osborn AM, Timmis KN, and McGenity TJ. (2007). Determining the identity and roles of oil-metabolizing marine bacteria from the Thames estuary, UK. Environ Microbiol 9, 165–176.

Mikheenko A, Prjibelski A, Saveliev V, Antipov D, and Gurevich A. (2018). Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 34, i142–i150.

Morcos SA. (1960). Die Verteilung des Salzgehaltes im Suez-Kanal. Kieler Meeresforsch (The Distribution of Salinity in the Suez Canal)16, 133–154.

NCBI Resource Coordinators. (2018). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 46, D8–D13.

Nikhil T, Deepa V, Rohan G, and Satish B. (2013). Isolation, characterization and identification of diesel engine oil degrading bacteria from garage soil and comparison of their bioremediation potential. Int Res J Environmental Sci (International Science Congress Association) 2, 48–52.

Overbeek R, Begley T, Butler RM, et al. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 33, 5691–5702.

Pacheco M, and Santos MA. (2001). Biotransformation, endocrine, and genetic responses of *Anguilla anguilla* L. to petroleum distillate products and environmentally contaminated waters. Ecotoxicol Environ Saf 49, 64–75.

Parales RE, and Resnick SM. (2006). Aromatic ring hydroxylating dioxygenases BT. In: *Pseudomonas: Volume 4 Molecular Biology of Emerging Issues*. Ramos JL, and Levesque RC, eds. Boston, MA: Springer US, 287–340.

Radwan SS, Barabás G, Sorkhoh NA, et al. (1998). Hydrocarbon uptake by *Streptomyces*. FEMS Microbiol Lett 169, 87–94.

Rahman KS, Thahira-Rahman J, Lakshmanaperumalsamy P, and Banat I. (2002). Towards efficient crude oil degradation by a mixed bacterial consortium. Bioresour Technol 85, 257–261.

Rivkin RB, Anderson MR, and Lajzerowicz C. (1996). Microbial processes in cold oceans. I. Relationship between temperature and bacterial growth rate. Aquat Microb Ecol 10, 243–254.

Said TO, and El Agroudy NA. (2006). Assessment of PAHs in water and fish tissues from Great Bitter and El Temsah lakes, Suez Canal, as chemical markers of pollution sources. Chem Ecol 22, 159–173.

Schaum J, Cohen M, Perry S, et al. (2010). Screening level assessment of risks due to dioxin emissions from burning oil from the BP Deepwater Horizon Gulf of Mexico spill. Environ Sci Technol 44, 9383–9389.

Shreadah MA, El Moneam NMA, Al-Assar SA, and Nabil-Adam A. (2018). Phytochemical and pharmacological screening of Sargassium vulgare from Suez Canal, Egypt. Food Sci Biotechnol 27, 963–979.

Sorkhoh N, Ghannoum M, Ibrahim AS, Stretton RJ, and Radwan SS. (1990). Crude oil and hydrocarbon-degrading strains of Rhodococcus rhodochrous isolated from soil and marine environments in Kuwait. Environ Pollut 65, 1–17.

Tamura K, Dudley J, Nei M, and Kumar S. (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24, 1596–1599.

Tanaka D, Tanaka S, Yamashiro Y, and Nakamura S. (2008). Distribution of oil-degrading bacteria in coastal seawater, Toyama Bay, Japan. Environ Toxicol 23, 563–569.

Thomsen PF, Kielgast J, Iversen LL, et al. (2012). Detection of a diverse marine fish fauna using environmental DNA from seawater samples. PLoS One 7, e41732.

Tundo P, Reda LA, Mosleh YY, and Ahmed MT. (2005). Residues of PCDD, PCDF, and PCBs in some marine organisms in Lake Temsah, Ismailia, Egypt. Toxicol Environ Chem 87, 21–30.

Vaillancourt FH, Bolin JT, and Eltis LD. (2006). The ins and outs of ring-cleaving dioxygenases. Crit Rev Biochem Mol Biol 41, 241–267.

Vázquez-Domínguez E, Vaqué D, and Gasol JM. (2007). Ocean warming enhances respiration and carbon demand of coastal microbial plankton. Glob Change Biol 13, 1327–1334.

White PA, Kalff J, Rasmussen JB, and Gasol JM. (1991). The effect of temperature and algal biomass on bacterial production and specific growth rate in freshwater and marine habitats. Microb Ecol 21, 99–118.

Xi L, Ruan J, and Huang Y. (2012). Diversity and biosynthetic potential of culturable actinomycetes associated with marine sponges in the China seas. Int J Mol Sci 13, 5917–5932.

Zaki MS, Authman MMN, Ata NS, Abdelzaher MF, and Hammam AM. (2014). Effects of environmental oil spills on commercial fish and shellfish in Suez Canal and Suez Gulf regions. Life Sci J 11, 269–274.

Zheng M, Ahuja M, Bhattacharya D, et al. (2014). Evaluation of differential cytotoxic effects of the oil spill dispersant Corexit 9500. Life Sci 95, 108–117.

Zhou H, Qu Y, Wu H, et al. (2009). Molecular phylogenies and evolutionary behavior of AhR (aryl hydrocarbon receptor) pathway genes in aquatic animals: Implications for the toxicology mechanism of some persistent organic pollutants (POPs). Chemosphere 78, 193–205.

Ziemert N, Podell S, Penn K, et al. (2012). The natural product domain seeker NaPDoS: A phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. PLoS One 7, e34064.

Zilli M, Del Borghi A, and Converti A. (2000). Toluene vapour removal in a laboratory-scale biofilter. Appl Microbiol Biotechnol 54, 248–254.

Address correspondence to:
Amro Hanora, PhD
Department of Microbiology and Immunology
Faculty of Pharmacy
Suez Canal University
Ismailia 41522
Egypt

E-mail: ahanora@yahoo.com

Abbreviations Used

ACP = acyl carrier protein
HCs = hydrocarbons
KS = ketosynthase
NCBI = National Center for Bioecology Information
NRPSs = nonribosomal peptide synthetases
PAHs = polycyclic aromatic hydrocarbons
PKSs = polyketide synthases