MOLECULAR CHARACTERIZATION OF

AROMATIC HYDROCARBON DEGRADATION

IN EXTREME HALOPHILES

By

SONAL DALVI

Bachelor of Science in Botany
University of Mumbai
Mumbai, Maharashtra, India
2004

Master of Science in Botany
University of Mumbai
Mumbai, Maharashtra, India
2006

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2014

MOLECULAR CHARACTERIZATION OF AROMATIC HYDROCARBON DEGRADATION
IN EXTREME HALOPHILES

Dissertation Approved:

Dr. Babu Fathepure

Dissertation Adviser

Dr. Mostafa Elshahed

Dr. Jeff Hadwiger

Dr. Marianna Patrauchan

Dr. Udaya Desilva

ACKNOWLEDGEMENTS

This great journey would not have been possible without the kind support, patience and guidance of the following people. I would like to take this opportunity to express my gratitude to all of them who in some or the other way helped me accomplish this challenging project. It is to them that I owe my deepest gratitude.

Firstly, I would like to thank my adviser Dr. Babu Fathepure who gladly accepted me as a PhD student in his lab. I greatly appreciate his guidance, his supportive and understanding nature during my time as a PhD student. Dr. Babu helped me improve on my presentation skills by coaching me and provided ample opportunities for me to attend regional and national conferences. Additionally, I would also like to thank my committee members Dr. Elshahed, Dr. Patrauchan, Dr. Hadwiger and Dr. DeSilva for their insightful comments and valuable suggestions during committee meetings.

During my time in the Fathepure lab, I have had the pleasure of working alongside past graduate students Madhu Prabhakaran and Abhaya Ranganathan as well as undergraduates Chris Henning, Elliott Haastrup, Brian Hercyk, Eve Adams and Colin Jackson. I would like to express my gratitude towards staff and other colleagues of the Department of Microbiology and Molecular Genetics for their welcoming and friendly nature. Special thanks to Sallie Robinson and Connie Budd. I am also grateful to Drs. Prade, Burnap and Hoff as well as their respective lab members for allowing me to use their lab facilities.

Finally, this would not have been possible without my family and friends. I have been blessed with a loving and supportive family. My parents and my sister, who have always supported, encouraged and believed in me, in all my endeavors. My mother, Reshma Dalvi, taught me the importance of education and shaped me to become the person that I am today. My father, Uday Dalvi, taught me about patience and the art of working around obstacles. Both have been wonderful role models in my life. I am also grateful to my sister, Dharitri Dalvi, who has been a source of constant and unconditional love to me for as long as I can remember. Rahul Mirani, my husband, without whose love and understanding this journey would not have been possible. He has always inspired and motivated me to realize my dreams. His patience and constant support throughout the study have been instrumental in the completion of this project. I would also like to thank my in-laws for their encouragement and profound understanding. I am indebted to all my friends who have supported me over the past few years. Special thanks to Dr. Rachana Rathod, Dr. Aparna Nagarajan, Jayalakshmi Nair and Dr. Sayali Saykhedkhar who despite having their own work pressures have been there for me throughout my tough times and always encouraged me to do my best.

Name: SONAL DALVI

Date of Degree: JULY, 2014

Title of Study: MOLECULAR CHARACTERIZATION OF AROMATIC HYDROCARBON DEGRADATION IN EXTREME HALOPHILES

Major Field: MICROBIOLOGY AND MOLECULAR GENETICS

Abstract: Many hypersaline environments such as natural saline lakes, salt marshes, saline industrial effluents, oil fields and coastal areas are often contaminated with high levels of petroleum hydrocarbons. Since non-halophiles do not operate efficiently at high salinity, halophilic and halotolerant hydrocarbon degraders are considered as potential candidates for bioremediation of hydrocarbon-impacted hypersaline sites. Several studies have reported the ability of microbial consortia and pure cultures of halophiles and halotolerants to degrade petroleum compounds. However, information regarding the genes and mechanisms of hydrocarbon degradation pathways at high salinity is scarce. This work describes the metabolic potential of halophilic bacteria and archaea to degrade aromatic hydrocarbons at high salinity. A combination of omics-based approaches was used to study the molecular basis of hydrocarbon degradation in *Arhodomonas* sp. strain Seminole and *Arhodomonas* sp. strain Rozel. Genomic analysis of strain Seminole predicted clusters of genes encoding enzymes involved in the metabolism of benzene, toluene, 4-hydroxybenzoic acid and phenylacetic acid to Krebs cycle intermediates. Many key enzymes of the predicted steps were identified in the cytosolic proteomes of hydrocarbon-grown cells by liquid chromatography-mass spectrometry, thus confirming the genomic data. Although these proteins have been described in non-halophiles, they differed from their non-halophilic homologs by exhibiting low-pI values, a unique feature known to maintain the stability and activity of halophilic proteins at high salinities. In addition, another aim of this project was to study the degradation potential of archaea in the presence of high salt. A highly enriched culture developed using sediment samples from Rozel Point; Great Salt Lake utilized benzoate as the growth substrate at salinity ranging from 2 to 5 M NaCl with highest rate of degradation at 4 M. Pyrosequencing of 16S rRNA gene sequences revealed that the enrichment composed solely of *Halobacteriaceae* members. Of these, *Halopenitus* was the most dominant member comprising of 91% of the enrichment. PCR amplification with degenerate primers revealed the presence of 4-hydroxybenzoate 3-monooxygenase and protocatechuate 3,4-dioxygenase genes, suggesting that the enrichment might degrade benzoate *via* protocatechuate. These results suggest the potential role of *Halopenitus* members in degrading oxygenated aromatic compounds at high salinity.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

1.1 Overview

Hypersaline ecosystems are defined as extreme environments with salt concentrations exceeding that of seawater (3.5%) and closer to saturation (35%). These include hypersaline lakes like the Great Salt Lake (GSL) in Utah, USA and the Dead Sea in Israel-Palestine-Jordon. The Wadi Natrun lakes of Egypt, Lake Magadi in Kenya, and the Great Basin lakes of the western United States (Mono Lake, Owens Lake, Searles Lake and Big Soda Lake) are examples of alkaline hypersaline soda brines that lack magnesium and calcium divalent cations (1). Other hypersaline environments include small evaporation ponds or sabkhas found near coastal areas like the Arabian Gulf or Guerrero Negro on the Baja California coast. Despite the extreme conditions found in hypersaline environments, they harbor diverse microbial communities that are able to survive and grow in such hostile environments. These environments are considered to be of considerable economic, ecological and scientific value (2).

Many hypersaline environments are often contaminated with high levels of petroleum hydrocarbons. For example, oil and natural gas exploration activities generate a large volume of produced waters all over the world. Approximately 10 barrels of produced waters are generated for every barrel of oil produced (http://www.epa.gov/radiation/tenorm/oilandgas.html) and these wastewaters contain high levels of salt (1000 to 250,000 ppm), oil and grease, heavy metals, trace elements and naturally occurring radioactive materials (3, 4). Presently, >95% of produced waters are re-injected, however prior to 1965-1970, a high percentage of produced waters were released to the surface and shallow subsurface (http://www.usgs.gov/water/). Remediation of produced water is required if it is intended for ocean disposal by meeting regulatory limits. However, treatment of produced water can be expensive for oil and gas production facilities (5). The Arabian Gulf with numerous offshore oil and gas platforms and major oil terminals with frequent oil spills unintentional or intentional (for example, military activities) have resulted in contamination of vast area of shorelines (6). Naturally occurring oil seeps in the northern arm of GSL at Rozel Point and the past oil production activities have led to abandoned leaking oil wells and storage tanks causing contamination of the lake (7). Also, industrial processes such as pesticide, chemical, and pharmaceutical production generate large amounts of highly saline wastewaters (8).

Clean-up of contaminated areas and treatment of industrial wastewaters can be achieved by bioremediation processes that are considered to be economical and environmentally safe (9). Bioaugmentation with hydrocarbon degrading organisms or biostimulation of indigenous hydrocarbon degrading organisms are two main approaches to bioremediation processes (10). Several reviews are available in the literature that report on the ability of many well-characterized bacteria belonging to the genera *Arthrobacter*, *Burkholderia*, *Mycobacterium*, *Pseudomonas*, *Sphingomonas*, *Rhodococcus*, *Acinetobacter*, *Actinopolyspora*, *Brevibacterium*, *Burkholderia*, *Corynebacterium*, *Flavobacterium*, *Nocardia*, *Nocardiodes*, and *Gordonia* to degrade hydrocarbons under oxic

conditions (10-13). Members of genera *Azoarcus, Dechloromonas, Rhodopseudomonas*, *Thaurea* and other organisms have been shown to degrade hydrocarbons anaerobically (14).

However, biological treatment of hypersaline industrial effluents and bioremediation of contaminated hypersaline environments is not possible by non-halophilic microorganisms, as they do not function efficiently at high salt concentrations (15, 16). Salt causes detrimental effects such as loss of cell wall integrity, denaturation of proteins affecting their ability to degrade hydrocarbons (17). Therefore, halophilic and halotolerant microorganisms that can tolerate high salt concentrations and degrade petroleum hydrocarbons are considered as potential candidates for bioremediation of contaminated hypersaline environments.

Halophiles are found in all domains of life: bacteria, archaea, and eukarya and are capable of carrying out metabolic functions at salinity ranging from 0.2 M-5.2 M. Halophiles can be classified based on their salt requirements for optimal growth: slight halophiles (0.2-0.5 M), moderate halophiles (0.5-2.5 M) and extreme halophiles (2.5-5.2 M) (18, 19). In the past two decades, there has been an increase in attention towards the diversity of halophiles that degrade hydrocarbons. Many reviews have contributed towards our knowledge on the ability of these microorganisms to degrade a range of hydrocarbons at varying salinities (2, 5, 20-23). The following section will provide an overview on the degradation of oxygenated and non-oxygenated hydrocarbons by bacteria and archaea at moderate to high salt conditions.

1.2 Hydrocarbon degradation by halophilic microbial consortia

Several studies have reported the ability of microbial consortia to degrade crude oil, aliphatic and polyaromatic as well as monoaromatic compounds such as benzene, toluene, ethylbenzene and xylenes (BTEX). For example, bacterial consortia isolated from North Sea (MPD-7) and sediments associated with the mangrove roots (MPD-M) were shown to effectively degrade both aliphatic and aromatic hydrocarbons in crude oil at salinity ranging from 0 to 22% NaCl. One of the consortium

mainly composed of species of genera *Marinobacter*, *Bacillus* and *Erwinia* (24). Similarly, another microbial consortium comprising of *Cellulomonas*, *Bacillus*, *Dietzia* and *Halomonas* developed using Argentinean saline soils degraded diesel at salinity up to 17.5% (25). A cyanobacterial-rich microbial mat from the Arabian Gulf coast of Saudi Arabia was shown to degrade petroleum compounds such as pristine, *n*-octadecane, phenanthrene and dibenzothiophene at salinity ranging between 3.5-12% (26). A halophilic bacterial consortium developed from soil samples from Shengli Oilfield in China degraded phenanthrene at 5-15% salinity (27). Another halophilic bacterial consortium (Qphe) degraded a wide range of polyaromatic hydrocarbons (PAHs) at salinity as high as 17% (28). Similar reports exist on the biodegradation of BTEX compounds at salt concentration ranging up to 29% NaCl by microbial consortia developed from soil samples from an oilfield in Oklahoma (29), samples from uncontaminated Great Salt Plains flat in Oklahoma (30) and sediment sample from Rozel Point, GSL (31). These studies clearly suggest the ability of bacterial consortia to degrade a range of hydrocarbons and their potential use in bioremediation of contaminated hypersaline environments. However, so far only one study has reported the ability of an archaeal community to degrade hydrocarbons in hypersaline conditions. Zviagintseva et al. developed enrichment form brines of Kalamkass oil fields in Kazakhstan that degraded a significant amount of isoprenoid and *n*-alkane fractions of crude oil (32).

1.3 Hydrocarbon degradation by halophilic and halotolerant bacteria

In the past two decades, many pure cultures of hydrocarbon degrading halophilic and halotolerant bacteria and archaea have been isolated. Among bacteria, members of the genera *Marinobacter*, *Alcanivorax*, *Halomonas* and few halophilic species of the order *Bacillales* and *Actinomycetales* were also shown to degrade a range of hydrocarbons at varying salinity.

Marinobacter spp. are commonly isolated from contaminated hypersaline sediments and oil reservoirs and have been shown to degrade hydrocarbons (23, 33). *Marinobacter*

hydrocarbonoclasticus strain SP17 was shown to utilize hexadecane, eicosane, heneicosane as the sources of carbon and energy at salinity ranging from 4 to 20% (34). *Marinobacter aquaeolei* isolated from oil producing well was shown to degrade crude oil in the presence of 0-20% salt (35). *Marinobacter vinifirmus* and *Marinobacter hydrocarbonoclasticus* degrade BTEX compounds at 3-15% salinity (36). Recently, Al-Mailem et al. isolated two *Marinobacters*, *M. falvimaris* and *M. sedimentalis* from hypersaline sabkhas in Kuwait with the ability to degrade crude oil, aliphatic compounds, as well as PAHs at 6% NaCl (37). *Marinobacter nanhaiticus* strain D15-8W was recently isolated from a phenanthrene-degrading enrichment that can degrade naphthalene, phenanthrene or anthracene optimally at 1-5% NaCl (38).

Alcanivorax spp. are mostly associated with oil-impacted environments and are known for their ability to efficiently degrade branched chain alkanes (39). *Alcanivorax borkumensis* SK2, one of the most well-known member of the genus *Alcanivorax*, can grow on *n*-alkanes at 1-12.5% NaCl (40, 41). *Alcanivorax dieselolei* strain B-5 utilizes *n*-alkanes with a chain length C_5-C_{36} at 1-15% salinity (42). *Alcanivorax* sp. HA03 isolated from Wadi El-Natrun Soda lakes degrades benzene and toluene at 7-15% salinity (43).

Members of the genera *Halomonas* and *Chromohalobacter* are known for their ability to degrade crude oil, phenolics and benzoates (5). Several strains of *Halomonas* spp. isolated from hypersaline areas in Southern Spain were shown to degrade phenol, benzoic acid, *p*-hydroxybenzoic acid (4-HBA), *p*-coumaric acid, salicylic acid, cinnamic acid and *p*-aminosalicyclic acid in the presence of 10% NaCl (44). *Halomonas* sp. strain IMPC isolated from table-olive fermentation brines degrades *p*-coumaric acid and lignin-related aromatic compounds at 8% NaCl (45). Another study reported the isolation of *Halomonas campisalis* from a haloalkaline Soap Lake that degrades benzoate and 4-HBA at 5-10% NaCl (46). *Halomonas shengliensis* and *Halomonas* sp. strain C2SS100 have also been shown to degrade crude oil as carbon source at salinity up to 15 % NaCl (47, 48). So far,

only one report exists on the ability of *Chromohalobacter* sp. strain HS-2 isolated from salted fermented clams to grow on benzoate and 4-HBA optimally at 10% NaCl (49).

A few halophilic and halotolerant members of *Actinomycetales* and *Bacillales* have also been reported to degrade hydrocarbons. *Streptomyces albiaxialis* strain K-3959 isolated from hypersaline brines of Perm Oblast oil field degraded crude oil at 30% NaCl (50). *Actinopolyspora* sp. DPD, an extremely halophilic actinomycete isolated from a oil field in the Sultanate of Oman degraded *n*-alkanes and flourene at 25% NaCl showing its metabolic versatility (51). Halotolerant bacteria of the genera *Cellulomonas*, *Dietzia, Rhodococcus* and *Gordonia* have been shown to degrade crude oil or *n*-alkanes in presence of salt up to 17.5% NaCl (25, 52-54). Halophilic and halotolerant members of *Bacillales* have been isolated with the ability to degrade crude oil, diesel and oxygenated hydrocarbons such as phenol, benzoate and *m*-hydroxybenzoate as sole sources of carbon (5, 23).

Overall, majority of these published reports indicate that *Halomonas* spp. are capable of degrading n-alkanes, phenolics and benzoates but little information exists about their capacity to degrade benzene, toluene, ethylbenzene, xylenes (BTEX) and polyaromatic hydrocarbons (PAHs). On the other hand, *Marinobacter* and *Alcanivorax* have been reported to degrade mainly aliphatic, BTEX and PAHs but not much is known about their ability to degrade phenols and benzoates in moderate to high salt conditions (5).

1.4 Hydrocarbon degradation by halophilic archaea

Hydrocarbon degrading archaea belonging to *Halobacteriaceae* have been isolated from extreme aquatic environments such as salt marshes, saline waters, salt flats, brines and coastal waters. Bertrand et. al were among the first to report isolation of a halophilic archaea strain EH_4 from a salt marsh in Southern France with the ability to degrade aliphatic and aromatic hydrocarbons at 20% NaCl (55). The strain was recently classified as *Haloarcula vallismortis* (56). *Halobacterium* strain H-352 identified based on its phenotypic characteristics was isolated with the ability to degrade *n*-

alkane (C_{10}-C_{30}) at 15-32% NaCl (57). *Haloferax* sp. D1227 isolated from oil brine soil in Michigan was reported to degrade aromatic acids such as benzoic acid, cinnamic acid, 3-hydroxybenzoic acid and 3-phenylpropionic acid ranging from 5 to 30% NaCl (58). *Haloarcula* sp. D1 was reported to degrade 4-hydroxybenzoic acid (4-HBA) as the sole carbon source (59). *Haloferax*, *Halobacterium* and *Halococcus* strains isolated from Arabian Gulf degraded aliphatic, mono- and polyaromatic hydrocarbons as the sole sources of carbon at 26% NaCl (60). Tapilatu et. al isolated archaeal strains closely related to *Haloarcula* and *Haloferax* from a shallow crystallizer pond in France with no known history of contamination and showed their potential to degrade heptadecane at 22.5% NaCl. Of these isolates, one strain also degraded phenanthrene (56). Ten haloarchaeal strains isolated from five different hypersaline locations degraded aromatic acids such as benzoic acid, 4-HBA, and salicylic acid at 20% NaCl. These strains also degraded a mixture of PAHs that included naphthalaene, anthracene, phenanthrene, pyrene and benzo[a]anthracene (61). Erdoğmuş et al. showed the degradation of 4-HBA, naphthalaene, phenanthrene and pyrene by archaea at 20% NaCl and were identified as *Halobacterium piscisalsi*, *Halorubrum ezzemoulense*, *Halobacterium salinarium*, *Haloarcula hispanica*, *Haloferax* sp., *Halorubrum* sp., and *Haloarcula* sp. by 16S rRNA gene sequences (62). Overall, these reports clearly suggest the metabolic potential of archaea belonging to *Haloferax*, *Haloarcula*, *Halobacterium, Halococcus* and *Halorubrum* to degrade a range of hydrocarbons including *n*-alkanes of varying length, PAHs and aromatic acids at high salinity (up to 32% NaCl).

1.5 Aerobic hydrocarbon degradation pathways

Pathways and enzymes involved in aerobic hydrocarbon metabolism in several non-halophiles have been studied extensively (12, 13, 63). In non-halophiles, hydroxylation of alkanes at one of the terminal or sub-terminal carbon atoms is initiated by monooxygenases. This is followed by oxidation by an aldehyde dehydrogenase to form corresponding fatty acid. This acid is further converted into acetyl-CoA *via* the β-oxidation pathway and assimilated into central metabolism

through the tricarboxylic acid (TCA) cycle (12). Degradation of aromatic hydrocarbons is initiated by breakdown of the energetically stable aromatic rings and their subsequent cleavage and utilization as growth substrates. Initial activation of the ring is catalyzed by multicomponent enzyme complexes such as soluble diiron monooxygenase or Rieske non-heme iron oxygenases (64, 65) to form a few central intermediates including catechols, protocatechuates, gentisates, or (hydroxyl)benzoquinols (63). Catechols and protocatechuates are further cleaved at *ortho*- or *meta*- position by intradiol or extradiol dioxygenases. During the *ortho*-cleavage, catechol (CAT) and protocatechuate (PCA) are cleaved between their hydroxyl groups by catechol 1,2-dioxygenase (1,2-CAT) or protocatechuate 3,4-dioxygenase (3,4-PCA) respectively. In the *meta*-cleavage, CAT and PCA undergo ring-cleavage adjacent to one of the hydroxyls catalyzed by catechol 2,3-dioxygenase (2,3-CAT) and protocatechuate 4,5-dioxygenase (4,5-PCA) respectively (63). The ring-cleavage products are further metabolized into intermediates that enter the Krebs cycle. Gentisate and substituted gentisates are cleaved by gentisate 1,2-dioxygenase (1,2-GDO) to yield methylpyruvate that is further hydrolytically cleaved to pyruvate and maleate and enter Krebs cycle (63). Alternatively, aromatic hydrocarbons can also be metabolized aerobically *via* non-oxygenolytic cleavage and formation of corresponding CoA thioesters (63). For example aromatic compounds such as styrene, ethylbenzene, tyrene, 2-phenylethylamine, phenylacetaldehyde, *n*-phenylalkanoates, tropic acid, phenylacetyl esters and amides are metabolized *via* phenylacetate (PAA) degradation pathway. PAA is transformed *via* CoA ligase to phenylacetyl-CoA that undergoes subsequent ring hydroxylation, hydrolytic ring-cleavage and further conversion to TCA cycle intermediates (66). Similar aerobic hybrid pathways for degradation of benzoate *via* benzoyl-CoA, anthranilate *via* 2-aminobenzoyl-CoA and salicylate *via* salicyl-CoA have been reported (63).

Little information exists about the pathways of hydrocarbon degradation at moderate to high salt conditions. Recently, a few studies have identified genes and enzymes involved in hydrocarbon degradation in halophilic organisms using molecular and biochemical approaches. For example, PCR

with degenerate primers revealed the presence of genes that code for ring-cleavage enzymes such as 1,2-CAT and 3,4-PCA in several strains of phenol- and benzoate degrading *Halomonas* spp. Activities of these enzymes was also detected in *Halomonas organivorans* when grown on various aromatic compounds (44). Moreno et al. identified the gene cluster *catRBCA* involved in catechol degradation in a moderate halophile, *Halomonas organivorans* G-16.1. Also found in contiguous with the *cat* operon were genes required for benzoate degradation (*benAB*). The study also demonstrated the expression of *cat* and *ben* genes by reverse transcriptase-PCR (RT-PCR) analysis in *H. organivorans* on induction with phenol and benzoic acid (67). Kim et al. elucidated benzoate and 4-HBA degradation pathways in *Chromohalobacter* sp. HS2 using RT-PCR and metabolite analyses. Their data indicated that benzoate 1,2-dioxygenase catalyzes benzoate to form *cis*-benzoate dihydrodiol that undergoes dehydrogenation to form CAT that is further cleaved *via ortho*-cleavage pathway while 4-HBA was catalyzed by *p*-hydroxybenzoate hydroxylase to form PCA (49). Few studies have elucidated hydrocarbon degradation pathways using gas chromatography (GC), mass spectrometry (MS) and high-performance liquid chromatography (HPLC). For example, CAT and *cis, cis*-muconate were detected as intermediates of benzoate and phenol degradation in *Halomonas* spp. using HPLC (46, 68, 69). Using GC-MS, *p*-hydroxybenzaldehyde and 4-HBA were detected as breakdown products of *p*-coumaric acid degradation in a moderate halophile; *Halomonas* strain IMPC (45). Metabolite analysis by HPLC and GC–MS showed that 2-hydroxy 1-naphthoic acid and 2-naphthol were among the major metabolites during phenanthrene degradation by Qphe-SubIV consortium (28).

A few recent studies have reported mechanisms of hydrocarbon degradation by archaea in high salt conditions. The enzyme, 1,2-GDO was purified and characterized from *Haloferax* sp. strain D1227 that utilizes a variety of aromatic compounds (58, 70). A closely related 1,2-GDO gene was also found in 4-HBA degrading *Haloarcula s*p. strain D1 (71). Along with gentisate being detected as an intermediate by GC-MS; 1,2-GDO enzyme activity was also measured during 4-HBA degradation

in strain D1 (59). A recent study detected enzyme activities of 1,2-CAT and 3,4-PCA in the archaeal isolates belonging to *Haloferax, Halorubrum,* and *Halobacterium* when grown on aromatic hydrocarbons (62). These reports have clearly suggested that halophilic bacteria and archaea degrade hydrocarbons using enzymes and pathways similar to those described in non-halophiles.

1.6 Current research focus

Inspite of the considerable amount of literature on hydrocarbon degradation potential of halophilic and halotolerant microorganisms, very little information is available on the genetics and biochemistry of hydrocarbon degradation in high salinity environments. Research efforts on characterization of halophilic hydrocarbon degraders and understanding the mechanisms of hydrocarbon degradation at high salt conditions are necessary. Such studies will help determine the factors that influence halophiles or halotolerants to utilize hydrocarbons as substrates and this in turn will help to develop cost-effective bioremediation technologies for clean-up of contaminated extreme environments. Recent advances in genomics and high throughput -omics has enabled researchers to determine the metabolic potential and cellular activity of microorganisms in various environmental conditions (13). Analysis of genome sequences provides us with valuable insights into the genetic basis of hydrocarbon degradation and predicts the functions of candidate genes (13). Proteomics is an effective method to confirm whether the candidate genes and enzymes predicted by genomics are actually expressed and involved in hydrocarbon metabolism (72). Metabolomics is another powerful tool that can be used to validate the pathways by detecting and identifying the intermediates of hydrocarbon degradation. Overall, genomic, proteomic and metabolomics approaches help to better understand the bacterial physiology and regulatory mechanisms during hydrocarbon degradation (13).

This study describes the physiological and molecular characterization of hydrocarbon degradation pathways at high salt conditions by bacteria and archaea. Chapter 2 investigates the initial

steps involved in benzene degradation pathway in a novel halophile, *Arhodomonas* sp. strain Rozel. The strain was isolated by Mr. Sei Azetsu previously in our lab from an enrichment developed using sediment obtained from GSL. Using a draft genome of closely related strain Seminole, a pathway for benzene degradation was predicted and then validated using a two-dimensional gel electrophoresis followed by liquid chromatography-mass spectrometry (LC-MS/MS). Proteogenomic data was corroborated using GC-MS where phenol was detected as an initial metabolite of benzene degradation. This chapter has been published in the Journal of Applied and Environmental Microbiology.

Chapter 3 describes the reconstruction of the catabolic steps involved in degradation of a variety of hydrocarbons in another halophile, *Arhodomonas* sp. strain Seminole. This strain was previously isolated by Ms. Carla Nicholson in our lab from an enrichment developed using brine soil obtained from oil production facility in Seminole, Oklahoma and shown to degrade benzene and toluene at 2.5 M NaCl. The metabolic versatility of the strain was tested by its ability to utilize other oxygenated aromatics like hexadecane, phenol, catechol, benzoate, PCA, 4-HBA, and PAA as carbon substrates. Among these, strain Seminole degraded PCA, PAA and 4-HBA as sole carbon sources. A draft genome sequence of strain Seminole was used to predict genes involved in the degradation steps of each of the above hydrocarbons. Enzymes involved in the predicted degradation pathways were confirmed by LC-MS/MS analysis of the cytosolic proteomes of hydrocarbon-grown cells. This work has been submitted to Journal of Applied and Environment Microbiology and is currently under revision.

Chapter 4 explores the degradation potential of a highly enriched archaeal consortium that degrades benzoate at salinities ranging from 2 to 5 M NaCl. The enrichment was also able to utilize 4-HBA as the sole carbon source at 4 M NaCl. Microbial diversity and community structure of the consortium was studied using the pyrosequencing technique. The enrichment comprised entirely of members belonging to the family *Halobacteriaceae*. In addition, the enrichment was screened for

presence of catabolic genes that may be involved in benzoate degradation using degenerate primers. This led to the identification of 4-hydroxybenzoate 3-monooxygenase and 3,4-PCA genes that are involved in initial steps of benzoate degradation.

1.7 REFERENCES:

1. **DasSarma S, and DasSarma P.** 2012. Halophiles, eLS. John Wiley & Sons, Ltd. DOI: 10.1038/npg.els.0000394.
2. **Le Borgne S, Paniagua D, and Vazquez-Duhalt R.** 2008. Biodegradation of organic pollutants by halophilic bacteria and archaea. J Mol Microbiol Biotechnol. **15:**74-92.
3. **Veil JA, Puder, M.G., Elock, D. and Redweik, R.J.Jr.** 2004. A White Paper describing produced water from production of crude oil, natural gas, and coal bed methane. US Department of Energy, National Energy Technology Laboratory 1-87.
4. **Cuadros-Orellana S, Pohlschröder M, and Durrant LR.** 2006. Isolation and characterization of halophilic archaea able to grow in aromatic compounds. Int Biodeterior Biodegradation. **57:**151-154.
5. **Fathepure BZ.** 2014. Recent studies in microbial degradation of petroleum hydrocarbons in hypersaline environments. Frontiers in microbiology. **5**:173. doi: 10.3389/fmicb.2014.00173
6. **Al-Saad H, and Salman N.** 2012. Status of oil pollution in the Arabian Gulf and Shatt Al-Arab Estuary: A Review. p. 36-46. *In* V. Subramanian (ed.), Coastal Environments: Focus on Asian Regions. Springer Netherlands.
7. **Baxter B, Litchfield C, Sowers K, Griffith J, Dassarma P, and Dassarma S.** 2005. Microbial diversity of Great Salt Lake. p. 9-25. *In* N. Gunde-Cimerman, A. Oren, and A. Plemenitaš (ed.), Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya, vol. 9. Springer Netherlands.

8. **Lefebvre O, and Moletta R.** 2006. Treatment of organic pollution in industrial saline wastewater: a literature review. Water Res. **40:**3671-3682.

9. **Margesin R, and Schinner F.** 2001. Bioremediation (natural attenuation and biostimulation) of diesel-oil-contaminated soil in an alpine glacier skiing area. Appl Environ Microbiol. **67:**3127-3133.

10. **Das N, and Chandran P.** 2011. Microbial Degradation of Petroleum Hydrocarbon Contaminants: An Overview. Biotechnology Research International. **2011**.doi:10.4061/2011/941810.

11. **Van Hamme JD, Singh A, and Ward OP.** 2003. Recent advances in petroleum microbiology. Microbiol Mol Biol Rev. **67:**503-549.

12. **Patzelt H.** 2001. Hydrocarbon Metabolism, eLS. John Wiley & Sons, Ltd, Chichester. http://www.els.net [doi: 10.1002/9780470015902.a0000473].

13. **Cao B, Nagarajan K, and Loh K-C.** 2009. Biodegradation of aromatic compounds: current status and opportunities for biomolecular approaches. Appl Microbiol Biotechnol. **85:**207-228.

14. **Widdel F, and Rabus R.** 2001. Anaerobic biodegradation of saturated and aromatic hydrocarbons. Curr Opin Biotechnol. **12:**259-276.

15. **Oren A.** 2002. Diversity of halophilic microorganisms: environments, phylogeny, physiology, and applications. J Ind Microbiol Biotechnol. **28:**56-63.

16. **Pieper DH, and Reineke W.** 2000. Engineering bacteria for bioremediation. Curr Opin Biotechnol. **11:**262-270.

17. **Pernetti M, and Di Palma L.** 2005. Experimental evaluation of inhibition effects of saline wastewater on activated sludge. Environ Technol. **26:**695-703.

18. **Oren A.** 2006. Life at high salt concentrations. p. 263-282. *In* M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt (ed.), The Prokaryotes. Springer New York.

19. **Kushner D.** 1978. Microbial life in extreme environments, p. xii, 465 p. Academic Press, London ;New York.

20. **Sorokin DY, Janssen AJH, and Muyzer G.** 2011. Biodegradation potential of halo(alkali)philic prokaryotes. Crit Rev Environ Sci Technol. **42:**811-856.

21. **Martins LF, and Peixoto RS.** 2012. Biodegradation of petroleum hydrocarbons in hypersaline environments. Braz J Microbiol. **43:**865-872.

22. **Patzelt H.** 2005. Hydrocarbon degradation under hypersaline conditions - some facts, some experiments and many open questions. p. 105-122. *In* N. Gunde-Cimerman, A. Oren, and A. Plemenitaš (ed.), Adaptation to life at high salt concentrations in Archaea, Bacteria, and Eukarya, vol. 9. Springer Netherlands.

23. **McGenity TJ.** 2010. Halophilic Hydrocarbon Degraders. p. 1939-1951. *In* K. N. Timmis (ed.), Handbook of Hydrocarbon and Lipid Microbiology. Springer-Verlag, Berlin Heidelberg.

24. **Diaz PM, Grigson SJ, Peppiatt CJ, and Burgess JG.** 2000. Isolation and characterization of novel hydrocarbon-degrading euryhaline consortia from crude oil and mangrove sediments. Mar Biotechnol (NY). **2:**522-532.

25. **Riis V, Kleinsteuber S, and Babel W.** 2003. Influence of high salinities on the degradation of diesel fuel by bacterial consortia. Can J Microbiol. **49:**713-721.

26. **Abed RM, Al-Thukair A, and de Beer D.** 2006. Bacterial diversity of a cyanobacterial mat degrading petroleum compounds at elevated salinities and temperatures. FEMS Microbiol Ecol. **57:**290-301.

27. **Zhao B, Wang H, Mao X, and Li R.** 2009. Biodegradation of phenanthrene by a halophilic bacterial consortium under aerobic conditions. Curr Microbiol. **58:**205-210.

28. **Dastgheib SM, Amoozegar MA, Khajeh K, Shavandi M, and Ventosa A.** 2011. Biodegradation of polycyclic aromatic hydrocarbons by a halophilic microbial consortium. Appl Microbiol Biotechnol. **95**:789-798.

29. **Nicholson CA, and Fathepure BZ.** 2004. Biodegradation of benzene by halophilic and halotolerant bacteria under aerobic conditions. Appl Environ Microbiol. **70:**1222-1225.

30. **Nicholson CA, and Fathepure BZ.** 2005. Aerobic biodegradation of benzene and toluene under hypersaline conditions at the Great Salt Plains, Oklahoma. FEMS Microbiol Lett. **245:**257-262.

31. **S A, and Fathepure BZ.** 2009. Biodegradation of BTEX at high salinity by an enrichment culture from hypersaline sediments of Rozel Point at Great Salt Lake. J Appl Microbiol. **107:**2001-2008.

32. **Zviagintseva IS, Beliaev SS, Borzenkov IA, Kostrikina NA, Milekhina EI, and Ivanov MV.** 1995. Halophilic archaebacteria from the Kalamkass oilfield. Microbiology. **64:**83-87.

33. **Duran R.** 2010. Marinobacter. p. 1725-1735. *In* K. Timmis (ed.), Handbook of Hydrocarbon and Lipid Microbiology. Springer Berlin Heidelberg.

34. **Gauthier MJ, Lafay B, Christen R, Fernandez L, Acquaviva M, Bonin P, and Bertrand JC.** 1992. *Marinobacter hydrocarbonoclasticus* gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. Int J Syst Bacteriol. **42:**568-576.

35. **Huu NB, Denner EB, Ha DT, Wanner G, and Stan-Lotter H.** 1999. *Marinobacter aquaeolei* sp. nov., a halophilic bacterium isolated from a Vietnamese oil-producing well. Int J Syst Bacteriol. **49 Pt 2:**367-375.

36. **Berlendis S, Cayol J-L, Verhé F, Laveau S, Tholozan J-L, Ollivier B, and Auria R.** 2010. First evidence of aerobic biodegradation of BTEX compounds by pure cultures of *Marinobacter*. Appl Biochem Biotechnol. **160:**1992-1999.

37. **Al-Mailem DM, Eliyas M, and Radwan SS.** 2013. Oil-bioremediation potential of two hydrocarbonoclastic, diazotrophic *Marinobacter* strains from hypersaline areas along the Arabian Gulf coasts. Extremophiles. **17:**463-470.

38. **Gao W, Cui Z, Li Q, Xu G, Jia X, and Zheng L.** 2013. *Marinobacter nanhaiticus* sp. nov., polycyclic aromatic hydrocarbon-degrading bacterium isolated from the sediment of the South China Sea. Antonie Van Leeuwenhoek. **103:**485-491.

39. **Hara A, Syutsubo K, and Harayama S.** 2003. *Alcanivorax* which prevails in oil-contaminated seawater exhibits broad substrate specificity for alkane degradation. Environ Microbiol. **5:**746-753.

40. **Schneiker S, Martins dos Santos VA, Bartels D, Bekel T, Brecht M, Buhrmester J, Chernikova TN, Denaro R, Ferrer M, Gertler C, Goesmann A, Golyshina OV, Kaminski F, Khachane AN, Lang S, Linke B, McHardy AC, Meyer F, Nechitaylo T, Puhler A, Regenhardt D, Rupp O, Sabirova JS, Selbitschka W, Yakimov MM, Timmis KN, Vorholter FJ, Weidner S, Kaiser O, and Golyshin PN.** 2006. Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. Nat Biotechnol. **24:**997-1004.

41. **Yakimov MM, Golyshin PN, Lang S, Moore ERB, Abraham W-R, Lünsdorf H, and Timmis KN.** 1998. *Alcanivorax borkumensis* gen. nov., sp. nov., a new, hydrocarbon-degrading and surfactant-producing marine bacterium. International Journal of Systematic Bacteriology. **48:**339-348.

42. **Liu C, and Shao Z.** 2005. *Alcanivorax dieselolei* sp. nov., a novel alkane-degrading bacterium isolated from sea water and deep-sea sediment. Int J Syst Bacteriol. **55:**1181-1186.

43. **Hassan H A NMR, M. A. Hefnawy, and Ahmad M. Awad.** 2012. Isolation and characterization of halophilic aromatic and chloroaromatic degrader from Wadi El-Natrun Soda lakes. Life Sci. **9:**1565-1570.

44. **Garcia MT, Ventosa A, and Mellado E.** 2005. Catabolic versatility of aromatic compound-degrading halophilic bacteria. FEMS Microbiol Ecol. **54:**97-109.

45. **Abdelkafi S, Labat M, Casalot L, Chamkha M, and Sayadi S.** 2006. Isolation and characterization of *Halomonas* sp strain IMPC, a p-coumaric acid-metabolizing bacterium

that decarboxylates other cinnamic acids under hypersaline conditions. FEMS Microbiol Lett. **255:**108-114.

46. **Oie CSI, Albaugh CE, and Peyton BM.** 2007. Benzoate and salicylate degradation by *Halomonas campisalis*, an alkaliphilic and moderately halophilic microorganism. Water Res. **41:**1235-1242.

47. **Wang YN, Cai H, Chi CQ, Lu AH, Lin XG, Jiang ZF, and Wu XL.** 2007. *Halomonas shengliensis* sp. nov., a moderately halophilic, denitrifying, crude-oil-utilizing bacterium. Int J Syst Evol Microbiol. **57:**1222-1226.

48. **Mnif S, Chamkha M, and Sayadi S.** 2009. Isolation and characterization of *Halomonas* sp. strain C2SS100, a hydrocarbon-degrading bacterium under hypersaline conditions. J Appl Microbiol. **107:**785-794.

49. **Kim D, Kim SW, Choi KY, Lee JS, and Kim E.** 2008. Molecular cloning and functional characterization of the genes encoding benzoate and *p*-hydroxybenzoate degradation by the halophilic *Chromohalobacter* sp strain HS-2. FEMS Microbiol Lett. **280:**235-241.

50. **Kuznetsov VD ZT, Vakulenko LV, Filippova SN** 1992. *Streptomyces albiaxialis* sp. nov.: a new petroleum hydrocarbon-degrading species of thermo- and halotolerant *Streptomyces*. Microbiology. **61:**62-67.

51. **Al-Mueini R, Al-Dalali M, Al-Amri IS, and Patzelt H.** 2007. Hydrocarbon degradation at high salinity by a novel extremely halophilic actinomycete. Environ Chem. **4:**5-7.

52. **Plotnikova EG, Altyntseva OV, Kosheleva IA, Puntus IF, Filonov AE, Gavrish EY, Demakov VA, and Boronin AM.** 2001. Bacterial degraders of polycyclic aromatic hydrocarbons isolated from salt-contaminated soils and bottom sediments in salt mining areas. Microbiology. **70:**51-58.

53. **Zvyagintseva IS, Poglazova MN, Gotoeva MT, and Belyaev SS.** 2001. Effect on the medium salinity on oil degradation by nocardioform bacteria. Microbiology. **70:**652-656.

54. **Borzenkov IA, Milekhina EI, Gotoeva MT, Rozanova EP, and Belyaev SS.** 2006. The properties of hydrocarbon-oxidizing bacteria isolated from the oilfields of Tatarstan, western Siberia, and Vietnam. Microbiology. **75:**66-72.

55. **Bertrand JC, Almallah M, Acquaviva M, and Mille G.** 1990. Biodegradation of hydrocarbons by an extremely halophilic archaebacterium. Lett Appl Microbiol. **11:**260-263.

56. **Tapilatu YH, Grossi V, Acquaviva M, Militon C, Bertrand JC, and Cuny P.** 2010. Isolation of hydrocarbon-degrading extremely halophilic archaea from an uncontaminated hypersaline pond (Camargue, France). Extremophiles. **14:**225-231.

57. **Kulichevskaya IS, Milekhina EI, Borzenkov IA, Zvyagintseva IS, and Belyaev SS.** 1991. Oxidation of petroleum-hydrocarbons by extremely halophilic archaebacteria. Microbiology. **60:**596-601.

58. **Emerson D, Chauhan S, Oriel P, and Breznak JA.** 1994. *Haloferax* Sp D1227, a Halophilic Archaeon Capable of Growth on Aromatic-Compounds. Arch Microbiol. **161:**445-452.

59. **Fairley DJ, Boyd DR, Sharma ND, Allen CCR, Morgan P, and Larkin MJ.** 2002. Aerobic metabolism of 4-hydroxybenzoic acid in Archaea via an unusual pathway involving an intramolecular migration (NIH shift). Appl Environ Microbiol. **68:**6246-6255.

60. **Al-Mailem DM, Sorkhoh NA, Al-Awadhi H, Eliyas M, and Radwan SS.** 2010. Biodegradation of crude oil and pure hydrocarbons by extreme halophilic archaea from hypersaline coasts of the Arabian Gulf. Extremophiles. **14:**321-328.

61. **Bonfa MRL, Grossman MJ, Mellado E, and Durrant LR.** 2011. Biodegradation of aromatic hydrocarbons by haloarchaea and their use for the reduction of the chemical oxygen demand of hypersaline petroleum produced water. Chemosphere. **84:**1671-1676.

62. **Erdoğmuş S, Mutlu B, Korcan S, Güven K, and Konuk M.** 2013. Aromatic hydrocarbon degradation by halophilic archaea isolated from Çamaltı Saltern, Turkey. Water, Air, Soil Pollut. **224:**1-9.

63. **Pérez-Pantoja D, González B, and Pieper DH.** 2010. Aerobic degradation of aromatic hydrocarbons. p. 799-837. *In* K. N. Timmis (ed.), Handbook of Hydrocarbon and Lipid Microbiology. Springer-Verlag Berlin Heidelberg.

64. **Leahy JG, Batchelor PJ, and Morcomb SM.** 2003. Evolution of the soluble diiron monooxygenases. FEMS Microbiol Rev. **27:**449-479.

65. **Gibson DT, and Parales RE.** 2000. Aromatic hydrocarbon dioxygenases in environmental biotechnology. Curr Opin Biotechnol. **11:**236-243.

66. **Navarro-Llorens JM, Patrauchan MA, Stewart GR, Davies JE, Eltis LD, and Mohn WW.** 2005. Phenylacetate catabolism in *Rhodococcus* sp. strain RHA1: a central pathway for degradation of aromatic compounds. J Bacteriol. **187:**4497-4504.

67. **Moreno MD, Sanchez-Porro C, Piubeli F, Frias L, Garcia MT, and Mellado E.** 2011. Cloning, characterization and analysis of *cat* and *ben* genes from the phenol degrading halophilic bacterium *Halomonas organivorans*. Plos One. **6**. doi: 10.1371/journal.pone.0021049.

68. **Hinteregger C, and Streichsbier F.** 1997. *Halomonas* sp., a moderately halophilic strain, for biotreatment of saline phenolic waste-water. Biotechnol Lett. **19:**1099-1102.

69. **Alva VA, and Peyton BM.** 2003. Phenol and catechol biodegradation by the haloalkaliphile *Halomonas campisalis*: influence of pH and salinity. Environ Sci Technol. **37:**4397-4402.

70. **Fu W, and Oriel P.** 1998. Gentisate 1,2-dioxygenase from *Haloferax* sp. D1227. Extremophiles. **2:**439-446.

71. **Fairley DJ, Wang G, Rensing C, Pepper IL, and Larkin MJ.** 2006. Expression of gentisate 1,2-dioxygenase (gdoA) genes involved in aromatic degradation in two haloarchaeal genera. Appl Microbiol Biotechnol. **73:**691-695.

72. **Kim S-J, Kweon O, and Cerniglia CE.** 2009. Proteomic applications to elucidate bacterial aromatic hydrocarbon metabolic pathways. Curr Opin Microbiol. **12:**301-309.

CHAPTER II

Proteogenomic Elucidation of the Initial Steps in the Benzene Degradation Pathway of a Novel Halophile, *Arhodomonas* sp. Strain Rozel, Isolated from a Hypersaline Environment

This chapter has been published in Appl. Environ. Microbiol. 2012, 78(20):7309-7316 (2012) and reprinted in the thesis with permission of the publisher.

2.1 ABSTRACT

Recently, there has been an increased interest in understanding the role of halophilic and halotolerant organisms for their ability to degrade hydrocarbons. The focus of this study was to investigate the genes and enzymes involved in the initial steps of the benzene degradation pathway in halophiles. The extremely halophilic bacteria *Arhodomonas* sp. strain Seminole and *Arhodomonas* sp. strain Rozel that degrade benzene and toluene as the sole carbon sources at high salinity (0.5 to 4 M NaCl), were isolated from enrichments developed from contaminated hypersaline environments. To obtain insights into the physiology of this novel group of organisms, a draft genome sequence of the strain Seminole was obtained. A cluster of 13 genes predicted to be functional in the hydrocarbon degradation pathway was identified from the genome. Two-dimensional (2D) gel electrophoresis and liquid chromatography-mass spectrometry were used to corroborate the role of the predicted open reading frames (ORFs). ORFs 1080 and 1082 were identified as the components of a multicomponent phenol hydroxylase complex, and ORF 1086 was identified as catechol 2, 3-dioxygenase (2,3-CAT).

Based on this analysis, it was hypothesized that benzene is converted to phenol and then to catechol by phenol hydroxylase components. The resulting catechol undergoes ring cleavage *via* the *meta* pathway by 2,3-CAT to form 2-hydroxymuconic semialdehyde, which enters the tricarboxylic acid cycle (TCA). To substantiate these findings, strain Rozel was grown on deuterated benzene, and gas chromatography-mass spectrometry detected deuterated phenol as the initial intermediate of benzene degradation. These studies establish the initial steps of the benzene degradation pathway in halophiles.

2.2 INTRODUCTION

Many hypersaline environments, such as natural saline lakes, salt flats, solar salterns, saline industrial effluents, oil fields, and salt marshes are often contaminated with high levels of petroleum hydrocarbons. The cleanup of such environments can only be accomplished by stimulating the growth of indigenous microorganisms capable of degrading petroleum hydrocarbons or through the bioaugmentation of halophilic or halotolerant organisms that degrade hydrocarbons.

There is growing evidence suggesting that microorganisms play a significant role in the fate of hydrocarbons in high-salinity environments. Microcosms established with contaminated soil and sediment samples, enrichment cultures, and pure cultures have convincingly shown the ability of halophiles and halotolerants to degrade a variety of hydrocarbons, including crude oil, aliphatics, and mono- and polyaromatic compounds, as well as phenolics and benzoates, at salinities as high as ≥ 30% (1-4). Despite such numerous reports, little information exists on the genes, pathways, and mechanisms of their degradation. A few recent studies indicated that the degradation of hydrocarbons at high salinity occurs using enzymes described for many nonhalophiles (see chapter I). However, in-depth studies are needed to obtain greater insights into

the molecular mechanisms, intermediates, and pathways of hydrocarbon degradation in hypersaline environments.

Benzene is a category A carcinogen and considered as a priority pollutant by the Environmental Protection Agency (EPA). It is a volatile compound that contaminates soil, sediments and groundwater *via* leakage from produced water storage tanks, and accidental spills at petroleum production facilities (5). Despite its toxicity and large resonance energy, the benzene ring is activated *via* monooxygenases or dioxygenases in non-halophilic microorganisms under oxic conditions (6). Monooxygenases such as toluene/*o*-xylene monooxygenase, toluene 4-monooxygenase and phenol hydroxylase (PH) catalyze the conversion of benzene to phenol by introduction of an oxygen atom (7-9). Phenol undergoes further hydroxylation by PH or other monooxygenases to form catechol (10). On the other hand, dioxygenases like benzene 1, 2-dioxygenase catalyzes the dihydroxylation of the benzene ring to yield benzene *cis*-dihydrodiol that undergoes dehydrogenation to form catechol (CAT) by *cis*-dihydrodiol dehydrogenases (10, 11). CAT is then cleaved by *ortho* or *meta*-pathway catalyzed by catechol 1, 2-dioxygenase (1,2-CAT) or 2,3-CAT respectively to form intermediates that can be assimilated *via* central metabolism (10). However, these steps of benzene degradation have been described only in non-halophilic hydrocarbon degraders. It is not known whether benzene is degraded under hypersaline conditions *via* novel genes and pathways compared to non-halophiles. Undoubtedly, the discovery of novel genes and pathways is significant, as it could lead to the development of alternative and cost-effective remediation strategies.

This study describes the initial steps of the benzene degradation pathway in novel halophilic strain *Arhodomonas* sp. strain Rozel. A high-quality draft genome sequence of the strain Seminole was used as a guide for elucidating the steps. Such information is important for understanding the rate-limiting initial steps that have to be overcome for the efficient removal of toxic compounds.

2.3 MATERIALS AND METHODS

2.3.1 Chemicals. Benzene, toluene, benzene-D_6, and phenol-D_6 were purchased from Sigma-Aldrich Co. All the chemicals were of analytical grade and were used without further purification.

2.3.2 Bacterial strains. The *Arhodomonas* sp. strain Seminole (GenBank accession no. JX099567) was isolated from an enrichment developed from an oil-brine soil obtained from an oil production facility in Seminole County, Oklahoma. 16S rRNA gene sequence analysis shows 96% sequence identity with *Arhodomonas aquaeolei* (GenBank accession no. NR_044676). The strain Seminole degraded benzene in the presence of 0.5 to 3 M NaCl, and no degradation occurred in the absence of salt, suggesting that the isolate is a strict halophile (12). *Arhodomonas* sp. strain Rozel (GenBank accession no. JX128266) was isolated from an enrichment developed from sediment obtained from Great Salt Lake near Rozel Point, Utah (13). 16S rRNA gene sequence analysis shows 99% sequence identity with *Arhodomonas aquaeolei* (GenBank accession no. NR_044676). The strain Rozel is able to degrade benzene over a wide range of salinity (0.5 to 4 M NaCl) with optimal degradation at 3 M NaCl (13). 16S rRNA gene sequence alignment of the two isolates shows 98% sequence similarity with 100% query coverage and an E-value of 0, suggesting that these isolates are closely related to each other. The strain Seminole requires >2 weeks to completely degrade 17 to 24 μmol of benzene, while the strain Rozel is able to completely degrade 20 to 25 μmol of benzene in <7 days. Both strains were maintained in 1-liter bottles with 500 ml of mineral salts medium (MSM). The composition of MSM (in grams/liter): $MgCl_2$, 0.5; KH_2PO4 , 0.45; K_2HPO4 , 0.9; NH_4Cl, 0.3; KCl, 0.3. The bottles were supplemented with 2.5 M NaCl, 20 to 25 μmol of neat benzene, 0.02% yeast extract for strain Rozel. These bottles served as the mother cultures for all the experiments performed in this study.

To understand the hydrocarbon degradation capacity of the *Arhodomonas* isolates, a draft genome sequence (7x coverage) of the strain Seminole was obtained by pyrosequencing.

However, the strain Rozel was selected for proteomic and other studies reported in this work because of its higher benzene degradation rate and broader range of salinity tolerance than those of strain Seminole. To generate sufficient biomass required for proteomic studies, strain Rozel was grown in two sets of 500-ml bottles containing 300 ml MSM supplemented with 2.5 M NaCl and lactate as the sole source of carbon. Bottles were fed 5 mM lactate 2 times consecutively each at the end of log-phase growth. Growth was monitored by measuring both the optical density at 600 nm (OD600) and the total protein by using the Lowry method (14). Once sufficient biomass (114 μg/ml) was generated, approximately 20-25 μmol/bottle benzene was added twice to one set of 3 bottles to induce benzene-degrading enzymes. No benzene was added to the other set of 3 bottles as a control. The benzene-amended bottles were closed with rubber septa and aluminum crimps, and headspace samples were withdrawn periodically and monitored for the consumption of added benzene using GC as described previously (15). Cells from both sets of bottles were harvested by centrifugation for 15 min at 10,000 x g at 4°C. The cell pellets were immediately frozen and stored at -80°C until further use.

2.3.3 Identification of genes. To identify the putative genes and proteins of interest, a locally installed and used a stand-alone BLAST software pack- age (version 2.2.6) was obtained from the NCBI (16). The selected protein sequences were searched for in BLAST against all predicted peptides in the genome of the *Arhodomonas* sp. strain Seminole using a cutoff E-value of 1e-45. The relative position within the contig and the putative transcription direction of the predicted genes were determined by using the GeneMark.hmm for Prokaryotes software (17).

2.3.4 Phylogenetic analysis. The predicted amino acid sequences of open reading frames (ORFs) 1080, 1082, and 1084 from the *Arhodomonas* sp. strain Seminole genome were aligned using the ClustalW option in MEGA5 (18) with closely related and well-characterized phenol hydroxylase, toluene monooxygenase, and benzene monooxygenase components from different aromatic hydrocarbon-degrading nonhalophiles. The sequences for alignment were obtained from the

GenBank and UniProtKB. Phylogenetic analysis was performed by using the neighbor-joining algorithm and the Poisson correction method in MEGA5 (19). Bootstrap values were calculated as a percentage of 1,000 replicates (20).

2.3.5 Preparation of cell extracts. Cell extracts were prepared essentially as described previously (21) with some modifications. Briefly, cell pellets were washed once with 0.14 M NaCl, washed once with Tris-EDTA (TE) buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8.0]), and then stored as aliquots at -80°C. The pelleted cells were suspended in TE buffer with complete Mini protease inhibitor cocktail (Roche) (1:100 [vol/vol]) and disrupted by using a FastPrep Bio 101 Thermo Savant bead beater (5 cycles of 15 s each). The protein concentration was determined using the 2D Quant kit (GE Healthcare).

2.3.6 Proteomic analysis. The cytosolic proteomes were resolved using 2D gel electrophoresis as described previously (21) with minor modifications. Briefly, aliquots of cell extracts containing 70 μg protein were separated in the first dimension by isoelectric focusing (IEF) in the Ettan IPGphor3 system (GE Healthcare) for a total of 73 kVh at 20°C using 24-cm nonlinear immobilized pH gradients with a pH gradient of 3 to 7. The rehydration solution contained 9.47 M urea, 2.63 M thiourea, 33.4 mM dithiothreitol, 2.4% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate, and 2% Pharmalyte (broad range, pH 3 to 10). Proteins were separated in the second dimension by using 12% SDS-PAGE and the Ettan DALTsix System (GE Healthcare). Gels were stained using SyproRuby and digitally imaged using a Typhoon 9400 (GE Healthcare). Spot detection, matching, abundance quantification, and normalization were performed using Progenesis Workstation software (Nonlinear Dynamics, Durham, NC). The protein patterns of each growth condition were based on gels from three independent cultures. The protein spots of interest were excised and denatured in urea, alkylated, and digested with trypsin, and their trypsinolytic peptides were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) using an LTQ-Orbitrap XL

hybrid mass spectrometer (Thermo Fisher Scien- tific, San Jose, CA). Proteins were identified by using Mascot (v2.2.2 from Matrix Science, Boston, MA) and a database generated by *in-silico* digestion of the strain Seminole proteome predicted from the genome. Search results were validated by using Scaffold 03 (Proteome Software Inc., Portland, OR), the Peptide Prophet algorithm (22), and Protein Prophet (23). The criteria for accepting each identification will conform to the "Paris" guidelines for proteomics results (http://www.mcponline.org/misc/ParisReport_Final.dtl). A set of stringent criteria for protein identification was used; only protein probability thresholds greater than 99% were accepted, and at least three peptides needed to be identified, each with 95% certainty. Protein candidates containing similar peptides were grouped to satisfy the principles of parsimony. The search results were assessed for false-discovery rates (FDR) using randomized sequence databases.

2.3.7 Identification of intermediates. The strain Rozel was grown in 160-ml serum bottles filled with 48 ml of MSM supplemented with 2.5 M NaCl and inoculated with 2 ml of actively growing culture from the mother bottle. All bottles were amended with 2 μl of deuterated benzene (benzene-D_6) and 2 μl of unlabeled benzene to achieve an approximate starting concentration of 46 μmol/bottle. These bottles were closed with rubber septa and aluminum crimps and incubated at 30°C for 10 days in an inverted position. Autoclaved control bottles were set up similarly with the labeled and unlabeled benzene. The benzene concentrations were monitored daily by GC. Triplicate active bottles were sacrificed each day (at the end of 24 h), and triplicate control bottles were sacrificed at the end of day 0 and day 10. Bacterial activity was stopped by acidifying the content with 5 N HCl (pH <2). The entire content of the bottles was extracted with ethyl acetate (10% [vol/vol]; 4 times). The extracts were dried over anhydrous Na_2SO_4, concentrated by rotary evaporation, and reduced further under a stream of N_2 to a volume of 50 μl. The extracts and the phenol-D_6 standard were derivatized with *N*, *O*-bis(trimethylsilyl)tri-fluoroacetamide (BSTFA) (Pierce Chemical Co., Rockford, IL) prior to the analyses of the resulting compounds using an

Agilent 6890 model GC coupled with an Agilent model 5973 mass spectrometer (MS) as previously described (24). All the identifications were made by comparison to the GC retention times and the mass spectral fragmentation profiles of commercial standards (Sigma-Aldrich, St. Louis, MO) that were similarly analyzed or by comparison with the National Institute of Standards and Technology Mass Spectral Library, version 2.0a.

2.3.8 Nucleotide sequence accession numbers. The nucleotide sequences of the six ORFs (1079 to 1084) corresponding to components of phenol hydroxylase were deposited in the GenBank database under accession numbers JX311705 to JX311710. The nucleotide sequences of ORFs 1078, 1085, 1086, 1087, 1088, 1089, and 1090 were deposited in the GenBank database under accession numbers JX311711 to JX311717.

2. 4 RESULTS

2.4.1 Genomic analysis of benzene-degrading genes. *In silico* analysis of the genome of strain Seminole revealed a number of ORFs predicted to encode enzymes for aromatic hydrocarbon degradation. The genes were clustered on a 32-kb contig. To infer possible catabolic functions of these putative ORFs, BLASTp analyses were performed (E-value < 1e-45) against the UniProtKB database (25). 13 putative genes were predicted that encode enzymes for benzene degradation. These proteins share 44 to 77% sequence identity with proteins previously described in non-halophilic organisms (Table 2-1).

Table 2-1. *In silico* identification of putative ORFs in the benzene degradation pathway in strain Seminole[a]

ORF	Query length[b] (aa)	Putative function[c]	GenBank accession no.	Domain	Organism	% identity[d]	E-value	UniProtKB accession no.
1078	567	Activator of phenol-degradative genes	JX311711	Transcriptional regulator containing PAS, AAA-type ATPase, DNA-binding domains	*Cupriavidus necator* N-1	56	e−174	F8GQJ3
1079	76	Phenol hydroxylase assembly protein (P0)	JX311705	Phenol hydroxylase subunit	*Acinetobacter* sp. RUH2624	46	4e−10	D0C567
1080	**334**	**Phenol hydroxylase P1 protein**	**JX311706**	**Aromatic and alkene monooxygenase hydroxylase, subunit B, ferritin-like diiron-binding domain**	***Pseudomonas* sp. CF600**	**51**	**1e−115**	**P19730**
1081	91	Phenol hydroxylase component (*phyB*)	JX311707	MmoB/DmpM family	*Ralstonia* sp. KN1	66	2e−25	Q9RAF7
1082	**513**	**Phenol hydroxylase component (*phtD*)**	**JX311708**	**Aromatic and alkene monooxygenase hydroxylase, subunit A, ferritin-like diiron-binding domain**	***Wautersia numazuensis***	**77**	**0**	**Q5KT19**
1083	56	Phenol hydroxylase component (*poxE*)	JX311709	Phenol hydroxylase conserved region	*Ralstonia* sp. E2	56	4e−35	O84962
1084	353	Phenol hydroxylase P5 protein	JX311710	Oxygenase reductase FAD/NADH binding domain	*Pseudomonas* sp. CF600	64	3e−161	P19734
1085	115	Plant type ferredoxin-like protein	JX311712	2Fe-2S iron-sulfur cluster binding domain	*Azoarcus* sp. strain BH72	44	3e−15	A1K6K5
1086	**317**	**Catechol 2,3-dioxygenase**	**JX311713**	**Catechol 2,3-dioxygenase**	***Ralstonia metallidurans***	**69**	**e−129**	**Q1LNR9**
1087	194	Uncharacterized protein	JX311714	No putative domains detected	*Magnetospirillum magneticum*	50	8e−139	Q2W7L9
1088	229	Transcriptional regulator	JX311715	Transcriptional regulators	*Azoarcus* sp. strain BH72	48	4e−49	A1K899
1089	141	Putative uncharacterized protein	JX311716	Domain of unknown function	*Pseudomonas putida*	59	2e−27	Q49KG4
1090	486	2-Hydroxymuconic semialdehyde dehydrogenase	JX311717	Aldehyde dehydrogenase family	*Marinobacter* sp. MnI7-9	70	0	G6YS35

[a] Shown are ORFs putatively involved in both upper and lower benzene degradation pathways identified by genomic analysis of the draft genome of *Arhodomonas* sp. strain Seminole.

[b] Amino acid length.

[c] The putative functions of ORFs were predicted using BLASTp analyses with the UniProtKB database. The identification of the proteins in boldface type was verified using proteomic analyses.

[d] Percentage identity was based on BLASTp hits against the UniProtKB database.

The products of selected ORFs likely to be involved in both the upper and lower benzene degradation pathways are shown in Fig. 2-1. For example, six ORFs (1079 to 1084) shared a significant identity with components of PH found in many hydrocarbon-degrading non-halophiles, including the P0 to P5 components that catalyze phenol-to-catechol conversion in the *Pseudomonas* sp. strain CF600 (26). The deduced amino acid sequence of ORF 1080 is 51%

identical to the P1 component in *Pseudomonas* sp. strain CF600, and ORF 1082 is 77% identical to the *phtD*-encoded PH component in *Wautersia numazuensis* (27), which has a binuclear iron center. ORF 1084 is similar to the P5 component of PH, which acts as an NADH-ferredoxin oxidoreductase in *Pseudomonas* sp. strain CF600 and other hydrocarbon-degrading organisms (28-30).

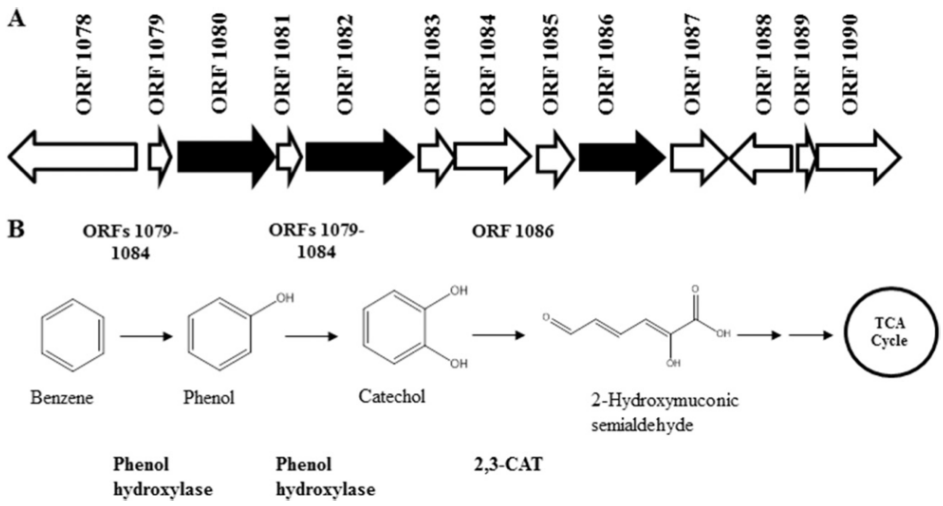

Fig 2-1. (A) Schematic diagram showing the genetic organization of benzene-degrading ORFs predicted in strain Seminole genome. These candidate ORFs are involved in the initial steps of the benzene degradation pathway. The putative functions of the candidate ORFs are listed in Table 2-1. The ORFs with dark arrows were identified by proteomic analysis. The arrowheads indicate the directions of transcription, and the gene sizes are not proportional to the sizes of the arrows.

(B) Proposed benzene degradation pathway by a multicomponent phenol hydroxylase-like enzyme in *Arhodomonas* sp. strain Seminole. ORFs and the corresponding putative enzymes in bold were identified by genomic and proteomic analyses. Phenol was confirmed as the initial intermediate of benzene degradation by GC-MS. TCA, tricarboxylic acid.

Phylogenetic analyses (Fig. 2-2) also confirmed that ORFs 1080, 1082, and 1084 are closely related to the P1, P3, and P5 components of PH in *Pseudomonas* sp. strain CF600. The remaining ORFs, 1086 and 1090, code for ring-cleaving 2,3-CAT and 2-hydroxymuconic semialdehyde dehydrogenase, respectively.

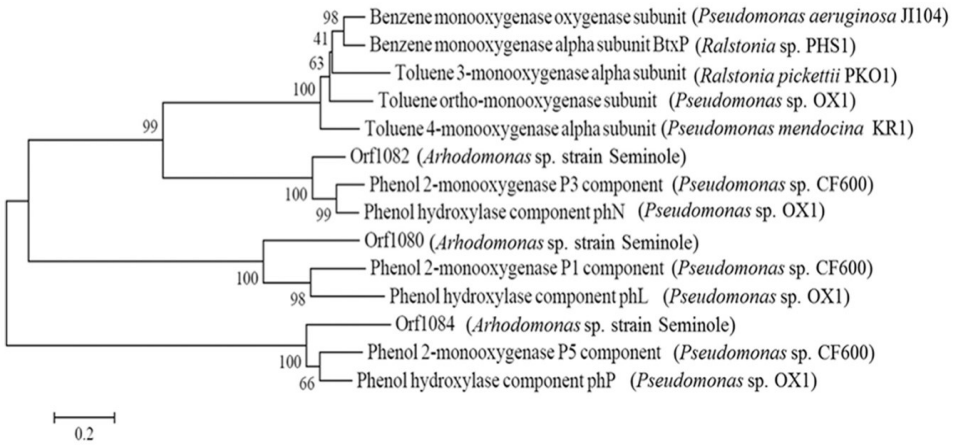

Fig 2-2 Phylogenetic analysis chart showing the relationships among various monooxygenase components in strain Seminole and other aromatic hydrocarbon-degrading nonhalophiles. The unrooted neighbor-joining tree was constructed in MEGA 5 by using predicted amino acid sequences of ORF 1080, ORF 1082, and ORF 1084 from the strain Seminole genome and closely related phenol hydroxylase, toluene monooxygenase, and benzene monooxygenase subunits from nonhalophiles. Sequences for the analysis were obtained from the GenBank and UniProtKB database. Bootstrap values were calculated as a percentage of 1,000 replicates and are shown next to the branches. The enzyme components (and the corresponding GenBank accession numbers) are benzene monooxygenase oxygenase subunit (BAA11761), benzene monooxygenase alpha subunit BtxP (ABG82181), toluene 3-monooxygenase alpha subunit (AAB09618), toluene *ortho*-monooxygenase subunit (CAA06654), toluene 4-monooxygenase alpha subunit (AAS66660), phenol hydroxylase phL component (AAO47356), phenol hydroxylase phN component (AAO47358), and phenol hydroxylase phP component (AAO47360). The enzyme components (and the UniProtKB accession numbers) are phenol 2-monooxygenase P1 component (P19730), phenol 2-monooxygenase P3 component (P19732), and phenol 2-monooxygenase P5 component (P19734).

2.4.2 Proteomic analyses. To identify the enzymes involved in the initial steps of benzene degradation, a proteomic approach was employed. 2D-gel electrophoresis resolved approximately 1,100 protein spots in the cytosolic proteomes of the strain Rozel cells grown on benzene or lactate as the growth substrate. Quantitative comparison of the resolved proteomes using the Progenesis algorithm revealed significant differences in the protein profiles of cells grown on benzene compared to cells grown on lactate. At least 15 additional proteins were present in benzene-grown cells than in lactate-grown cells. Considering the isoelectric point (pI) and the molecular weight (MW) of the predicted enzymes involved in aromatic catabolism, 7 of these proteins were chosen for identification, 3 of which are shown in Fig. 2-3.

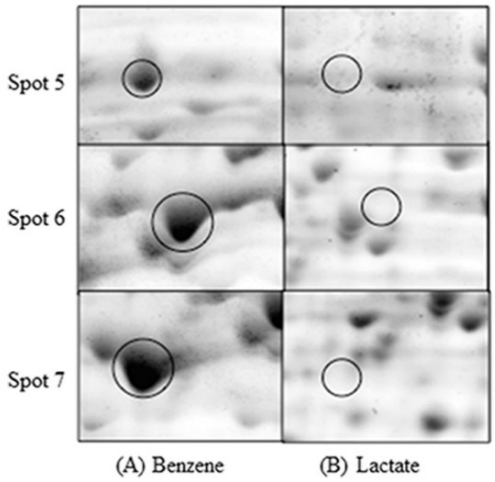

Fig 2-3. A 2D gel image showing candidate protein spots. Progenesis Workstation software was used for protein spot detection, matching, and abundance quantification. (A and B) Sections of the 2D gels with protein spots of the cytosolic proteome of cells grown on benzene (A) and lactate (B). The protein spots detected only in benzene-degrading cells are circled (lane A). The spots were in-gel digested and analyzed by LTQ Orbitrap LC-MS/MS to create peptide mass fingerprints (PMFs). The PMFs were identified by using MASCOT and the translated protein database of the *Arhodomonas* sp. strain Seminole genome.

These protein spots were excised and analyzed by using LTQ Orbitrap LC-MS/MS, and the generated peptide mass fingerprints were used for protein identification by using MASCOT and the translated protein database of the strain Seminole genome. Spot 5 was identified as the product of ORF 1086, while spots 6 and 7 matched the products of ORFs 1080 and 1082, respectively, in the strain Seminole genome. ORF 1086 was predicted to encode 2,3-CAT, and ORFs 1080 and 1082 encode PH components (Table 2-2). All three identified proteins were among the most abundant proteins in the cytosolic proteome of the benzene-grown cells. The products of the other predicted ORFs listed in Table 2.1 were either not resolved well or not translated in appreciable quantities under the conditions used for growing the strain Rozel.

Table 2-2 Identification of proteins induced in the proteome of strain Rozel grown on benzene using LC-MS/MS.

Protein spot[a]	ORF	pI/MW		Protein	Organism	No. of unique spectra	Sequence coverage (%)	Average signal intensity[c] (SD)	
		Experimental	Theoretical[b]					Lactate	Benzene
5	1086	5.2/35	4.9/35	Catechol 2,3-dioxygenase	*Ralstonia metallidurans*	12	19	ND[d]	2.339 (0.198)
6	1080	5.5/38	5.02/38	Phenol hydroxylase P1 protein	*Pseudomonas* sp. CF600	27	25	ND	1.224 (0.049)
7	1082	6.0/63	5.34/61	Phenol hydroxylase component, phtD	*Wautersia numazuensis*	40	32	ND	0.181 (0.058)

[a] Spot numbers correspond to those in Fig. 2-3. The protein spots 5, 6, and 7 on the 2D gel were identified using LC-MS/MS, MASCOT, and the translated genome of *Arhodomonas* sp. strain Seminole.
[b] Theoretical isoelectric points and molecular weights of the proteins of interest were calculated using the Compute pI/MW tool from EXPASY. These values were consistent with the experimentally determined pI values and molecular weights.
[c] Spot signal intensities were normalized and averaged over three replica gels (each from independent experiment). These values and SD were calculated using Progenesis algorithms.
[d] Not detected.

2.4.3 Metabolite detection using GC-MS. During benzene degradation by the strain Rozel, culture fluids collected periodically were extracted, derivatized, and analyzed by GC-MS for the detection of intermediates. Metabolic intermediates were identified by comparing their molecular mass ions (*m*/*z*) and retention times to those of authentic standards. Based on the proteogenomic studies, formation of phenol as a metabolic intermediate was hypothesized. To confirm that phenol is indeed produced as an intermediate, strain Rozel was grown in the presence of deuterated benzene (benzene-D6) and monitored for the formation of deuterated phenol (phenol-D_6) using GC-MS. A mass spectral profile of a BSTFA-derivatized phenol-D_6 was detected in the inoculated bottles (Fig. 2-4), and no such peak was detected in the control bottles (data not shown). Formation of phenol-D_6 was monitored on a daily basis for 10 days and was detected by GC-MS in the culture fluids. However, CAT was not detected in the analyses. The formation of phenol-D_6 was confirmed by injecting the BSTFA-derivatized authentic phenol-D_6 (Fig. 2-4) and by comparing its mass spectrum and retention time. The mass ions occurring at *m*/*z* 171 and 156 correspond to trimethylsilyl- and dimethylsilyl (formed by the loss of one methyl group from the derivatizing group) derivatized authentic phenols, respectively. This finding conclusively shows that phenol is an initial intermediate of aerobic benzene degradation by *Arhodomonas* sp. strain Rozel under hypersaline conditions.

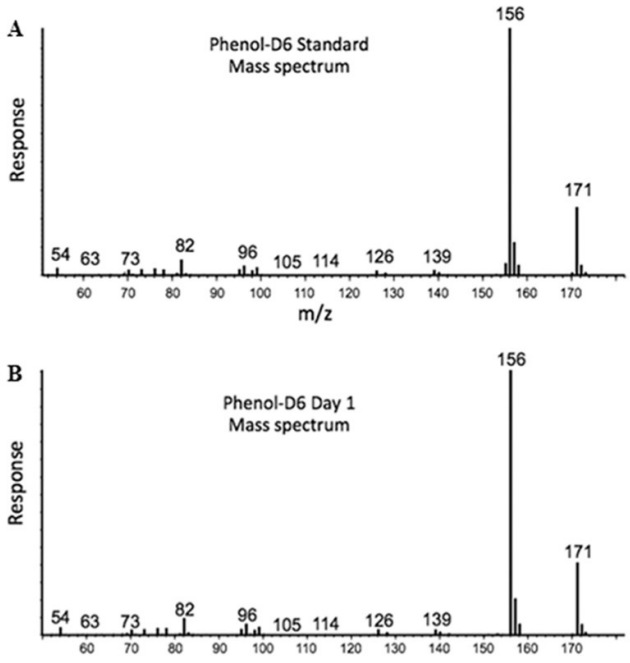

Fig 2-4 (A) Mass spectrum of a BSTFA-derivatized phenol-D6 standard. **(B)** A similar mass spectrum of a BSFTA-derivatized metabolite was found in a deuterated benzene (benzene-D6)-fed strain Rozel culture on day 1, confirming the metabolite as phenol.

2. 5 DISCUSSION

Genomic studies provide valuable information regarding degradation pathways as well as the general physiological potential of microorganisms. This study clearly demonstrates how genomic information from one organism can be a useful tool for studying physiology in phylogenetically related organisms. A draft genome sequence of the strain Seminole was used as a basis for predicting the early steps of the benzene degradation pathway in a closely related halophile, the strain Rozel. *In silico* analysis of the genome revealed a cluster of 13 genes that encode upper and lower pathway enzymes for aromatic compound degradation located on a 32-kb DNA fragment (Fig. 2-1A; Table 2-1). These enzymes share high (44 to 70%) amino acid sequence identity with the enzymes from non-halophilic microorganisms, thus suggesting that these halophiles use enzymes similar to those found in non-halophiles. In this study, genome-based predictions were used to target the experimental proteomic analyses and identify the enzymes involved in the initial steps of the benzene degradation pathway. Proteomic analyses showed significant differences in the protein profiles of benzene-grown and lactate-grown cells and identified several proteins that were highly abundant and present only in benzene-grown cells. These included two putative PH components with ferritin-like di-iron-binding domains and a 2,3-CAT. These results indicate that benzene is converted to phenol and then to CAT in two steps by monooxygenase-like enzymes closely related to PH. Thus, formed catechol is most likely further degraded to 2-hydroxymuconic semialdehyde via *meta*-cleavage by ORF 1086, identified as 2,3-CAT (Fig. 2-1B).

Phenol hydroxylases are multicomponent soluble diiron containing monooxygenases that hydroxylate a variety of aromatic hydrocarbons (28). These enzymes are similar to toluene monooxygenases and have been shown to catalyze the insertion of one oxygen atom not only in phenols but also in a number of hydrocarbons, including benzene, toluene, naphthalene, and trichloroethylene. In this study, genomic analyses predicted multicomponent PH-like proteins,

including ORFs 1079 to 1084, that are closely related to the *dmp* phenol hydroxylase components P0 to P5, respectively, from the *Pseudomonas* sp. strain CF600 (26, 30, 31). Also, a similar arrangement has been observed in plasmid- or chromosomally encoded PH complexes, such as the *phy* operon from *Ralstonia* sp. KN1 (32), the *pox* operon from *Ralstonia eutropha* strain E2 (33), the *pht* operon from *Wautersia numazuensis* (27), the *ph* operon from *Pseudomonas stutzeri* OX1 (8), and the gene identifier MpeA2280-85 in *Methylibium petroleiphilum* (34). In addition to those in multicomponent PHs, ORFs 1079 to 1084 also showed significant sequence identity with the polypeptides from toluene/*o*-xylene monooxygenase in *Pseudomonas stutzeri* OX1 (7). Phylogenetic analysis (Fig. 2-2) of the three major proteins, including ORFs 1080, 1082, and 1084, also showed these ORFs to be closely related (51 to 77% sequence identity) to the major components (P1, P3, and P5) of PH in *Pseudomonas* sp. strain CF600 and other organisms in comparison to benzene monooxygenases (28% sequence identity). All together, these analyses suggest that ORFs 1079 to 1084 encode functional components of PH-like proteins in the *Arhodomonas* strains and that these enzyme complexes are composed of hydroxylase, reductase [iron-sulfur flavoprotein that transfers electrons from NAD(P)H to the oxygenase], and a regulator protein that is required for catalysis (28, 29).

The proteogenomic data was further supported by the GC-MS detection of phenol as an intermediate during benzene degradation by the strain Rozel. The analysis showed the formation of phenol within 24 h of incubation, suggesting that benzene is quickly converted to phenol. Thus, these results indicate that the formation of phenol is a result of enzymatic hydroxylation of the benzene ring by PH-like enzymes. GC-MS analysis did not detect the formation of CAT. This could be due to many reasons, including rapid turnover rate and/or the low concentration of catechol. However, the proteogenomic data clearly suggests catechol as the intermediate. These observations suggest that monooxygenases are versatile catalysts that are not restricted to non-halophiles but also are employed by halophiles for similar functions, underscoring their diversity,

versatility, and ecological spread in microorganisms. This study reports the biochemical mechanism underlying the initial steps of benzene degradation in the novel Seminole and Rozel halophilic strains.

2.6 REFERENCES

1. **Le Borgne S, Paniagua D, and Vazquez-Duhalt R.** 2008. Biodegradation of organic pollutants by halophilic bacteria and archaea. J Mol Microbiol Biotechnol. **15:**74-92.

2. **McGenity TJ.** 2010. Halophilic hydrocarbon degraders. p. 1939-1951. *In* K. N. Timmis (ed.), Handbook of Hydrocarbon and Lipid Microbiology. Springer-Verlag, Berlin Heidelberg.

3. **Sorokin DY, Janssen AJH, and Muyzer G.** 2011. Biodegradation potential of halo(alkali)philic prokaryotes. Crit Rev Environ Sci Technol. **42:**811-856.

4. **Fathepure BZ.** 2014. Recent studies in microbial degradation of petroleum hydrocarbons in hypersaline environments. Frontiers in Microbiology. **5**. :173. doi: 10.3389/fmicb.2014.00173.

5. **Fries MR, Zhou J, Chee-Sanford J, and Tiedje JM.** 1994. Isolation, characterization, and distribution of denitrifying toluene degraders from a variety of habitats. Appl Environ Microbiol. **60:**2802-2810.

6. **Jindrova E, Chocova M, Demnerova K, and Brenner V.** 2002. Bacterial aerobic degradation of benzene, toluene, ethylbenzene and xylene. Folia Microbiol (Praha). **47:**83-93.

7. **Bertoni G, Bolognese F, Galli E, and Barbieri P.** 1996. Cloning of the genes for and characterization of the early stages of toluene and o-xylene catabolism in *Pseudomonas stutzeri* OX1. Appl Environ Microbiol. **62:**3704-3711.

8. **Cafaro V, Izzo V, Scognamiglio R, Notomista E, Capasso P, Casbarra A, Pucci P, and Di Donato A.** 2004. Phenol hydroxylase and toluene/o-xylene monooxygenase from *Pseudomonas stutzeri* OX1: interplay between two enzymes. Appl Environ Microbiol. **70:**2211-2219.

9. **Fishman A, Tao Y, and Wood TK.** 2004. Toluene 3-monooxygenase of *Ralstonia pickettii* PKO1 is a *para*-hydroxylating enzyme. J Bacteriol. **186:**3117-3123.

10. **Pérez-Pantoja D, González B, and Pieper DH.** 2010. Aerobic degradation of aromatic hydrocarbons. p. 799-837. *In* K. N. Timmis (ed.), Handbook of Hydrocarbon and Lipid Microbiology. Springer-Verlag Berlin Heidelberg.

11. **Bagnéris C, Cammack R, and Mason JR.** 2005. Subtle difference between benzene and toluene dioxygenases of *Pseudomonas putida*. Appl Environ Microbiol. **71:**1570-1580.

12. **Nicholson CA.** 2005. Biodegradation of petroleum hydrocarbons by halophilic and halotolerant microorganisms (Master's Thesis). Oklahoma State University, Stillwater, OK. Retrieved from http://digital.library.okstate.edu/etd/umi-okstate-1536.pdf.

13. **Azetsu S.** 2009. BTEX degradation at high salinity in Rozel Point (Master's thesis). Oklahoma State University, Stillwater, OK. Retrieved from http://digital.library.okstate.edu/etd/Azetsu_okstate_0664M_10581.pdf.

14. **Lowry OH, Rosebrough NJ, Farr AL, and Randall RJ.** 1951. Protein measurement with the Folin phenol reagent. J Biol Chem. **193:**265-275.

15. **Nicholson CA, and Fathepure BZ.** 2004. Biodegradation of benzene by halophilic and halotolerant bacteria under aerobic conditions. Appl Environ Microbiol. **70:**1222-1225.

16. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389-3402.

17. **Lukashin AV, and Borodovsky M.** 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. **26:**1107-1115.

18. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S.** 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol. **28:**2731-2739.

19. **Saitou N, and Nei M.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. **4:**406-425.

20. **Felsenstein J.** 1985. Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. Evolution. **39:**783-791.

21. **Patrauchan MA, Florizone C, Dosanjh M, Mohn WW, Davies J, and Eltis LD.** 2005. Catabolism of benzoate and phthalate in *Rhodococcus* sp. strain RHA1: redundancies and convergence. J Bacteriol. **187:**4050-4063.

22. **Keller A, Nesvizhskii AI, Kolker E, and Aebersold R.** 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. **74:**5383-5392.

23. **Nesvizhskii AI, Keller A, Kolker E, and Aebersold R.** 2003. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem. **75:**4646-4658.

24. **Aktas DF, Lee JS, Little BJ, Ray RI, Davidova IA, Lyles CN, and Suflita JM.** 2010. Anaerobic metabolism of biodiesel and its impact on metal corrosion. Energy & Fuels. **24:**2924-2928.

25. **Consortium TU.** 2012. Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Res. **40:**D71-D75.

26. **Nordlund I, Powlowski J, and Shingler V.** 1990. Complete nucleotide sequence and polypeptide analysis of multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CF600. J Bacteriol. **172:**6826-6833.

27. **Kageyama C, Ohta T, Hiraoka K, Suzuki M, Okamoto T, and Ohishi K.** 2005. Chlorinated aliphatic hydrocarbon-induced degradation of trichloroethylene *Wautersia numadzuensis* sp. nov. Arch Microbiol. **183:**56-65.

28. **Leahy JG, Batchelor PJ, and Morcomb SM.** 2003. Evolution of the soluble diiron monooxygenases. FEMS Microbiol Rev. **27:**449-479.

29. **Pessione E, Divari S, Griva E, Cavaletto M, Rossi GL, Gilardi G, and Giunta C.** 1999. Phenol hydroxylase from *Acinetobacter radioresistens* is a multicomponent enzyme. Purification and characterization of the reductase moiety. Eur J Biochem. **265:**549-555.

30. **Powlowski J, and Shingler V.** 1994. Genetics and biochemistry of phenol degradation by *Pseudomonas* sp. CF600. Biodegradation. **5:**219-236.

31. **Shingler V, Franklin FC, Tsuda M, Holroyd D, and Bagdasarian M.** 1989. Molecular analysis of a plasmid-encoded phenol hydroxylase from *Pseudomonas* CF600. J Gen Microbiol. **135:**1083-1092.

32. **Nakamura K, Ishida H, and Iizumi T.** 2000. Constitutive trichloroethylene degradation led by tac promoter chromosomally integrated upstream of phenol hydroxylase genes of *Ralstonia* sp. KN1 and its nucleotide sequence analysis. J Biosci Bioeng. **89:**47-54.

33. **Hino S, Watanabe K, and Takahashi N.** 1998. Phenol hydroxylase cloned from *Ralstonia eutropha* strain E2 exhibits novel kinetic properties. Microbiology. **144:**1765-1772.

34. **Kane SR, Chakicherla AY, Chain PS, Schmidt R, Shin MW, Legler TC, Scow KM, Larimer FW, Lucas SM, Richardson PM, and Hristova KR.** 2007. Whole-genome analysis of the methyl tert-butyl ether-degrading beta-proteobacterium *Methylibium petroleiphilum* PM1. J Bacteriol. **189:**1931-1945.

CHAPTER III

Genomic Analysis of Hydrocarbon Degradation Pathways in *Arhodomonas* sp. strain Seminole at High Salinity

Contents from this chapter is a part of a manuscript submitted to Journal of Applied and Environmental Microbiology titled "Isolation and Characterization of a Novel *Arhodomonas* sp. Strain Seminole and its Genetic Potential to Degrade Hydrocarbons at High Salinity" and is currently under revision.

3.1 ABSTRACT

The focus of this study is to understand the physiology and genomics of hydrocarbon degradation in a novel halophile, *Arhodomonas* sp. strain Seminole isolated from a crude oil-impacted brine soil. Apart from benzene and toluene, the strain was able to degrade oxygenated aromatic compounds such as 4-hydroxybenzoic acid (4-HBA), protocatechuic acid (PCA), and phenylacetic acid (PAA) as the sole sources of carbon at high salinity. The draft genome of strain Seminole was analyzed to gain insight into the various hydrocarbon degradation pathways. Analysis of the genome predicted a number of catabolic genes for the complete metabolism of benzene, toluene, 4-HBA and PAA. The predicted pathways were corroborated by identification of enzymes expressed in the cytosolic proteome of hydrocarbon-grown cells using liquid chromatography-mass spectrometry (LC-MS/MS).

Genome analysis predicted a cluster of 19 genes necessary for the breakdown of benzene or toluene to acetyl-CoA and pyruvate. Of these, 12 gene products (enzymes) were expressed in benzene or toluene-grown cells compared to lactate-grown cells. Genomic analysis revealed the presence of 11 genes required for 4-HBA degradation to form intermediates that will be assimilated *via* TCA cycle. Of these, proteomic analysis of 4-HBA grown cells identified 6 key enzymes involved in 4-HBA degradation pathway. Similarly, 15 genes needed for the degradation of PAA to TCA cycle intermediates were predicted *in-silico*. Of these, 9 enzymes of the PAA degradation pathway were up-regulated and identified by LC-MS/MS only in PAA-grown cells and not in lactate-grown cells. Overall, using genomics and proteomic approaches we were able to reconstruct catabolic steps for the breakdown of a variety of hydrocarbons in an extreme halophile, *Arhodomonas* sp. strain Seminole. Such knowledge is important for understanding its role in the natural attenuation of hydrocarbons in hypersaline environments.

3.2 INTRODUCTION

Hypersaline environments such as oil fields, industrial effluents, and coastal ecosystems are often contaminated with high levels of petroleum hydrocarbons (1, 2). Among these, oil and gas production sites pose the greatest risk of soil and groundwater contamination due their large numbers all over the world. They are not only contaminated with a complex mixture of hydrocarbons but also display a wide range of salinities, from very low up to saturated brines. The arid coastlines of Gulf countries are highly saline and susceptible to oilspill and environmental damage. Waste streams from pesticide, chemical, and pharmaceutical also generate large quantities of highly saline waste waters that are contaminated with high levels of petroleum compounds (3).

Halophilic and halotolerant organisms that degrade hydrocarbons have received considerable attention in recent years due to their potential use in the remediation of hydrocarbon-impacted high salinity environments. Studies documenting the ability of bacteria, archaea and a

few eukaryotes to degrade hydrocarbons at moderate to high salinity conditions have been reviewed recently (4-7). Among bacteria, members of the genera *Halomonas*, *Marinobacter* and *Alcanivorax* have been widely reported to degrade hydrocarbons at varying salinity (for detailed information see Chapter I). Among archaea, members of *Haloferax* (8-11), *Haloarcula* (11, 12), *Halobacterium* (10, 11, 13) and *Halorubrum* (14) have been shown to degrade both aliphatic and aromatic hydrocarbons in high salinity environments. Although these studies report the ability of halophilic and halotolerant organisms to degrade hydrocarbons in moderate and high salinity environments, little is known about molecular mechanisms, degradation pathways and steps leading to intermediates that enter the tricarboxylic acid cycle (TCA).

Arhodomonas sp. strain Seminole was isolated from a BTEX-degrading enrichment developed using brine samples from an oil production facility in Seminole, OK. The strain can degrade benzene at salinity ranging from 1 to 3 M NaCl. This work was previously performed by Ms. Carla Nicholson in our laboratory (15). The focus of the current study is to further characterize strain Seminole's physiology, ecology and molecular mechanism of hydrocarbon degradation at high salinity. In addition to benzene and toluene, the strain also degrades oxygenated aromatic compounds such as PCA, 4-HBA, PAA as the sole sources of carbon and energy. A high-quality draft genome sequence of the strain was used to predict genes involved in the degradation steps of the above hydrocarbons. The predicted pathways were corroborated by semi-quantitative identification of enzymes expressed in the cytosolic proteome of hydrocarbon-grown cells using liquid chromatography and mass spectrometry (LC-MS/MS).

3.3 MATERIALS AND METHODS

3.3.1 Chemicals and Media. Hexadecane and sodium benzoate were purchased from Sigma-Aldrich, MO and Fisher Scientific, PA, respectively. Catechol (CAT), PAA and 4-HBA were purchased from Alfa Aesar, MA. Gentisic acid (GA) and protocatechuic acid (PCA) were

purchased from MP Biomedicals, OH. All the chemicals were of analytical grade and were used without further purification.

The composition of mineral salts medium (MSM) used in this study is previously described (16). MSM was supplemented with 2 M NaCl and yeast extract (0.01%) for all experiments unless mentioned otherwise. For solidified medium, 1.5% (w/v) agar was added prior to autoclaving.

3.3.2 Ecological distribution of Strain Seminole. To screen for *Arhodomonas* sp-like phylotypes in samples from different locations, specific primer pairs 1465R (5'-GTCTCGACCACACCGTGG -3') and 206F (5'-GTTTCATGGTCACGCCGA -3') were designed using probe design function in the ARB software package (17). The primers were validated by querring them using the probe match function in the Ribosomal database project (18). Contaminated as well as uncontaminared soil, water, or sediment samples with varying level of salinity were collected from different locations. These include, (i) soil from a brine pit at site B, Skiatook, Oklahoma, (ii) soil from Great Salt Plains National Wildlife Refuge, Oklahoma, (iii) water sample from Dead Sea Israel, (iv) soil contaminated by produced water spill, Tall Grass Prairie Preserve, Oklahoma (19), (v) hydrocarbon-impacted saline soil, Kuwait, (v) soil from East Texas, Chevron site, (vii) sediment and water samples from the base of a mangrove tree, Cabo Rojo, Puerto Rico (20), (viii) soil from a salt-manufacturing plant, Freedom, Oklahoma (20) and (ix), wastewater from Ramat-Hovav Industrial Park, Negev Desert, Israel (21).

3.3.3 DNA extraction, PCR amplification, cloning and sequencing for ecological distribution study. Genomic DNA was extracted using Ultraclean Soil DNA Isolation Kit (MO BIO Laboratories, Inc. CA) and FASTDNA Spin Kit for soil (MP Biomedicals, OH). Genomic DNA was pooled and nested PCR was performed with bacterial primers 27F (5'-AGAGTTTGATCMTGGCTCAG-3') and 1492R (5'-TACGGYTACCTTGTTACGACTT-3') using protocol described previously (22) followed by *Arhodomonas* sp-specific primers 206F

(5'-GTTTCATGGTCACGCCGA-3') and 1465R (5'-GTCTCGACCACACCGTGG-3') to screen for 16S rRNA genes of *Arhodomonas* spp. PCR was performed in 50-µl reaction mixture that contained 2 µl of the first-reaction PCR product, 1x PCR buffer (Teknova, CA), 2.5 mM $MgSO_4$, 10 mM deoxynucleoside triphosphate (dNTP) mixture, 1 U of Taq DNA polymerase and 10 µM of each forward and reverse primers. PCR with *Arhodomonas* sp-specific primers was carried out according to the following protocol: initial denaturation at 94°C for 3 min, followed by 30 cycles of denaturation at 94°C for 1 min, annealing at 54°C for 55 seconds, and elongation at 72°C for 2 min. A final elongation step at 72°C for 8 min was included. All samples were PCR amplified in triplicate. The resulting PCR products of the expected size (approximately 1250 bp) were gel purified using QIAquick gel extraction kit (QIAGEN, CA). The purified products were then cloned in to TOP10 *Escherichia coli* using a TOPO-TA cloning kit (Invitrogen, CA) and sequenced at Oklahoma State University Recombinant DNA/Protein Core Facility (Stillwater, OK). The sequences obtained were compared with GenBank nr database using blastn program in NCBI (23). Phylogenetic tree was constructed in MEGA6 to highlight the phylogenetic affiliation of the clones obtained in this study with sequences of closely related *Arhodomonas* spp. in GenBank (24).

3.3.4 Growth of strain Seminole on benzoates and phenolics. Degradation of oxygenated aromatic compounds including benzoate, phenol, 4-HBA, CAT, PCA, GA, PAA and hexadecane was studied in 250 ml-Erlenmeyer flasks containing 100 ml of MSM supplemented with 2 M NaCl and filter-sterilized individual aromatic compounds incubated separately at 30^{o}C without shaking under aerobic conditions. Growth on hexadecane was monitored by measuring the O.D at 600 nm. Degradation of other compounds was monitored as depletion of the added compound using a UV-VIS-spectrophotometer (Ultraspec 2000, Pharmacia Biotech) and scanning over specific wavelengths ranging between 200 and 400 nm. Briefly, samples were withdrawn periodically and centrifuged at 10,000 rpm for 10 min. The supernatant was appropriately diluted

with distilled water and analyzed by measuring absorbance at 228 nm for benzoate, 269 nm for phenol, 245 nm for 4-HBA, 288 nm for PCA, 318 nm for GA, 274 nm for catechol and 257 nm for PAA.

3.3.5 Genome sequencing and assembly. The detailed procedure for the sample preparations were as described previously (25). DNA was sheared to the size of ~2 Kbp using nebulization at 30 psi and -20^{0}C for 2.5 minutes as described (26). Smaller fragments were removed using SPRI beads (cat# 000130) (27). DNA fragments were then end-repaired (polished) by treating them with DNA polymerase and T4 polynucleotide kinase as described previously (25). Adaptors were then ligated using DNA ligase and end-repaired using DNA polymerase. Single-stranded DNA molecules are then captured using DNA capture beads and emulsion PCR was performed as described (25) and the resulting amplified products are then run on 454 GS20 according to the recommendation of the manufacturer. Flows from the 454 are assembled using Newbler, the 454 assembly software. Three different trimming lengths were used from the 454 to reduce the number of artificial contigs produced due to poor qualities at the end of the contigs. The results were then utilized to assemble using Phrap.

3.3.6 Genomic analysis of hydrocarbon degradation potential. The genome was uploaded to the Integral Microbial Genomes (IMG) server (http.//img.jgi.doe.gov) of the Joint Genome Institute for genome predictions, comparisons and analysis. The Clusters of Orthologous Groups (COGs) of protein sequences from strain Seminole were analyzed using the function category tool of the IMG. A one-sample *t*-test was used to evaluate the statistically significant differences of gene abundances in each COG category between strain Seminole and other genomes of hydrocarbon-degrading halophiles deposited in the IMG bacteria genome database. A total of 10 genomes of hydrocarbon-degrading halophiles were used for comparative COG analysis.

The draft genome sequence of strain Seminole was screened for hydrocarbon degrading genes using a locally installed stand-alone BLAST software package (version 2.2.6). Genes encoding enzymes involved in various aromatic hydrocarbon degradation pathways that have been characterized in other organisms and deposited in NCBI database were searched using blastp against the predicted peptides in strain Seminole genome (23). The putative functions of the open reading frames (ORFs) predicted to be involved in hydrocarbon degradation were confirmed by blastp search in Uniprot database (28). The direction of transcription of the predicted genes and their relative positions on the contigs were determined by using GeneMark.hmm for Prokaryotes software (29).

3.3.7 Preparation of cell extracts. For proteomic studies, sufficient biomass was obtained by growing strain Seminole on lactate as described previously in Chapter II, Section 2.3. Briefly, strain Seminole was grown on 5 mM lactate, fed twice, as the sole source of carbon. At the end of log growth phase, cells were centrifuged at 8,000 rpm for 15 mins at 4^{o}C and the pellet was washed twice with sterile 0.14 M NaCl. The pellet was resuspended in MSM containing 2 M NaCl and used as an inoculum for the induction of enzymes involved in toluene, 4-HBA, or PAA degradation. Lactate-grown cells devoid of the respective hydrocarbon served a control. Cells were harvested by centrifuging at 8,000 rpm for 15 mins at 4^{o}C. The cell pellets were washed once with 0.14 M NaCl and once with Tris-EDTA (TE) buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8.0]). The pelleted cells were resuspended in TE buffer and disrupted using sonication (5 cycles of 15 s each). The protein concentration was determined using Bradford (30).

3.3.8 Proteome analyses. Protein extracts were pre-fractionated on 12% SDS-PAGE gels, stained with Coomassie blue, and each lane excised into 3 slices. Briefly, 20 μg of protein extracts obtained from each hydrocarbon-grown actives and lactate-grown actives was loaded on the SDS-PAGE in triplicates. Gel slices were destained by extensive washing with 50% acetonitrile/50 mM ammonium bicarbonate pH 8, dehydrated with 100% acetonitrile, and dried

briefly. Dried acrylamide pieces were rehydrated with 10 mM tris (2-carboxyethyl) phosphine, 50 mM ammonium bicarbonate, and reduced with Tris(2-carboxyethyl)phosphine hydrochloride for 1 hr at room temperature. After incubation, the reducing buffer was replaced with 55 mM iodoacetamide in 50 mM ammonium bicarbonate, and alkylated for 1 hr at room temperature in the dark. Samples were then rinsed with ammonium bicarbonate, dehydrated with acetonitrile, and rehydrated/infiltrated with trypsin solution containing 8 µg trypsin per ml of 50 mM ammonium bicarbonate. After overnight digestion at 37^{o}C, the trypsinolytic peptide products were extracted with 0.5% Trifluoroacetic acid, and analyzed on a hybrid LTQ-OrbitrapXL mass spectrometer (ThermoFisher Scientific, MA) basically as described in Vorungati et al (31) but using a 40 cm column packed with 3-micron Magic C18 AQ particles (Bruker, MA). Proteins were identified using a database generated by *in silico* digestion of the strain Seminole proteome predicted from the genome as well as sequences for 114 common adventitious laboratory contaminants. (http://www.thegpm.org). Peptide and protein identifications were validated using Scaffold (v4.0.6.1; Proteome Software Inc., Portland, OR) using the ProteinProphet and PeptideProphet algorithms (32), with minimum of 2 peptides identified and at protein thresholds that yielded a 1% False Discovery Rate for the dataset. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Differences in protein expression were analyzed by spectrum counting (33), comparing each protein's spectrum counts from extracts of cells cultured in lactate, toluene, 4-HBA, or PAA as indicated. Statistical significances in the total spectral counts was determined by Student's t-test (2-tailed, equal variance) using a p-value < 0.05 as the signficance cut off.

3.3.9 Nucleotide sequence accession numbers. The 16S rRNA genes sequence of *Arhodomonas* spp. clones obtained from this study have been deposited in the GenBank database under accession numbers KJ829488-KJ829496. The protein sequences of benzene-, toluene-, 4-HBA-

and PAA-degrading ORFs were deposited in the GenBank database under accession numbers JX311705-JX311717 and KJ829497-KJ829529.

3.4 RESULTS

3.4.1 Ecological distribution of strain Seminole-like organisms. In order to understand the diversity and adaptability of *Arhodomonas* spp. in different ecological niches; soil, sediment and water samples from both hydrocarbon-contaminated and uncontaminated sites with varying salinity were analyzed. The presence of *Arhodomonas* sp-like organisms was confirmed by using *Arhodomonas*-specific primers, cloning and sequencing of the PCR products. The limited survey showed that the organisms were found in both hydrocarbon contaminated as well as uncontaminated geographically distant saline sites (Table 3-1). A phylogenetic tree using 16S rRNA gene sequences of *Arhodomonas* spp. clones obtained in this study as well as sequences from the NCBI database was constructed. The analysis revealed that the sequences formed two major clusters, each cluster harboring organisms from both hydrocarbon contaminated as well as uncontaminated sites (Fig. 3-1). Interestingly, *Arhodomonas* sp. 50B226 a3 (accession EU308280) isolated from a Greek solar saltern did not cluster with other *Arhodomonas* spp. Also, *Arhodomonas* sp. SP71 (accession JF798749) isolated from Santa Pola solar salterns and *Arhodomonas* sp. clone (accession KJ829495) from sediment from mangroove roots formed a distinct cluster.

Table 3-1. List of sites screened for presence of Arhodomonas sp-like phylotypes

Location	Sample Information (Salinity)	Presence of *Arhodomonas* sp. [a]
Skiatook pit, Oklahoma	Soil from brine pit at site B (15%)	Yes [b]
Great Salt Plains, Oklahoma	Soil sample (20%)	No
Dead Sea, Israel	Water sample (33%)	No
Tall Grass Prairie, Oklahoma	Soil contaminated by produced water spill (0.4%	Yes
Kuwait	Soil from crude oil contamination (8 and 16%)	Yes
Chevron, Texas	Soil from crude oil contamination (0.33%)	Yes
Cabo Rojo, Puerto Rico	Mangrove sediment (1.5%) and water (7%)	Yes
Freedom, Oklahoma	Soil from salt-manufacturing plant (7%)	No
Ramat-Hovav industrial park, Isra	Water from Industrial Wastewater evaporation ponds (12%)	No

[a] *Arhodomonas* sp-like phylotypes were screened using specific primer pairs 1465R and 206F.

[b] The presence of *Arhodomonas* sp-like organisms was confirmed by cloning and sequencing of the PCR products.

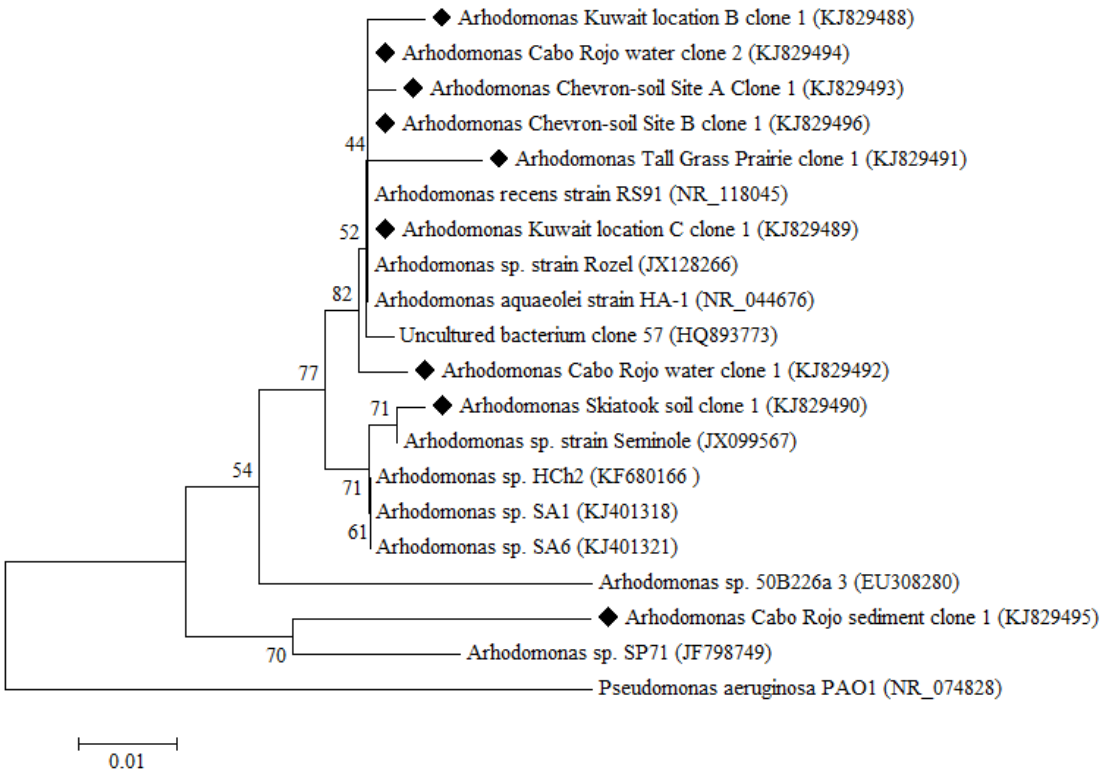

Fig. 3-1 Phylogenetic relationship among *Arhodomonas* spp. isolated from soil, water, and sediment samples collected from geographically different locations. Neighbor-joining tree was constructed using 16S rRNA gene sequences from *Arhodomonas* spp. clones obtained in this study (marked with dark diamonds) as well as sequences from GenBank. The tree was constructed using p-distance model and rooted with 16S rRNA gene sequence of *Pseudomonas aeruginosa* PAO1. Bootstrap values are shown next to the branches and were calculated as a percentage of 1000 replicates. Bar represents 1% sequence difference.

3.4.2 Biodegradation of benzoates and phenolics by strain Seminole. Strain Seminole was able to utilize 4-HBA, PCA, and PAA as carbon sources (Table 3-2). A lag period of 7 days was consistently observed when strain was grown on 4-HBA, whereas no lag was observed when grown on PCA. The strain was able to grow on PAA, but only 45% disappearance of the substrate was observed over a period of 35 days of incubation. No growth was observed with benzoate, CAT, GA and hexadecane even after 35 days of incubation.

Table 3-2 Utilization of aromatic compounds as growth substrates by Strain Seminole

Substrates	% Degraded	Initial concentration
Benzene	100	22-25 µmol/bottle
Toluene	100	22-25 µmol/bottle
Ethylbenzene	0	22-25 µmol/bottle
Xylene	0	22-25 µmol/bottle
Hexadecane	0	1%
Benzoate	0	2 mM
4-hydroxybenzoic acid (4-HBA)	100	3 mM
Phenylacetic acid (PAA)	45	4.5 mM
Phenol	60	0.5 mM
Catechol	0	0.5 mM
Protocatechuic acid (PCA)	100	2.5 mM
Gentisic acid (GA)	0	2.5 mM

Strain Seminole was tested for its ability to utilize oxygenated and non-oxygenated aromatic compounds as sole sources of carbon. Growth was measured either by GC or UV scans or total protein.

3.4.3 General genome features. The draft genome of Strain Seminole has a single chromosome of 5,026,701 bp with a G+C content of 66.27%. The chromosome contains 5180 predicted protein coding genes (CDS) with an average size of 955 bp, giving a coding intensity of 86.87%. Analysis revealed 48 tRNA genes and 1 rRNA operon in the chromosome (Table 3-3).

Table 3-3 General features of *Arhodomonas* sp. strain Seminole genome[a]

Characteristic	Chromosome
Size (bp)	5026701
G+C content (%)	66.27
Total number of genes	5262
Number of protein coding sequences	5180
Number of proteins with function prediction	4060
DNA coding intensity (%)	86.87
Average gene length (bp)	955.28
Proteins assigned to COGs (%)	2978 (56.59%)
Number of rRNA operon	1
Number of tRNA genes	48

[a]Derived in part from the DOE-JGI IMG server (https://img.jgi.doe.gov/)

Of the 5180 CDS, 2978 could be assigned to 22 different categories of clusters of orthologous groups (COGs) (Table 3-4, Fig 3-2). Comparative analysis of the gene abundances of each COG categories between the genome of strain Seminole and the genomes of 10 other hydrocarbon-degrading halophiles in IMG (http.//img.jgi.doe.gov) was evaluated using one-sample t-test. The abundance in genes related to amino acid transport and metabolism (11.34%), carbohydrate transport and metabolism (6.71%) and secondary metabolite biosynthesis, transport and metabolism (4.17%) were significantly higher ($P<0.05$) in the Seminole genome compared to the average abundances of 9.92%, 4.93%, and 3.25%, respectively in the genomes of other ten hydrocarbon degrading halophiles (Table 3-4).

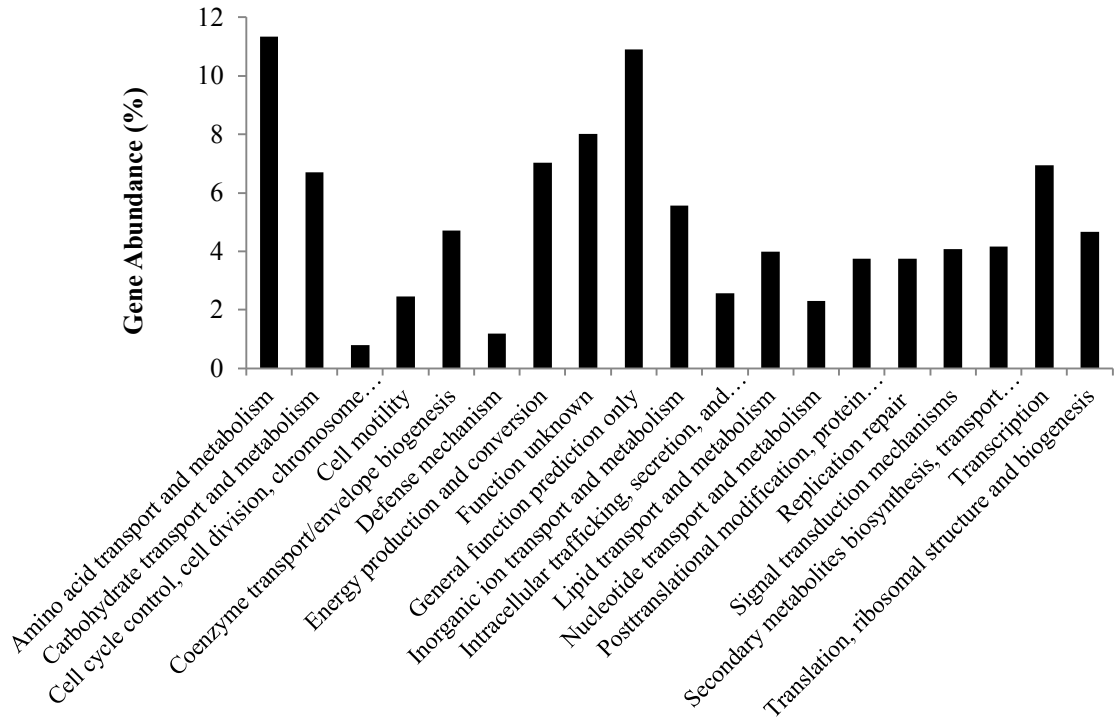

Fig. 3-2 Abundance of genes (%) belonging to different COG functional classes in strain Seminole genome. The % of gene abundances is listed in Table 3-4.

Table 3-4 Comparative analysis of COG categories between *Arhodomonas* sp. strain Seminole and other genomes of hydrocarbon-degrading halophilic bacteria in IMG bacteria genome database

	Gene Abundance (%)		Std. Deviation (%)	Std. error mean (%)	t-score[a]	P-value[b]
	Seminole	Mean				
COG category						
Amino acid transport and metabolism	11.34	9.92	1.50	0.48	-2.978	0.016
Carbohydrate transport and metabolism	6.71	4.93	1.87	0.59	-3.004	0.015
Cell cycle control, cell division, chromosome partitioning	0.79	0.98	0.17	0.05	3.491	0.007
Cell motility	2.45	2.35	1.01	0.32	-0.319	0.757
Coenzyme transport/envelope biogenesis	4.72	4.57	0.32	0.10	-1.507	0.166
Defense mechanism	1.18	1.20	0.33	0.10	0.198	0.847
Energy production and conversion	7.04	6.85	0.50	0.16	-1.203	0.26
Function unknown	8.01	8.14	0.67	0.21	0.606	0.559
General function prediction only	10.91	11.59	0.96	0.30	2.241	0.052
Inorganic ion transport and metabolism	5.56	5.69	0.47	0.15	0.892	0.396
Intracellular trafficking, secretion, and vesicular transport	2.57	2.45	0.70	0.22	-0.525	0.612
Lipid transport and metabolism	3.99	4.92	1.41	0.44	2.093	0.066
Nucleotide transport and metabolism	2.3	2.26	0.30	0.09	-0.423	0.682
Posttranslational modification, protein turnover, chaperones	3.75	4.12	0.64	0.20	1.849	0.098
Replication repair	3.75	4.69	1.52	0.48	1.95	0.08
Signal transduction mechanisms	4.08	4.74	1.02	0.32	2.04	0.07
Secondary metabolites biosynthesis, transport and catabolism	4.17	3.25	1.00	0.31	-2.929	0.017
Transcription	6.95	7.02	1.16	0.37	0.187	0.856
Translation, ribosomal structure and biogenesis	4.66	5.23	0.70	0.22	2.551	0.031

The genomes of *Alcanivorax borkumensis* SK2, *Arhodomonas aquaeolei*, *Chromohalobacter salexigens*, *Halomonas elongata*, *Marinobacter algicola*, *Marinobacter aquaeolei* VT8, *Amycolicicoccus subflavus*, *Polymorphum gilvum*, *Marinobacter* sp. BSs20148 and *Marinobacter* sp. ELB17 were used for COG analysis.

[a]One sample t-test was used to determine significant differences between genes abundances in each COG category of strain Seminole and ten other genomes of hydrocarbon-degrading bacteria obtained from IMG database.

[b]Gene abundances of the different COG categories in strain Seminole were considered to be significantly lower or higher than the average values if $P < 0.05$

3.4.4 Genomic analysis of aromatic hydrocarbon degradation in strain Seminole. Analysis of the draft genome of strain Seminole predicted a number of catabolic ORFs encoding enzymes involved in the upper and lower pathways of aromatic hydrocarbon degradation. Most of these putative ORFs were found clustered together on the chromosome. Blastp analyses of these ORFs against UniProt database were performed to infer their catabolic functions (28).

(i) Degradation of benzene and toluene *via* catechol ring cleavage pathway: Using the draft genome we identified ORFs 1079-1097 that code for proteins required for complete degradation of both benzene and toluene to acetaldehyde and pyruvate that can eventually enter the central metabolism (Table 3-5).

TABLE 3-5 Genomic and proteomic identification of putative ORFs and proteins involved in toluene degradation pathway in Strain Seminole

ORF	Accession number[a]	Putative Function[b]	Average spectral counts[c]		Student t-test (p -value)[d]	Organism	% Identity[e]	E-Value	Uniprot Accession number
			Lactate	Toluene					
1079	JX311705	Phenol hydrolase assembly protein	0	0		*Acinetobacter nosocomialis* 28F	46	1e-09	U4Q6B8
1080	**JX311706**	**Phenol hydrolase beta subunit**	**6**	**61**	**<0.00010**	***Methylibium petroleiphilum***	**52**	**e-102**	**A2SI51**
1081	JX311707	Phenol hydroxylase component 2	0	1	0.37	*Ralstonia* sp. KN1	66	6e-25	Q9RAF7
1082	**JX311708**	**Phenol hydroxylase component 3**	**6**	**60**	**<0.00010**	***Wautersia numazuensis***	**77**	**0**	**Q5KT19**
1083	**JX311709**	**Phenol hydroxylase component 4**	**0**	**5**	**0.00013**	***Ralstonia* sp. E2**	**56**	**8e-35**	**O84962**
1084	**JX311710**	**Ferredoxin oxidoreductase**	**0**	**35**	**<0.00010**	***uncultured bacterium***	**66**	**e-136**	**C6KUI9**
1085	JX311712	Plant type ferredoxin like protein	0	0		*Azoarcus* sp. (strain BH72)	44	7e-15	A1K6K5
1086	**JX311713**	**Catechol 2,3-dioxygenase**	**10**	**66**	**<0.00010**	***Ralstonia metallidurans***	**69**	**e-128**	**Q1LNR9**
1087	JX311714	Uncharacterized protein	0	2	0.13	*Magnetospirillum* sp. SO-1	50	2e-39	M2ZC26
1088	JX311715	Transcriptional regulator	0	2	0.13	*Azoarcus* sp. (strain BH72)	48	9e-49	A1K899
1089	JX311716	Putative uncharacterized protein	0	0		*Thauera* sp. 63	59	9e-29	N6YI61
1090	**JX311717**	**2-hydroxymuconic semialdehyde dehydrogenase**	**0**	**62**	**<0.00010**	***Pseudomonas pseudoalcaligenes* CECT 5344**	**74**	**0**	**I7J281**
1091	**KJ829529**	**2-hydroxymuconic semialdehyde hydrolase**	**0**	**21**	**<0.00010**	***Ralstonia metallidurans***	**67**	**e-107**	**Q1LNT5**
1092	**KJ829528**	**2-hydroxypenta-2,4-dienoate hydratase**	**2**	**26**	**0.00017**	***uncultured bacterium***	**72**	**e-106**	**C6L0Y6**
1093	**KJ829527**	**Acetaldehyde dehydrogenase**	**6**	**44**	**0.00011**	***Marinobacter algicola* DG893**	**76**	**e-122**	**A6EWL7**
1094	**KJ829526**	**4-hydroxy-2-oxovalerate aldolase**	**1**	**59**	**<0.00010**	***Cupriavidus necator* N-1**	**84**	**e-162**	**F8GQT8**
1095	**KJ829525**	**4-oxalocrotonate decarboxylase**	**2**	**26**	**0.00037**	***Thauera linaloolentis***	**69**	**1e-92**	**N6XRY3**
1096	**KJ829524**	**Uncharacterized protein precursor**	**1**	**15**	**0.0011**	***Alicycliphilus denitrificans***	**42**	**2e-54**	**E8TUN5**
1097	KJ829523	4-oxalocrotonate tautomerase family enzyme precursor	0	0		*Pseudomonas sp.* GM18	60	1e-17	J2WLU6

[a]Accession number assigned to each ORF in NCBI.

[b]The putative functions of ORFs was predicted using blastp with UniProt Knowledgebase (Swiss-Prot + TrEMBL) database. ORFs in boldface type were corroborated using LC-MS/MS analysis.

[c]Average of total spectral counts obtained from SDS-PAGE gels ran with lactate or toluene-grown cell extracts from three bottles.

[d]Student t-test used to determine significant difference between total spectral counts of lactate-induced cells and toluene-induced cells obtained from Scaffold.

[e]Percent identity was based on blastp hits against the UniProt Knowledgebase (Swiss-Prot + TrEMBL) database.

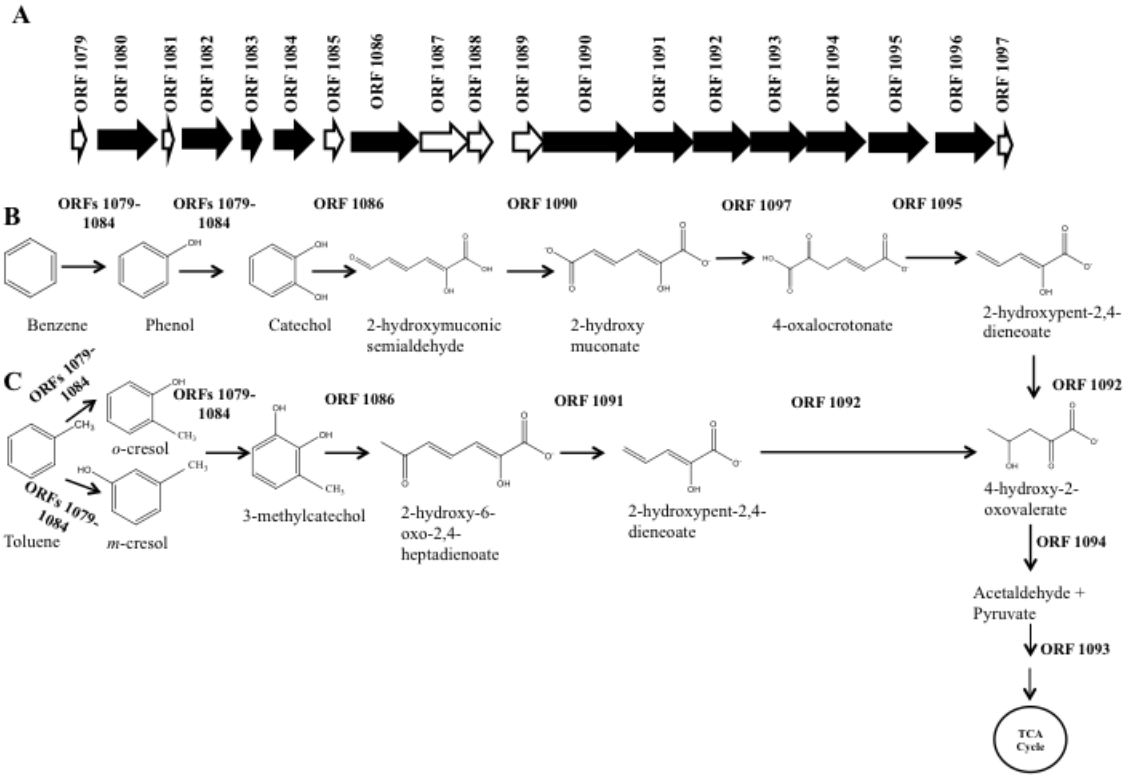

Fig 3-3 (A) Genetic organization of ORFs involved in degradation of benzene and toluene predicted in the genome of strain Seminole. Arrows represent the predicted ORFs and the arrowheads indicate the direction of their transcription. The gene sizes are not proportional to the arrows. ORFs represented by dark arrows were identified by LC-MS/MS in the toluene-grown cells. (B & C) Proposed benzene and toluene degradation pathways based on genomic and proteomic analysis. Putative functions of the ORFs are listed in Table 3-5. Multiple arrows indicate two or more steps.

Our analysis showed that ORFs 1079-1084 encode a multicomponent phenol hydroxylase like enzyme (PH) predicted to be involved in ring-hydroxylation of benzene and toluene to form CAT and 3-methylcatechol, respectively. Thus formed CAT and 3-methylcatechol can further undergo *meta*-cleavage by ORF 1086 encoding catechol 2,3-dioxygenase (2,3-CAT) to form a

ring cleavage intermediate that is further catabolized by a series of intermediates that ultimately become part of the central metabolism (Fig. 3-3A & B).

In the case of benzene metabolism, ORF 1086 (2,3-CAT) catalyzes the ring fission of CAT to 2-hydroxymuconic semialdehyde and thus formed ring-cleavage product is further converted to 2-hydroxypent-2,4-dieneoate by the activities of ORF 1090, ORF 1097, and ORF 1095 that code for 2-hydroxymuconic semialdehyde dehydrogenase, 4-oxalocrotonate tautomerase (4-oxalocrotonate isomerase), and 4-oxalocrotonate decarboxylase, respectively. In the next step, 2-hydroxypent-2,4-dieneoate is hydroxylated to 4-hydroxy-2-oxovalerate by ORF 1092 that codes for 2-hydroxypenta-2, 4-dienoate hydratase. Thus formed 4-hydroxy-2-oxovalerate is split into acetaldehyde and pyruvate by ORF 1094 that codes for 4-hydroxy-2-oxovalerate aldolase. Acetaldehyde is converted to acetyl-CoA by ORF 1093 that codes for acetaldehyde dehydrogenase.

Similar enzymes are predicted to participate in the breakdown of 3-methycatechol formed during toluene degradation. ORF 1086 (2,3-CAT) catalyzes the ring cleavage of 3-methylcatechol to 2-hydroxy-6-oxo-2,4-heptadienoate. Thus formed ring cleavage product is a ketone, rather than an aldehyde; it cannot be further oxidized by the 2-hydroxymuconic semialdehyde dehydrogenase and is exclusively metabolized *via* the hydrolytic route (34). ORF 1091 codes for 2-hydroxymuconic semialdehyde hydrolase that catalyzes the hydrolysis of 2-hydroxy-6-oxo-2,4-heptadienoate to 2-hydroxypent-2,4-dienoate which is further converted to acetyl-CoA and pyruvate by steps identical to benzene metabolism pathway (Fig. 3-3B & C).

(ii) 4-hydroxybenzoate degradation pathway *via* protocatechuate: Genomic analysis revealed the presence of *pob* and *pca* genes that are required for 4-HBA degradation (Table 3-6).

Table 3-6 Genomic and proteomic identification of putative ORFs and proteins involved in 4-HBA degradation pathway in Strain Seminole

ORF	Accession number[a]	Putative Function[b]	Average spectral counts[c]		Student t-test (*p* -	Organism	% Identity[e]	E-Value	Uniprot Accession number
			Lactate	4-HBA					
3910	KJ829514	3-carboxy-*cis,cis*-muconate	0	0		*Deinococcus geothermalis*	44	1e-8	Q1J3Z8
3911	**KJ829513**	**β-ketoadipyl succinyl-CoA transferase, A subunit**	**0**	**7**	**0.00081**	***Azospirillum lipoferum* strain 4B**	**74**	**3e-176**	**G7ZFP6**
3912	**KJ829512**	**β-ketoadipyl succinyl-CoA transferase, subunit B**	**0**	**5**	**0.0042**	***Xanthomonas citri* pv. *mangiferaeindicae***	**69**	**5e-150**	**H8FEZ7**
3913	KJ829511	β-ketoadipyl CoA thiolase	0	0		*Marinobacter manganoxydans*	69	e-147	G6YUR0
3914	**KJ829510**	**3-oxoadipate enol-lactonase**	**0**	**5**	**0.00037**	***Xanthobacter autotrophicus***	**46**	**2e-59**	**A7IEC4**
3915	KJ829509	4-carboxy muconolactone decarboxylase	0	0		*Microvirga* sp. WSM3557	51	4e-27	I4Z0N1
3916	**KJ829508**	**Protocatechuate 3,4-dioxygenase, alpha subunit**	**0**	**53**	**<0.0001**	***Pseudomonas aeruginosa* 39016**	**52**	**e-119**	**E3A6W3**
3917	**KJ829507**	**Pca regulon regulatory protein**	**0**	**7**	**<0.0001**	***Marinobacter* sp. ELB17**	**77**	**e-109**	**A3JCK5**
3919	**KJ829497**	**4-hydroxybenzoate 3-monooxygenase**	**0**	**7**	**0.007**	***Chromohalobacter* sp. HS2**	**82**	**0**	**A8I4C8**
3920	KJ829498	AraC family of transcriptional regulator	0	0		*Halomonas* sp. A3H3	58	5e-91	T2L640
3921	KJ829506	Transcriptional regulator, AraC family	0	0		*Halomonas elongata*	69	2e-34	E1V9E2

[a]Accession number assigned to each ORF in NCBI.

[b]The putative functions of ORFs was predicted using blastp with UniProt Knowledgebase (Swiss-Prot + TrEMBL) database. ORFs in boldface type were corroborated using LC-MS/MS analysis.

[c]Average of total spectral counts obtained from SDS-PAGE gels ran with lactate or 4HBA-grown cell extracts from three flasks.

[d]Student t-test used to determine significant difference between total spectral counts of lactate-induced cells and toluene-induced cells obtained from Scaffold.

[e]Percent identity was based on blastp hits against the UniProt Knowledgebase (Swiss-Prot + TrEMBL) database.

In silico analysis showed that *pob* and *pca* genes are clustered together on a 21.6 kb contig of the genome (Fig. 3-4A). The *pob* genes were shown to encode initial enzymes required for 4-HBA degradation pathway and *pca* genes encode enzymes required for the downstream protocatechuate branch of the β-ketoadipate pathway (Fig. 3-4B).

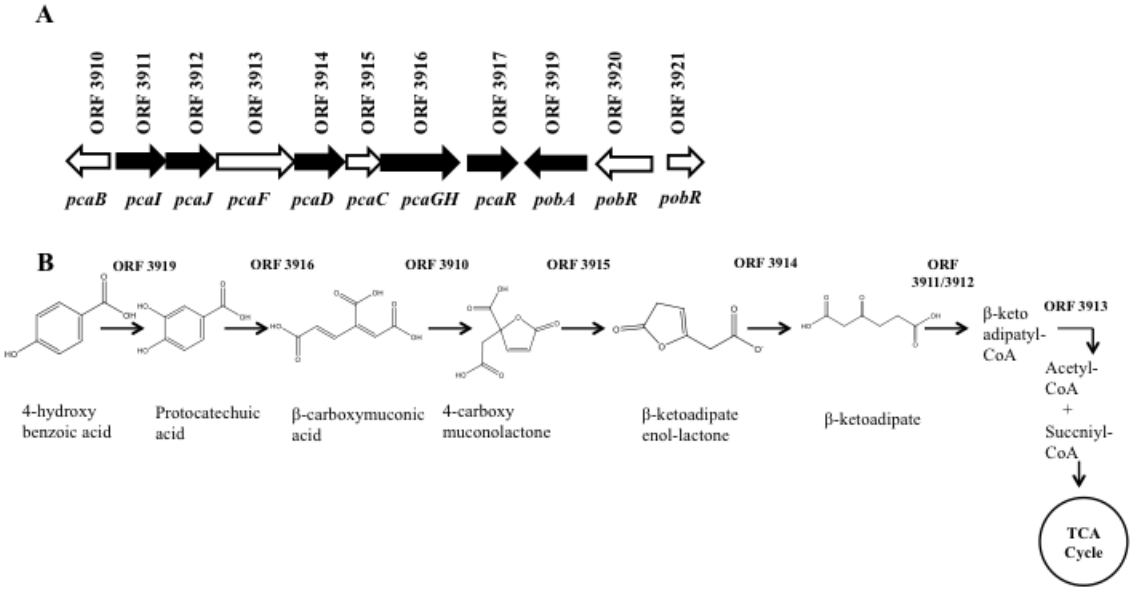

Fig 3-4 (A) Genetic organization of ORFs involved in degradation of 4-HBA predicted in strain Seminole genome. Arrows represent the predicted ORFs and the arrowheads indicate the direction of their transcription. ORFs represented by dark arrows were identified by proteomic analysis of 4-HBA proteome. The gene sizes are not proportional to the arrows. (B) Proposed 4-HBA degradation pathway based on genomic and proteomic analysis. Putative functions of the ORFs are listed in Table 3-6.

The *pob* genes include *pobA* gene (ORF 3919) that encodes the 4-hydroxybenzoate 3-monooxygenase enzyme catalyzing the conversion of 4-HBA to PCA and *pobR* gene (ORF 3920, 3921) encodes the AraC family transcriptional regulatory enzyme that activates the expression of *pobA* gene in response to 4-HBA (35). The *pca* genes responsible for the conversion of PCA to β-

ketoadipate included ORF 3916 that encode α and β subunits of the protocatechuate 3,4-dioxygenase (3,4-PCA) (PcaGH) that catalyzes the conversion of PCA to β-carboxymuconic acid, which is transformed into 4-carboxymuconolactone by ORF 3910 encoding 3-carboxy-*cis*,*cis*-muconate cycloisomerase *(*PcaB). ORF 3915 that codes for 4-carboxymuconolactone decarboxylase (PcaC) converts 4-carboxymuconolactone into β-ketoadipate enol-lactone, which is hydrolyzed by ORF 3914 encoding 3-oxoadipate enol-lactonase (PcaD) into β-ketoadipate (Fig. 3-4). ORFs 3911 and 3912 encoding β-ketoadipyl succinyl-CoA transferase (PcaIJ) catalyzes the conversion of β-ketoadipate to β-ketoadipyl-CoA. The cluster also includes *pcaF* gene that code for β-ketoadipyl CoA thiolase (ORF 3913) which catalyzes the last step transforming β-ketoadipyl-CoA to succinyl-CoA and acetyl-CoA (36) that can enter the TCA cycle. The *pcaR* gene (ORF 3917) encoding Icl-R type family transcriptional regulator required for inducible expression of PcaBDCIJF for conversion of β-carboxy-*cis*, *cis*-muconate to TCA intermediates was also clustered on the same contig (36).

(iii) Phenylacetate degradation pathway: *In-silico* genomic analysis showed that the PAA catabolic pathway is organized in five putative functional units: (i) a substrate activating enzyme, phenylacetyl-CoA ligase (PaaF), (ii) a ring-hydroxylation complex (PaaGHIJK), (iii) a ring-opening protein (PaaN*),* (iv) a β-oxidation-like system (PaaACDE*),* and (v) two regulatory proteins (PaaX and PaaY) (37). Several investigators have used different nomenclature for describing the *paa* genes. This study has adapted the system used by Luengo et. al (38). Analysis of the draft genome of Seminole revealed 15 ORFs putatively involved in complete PAA catabolism organized in distinct clusters on three different contigs (Table 3-7, Fig. 3-5A).

Table 3-7 Genomic and proteomic identification of putative ORFs and proteins involved in PAA degradation pathway in Strain Seminole

ORF	Accession number[a]	Putative Function[b]	Average spectral counts[c]		Student t-test (p -value)[d]	Organism	% Identity[e]	E-Value	Uniprot Accession number
			Lactate	PAA					
2877	**KJ829522**	**Ring-hydroxylation complex protein 4**	**0**	**25**	**<0.00010**	***Nitrococcus mobilis* Nb-231**	**59**	**e-119**	**A4BNW0**
2878	KJ829521	Phenylacetate-CoA oxygenase, PaaJ subunit	0	0		*Microvirga* sp. WSM3557	63	7e-54	I4YLN6
2879	**KJ829520**	**Phenylacetate-CoA oxygenase**	**0**	**47**	**<0.00010**	***Nitrococcus mobilis* Nb-231**	**65**	**2e-91**	**A4BNV8**
2880	**KJ829519**	**Phenylacetic acid degradation protein**	**0**	**7**	**0.002**	***Nitrococcus mobilis* Nb-231**	**80**	**1e-40**	**A4BNV7**
2881	**KJ829518**	**Phenylacetate-CoA oxygenase, PaaG**	**0**	**54**	**<0.00010**	***Cupriavidus pinatubonensis***	**75**	**e-144**	**Q46UU1**
2882	KJ829517	Uncharacterized protein	0	0		*Pseudomonas* sp. HPB0071	48	2e-52	N2J7Q2
2883	KJ829516	PaaX family transcriptional regulator	0	4	0.12	*Alcanivorax dieselolei*	51	2e-86	K0CBT9
2986	**KJ829515**	**Phenylacetate-CoA ligase**	**1**	**17**	**0.00035**	***Herbaspirillum* sp. CF444**	**72**	**e-179**	**J2LFC3**
4127	**KJ829505**	**Phenylacetic acid degradation protein**	**0**	**10**	**<0.00010**	***Thauera* sp. 28**	**71**	**5e-78**	**N6YDP9**
4128	**KJ829504**	**Phenylacetic acid degradation protein paaN2 (ring opening**	**0**	**21**	**<0.00010**	***Nitrococcus mobilis* Nb-231**	**68**	**0**	**A4BNX0**
4129	KJ829503	Putative enoyl-CoA hydratase I [paaA2]	0	0		*Pseudomonas* sp. Y2	59	3e-71	Q70IM9
4130	**KJ829502**	**1, 2-epoxyphenylacetyl-CoA isomerase**	**0**	**5**	**0.00016**	***Halomonas* sp. A3H3**	**61**	**4e-85**	**T2L8B6**
4131	**KJ829501**	**3-hydroxybutyryl-CoA dehydrogenase**	**0**	**44**	**<0.00010**	***Nitrococcus mobilis* Nb-231**	**61**	**e-174**	**A4BNX2**
4132	KJ829500	Probable phenylacetic acid degradation protein	0	0		*Nitrococcus mobilis* Nb-231	67	3e-50	A4BNX3
4133	**KJ829499**	**β-ketoadipyl CoA thiolase**	**0**	**32**	**0.00035**	***Azospirillum lipoferum* strain 4B**	**77**	**0**	**G7ZG05**

[a]Accession number assigned to each ORF in NCBI.

[b]The putative functions of ORFs was predicted using blastp with UniProt Knowledgebase (Swiss-Prot + TrEMBL) database. ORFs in boldface type were corroborated using LC-MS/MS analysis.

[c]Average of total spectral counts obtained from SDS-PAGE gels ran with lactate or PAA-grown cell extracts from three flasks.

[d]Student t-test used to determine significant difference between total spectral counts of lactate-induced cells and PAA-induced cells obtained from Scaffold.

[e]Percent identity was based on blastp hits against the UniProt Knowledgebase (Swiss-Prot + TrEMBL) database.

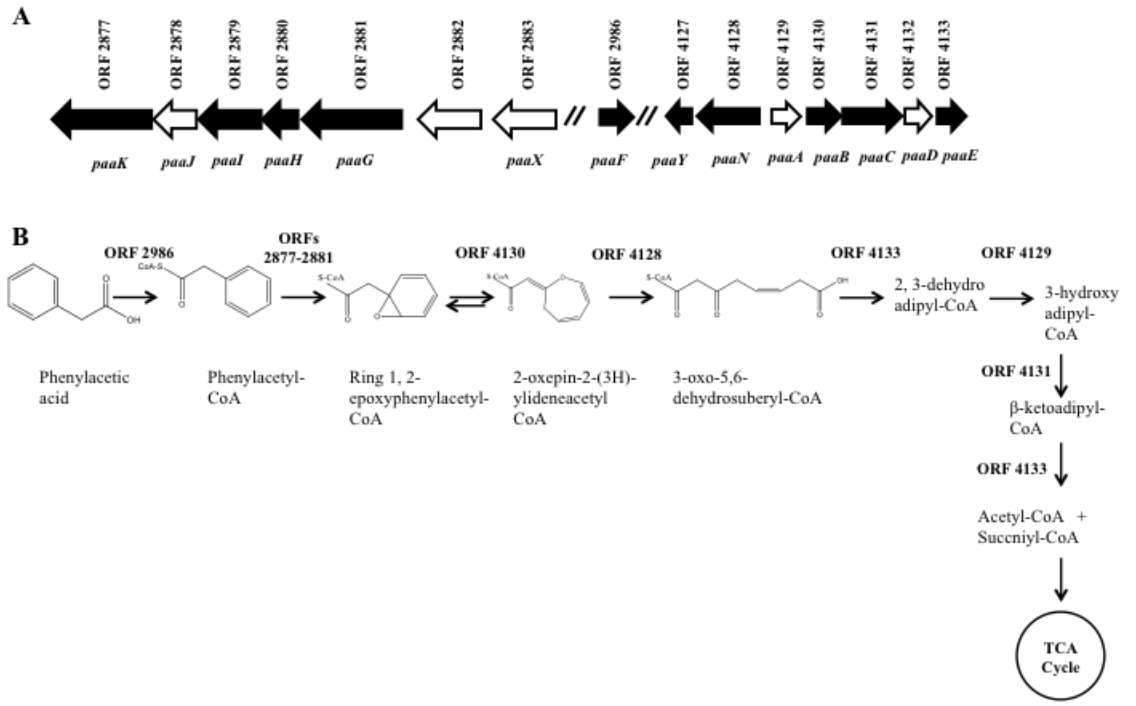

Fig 3-5 (A). Genetic organization of ORFs involved in degradation of PAA predicted in *Arhodomonas* strain sp. Seminole genome. The predicted ORFs are represented by arrows and the arrowheads indicate the direction of their transcription. ORFs represented by dark arrows were identified by proteomic analysis of PAA proteome. The gene sizes are not proportional to the arrows. (B) Proposed PAA degradation pathway based on genomic and proteomic analysis. Putative functions of the ORFs are listed in Table 3-7. Multiple arrows indicate two or more steps.

ORF 2986 encodes phenylacetyl-CoA ligase (PaaF) that catalyzes the initial step of activation of PAA into phenylacetyl-CoA (PA-CoA) (39). PA-CoA is converted to ring 1, 2-epoxyphenylacetyl-CoA by ORFs 2877-2881 that code a ring-oxygenase/reductase multicomponent complex (PaaK, PaaJ, PaaI, PaaH and PaaG). ORFs 2877-2881 were found clustered together on contig 670 of the draft genome along with ORF 2883 that codes for a transcriptional regulator (PaaX) that controls the regulation of PAA degradative genes (40). ORF 4130 encodes for ring 1, 2-epoxyphenylacetyl-CoA isomerase that rearranges the epoxide ring 1,

2-epoxyphenylacetyl-CoA into 2-oxepin-2-(3H)-ylideneacetyl CoA (Oxepin-CoA). ORF 4128 (*paaN*) encoding a putative ring-opening enzyme catalyzes the hydrolytic cleavage of the oxepin-CoA ring to form the intermediate, 3-oxo-5, 6-dehydrosuberyl-CoA (40). ORFs 4129, 4131, and 4133 (*paaA, C, and E*) code for enoyl-CoA hydratase, 3-hydroxybutyrl-CoA dehydrogenase, and β-ketoadipyl CoA thiolase, respectively. These proteins are similar to fatty acid β-oxidation enzymes that catalyze the β-oxidation of the ring-opened intermediate leading to the formation of acetyl-CoA and succinyl-CoA (41).

3.4.5 Semiquantitative proteomic analysis of strain Seminole. Using the draft genome of the strain Seminole, complete degradation pathways for benzene, toluene, 4-HBA, and PAA were predicted. To further validate these predictions, 1-dimensional (1-D) gel electrophoresis based proteomic analysis (SDS-PAGE in combination with LC/MS-MS) of cytosolic proteomes of cells grown on lactate (control), toluene, 4-HBA and PAA was performed to identify the enzymes involved in the degradation pathways and quantify their induction. In this work, proteomic confirmation of proteins involved in benzene degradation was not performed since this degradation pathway has already been recently elucidated in a closely related organism, *Arhodomonas* sp. strain Rozel using strain Seminole genome as a guide (Described in Chapter II) (42). In the present study, spectral counting was used to indicate differences in protein abundance in the proteomes of lactate-grown and the hydrocarbon-grown samples. Student t-test was used to determine whether the total spectral counts were significantly different between the above samples.

Analysis of the genome predicted 19 ORFs encoding enzymes involved in toluene degradation. Of these, 12 gene products were identified that were up-regulated in the toluene-induced cells compared to lactate-induced cells using proteomics. The total spectral counts for these 12 proteins were significantly higher (p-value < 0.05) in toluene-induced cells compared to lactate-induced cells (Table 3-5). These proteins included PH subunits (ORF 1080, 1082, 1083

and 1084) responsible for the conversion of toluene to *o*- or *m*-cresol that is further converted to 3-methylcatechol. Also identified was 2, 3-CAT (ORF 1086) responsible for the *meta*-cleavage of 3-methylcatechol and the downstream enzymes of responsible for conversion of ring cleavage intermediates to central metabolism. These include 2-hydroxymuconic semialdehyde dehydrogenase (ORF 1090), 2-hydroxymuconic semialdehyde hydrolase (ORF 1091), 2-hydroxypenta-2, 4-dienoate hydratase (ORF 1092), acetaldehyde dehydrogenase (ORF 1093), 4-hydroxy-2-oxovalerate aldolase (ORF 1094) and 4-oxalocrotonate decarboxylase (ORF 1095). Some of these proteins were also detected in the lactate-grown cells but at significantly low levels (Table 3-5). The presence of low levels of these proteins could be due to the residual toluene in lactate grown cells or induced by the presence of structurally similar metabolic intermediates of aromatic amino acids in the cells.

Genomic analysis predicted 11 ORFs encoding proteins for 4-HBA degradation and of these, 6 proteins were identified using proteomics (Table 3-6). These included 4-hdyroxybenzoate 3-monooxygenase (ORF 3919) responsible for the conversion of 4-HBA into PCA, 3,4-PCA (ORF 3916) responsible for *ortho*-cleavage of PCA into β-carboxymuconic acid, 3-oxoadipate enol-lactonase (ORF 3914) responsible for conversion of β-ketoadipate enol lactone into β-ketoadipate, 3-oxoadipate succinyl-CoA transferase, A and B subunit (ORFs 3911 and 3912, respectively) responsible for further conversion of β-ketoadipate into β-ketoadipyl CoA. All these proteins were detected only in the 4-HBA-grown cells and not in the lactate-grown cells. The differences in spectral counts between the two samples were also significantly different according to t-test (Table 3-6).

In silico analysis of the Seminole genome identified 15 ORFs encoding enzymes involved in the degradation of PAA. Of these, a total of 10 gene products of the *paa* gene cluster (*paa BCEFGIKN*) were upregulated and identified by LC-MS/MS only in the PAA-grown cells and not in lactate-grown cells (Table 3-7). These proteins included *paaF*-encoded protein

phenylacetyl-CoA ligase (ORF 2986), and *paaGHIK*-encoded proteins that are known to be the core subunits of the phenylacetyl-CoA epoxidase (ORFs 2881, 2880, 2879 and 2877). Also identified was PaaY, a putative regulator (ORF 4127), PaaN (ORF 4128) the ring-opening protein, PaaB (ORF 4130) the oxepin-CoA forming, and PaaCE (ORFs 4131 and 4133) that are involved in the β-oxidation like steps of PAA degradation (Table 3-7).

3.5 Discussion

Our understanding of the biology and genetic potential of halophilic and halotolerant microorganisms that degrade hydrocarbons in moderate to high salinity environments is severely lacking, consequently our ability to use the natural potential of these organisms in the cleanup of contaminated hypersaline environments is not available. This chapter describes the metabolic potential of strain Seminole by studying its physiology and genome. The limited survey revealed that the sequences formed two major clusters, despite their diverse habitats and geographic origin, each cluster harboring organisms from both hydrocarbon contaminated as well as uncontaminated saline sites (Table 3-1, Fig. 3-1). Since most of these organisms' are not available in pure culture, determination of their ecophysiology and phylogenetic diversity cannot be fully interpreted.

Strain Seminole was capable of metabolizing oxygenated hydrocarbons such as 4-HBA, phenol, PCA, and PAA as the sole growth substrates in the presence of high salt. However, no degradation of benzoate, gentisate, and catechol was observed at 2 M NaCl (Table 3-2). These results are consistent with the genome analysis in which no specific genes that code for enzymes catalyzing the initial step in the breakdown of benzoate, GA, and hexadecane pathways were detected. The organism's inability to grow on catechol is puzzling since it was predicted to be an intermediate in benzene degradation pathway indicating that the substrates could be substituted variants and the conditions used for testing were not optimal.

In this study, bioinformatic analysis of the genome was performed to investigate the presence of ORFs encoding enzymes required for the degradation of various aromatic and aliphatic hydrocarbons. A comparative analysis of the COG categories between the Seminole genome and the genomes of other ten hydrocarbon degrading halophiles in IMG revealed the abundance in genes related to amino acid transport and metabolism (11.34%), carbohydrate transport and metabolism (6.71%) and secondary metabolite biosynthesis, transport and metabolism (4.17%) were significantly higher ($P<0.05$) in the Seminole genome compared to the average abundances of 9.92%, 4.93%, and 3.25%, respectively in the genomes of other ten hydrocarbon degrading halophiles (Table 3-4, Fig 3-2). These results clearly suggest the organism's efficient carbohydrate and amino acid acquisition and metabolism for energy.

Genome analysis of strain Seminole predicted the existence of complete degradation pathways for benzene, toluene, 4-HBA and PAA. The genomic findings were further corroborated by proteomic analysis where majority of the enzymes involved in the degradation pathways were identified. The genome analysis of strain Seminole identified a cluster of 19 genes arranged in a sequential fashion that code for both upper and lower pathway of benzene and toluene degradation. Genes encoding a multicomponent PH-like protein were contiguous genes that encode downstream enzymes for *meta*-cleavage of CAT or 3-methylcatechol (Fig. 3-3A, Table 3-5).

A similar genetic organization of PH genes coupled with CAT *meta*-cleavage pathway has been described in *Pseudomonas* sp. CF600 (43), *Comamonas testosteroni* TA441 (44) and *Cupriavidus necator* JMP134 (45). No genes homologous to benzene monooxygenases, benzene dioxygenases, toluene monooxygenases, and toluene dioxygenases that catalyze the initial ring oxidation steps were identified in the genome of strain Seminole further supporting the assertion that enzymes homologous to PH are responsible for the initial steps in benzene and toluene degradation in strain Seminole. The proteomic analysis corroborated many of the predicted

enzymes. Four of the putative components of PH complex (ORFs 1080, 1082, 1083 and 1084) were highly expressed (average spectral counts of 61, 60, 5 and 35 respectively) in the proteome of toluene-grown cells. Based on these results, it was predicted that PH catalyzes the oxidation of benzene and toluene to corresponding one hydroxylated intermediates such as phenol and *o*- or *m*- cresol, respectively, in the first step. Thus formed phenol and *o*- or *m*- cresols are converted to CAT and 3-methylcatechol, respectively in the second step of oxidation (Fig. 3-3B and C). Multicomponent PHs are soluble diiron monooxygenase that can hydroxylate a variety of hydrocarbons including benzene, toluene, naphthalene, phenol, and methyl-substituted phenols (46). Chapter II described the elucidation of benzene degradation pathway in the phylogenetically closely related organism, *Arhodomonas* sp. strain Rozel (98% 16S rRNA-gene identity to strain Seminole) where benzene is converted to phenol and then to catechol by PH-like enzymes. The formation of phenol was confirmed as the first intermediate using GC-MS (42). Besides PH-like enzymes, the proteome of toluene-grown cells also contained all the downstream enzymes that can catalyze the conversion of CAT or 3-methylcatechol to acetaldehyde and pyruvate that can enter the TCA cycle. These findings suggest that both benzene and toluene were degraded *via* the same set of enzymes.

Strain Seminole was able to utilize 4-HBA as the sole source of carbon in the presence of 2 M NaCl. Genomic analysis revealed the presence of a complete set of genes required for 4-HBA degradation pathway, where 4-HBA is converted to PCA that undergoes *ortho*-ring cleavage and enters the β-ketoadipate pathway to form intermediates that will be assimilated *via* the TCA cycle (Table 3-6 and Fig. 3-4). Analysis of the cytosolic proteome of 4-HBA grown cells has identified key proteins involved in 4-HBA degradation pathway substantiating the genome predicted pathway. 4-HBA is hydroxylated to PCA by 4-hydroxybenzoate 3-monooxygenase encoded by ORF 3919, a homolog of *pobA* gene in *Chromohalobacter* sp. HS2 (47). The PCA undergoes ring-cleavage at *ortho* position to form β-carboxymuconic acid and is

catalyzed by 3,4-PCA encoded by ORF 3916. This is an important enzyme because PCA is a key intermediate produced from decaying lignin and other industrial aromatic compounds (36). Our proteomic analysis showed that a significant amount (53 average spectral counts) of this enzyme is produced only in 4-HBA grown cells compared to lactate-grown cells (Table 3-6). The ring cleavage product of PCA, β-carboxymuconic acid is further converted in a stepwise fashion to 4-carboxymuconolactone, β-ketoadipate enol-lactone, β-ketoadipate, β-ketoadipyl CoA, and finally to central intermediates, succinyl-CoA and acetyl-CoA (Table 3-6 and Fig. 3-4B). These results clearly show the physiological and genetic ability of an extreme halophile to completely assimilate aromatic compounds originating from decaying plant material and environmental pollution in high salinity environments. The genetic organization of the upper pathway (*pob* genes*)* and lower pathway (*pca*) varies in different 4-HBA degrading organisms. They are either found clustered together in a contiguous pattern or scattered over several portions of the genome. For example, the *pob* and *pca* genes were found clustered in two different locations in *Cupriavidus necator* JMP134 (45), three clusters in *Burkholderia xenovorans* LB400 (48), and dispersed throughout the chromosome in *Polaromonas* sp. strain JS666 (49). However, in Seminole genome, all the genes needed for 4-HBA degradation are located on one contig clustered together. Similar organization of *pob* and *pca* genes into one supraoperonic cluster was also reported in *Acinetobacter* sp. strain ADP1 (50).

Strain Seminole was also capable of utilizing PAA as the sole source of carbon in the presence of 2 M NaCl. Genomic analysis of strain Seminole identified 15 genes involved in the degradation of PAA. PAA degradation pathway is the central route where peripheral pathways from many aromatic compounds including phenylalanine, tropic acid, phenylethylamine, phenylalkanoic acids, phenyl acetaldehyde, styrene, and ethylbenzene converge at PAA and then funnel into the TCA cycle (38). Our analysis shows that like many non-halophiles, strain Seminole degrades PAA using an aerobic hybrid strategy that incorporates features of both

aerobic and anaerobic degradation steps (38, 51). The first step is the Co-A dependent activation step in which PAA is converted to phenylacetate-CoA (PA-CoA) by a PA-CoA ligase (PaaF) and thus formed PA-CoA is transformed by a multicomponent oxygenase into its ring-1, 2-epoxide that becomes oxepin-CoA intermediate which undergoes hydrolytic cleavage to form an aliphatic compound that is channeled through β-oxidation like mechanism to form acetyl-CoA and succinyl-CoA (41). The *paa* genes are widely found in many different and phylogenetically unrelated bacteria (38). Our observations that PAA is degraded by a similar hybrid pathway in strain Seminole extends this mechanism to organisms living in high salinity environments. This is important since degradation of hydrocarbons in high salinity environment can only be achieved by halophiles or halotolerants since other organisms cannot survive high salinity. *In-silico* analysis shows that all 15 genes are arranged in clusters and distributed on three different contigs in the Seminole genome. Similar studies in non-halophiles indicate that even though the arrangement of *paa* genes differs among different organisms, some features remain common to most described *paa* clusters (51). For example, genes encoding the ring-hydroxylating system (PaaGHIJK) and beta-oxidation system (PaaACE) are usually found to consistently cluster together in most *paa* systems either on the same or different locations on chromosome (51). Similar organization was also predicted in Seminole genome, where ORFs 2877-2882 (ring hydroxylating) and ORFs 4129, 4131 and 4133 (β-oxidation system) are clustered on two different contigs (Fig. 3-5).

In the current study we have used the draft genome to predict putative steps and requisite enzymes that catalyze degradation pathways of benzene, toluene, 4-HBA, and PAA. Furthermore using proteomic tools, many of the predicted enzymes were identified upon their expression when grown on the respective hydrocarbons. It is important to point out that some of the predicted genes were not detected in the proteomic analysis. This could be due to low expression levels or expressed at different times during the degradation process. Overall, this study suggests that

Arhodomonas sp. strain Seminole degrades aromatic hydrocarbons using enzymes similar to non-halophilic hydrocarbon degraders. Although these proteins could perform identical reactions as their non-halophilic homologs, they differed by exhibiting low-pI values (theoretical) when compared to their non-halophilic homologs. For example, pI values ranging from 4.24-5.34 were computed for proteins identified in benzene and toluene degradation. However, their non-halophilic homologs displayed pI values ranging between 5.07-5.93. Similarly, proteins identified in 4-HBA pathway showed pI values ranging between 4.6-6.53, while their non-halophilic homologs showed pI values between 5.16-6.66. In case of PAA pathway, Seminole proteins showed pI values ranging between 4.72-5.94 compared to their non-halophilic homologs that showed pI values from 5.16-6.37. Low pI values of halophilic proteins are mostly due to high content of acidic residues such as glutamate and aspartate on the surface of proteins (52). Excessive acidic residues protect the enzyme from aggregation, binding of essential water molecules and provide flexibility within a protein through electrostatic repulsion under high salinity (53, 54). These unique features allow halophilic proteins maintain their stability and activity at high salinity (1-4M). In contrast, non-halophilic proteins precipitate and aggregate at high salinities thus losing their activity.

The current work shows that *Arhodomonas* are found to be widely distributed in contaminated as well as uncontaminated environments with varying salinity and are capable of degrading various aromatic compounds notably benzene and toluene suggesting their important role in the natural attenuation of petroleum compounds. A combination of genomic and proteomic approaches was utilized to elucidate hydrocarbon degradation pathways in a novel halophile, *Arhodomonas* sp. strain Seminole and such attempts to reconstruct metabolic steps represents a reliable platform for developing strategies aimed at the exploitation of halophiles.

3.6 ACKNOWLEDGEMENTS

This work was supported by funding from the Oklahoma Transportation Center (grant DTRT06-G-0014) and by the National Science Foundation (grant OCE1049301). Mass spectrometry analyses were performed in the DNA/Protein Resource Facility at Oklahoma State University, using resources supported by the NSF MRI and EPSCoR programs (award DBI/0722494).

We thank Janet Rogers from the DNA/Protein Resource Facility at Oklahoma State University for her assistance with the LC-MS/MS analysis. We also thank Dr. Mostafa Elshahed for helping with designing Arhodomonas specific primers. We are grateful to Drs. Aharon Oren, Kathleen Duncan, Frank Loeffler, Mostafa Elshahed, Rene F. Bernier, and Eitan Ben-Dov for their assistance in obtaining samples for ecological distribution study.

3.7 REFERENCES

1. **Oren A, Gurevich P, Azachi M, and Henis Y.** 1992. Microbial degradation of pollutants at high salt concentrations. Biodegradation. **3:**387-398.

2. **Margesin R, and Schinner F.** 2001. Biodegradation and bioremediation of hydrocarbons in extreme environments. Appl Microbiol Biotechnol. **56:**650-663.

3. **Lefebvre O, and Moletta R.** 2006. Treatment of organic pollution in industrial saline wastewater: a literature review. Water Res. **40:**3671-3682.

4. **Sorokin DY, Janssen AJH, and Muyzer G.** 2011. Biodegradation potential of halo(alkali)philic prokaryotes. Crit Rev Environ Sci Technol. **42:**811-856.

5. **Fathepure BZ.** 2014. Recent studies in microbial degradation of petroleum hydrocarbons in hypersaline environments. Frontiers in microbiology. **5:**173. doi: 10.3389/fmicb.2014.00173.

6. **Martins LF, and Peixoto RS.** 2012. Biodegradation of petroleum hydrocarbons in hypersaline environments. Braz J Microbiol. **43:**865-872.

7. **Patzelt H.** 2005. Hydrocarbon degradation under hypersaline conditions - some facts, some experiments and many open questions. p. 105-122. *In* N. Gunde-Cimerman, A. Oren, and A. Plemenitaš (ed.), Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya, vol. 9. Springer Netherlands.

8. **Bertrand JC, Almallah M, Acquaviva M, and Mille G.** 1990. Biodegradation of hydrocarbons by an extremely halophilic archaebacterium. Lett Appl Microbiol. **11:**260-263.

9. **Emerson D, Chauhan S, Oriel P, and Breznak JA.** 1994. *Haloferax* Sp D1227, a halophilic archaeon capable of growth on aromatic-compounds. Arch Microbiol. **161:**445-452.

10. **Al-Mailem DM, Sorkhoh NA, Al-Awadhi H, Eliyas M, and Radwan SS.** 2010. Biodegradation of crude oil and pure hydrocarbons by extreme halophilic archaea from hypersaline coasts of the Arabian Gulf. Extremophiles. **14:**321-328.

11. **Tapilatu YH, Grossi V, Acquaviva M, Militon C, Bertrand JC, and Cuny P.** 2010. Isolation of hydrocarbon-degrading extremely halophilic archaea from an uncontaminated hypersaline pond (Camargue, France). Extremophiles. **14:**225-231.

12. **Fairley DJ, Boyd DR, Sharma ND, Allen CCR, Morgan P, and Larkin MJ.** 2002. Aerobic metabolism of 4-hydroxybenzoic acid in Archaea via an unusual pathway involving an intramolecular migration (NIH shift). Appl Environ Microbiol. **68:**6246-6255.

13. **Cuadros-Orellana S, Pohlschröder M, and Durrant LR.** 2006. Isolation and characterization of halophilic archaea able to grow in aromatic compounds. Int Biodeterior Biodegradation. **57:**151-154.

14. **Erdoğmuş S, Mutlu B, Korcan S, Güven K, and Konuk M.** 2013. Aromatic Hydrocarbon Degradation by Halophilic Archaea Isolated from Çamaltı Saltern, Turkey. Water, Air, Soil Pollut. **224:**1-9.

15. **Nicholson CA.** 2005. Biodegradation of petroleum hydrocarbons by halophilic and halotolerant microorganisms (Master's Thesis). Oklahoma State University, Stillwater, OK. Retrieved from http://digital.library.okstate.edu/etd/umi-okstate-1536.pdf.

16. **Nicholson CA, and Fathepure BZ.** 2004. Biodegradation of benzene by halophilic and halotolerant bacteria under aerobic conditions. Appl Environ Microbiol. **70:**1222-1225.

17. **Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, Forster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, Konig A, Liss T, Lussmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, and Schleifer KH.** 2004. ARB: a software environment for sequence data. Nucleic Acids Res. **32:**1363-1371.

18. **Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, and Tiedje JM.** 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. **42:**D633-642.

19. **Sublette KL MA, Ford L, Duncan K, Thoma G, and Brokaw J.** 2005. Remediation of a spill of crude oil and brine without gypsum. Environ Geosci. **12:**115-125.

20. **Youssef NH, Ashlock-Savage KN, and Elshahed MS.** 2012. Phylogenetic diversities and community structure of members of the extremely halophilic Archaea (order *Halobacteriales*) in multiple saline sediment habitats. Appl Environ Microbiol. **78:**1332-1344.

21. **Ben-Dov E, Kushmaro A, and Brenner A.** 2009. Long-term surveillance of sulfate-reducing bacteria in highly saline industrial wastewater evaporation ponds. Saline Systems. **5:**1-5.

22. **Stackebrandt E, and Goodfellow M.** 1991. Nucleic acid techniques in bacterial systematics, p. xxix, 329 p., Modern microbiological methods. Wiley, Chichester; New York.

23. **Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ.** 1990. Basic local alignment search tool. J Mol Biol. **215:**403-410.

24. **Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S.** 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. **30:**2725-2729.

25. **Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, and Rothberg JM.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature. **437:**376-380.

26. **Bodenteich A, S. L. Chissoe, Y. F. Wang, and B. A. Roe.** 1993. Shotgun cloning as the strategy of choice to generate templates for high-throughput dideoxynucleotide sequencing. p. 42-50. *In* C. F. Mark D. Adams, J. Craig Venter (ed.), Automated DNA sequencing and analysis techniques. Academic Press, London.

27. **DeAngelis MM, Wang DG, and Hawkins TL.** 1995. Solid-phase reversible immobilization for the isolation of PCR products. Nucleic Acids Res. **23:**4742-4743.

28. **Consortium TU.** 2012. Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Res. **40:**D71-D75.

29. **Lukashin AV, and Borodovsky M.** 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. **26:**1107-1115.

30. **Bradford MM.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal biochem. **72:**248-254.

31. **Voruganti S, LaCroix JC, Rogers CN, Rogers J, Matts RL, and Hartson SD.** 2013. The Anticancer Drug AUY922 Generates a Proteomics Fingerprint That Is Highly Conserved among Structurally Diverse Hsp90 Inhibitors. J Proteome Res. **12:**3697-3706.

32. **Keller A, Nesvizhskii AI, Kolker E, and Aebersold R.** 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem. **74:**5383-5392.

33. **Lundgren DH, Hwang S-I, Wu L, and Han DK.** 2010. Role of spectral counting in quantitative proteomics. Expert Rev Proteomics. **7:**39-53.

34. **Powlowski J, and Shingler V.** 1994. Genetics and biochemistry of phenol degradation by *Pseudomonas* sp. CF600. Biodegradation. **5:**219-236.

35. **DiMarco AA, and Ornston LN.** 1994. Regulation of p-hydroxybenzoate hydroxylase synthesis by PobR bound to an operator in *Acinetobacter calcoaceticus*. J Bacteriol. **176:**4277-4284.

36. **Harwood CS, and Parales RE.** 1996. The beta-ketoadipate pathway and the biology of self-identity. Annu Rev Microbiol. **50:**553-590.

37. **Olivera ER, Minambres B, Garcia B, Muniz C, Moreno MA, Ferrandez A, Diaz E, Garcia JL, and Luengo JM.** 1998. Molecular characterization of the phenylacetic acid catabolic pathway in *Pseudomonas putida* U: the phenylacetyl-CoA catabolon. Proc Natl Acad Sci U S A. **95:**6419-6424.

38. **Luengo JM, Garcia JL, and Olivera ER.** 2001. The phenylacetyl-CoA catabolon: a complex catabolic unit with broad biotechnological applications. Mol Microbiol. **39:**1434-1442.

39. **Ferrandez A, Minambres B, Garcia B, Olivera ER, Luengo JM, Garcia JL, and Diaz E.** 1998. Catabolism of phenylacetic acid in Escherichia coli. Characterization of a new aerobic hybrid pathway. J Biol Chem. **273:**25974-25986.

40. **Ferrandez A, Garcia JL, and Diaz E.** 2000. Transcriptional regulation of the divergent paa catabolic operons for phenylacetic acid degradation in *Escherichia coli*. J Biol Chem. **275:**12214-12222.

41. **Teufel R, Mascaraque V, Ismail W, Voss M, Perera J, Eisenreich W, Haehnel W, and Fuchs G.** 2010. Bacterial phenylalanine and phenylacetate catabolic pathway revealed. Proc Natl Acad Sci U S A. **107:**14390-14395.

42. **Dalvi S, Azetsu S, Patrauchan MA, Aktas DF, and Fathepure BZ.** 2012. Proteogenomic elucidation of the initial steps in the benzene degradation pathway of a novel halophile, *Arhodomonas* sp. strain Rozel, isolated from a hypersaline environment. Appl Environ Microbiol. **78:**7309-7316.

43. **Shingler V, Franklin FC, Tsuda M, Holroyd D, and Bagdasarian M.** 1989. Molecular analysis of a plasmid-encoded phenol hydroxylase from Pseudomonas CF600. J Gen Microbiol. **135:**1083-1092.

44. **Arai H, Ohishi T, Chang MY, and Kudo T.** 2000. Arrangement and regulation of the genes for meta-pathway enzymes required for degradation of phenol in *Comamonas testosteroni* TA441. Microbiology. **146:**1707-1715.

45. **Perez-Pantoja D, De la Iglesia R, Pieper DH, and Gonzalez B.** 2008. Metabolic reconstruction of aromatic compounds degradation from the genome of the amazing pollutant-degrading bacterium *Cupriavidus necator* JMP134. FEMS microbiology reviews. **32:**736-794.

46. **Leahy JG, Batchelor PJ, and Morcomb SM.** 2003. Evolution of the soluble diiron monooxygenases. FEMS Microbiol Rev. **27:**449-479.

47. **Kim D, Kim SW, Choi KY, Lee JS, and Kim E.** 2008. Molecular cloning and functional characterization of the genes encoding benzoate and *p*-hydroxybenzoate degradation by the halophilic *Chromohalobacter* sp strain HS-2. FEMS Microbiol Lett. **280:**235-241.

48. **Perez-Pantoja D, Donoso R, Agullo L, Cordova M, Seeger M, Pieper DH, and Gonzalez B.** 2012. Genomic analysis of the potential for aromatic compounds biodegradation in *Burkholderiales*. Environ Microbiol. **14:**1091-1117.

49. **Mattes TE, Alexander AK, Richardson PM, Munk AC, Han CS, Stothard P, and Coleman NV.** 2008. The genome of *Polaromonas* sp. strain JS666: insights into the evolution of a hydrocarbon- and xenobiotic-degrading bacterium, and features of relevance to biotechnology. Appl Environ Microbiol. **74:**6405-6416.

50. **Dal S, Trautwein G, and Gerischer U.** 2005. Transcriptional organization of genes for protocatechuate and quinate degradation from *Acinetobacter* sp. strain ADP1. Appl Environ Microbiol. **71:**1025-1034.

51. **Navarro-Llorens JM, Patrauchan MA, Stewart GR, Davies JE, Eltis LD, and Mohn WW.** 2005. Phenylacetate catabolism in *Rhodococcus* sp. strain RHA1: a central pathway for degradation of aromatic compounds. J Bac. **187:**4497-4504.

52. **Reed CJ, Lewis H, Trejo E, Winston V, and Evilia C.** 2013. Protein adaptations in archaeal extremophiles. Archaea. **2013:**373275.

53. **Sinha R, and Khare SK.** 2014. Protective role of salt in catalysis and maintaining structure of halophilic proteins against denaturation. Front Microbiol. **5:**165.

54. **Mevarech M, Frolow F, and Gloss LM.** 2000. Halophilic enzymes: proteins with a grain of salt. Biophys Chem. **86:**155-164.

CHAPTER IV

Benzoate Degradation by Archaeal Enrichment Under Extreme Salinity

4.1 ABSTRACT

This study explores the hydrocarbon degradation potential of benzoate-degrading archaeal enrichment and examines its diversity and community structure using pyrosequencing. The enrichment was developed using sediment samples from Rozel Point at Great Salt Lake, UT. The enrichment degrades benzoate as the sole carbon source at salinity ranging from 2-5 M NaCl with highest rate of degradation at 4 M. The enrichment was also tested for its ability to grow on other aromatic compounds such as 4-hydroxybenzoic acid (4-HBA), gentisic acid, protocatechuic acid (PCA), catechol, benzene and toluene as the sole carbon and energy sources. Of these, the culture only utilized 4-HBA as the carbon source. A survey using PCR and degenerate primers showed the presence of 4-hydroxybenzoate 3-monooxygenase (4-HBMO) and protocatechuate 3,4-dioxygenase (3,4-PCA) genes suggesting that the archaeal enrichment might degrade benzoate to 4-HBA which is further converted via 4-HBMO to form PCA which undergoes ring-cleavage via 3,4-PCA to form intermediates that enter the Krebs cycle.

Microbial community analysis using 16S rRNA gene sequences revealed that the enrichment consisted entirely of *Halobacteriaceae* members belonging to the genera *Halopenitus, Halosarcina, Natronomonas*, *Halosimplex*, *Halorubrum*, *Salinarchaeum* and *Haloterrigena.* Of these, members of *Halopenitus* were dominant accounting for almost 91% of the total sequences suggesting their potential role in degrading oxygenated aromatic compounds at extreme salinity.

4.2 INTRODUCTION

Halophilic archaea are included in the phylum *Euryarchaeaota*. Grant et al. proposed that extremely halophilic archaea be designated as haloarchaea and are currently included in the family *Halobacteriaceae* (1). *Halobacteriaceae* members are aerobic heterotrophs with diverse nutritional demands and metabolic pathways (2). Most halophilic archaea prefer amino acids as carbon and energy source. However, species requiring complex ingredients such as yeast extract and simple carbon sources like acetate, succinate, or pyruvate etc. have also been reported (3). Members of the genera *Haloferax, Haloarcula* and *Halococcus* are carbohydrate-utilizing that can metabolize hexoses, pentoses, sucrose and lactose (2). Reports have shown that few representatives of *Haloferax, Halobacterium, Halococcus, Halorubrum* and *Haloarcula* can metabolize components of crude oil, aliphatic, aromatic hydrocarbons, benzoate and 4-HBA (4). Cuadros et al. have also suggested that the ability to degrade 4-HBA acid could be a common feature of the family *Halobacteriaceae* (5). Given the potential to degrade a wide variety of hydrocarbons at high salinities (20-30% NaCl), haloarchaea can be considered as potential candidates for bioremediation of hydrocarbon-contaminated hypersaline sites.

To date only few studies have reported the archaeal ability to degrade hydrocarbons at high salinities. However, much fewer studies exist on the genetics and biochemical mechanisms of hydrocarbon degradation in halophilic archaea. Gentisate 1,2-dioxygenase (1,2-GDO) enzyme

activity was detected in cell extracts of *Haloferax* sp. D1227 grown on 3-phenylpropionic acid, cinnamic acid and benzoic acid (6). Fairley et. al detected 1,2-GDO enzyme activity in 4-HBA grown cells of *Haloarcula* sp. D1 (7). In the same study using 2, 6-dideutero-4HBA as a substrate, 3-deutero-gentisate was detected as an intermediate by GC-MS and proton nuclear magnetic resonance spectroscopy suggesting that 4-HBA is degraded *via* GA in strain D1 rather than the usual PCA, hydroquinone or CAT (7). A recent study reported the activity of catechol 1,2-dioxygenase (1,2-CAT) and protocatechuate 3,4-dioxygenase (3,4-PCA) enzymes of the *ortho*-cleavage pathway in *Halorubrum, Haloarcula, Halobacterium* and *Haloferax* that can degrade 4-HBA, naphthalene, phenanthrene and pyrene (8). Though these few reports show the presence of ring-cleavage genes, almost nothing is known about the initial hydroxylation steps during hydrocarbon degradation in halophilic archaea.

In the present study, a benzoate-degrading archaeal consortium was developed using sediment samples collected from the Great Salt Lake (GSL) near Rozel Point. The site represents two main characteristics: high levels of salinity ranging from 5-30% w/v and numerous oil seeps. In spite of these two extreme conditions, only a few studies have reported the hydrocarbon degradation potential of microorganisms indigenous to GSL. Ward et al. showed degradation of hexadecane at in the presence of 0 to 20% NaCl by a microbial consortium enriched using samples from GSL (9). Another study developed an enrichment culture using sediment samples from GSL near Rozel Point with the ability to degrade benzene and toluene optimally at 14.5 % salinity. Molecular analysis of the enrichment showed that members of *Gammaproteobacteria* such as *Arhodomonas* spp., *Sphingomonas* spp. and *Halomonas* spp. were the dominant members during benzene metabolism (10).

This study makes an attempt to enrich an archaeal community that degrades benzoate at extreme salinity. Efforts were made to predict the mechanism of benzoate degradation by

screening for the presence of key ring opening and ring cleaving catabolic genes. Pyrosequencing was used to investigate the archaeal diversity and community structure.

4.3 MATERIALS AND METHODS

4.3.1 Chemicals. All chemicals used in this study are already described in Chapter III, Section 3.3.1. The composition of mineral salts medium (MSM) used in this study is previously described (11).

4.3.2 Development of enrichment culture. Microcosms were set up using 250 ml capacity flasks, 10 grams of sediment samples from Rozel point and 90 ml of MSM containing 2 M NaCl. Each flask was fed 2mM benzoate as the only source of carbon and energy. To inhibit the growth of bacteria during the enrichment process, the flasks were amended with 100 µg/ml of filter-sterilized chloramphenicol. Un-inoculated autoclaved flasks were prepared similarly and used as controls. All flasks were incubated without shaking at 30^{o}C. Biodegradation of benzoate was monitored by UV-VIS-spectrophotometer (Ultraspec 2000, Pharmacia Biotech). On utilization of the substrate, the flasks were repeatedly fed 2 mM benzoate. A sediment-free enrichment was obtained by transferring 50% of the culture periodically to fresh MSM containing 2 M NaCl, 2mM benzoate and 100 µg/ml of chloramphenicol.

4.3.3 Analytical methods. Utilization of aromatic compounds provided as substrates was monitored by measuring their absorbance using a UV-VIS spectrophotometer. Approximately 200-500 µl of culture sample was withdrawn from flasks and centrifuged at 10,000 rpm for 10 min. The supernatant was appropriately diluted with distilled water and analyzed by measuring absorbance at 223 nm for benzoate, 245 nm for 4-HBA, 288 nm for PCA, 274 nm for CAT and 318 nm for GA. Degradation of benzene and toluene were analyzed using gas chromatography (GC) as described previously (11).

4.3.4 Benzoate degradation at different salt concentrations. To study the ability of the enrichment culture to degrade benzoate at different salt concentrations, 250 ml flasks were set up containing 45ml of MSM supplemented with 0, 1, 2, 3, 4 and 5 M NaCl, 50 µg/ml of chloramphenicol, and 2mM benzoate as the sole source of carbon and energy. Triplicate active flasks were prepared for each salt concentration and inoculated with 5 ml of actively growing enrichment. Un-inoculated autoclaved flasks (triplicate) were treated as controls. All flasks were incubated at 30^{o}C. Biodegradation of benzoate was monitored periodically at 223 nm using the UV-VIS spectrophotometer.

4.3.5 Growth of the enrichment on other aromatic compounds. The benzoate-degrading enrichment maintained at 4 M NaCl (4M-enrichment) was tested for its ability to utilize other aromatic compounds such as 4-HBA (2 mM), GA (3 mM), PCA (1 mM), and CAT (1mM) as the sole sources of carbon and energy. These experiments were carried out in 250 ml capacity flasks supplemented with 45 ml of 4M NaCl-MSM, 50 µg/ml of chloramphenicol and filter-sterilized test compounds. Triplicate active and duplicate control flasks were prepared for each test compound. All active flasks were inoculated with 5 ml of actively growing culture, whereas the control flasks were uninoculated. All the flasks were incubated at 30^{o}C and analyzed periodically for the removal of substrates by UV-VIS spectrophotometer.

The enrichment was also tested for its ability to degrade non-oxygenated hydrocarbons such as benzene and toluene. The enrichment was grown in 120-ml serum bottles filled with 45 ml of 4M NaCl-MSM and inoculated with 5 ml of actively growing culture. Bottles were amended with 17-22 µmoles of neat benzene and toluene as the sole sources of carbon and energy. These bottles were closed with Teflon-coated rubber septa and aluminum crimps and incubated in an inverted position at 30^{o}C. Degradation of benzene and toluene was analyzed on a weekly basis using GC.

4.3.6 Microbial community structure analysis.

(i) 454 pyrosequencing. Genomic DNA was extracted using FastDNA Spin kit for Soil (MP Biomedicals, OH) using the manufacturer's protocol from the 4M-enrichment. 16S rRNA-genes were amplified using universal primers, 515F (5'-GTG CCA GCM GCC GCG GTA A-3') and 1391R (5'-GAC GGG CGG TGW GTR CA-3'). Amplification was carried out as described previously (10). The PCR product was purified and cleaned using ExoSAP-IT following the manufacturer's protocol (USB, OH). 454 pyrosequencing was performed at MR DNA Laboratory (Molecular Research LP, TX). A single-step 30 cycle PCR using HotStarTaq Plus Master Mix Kit (Qiagen, CA) and 16S universal Eubacterial primers (515F and 806R) were used under the following conditions: 94^{o}C for 3 minutes, followed by 28 cycles of 94^{o}C for 30 seconds; 53^{o}C for 40 seconds and 72^{o}C for 1 minute; followed by final elongation step at 72^{o}C for 5 minutes. After PCR, the amplicon product was mixed in equal concentrations and purified using Agencourt Ampure beads (Agencourt Bioscience Corporation, MA, USA). Sequencing was performed using Roche 454 FLX titanium instruments and reagents and following manufacturer's guidelines.

(ii) Data processing. The data derived from the sequencing process was processed at MR DNA Laboratory (Molecular Research LP, Texas) using a proprietary analysis pipeline. Sequences were depleted of barcodes and primers. Sequences were filtered by removing short sequences of < 200 bp or with ambiguous base calls. Sequences with homopolymer runs exceeding 6 bp were also removed. This was followed by deionizing the sequences and removing chimeras. Operational taxonomic units (OTUs) were defined after removal of singleton sequences, clustering at 3% divergence (97% similarity) against a curated Greengenes database (12).

(iii) Phylogenetic analyses of 16S rRNA gene sequences. OTUs were queried using BLASTn against 16S Bacterial and Archaeal database in NCBI and assigned to a specific genus if it was ≥95% similar to a reference 16S rRNA gene sequence belonging to that genus (13).

Representative OTUs from the study and reference 16S rRNA gene sequences of known archaeal members were chosen to construct phylogenetic tree in MEGA6 (14). The selected sequences were aligned by Clustalw in MEGA6. The aligned sequences were then used to construct tree using maximum-likelihood method based on the Tamura 3-parameter model (15) with a discrete gamma distribution (+*G*, parameter=0.4782) and invariable site rate ([+*I*], 25.2751%). Bootstrap values were calculated based on 1000 replicates (16).

4.3.7 Design of 4-HBMO and 1, 2-GDO degenerate primers. Degenerate primers for 4-hydroxybenzoate 3-monooxygenase (4-HBMO) were designed using the following homologous protein sequences deposited in GenBank: WP_004593222 (*Haloarcula japonica*), YP_003536666 (*Haloferax volcanii* DS2), WP_008575994 (*Haloferax* sp.), WP_007275258 (*Haloferax sulfurifontis*), WP_006601488 (*Haloferax alexandrinus*), WP_004062322 (*Haloferax lucentense*), WP_004974749 (*Haloferax gibbonsii*), WP_004969491 (*Haloferax denitrificans*), WP_004962558 (*Haloarcula sinaiiensis*), WP_007188044 (*Haloarcula californiae*), WP_005535142 (*Haloarcula argentinensis*), AAV45405 (*Haloarcula marismortui*). The protein sequences were back translated to obtain nucleotide sequences by using the EMBOSS Backtranseq tool (http://www.ebi.ac.uk/Tools/st/emboss_backtranseq/). The nucleotide sequences were then aligned by ClustalW program in MEGA5 (17). The primers 4-HBMO-F (5'-CAACCACATCGCCTACCA-3') and 4-HBMO-R (5'- CTCGTCGCCGATGGC -3') were designed based on the two conserved regions observed in the alignment, spanning the residues 591-608 and 1132-1146 in 4-HBMO gene from *Haloferax volcanii* DS2 (Accession no. YP_003536666). Similarly, a set of degenerate primers for PCR amplification of 1,2-GDO gene were designed using the following homologous protein sequences deposited in GenBank: YP_003533551 (*Haloferax volcanii* DS2), WP_004967945 (*Haloferax denitrificans*), WP_004043351 (*Haloferax volcanii*), WP_006107340 (*Natrialba asiatica*), WP_006667521 (*Natrialba aegyptia*), WP_007260867 (*Natronolimnobius innermongolicus*), WP_007274074

(*Haloferax sulfurifontis*), WP_007694513 (*Halococcus hamelinensis*), WP_008388533 (*Halosarcina pallida*), WP_008422437 (*Natronococcus jeotgali*), WP_008892413 (*Haloterrigena salina*). The primers 1,2-GDO-F (5'- GTGGCACGACCACRTCAA -3') and 1,2-GDO-R (5'- TCSGTCATSCCGAGGAG -3') were designed based on the two conserved regions observed in the alignment, spanning the residues 438-455 and 1006-1019 in 1, 2-GDO gene from *Haloferax volcanii* DS2 (Accession no. YP_003533551). All the primers were then tested *in-silico* against known 4-HBMO and 1,2-GDO sequences using BLASTn.

4.3.8 Detection of catabolic genes using PCR, cloning and sequencing. Genomic DNA was extracted from the 4M-enrichment using method described in Section 4.3.6. The enrichment was screened for the presence of catabolic genes encoding key ring-hydroxylating and ring-cleaving enzymes involved in benzoate degradation pathway.

Table 4-1. List of primers used in this study for amplification of catabolic genes

Primer	Target gene	Sequence	Reference study
Ben 1,2-D-F	Benzoate 1,2-dioxygenase	5'- CGTTYCAYGGCTGGACVTTC	(18)
Ben 1,2-D-R	(*benA*)	5'- CGGAAYTCYTCSAGRTCGTC	
4-HBMO-F	4-hydroxybenzoate 3-monooxygenase	5'- CAACCACATCGCCTACCA	This
4-HBMO-R	(4-HBMO)	5'- CTCGTCGCCGATGGC	
1,2-GDO-F	Gentisate 1,2-dioxygenase	5'- GTCGCACGACCACRTCAA	This
1,2-GDO-R	(1,2-GDO)	5'- TCSGTCATSCCGAGGAG	
3,4-PCA-F	Protocatechuate 3,4-dioxygenase	5'- GAGRTSTGGCARGCSAAY	(19)
3,4-PCA-R	(3,4-PCA)	5'- CCGYSSAGCACGATGTC	
1,2-CAT-F	Catechol 1,2-dioxygenase	5'- ACCATCGARGGYCCSCTSTAY	(19)
1,2-CAT-R	(1,2-CAT)	5'- GGTRATCTGGGTGGTSAG	
2,3-CAT-F	Catechol 2,3-dioxygenase	5'- GARCTSTAYGCSGAYAAGGAR	(19)
2,3-CAT-R	(2,3-CAT)	5'- RCCGCTSGGRTCGAAGAARTA	

Each PCR reaction contained 2.5 µl of 10X reaction buffer (Teknova), 2.5 µl of 25 mM $MgCl_2$, 1 µl of dNTPs (200 µM each), 20 *p*mol of the appropriate primers, 1 µl of Taq polymerase and sterile distilled water to adjust the total volume to 25 µl. PCR conditions for amplification of 4-HBMO gene were as follows: 95°C for 8 min, followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 55.5°C for 1 min, and extension at 72°C for 1 min, followed by final elongation at 72°C for 8 min. Similar PCR conditions were used for amplification of 1,2-GDO gene and benzoate 1,2-dioxygnease (*benA*) gene (18) except for annealing temperature of 57°C and 54.3°C, respectively. PCR amplification of other ring-cleaving genes such as 1,2-CAT, 2,3-CAT, and 3,4-PCA genes was performed using degenerate primers and conditions described previously by Garcia et al. (19) with annealing temperatures of 50°C, 43.5°C and 60°C, respectively .

The PCR products were separated on 1.5% agarose gel, excised and purified using QIAquick Gel Extraction Kit (QIAGEN, CA) following the manufacturer's protocol. The purified PCR fragments were ligated into pGEM-T vector System I (Promega, WI) and transformed into One Shot chemically competent TOP10 *Escherichia coli* cells (Invitrogen, CA) following the manufacturer's protocol. The transformed cells were then plated on to Luria Broth (LB) plates with 100 µg/ml of ampicillin. The colonies were randomly picked and checked for inserts with M13 primers using similar PCR conditions, except the annealing temperature was 55°C. The PCR products were then cleaned prior to sequencing using ExoSAP-IT (USB, OH) following manufacturer's protocol. Sequencing was performed using Applied Biosystems 3730 DNA analyzer in the Core Facility at Oklahoma State University, Stillwater, OK. The sequences obtained were then searched against nucleotide database in NCBI using BLASTn.

4.4 RESULTS

4.4.1 Development of benzoate-degrading enrichment. The enrichment culture consistently utilized benzoate as the sole carbon source in the presence of 2 M NaCl and will be referred to as 2M-enrichment from now onwards. Initially, the enrichment degraded 2 mM benzoate in a period of 5 weeks. However, the rate increased with repeated feeding and complete degradation of the added benzoate occurred in 3 weeks after 9 weeks of enrichment (Fig 4-1).

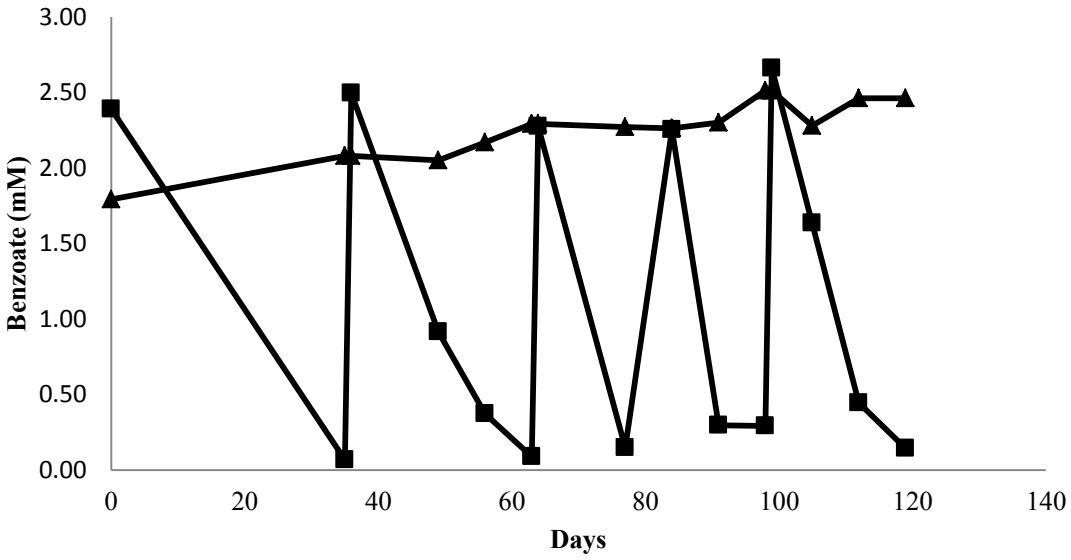

Fig 4-1. Degradation of benzoate by the enrichment at 2 M NaCl. Repeated use of benzoate by the enrichment at 2 M NaCl (■) developed with sediment sample from Rozel Point, GSL. No degradation was seen in control flasks (▲). Both active and control flasks were incubated without shaking at 30°C.

The enrichment was also evaluated for its ability to degrade benzoate at varying salinity ranging from 0 to 5 M NaCl. Results show that degradation of benzoate occurred only at 2, 3, 4 and 5 M NaCl (Fig 4-2). Degradation occurred best in the presence of both 4 and 5 M NaCl. More than 90% of the initially added benzoate was degraded within 17 days of incubation in the presence of 4 and 5 M NaCl. However, degradation proceeded with a lag period of 7 days at 5 M

NaCl. The data also revealed that only 57% and 82% of the added benzoate was degraded when grown in the presence 2 and 3 M NaCl, respectively within the same time period. No degradation of benzoate was observed at 0 and 1 M NaCl even after 21 days of incubation suggesting that the enrichment required much higher concentrations of NaCl for growth and degradation. Overall, these results suggested that 4 M NaCl was optimum for benzoate degradation. Therefore, all the experiments presented in this chapter were carried out in the presence of 4 M NaCl.

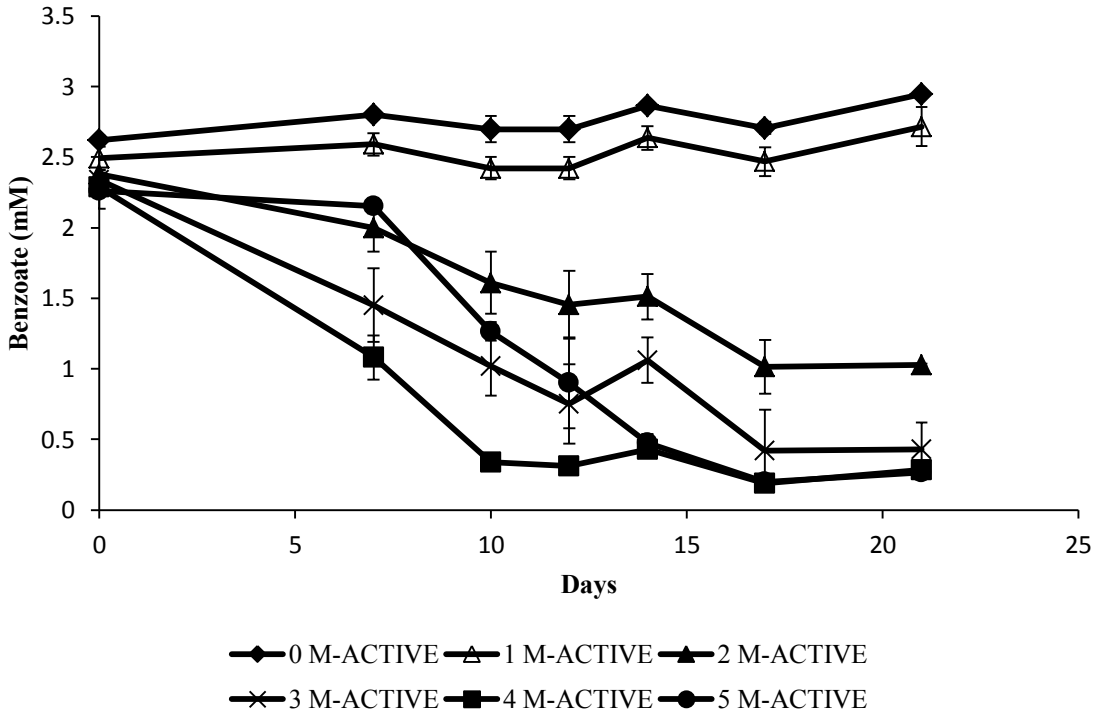

Fig 4-2. Degradation of benzoate by the enrichment at different salt concentration including, 0 M NaCl (♦), 1 M NaCl (△), 2 M NaCl (▲), 3 M NaCl (✕), 4 M NaCl (■), 5 M NaCl (●). Active flasks were set up for each salt concentration with 45 ml of MSM with 50 ug/ml of chloramphenicol and 2 mM benzoate. Each active flask was inoculated with 5 ml of actively growing enrichment culture. Highest rate of degradation was seen at 4 M NaCl. No degradation was seen at 0 and 1 M NaCl. No degradation occurred in controls (not shown). Data is mean of triplicate flasks and bars indicate ± standard deviation.

4.4.2 Growth on other aromatic compounds. The 4M-enrichment was tested for its ability to degrade other oxygenated aromatic compounds such as 4-HBA, GA, PCA, CAT and non-oxygenated aromatics like benzene and toluene as sole sources of carbon. The enrichment was able to utilize 4-HBA (2mM) within a period of 28 days with a lag period of 12 days (Fig. 4-3). However, the enrichment failed to utilize GA, CAT, PCA, benzene and toluene as the growth substrates even after 28 days of incubation (Data not shown).

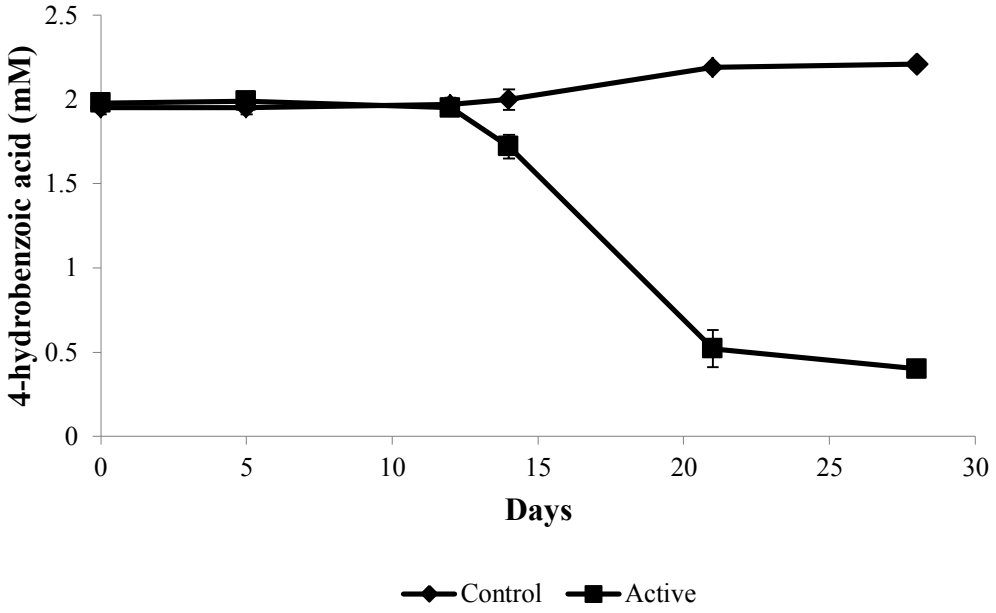

Fig 4-3. Degradation of 4-hydroxybenzoic acid (4-HBA) by the 4M-enrichment at high salinity (4M). A lag period of ten days observed in the active flasks (■). No degradation was seen in the control flasks (♦). Data is mean of triplicate flasks and bars indicate ± standard deviation.

4.4.3 Detection of catabolic genes. Degenerate primers were used to screen for the presence of ring-opening and ring-cleaving catabolic genes that could be involved in benzoate degradation. PCR products of expected sizes were excised, purified, cloned and sequenced to confirm their identity (Fig 4-4). PCR with *benA* and 1,2-GDO primers amplified PCR products of expected sizes, 800 and 850 bp repsectively, while PCR with 4-HBMO primers amplified a product of expected size between 500 and 600 bp. Similarly, PCR amplification with 1,2-CTD, 2,3-CTD and 3,4-PCA degenerate primers yielded products each of size 400-bp. However, cloning and sequencing confirmed the identity of only 4-HBMO and 3,4-PCA genes. Blast analysis of 4-HBMO gene showed 73% sequence identity to its homolog in *Haloarcula marismortui* ATCC 43049 (GenBank Accession no. AY596297). Similarly, analysis showed that 3,4-PCA gene was 77% identical to a homolog in *Burkholderia gladioli* (GenBank Accession number. U33634).

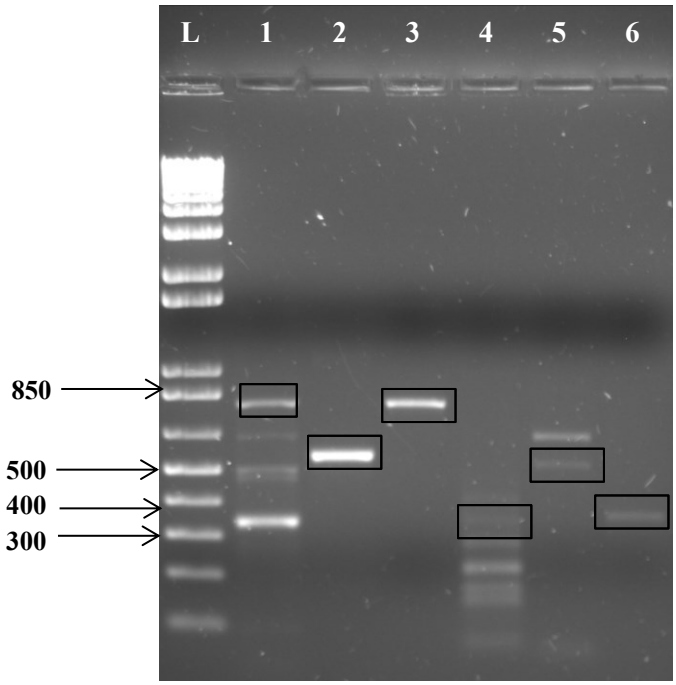

Fig. 4-4. PCR amplification of catabolic genes using primers listed in Table 4-1. Symbols: L (1 kb plus ladder), Lane 1 (*benA*), Lane 2 (4-HBMO), Lane 3 (1,2-GDO), Lane 4 (1,2-CAT), Lane 5 (2,3-CAT), Lane 6 (3,4-PCA). Bands shown in boxes were excised, purified, cloned and sequenced to confirm their identity.

4.4.4 Diversity of archaeal 16S rRNA sequences. To investigate the community structure of the 4-M enrichment, 16S rRNA gene sequences were obtained by 454 pyrosequencing. After implementation of stringent quality control screening process, a total of 15,494 sequences were obtained and included in phylogenetic analysis. The sequences were binned into 196 OTUs at 3% cut off. A BLASTn analysis of OTUs against the 16S Bacterial and Archaeal database in NCBI was performed to assign them to the currently described archaeal genera in the database. This analysis revealed that the culture was highly enriched and consisted solely of members belonging to the family *Halobacteriaceae*. OTUs were assigned to a specific genus based on 95% cut off value. Analysis showed that OTUs belonged to 7 of the currently recognized genera in NCBI database and these include *Halopenitus, Halosarcina, Natronomonas*, *Halosimplex*, *Halorubrum*,

Salinarchaeum and *Haloterrigena* (Fig. 4-5). OTUs showing less than 95% sequence similarity to any of the recognized *Halobacteriaceae* genera were grouped as unclassified genera forming 6.96% of the enrichment.

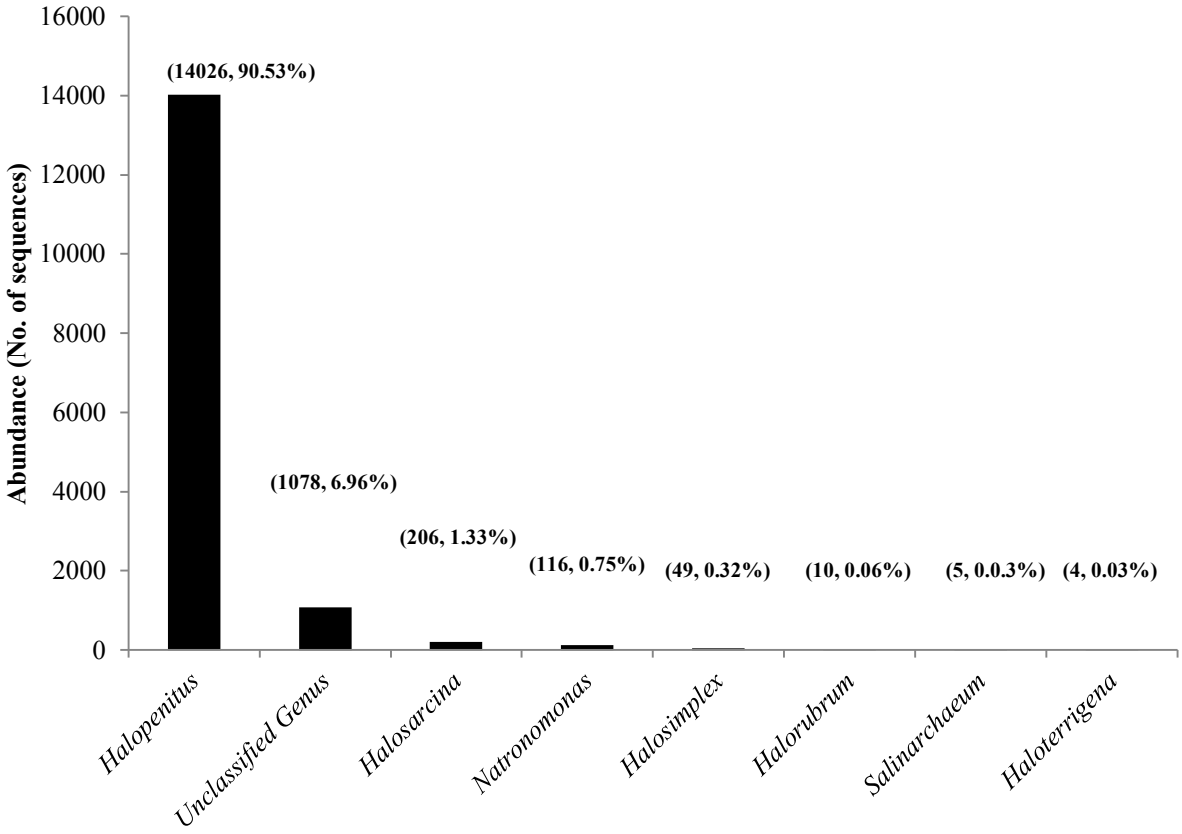

Fig 4-5. Abundance of archaeal 16S rRNA gene sequences obtained from benzoate-degrading enrichment at 4 M NaCl. A total of 15,494 sequences were obtained from pyrosequencing. Among these, *Halopenitus* was identified as the most abundant genus.

Of the total 196 OTUs identified from the enrichment, 88 OTUs belonged to *Halopenitus* making it the most dominant genus of the *Halobacteriaceae* community in the enrichment and accounted for 90.53% (n=14,026 sequences) of the enrichment. *Halopenitus* members showed 95-99% sequence similarity to 16S rRNA gene sequence of an extremely halophilic archaeon, *Halopenitus malekzadehii* strain CC65 (20). Amongst these, OTU 49 (n=7530 sequences) was highly abundant in number compared to others.

In addition to *Halopenitus*, other genera that were identified in the enrichment included *Halosarcina* (1.33%), *Natronomonas* (0.75%), and *Halosimplex* (0.32%). A total of 206 sequences were affiliated to the genus *Halosarcina*. OTU 11 and OTU 89 showed 97% and 95% sequence similarity, respectively to *Halosarcina pallida* strain BZ256T (21). OTUs 21, 51, 78, 81, and 147 showed 95-96% sequence similarity to *Halosarcina limi* strain RO1-6 (22) suggesting that this group may represent novel species of genera *Halosarcina.* Sequences affiliated to the genus *Natronomonas* (n=116 sequences) showed 95-97% sequence similarity to *Natronomonas moolapensis* strain 8.8.11, an extremely halophilic archaeaon isolated recently from a marine solar saltern crystallizer (23). Only, OTU 22 showed 97% sequence similarity whereas OTUs 26, 48, 103, 106 and 139 showed < 97% sequence similarity to *Natronomonas moolapensis* strain 8.8.11 suggesting that they may represent novel species in *Natronomonas* genera. Only 49 of the total sequences were affiliated to genus *Halosimplex* and showed 95-96% sequence similarity to *Halosimplex carlsbadense* type strain 2-9-1 (24).

The genera *Halorubrum, Salinarchaeum* and *Haloterrigena* formed relatively small fractions of the enrichment, 0.06%, 0.03% and 0.03% respectively. OTU 69 (n=10 sequences) was affiliated to genus *Halorubrum* and showed 97% sequence identity to *Halorubrum saccharovorum* (NR_113484). OTU 132 (n=5 sequences) and OTU 109 (n=4 sequences) showed 95% sequence similarity to *Salinarchaeum laminariae* strain JCM 17267 (NR_114364) and *Haloterrigena turkmenica* DSM 5511 (NR_074238), respectively suggesting that they may

represent novel species of their respective genera. Unclassified genera accounted for 6.95% of the enrichment with sequences (n=1078 sequences) showing less than 95% sequence similarity to currently described genera in *Halobacteriaceae* family.

Phylogenetic analysis showed that the OTU representatives of genera *Halopenitus, Halorubrum, Halosarcina, Natronomonas, Haloterrigena, Halosimplex* and *Salinarchaeum* formed groups with known archaeal members from their respective genera. According to BLASTn analysis, OTU 17 belonged to unclassified genera, however phylogenetic analysis

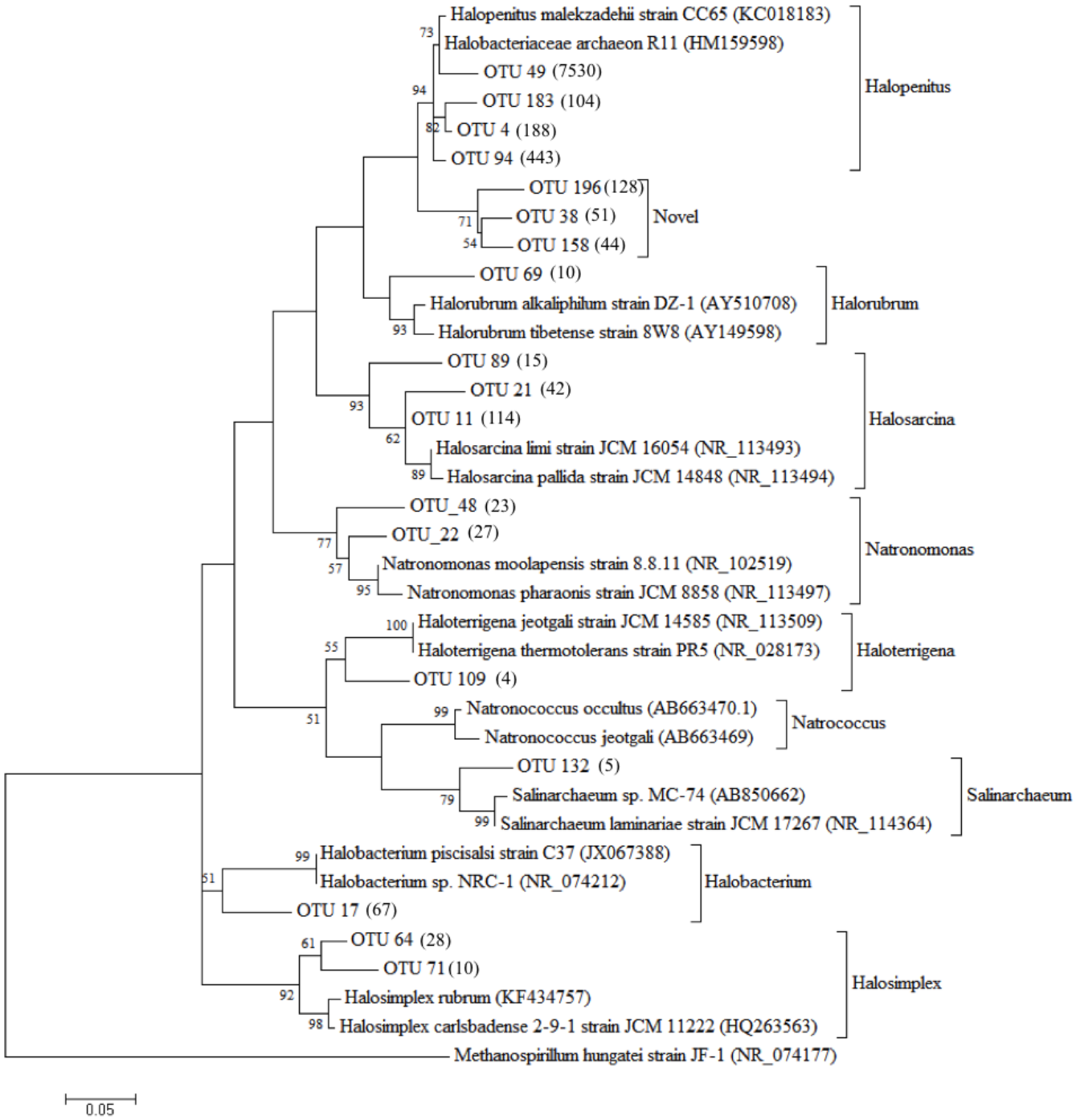

Fig 4-6. Phylogenetic relationship of OTUs obtained from 4M-enrichment with known archaeal members from NCBI database. The tree was inferred by maximum-likelihood method based on the Tamura 3-parameter model with a discrete gamma distribution (+*G*, parameter=0.4782) and invariable site rate ([+*I*], 25.2751%). Bootstrap values (%) are based on 1000 replicates and are shown for branches with more than 50% support. GenBank accession numbers of 16S rRNA sequences are given in parentheses. Number of sequences affiliated to the OTUs is shown in brackets.

4.5 DISCUSSION

Halophilic archaea with the potential to degrade hydrocarbons have become the subject of growing attention. Extremely halophilic archaeal strains belonging to *Haloferax*, *Halococcus*, *Halobacterium*, *Halorubrum*, and *Haloarcula* have been isolated for their ability to degrade crude oil, *n*-alkanes, polyaromatic hydrocarbons and benzoates at high salinities from 20-30% NaCl (4) suggesting that halophilic archaea can be used for bioremediation of hydrocarbon-impacted hypersaline environments. Although a few reports have shown the presence of genes encoding ring-cleaving dioxygenases such as 1,2-GDO, 1,2-CAT and 3,4-PCA either by PCR or measuring their expressions spectrophotometrically (4), no information exists about the genes involved during the initial degradation pathways.

GSL represents one of the largest terminal lakes with 5-30% salinity divided into South and North Arm due to the construction of railroad tracks in the 1960s. The North Arm of the lake is characterized by high levels of salinity (30% w/v) and numerous oil seeps near Rozel Point. Some culture-independent studies surveying bacterial and archaeal communities have shed some light on the biodiversity in GSL (25-28). Only a few reports exist on the potential of halophiles indigenous to GSL to degrade hydrocarbons (9, 10), especially the degradation potential of archaeal communities indigenous to GSL. This study describes the molecular characterization of a benzoate-degrading archaeal enrichment that was developed using sediment samples from Rozel Point at GSL.

Hypersaline environments are constantly subjected to fluctuating salinity. To test the robustness of the enrichment to adapt to fluctuating salinities, benzoate degradation was monitored at varied salt concentrations ranging from 0 to 5 M NaCl. The enrichment was able to degrade benzoate at high salt concentrations ranging from 2 to 5 M NaCl. No degradation was observed at 0 and 1 M NaCl. The rates of degradation also increased with increase in salinity,

optimal degradation was seen at 4 M. These results contradict previous studies that showed a decrease in rate of hydrocarbon degradation with increasing salinity using microorganisms native to GSL (9, 10). The enrichment was also tested for its ability to degrade other oxygenated aromatic compounds such as 4-HBA, PCA, CAT and GA. These compounds were selected since these are typically known intermediates of benzoate degradation (29). Of these, the enrichment was able to utilize only 4-HBA as carbon source and failed to grow on the other tested aromatic compounds.

Multiple pathways for aerobic degradation of benzoate have been identified in non-halophiles. Benzoate can be initially oxidized by monooxygenases or dioxygenases to form central intermediates like CAT, PCA or GA. CAT is cleaved by either *ortho* or *meta*-pathway by 1,2-CAT or 2,3-CAT respectively while, PCA undergoes *ortho*- or *meta*-cleavage by 3,4-PCA and protocatechuate 4,5-dioxygenase to form intermediates that can be assimilated through central metabolism. GA is cleaved by 1,2-GDO to form methylpyruvate that also gets assimilated through central metabolism. To predict the steps of benzoate and 4-HBA metabolism by the enrichment, genes encoding ring-hydroxylating and ring-cleavage enzymes were amplified using degenerate primers either designed in this study or previous studies by other investigators. Our analysis confirmed the presence of only 4-HBMO and 3,4-PCA genes. Although PCR with *benA*, and 1,2-CAT primers amplified products of expected sizes, cloning and sequencing did not confirm the validity of the products. The degenerate primers for *benA*, 1,2-CAT and 2,3-CAT were designed from conserved regions from non-halophilic bacteria. Therefore, some of the primers might not be suitable to amplify catabolic genes in halophilic archaea. Based on these limited findings, it can be hypothesized that benzoate could undergo hydroxylation to form 4-HBA that undergoes further hydroxylation *via* 4-HBMO to form the central intermediate PCA. PCA is further cleaved *via ortho*-pathway catalyzed by 3, 4-PCA to form intermediates that eventually enter the central metabolism. The few reports that exist on hydrocarbon degradation

pathways in halophilic archaea have shown induction of 1, 2-GDO during benzoate or 4-HBA metabolism (7, 30, 31). To our knowledge, the presence of 4-HBMO or 3,4-PCA during benzoate metabolism by halophilic archaea has not been reported previously.

To gain a better insight into the microbial community structure of the 4M-enrichment, pyrosequencing was performed. The culture-independent characterization of the enrichment revealed the predominance of *Halobacteriaceae* family (100%) with OTUs belonging to seven currently identified archaeal genera in NCBI. In addition to the different genera identified in the enrichment, 6.95% of the *Halobacteriaceae* community may be considered as putatively novel genera since the OTUs showed less than 95% 16S rRNA gene sequence similarity to the currently described genera in NCBI. OTUs affiliated to genus *Halopenitus* represented 90.53% of the *Halobacteriaceae* community forming the most dominant genus of the 4M-enrichment. *Halopenitus* is a newly described genus with only two studies reporting the isolation and taxonomic characterization of isolates; *Halopenitus malekzadehii* strain CC65 (20) and *Halopenitus persicus* strain DC30 from Aran-Bidgol hypersaline lake in Iran (32). Both strains can grow at 10-30% NaCl and utilize carbohydrates as the sole carbon sources (20, 32). However, no reports exist so far on the metabolic potential of *Halopenitus* to utilize hydrocarbons as the carbon sources.

Also identified in the enrichment were OTUs belonging to genera *Halosarcina* and *Natronomonas* representing 1.33% and 0.75% of the archaeal community. OTUs affiliated to *Halosarcina* showed 16S rRNA gene sequence similarity to *Halosarcina pallida* strain BZ256 isolated form Zodletone Spring, OK (21) and *Halosarcina limi* strain JCM 16054 isolated from a marine solar saltern in eastern China (33). OTUs affiliated to *Natronomonas* showed similarity to *Natronomonas pharaonis* strain 8.8.11, an extremely halophilic archaeon isolated from marine solar saltern crystallizer in southeastern Australia and requiring 18-20% NaCl for optimum growth (23). Recently, Youssef et al. also identified members of *Halosarcina* and *Natronomonas*

in samples obtained from Rozel Point at GSL. Using primers designed specific to the order *Halobacteriales*, the investigators showed that *Halosarcina* represented 0.67% while *Natronomonas* represented 12.8% of the *Halobacteriales* community *Natronomonas* was identified as one of the abundant genera in a heavily contaminated saline-alkali soil from Dagang Oilfied in China by Wang et al. The authors suggested that members belonging to the genera *Haloferax* and *Natronomonas a*re likely to play role the degradation of hydrocarbons in the saline-alkali soil (34).

OTUs affiliated to genera *Halosimplex, Halorubrum, Salinarchaeum and Haloterrigena* were also identified in the 4M-enrichment but at relatively low abundances at 0.32%, 0.06%, 0.03% and 0.03% respectively. Youssef at al. also identified members of *Halorubrum, Halosimplex, and Haloterrigena* from GSL samples however, at different abundances of 24.1%, 0.1% and 0.01%, respectively.

Some of the OTUs belonging to their respective genera identified in this study were chosen for phylogenetic analysis. OTUs belonging to *Halopenitus, Halorubrum, Natronomonas*, *Halosimplex, Halorubrum, Salinarchaeum and Haloterrigena* formed clusters with known archaeal members from the genera. The analysis also indicated that a few OTUs (158, 196, 36) that were not affiliated with known archaeal genera of *Halobacteriaceae* grouped together as putatively novel genera.

Overall, this study demonstrates the potential of highly enriched archaeal enrichment developed to degrade benzoate over a wide range of salinity. Rates of benzoate degradation increased with increase in salinity in contrast to the results observed in degradation studies at GSL (9, 10). The 4M-enrichment is entirely comprised of members of the *Halobacteriaceae* family with *Halopenitus* as the dominant genus. Phylogenetic analysis also revealed that close to 7% of the abundant OTUs belonged to putatively novel genera. Also identified in this study were

4-HBMO and 3,4-PCA genes suggesting that benzoate might be converted to 4-HBA which in turn is converted to PCA via 4-HBMO. PCA then undergoes *ortho* ring-cleavage *via* 3,4-PCA to form intermediates that will be assimilated *via* central metabolism. However, further studies identifying the intermediates of benzoate degradation by GC-MS can confirm the proposed pathway. Although, members of *Halopenitus* genus were identified as the most dominant in the enrichment, their degradation ability cannot be confirmed. Therefore, further efforts focusing on isolation of pure cultures from this group from enrichment are needed to determine the degradation capacity of *Halopenitus* members.

4.6 REFERENCES:

1. **Grant WD, Kamekura M, McGenity TJ, and Ventosa A.** 2001. Class III Halobacteria class. nov. Bergey's Manual of Systematic Bacteriology. **1:**294 - 334.
2. **Andrei AS, Banciu HL, and Oren A.** 2012. Living with salt: metabolic and phylogenetic diversity of archaea inhabiting saline ecosystems. FEMS Microbiol Lett. **330:**1-9.
3. **Oren A.** 2002. Halophilic Microorganisms and their Environments. International Microbiology, Volume 6, Issue 2, pp 151-152, Springer-Verlag.
4. **Fathepure BZ.** 2014. Recent studies in microbial degradation of petroleum hydrocarbons in hypersaline environments. Frontiers in microbiology. **5:**173. doi: 10.3389/fmicb.2014.00173.
5. **Cuadros-Orellana S, Pohlschröder M, and Durrant LR.** 2006. Isolation and characterization of halophilic archaea able to grow in aromatic compounds. Int Biodeterior Biodegradation. **57:**151-154.

6. **Fu W, and Oriel P.** 1999. Degradation of 3-phenylpropionic acid by *Haloferax* sp. D1227. Extremophiles. **3:**45-53.

7. **Fairley DJ, Boyd DR, Sharma ND, Allen CCR, Morgan P, and Larkin MJ.** 2002. Aerobic metabolism of 4-hydroxybenzoic acid in Archaea via an unusual pathway involving an intramolecular migration (NIH shift). Appl Environ Microbiol. **68:**6246-6255.

8. **Erdoğmuş S, Mutlu B, Korcan S, Güven K, and Konuk M.** 2013. Aromatic hydrocarbon degradation by halophilic archaea isolated from Çamaltı Saltern, Turkey. Water, Air, Soil Pollut. **224**:1-9.

9. **Ward DM, and Brock TD.** 1978. Hydrocarbon biodegradation in hypersaline environments. Appl Environ Microbiol. **35:**353-359.

10. **Azetsu SF, B.Z.** December 2009. Biodegradation of BTEX at high salinity by an enrichment culture from hypersaline sediments of Rozel Point at Great Salt Lake. J Appl Microbiol. **107:**pp. 2001-2008(2008).

11. **Nicholson CA, and Fathepure BZ.** 2004. Biodegradation of benzene by halophilic and halotolerant bacteria under aerobic conditions. Appl Environ Microbiol. **70:**1222-1225.

12. **DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, and Andersen GL.** 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. **72:**5069-5072.

13. **Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ.** 1990. Basic local alignment search tool. Journal of molecular biology. **215:**403-410.

14. **Tamura K, Stecher G, Peterson D, Filipski A, and Kumar S.** 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. **30:**2725-2729.

15. **Tamura K.** 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. Molecular Biology and Evolution. **9:**678-687.

16. **Felsenstein J.** 1985. Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. Evolution. **39:**783-791.

17. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S.** 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol. **28:**2731-2739.

18. **Morimoto SHO, Togami K, Ogawa N, Hasebe A, and Fujii T.** 2005. Analysis of a bacterial community in 3-chlorobenzoate-contaminated Soil by PCR-DGGE targeting the 16S rRNA gene and benzoate 1,2-dioxygenase gene (benA). Microbes Environ. **20:**151-159.

19. **Garcia MT, Ventosa A, and Mellado E.** 2005. Catabolic versatility of aromatic compound-degrading halophilic bacteria. FEMS Microbiol Ecol. **54:**97-109.

20. **Amoozegar MA, Makhdoumi-Kakhki A, Mehrshad M, Fazeli SA, and Ventosa A.** 2013. *Halopenitus malekzadehii* sp. nov., an extremely halophilic archaeon isolated from a salt lake. Int J Syst Evol Microbiol. **63:**3232-3236.

21. **Savage KN, Krumholz LR, Oren A, and Elshahed MS.** 2008. *Halosarcina pallida* gen. nov., sp. nov., a halophilic archaeon isolated from a low-salt, sulfide-rich spring. Int J Syst Evol Microbiol. **58**: 856-860.

22. **Cui H-L, Gao X, Li X-Y, Xu X-W, Zhou Y-G, Liu H-C, and Zhou P-J.** 2010. *Halosarcina limi* sp. nov., a halophilic archaeon from a marine solar saltern, and emended description of the genus Halosarcina. Int J Syst Evol Microbiol. **60:**2462-3466.

23. **Burns DG, Janssen PH, Itoh T, Minegishi H, Usami R, Kamekura M, and Dyall-Smith ML.** 2010. *Natronomonas moolapensis* sp. nov., non-alkaliphilic isolates

recovered from a solar saltern crystallizer pond, and emended description of the genus Natronomonas. Int J Syst Evol Microbiol. **60:**1173-1176.

24. **Vreeland RH, Straight S, Krammes J, Dougherty K, Rosenzweig WD, and Kamekura M.** 2002. *Halosimplex carlsbadense* gen. nov., sp. nov., a unique halophilic archaeon, with three 16S rRNA genes, that grows only in defined medium with glycerol and acetate or pyruvate. Extremophiles. **6:**445-452.

25. **Baxter B, Litchfield C, Sowers K, Griffith J, Dassarma P, and Dassarma S.** 2005. Microbial diversity of Great Salt Lake. p. 9-25. *In* N. Gunde-Cimerman, A. Oren, and A. Plemenitaš (ed.), Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya, vol. 9. Springer Netherlands.

26. **Meuser JE, Baxter BK, Spear JR, Peters JW, Posewitz MC, and Boyd ES.** 2013. Contrasting patterns of community assembly in the stratified water column of Great Salt Lake, Utah. Microb Ecol. **66:**268-280.

27. **Parnell JJ, Rompato G, Crowl TA, Weimer BC, and Pfrender ME.** 2011. Phylogenetic distance in Great Salt Lake microbial communities. Aquat Microb Ecol. **64:**267-273.

28. **Tazi L, Breakwell DP, Harker AR, and Crandall KA.** 2014. Life in extreme environments: microbial diversity in Great Salt Lake, Utah. Extremophiles. **18:**525-535.

29. **Chen DW, Zhang Y, Jiang CY, and Liu SJ.** 2014. Benzoate metabolism intermediate benzoyl coenzyme A affects gentisate pathway regulation in *Comamonas testosteroni*. Appl Environ Microbiol. **80:**4051-4062.

30. **Fairley DJ, Wang G, Rensing C, Pepper IL, and Larkin MJ.** 2006. Expression of gentisate 1,2-dioxygenase (gdoA) genes involved in aromatic degradation in two haloarchaeal genera. Appl Microbiol Biotechnol. **73:**691-695.

31. **Fu W, and Oriel P.** 1998. Gentisate 1,2-dioxygenase from *Haloferax* sp. D1227. Extremophiles. **2:**439-446.

32. **Amoozegar MA, Makhdoumi-Kakhki A, Shahzadeh Fazeli SA, Azarbaijani R, and Ventosa A.** 2012. *Halopenitus persicus* gen. nov., sp. nov., an archaeon from an inland salt lake. Int J Syst Evol Microbiol. **62:**1932-1936.

33. **Minegishi H, Kamekura M, Kitajima-Ihara T, Nakasone K, Echigo A, Shimane Y, Usami R, Itoh T, and Ihara K.** 2012. Gene orders in the upstream of 16S rRNA genes divide genera of the family Halobacteriaceae into two groups. Int J Syst Evol Microbiol. **62:**188-195.

34. **Wang X, Han Z, Bai Z, Tang J, Ma A, He J, and Zhuang G.** 2011. Archaeal community structure along a gradient of petroleum contamination in saline-alkali soil. J Environ Sci (China). **23:**1858-1864.

CHAPTER V

SUMMARY

In the last two decades, there have been many reports in the literature exploring the ability of halophilic and halotolerant microorganisms to degrade petroleum compounds in moderate to high salt conditions. However, the same cannot be said about the genetics and mechanisms of hydrocarbon degradation at high salinity. To design new strategies for effective bioremediation and biotreatment of contaminated hypersaline environments and saline industrial wastewaters using halophilic microorganisms, knowledge of their metabolic capacities and hydrocarbon degradation pathways is necessary.

This work provides insights into the metabolic potential of novel halophiles, *Arhodomonas* sp. strain Seminole and *Arhodomonas* sp. strain Rozel. A combination of –omics approach was used to identify genes and to elucidate hydrocarbon degradation pathways in both the strains (described in Chapter II and Chapter III). In Chapter II, using strain Seminole's genome as a guide, initial steps in benzene degradation pathway were predicted in strain Rozel. Phenol hydroxylase and catechol 2,3-dioxygenase (2,3-CAT) were identified as major enzymes involved in the initial steps using 2D-gel electrophoresis and LC-MS/MS. Further verification of the data was done by detection of phenol as an initial intermediate of benzene degradation using GC-MS. These findings suggested that the multicomponent phenol hydroxylase converts benzene

to phenol and then to catechol in two steps of oxidation, which undergoes *meta*-cleavage *via* 2,3-CAT to form central metabolism intermediates. In Chapter III, *in silico* analysis of the draft genome of strain Seminole revealed the presence of complete degradation pathways for benzene, toluene, 4-hydroxybenzoic acid (4-HBA) and phenylacetic acid (PAA). The proteomics data complemented the predicted pathways and enabled us to identify enzymes in cytosolic proteomes of hydrocarbon-grown cells using LC-MS/MS. Majority (>50%) of the key enzymes involved in the degradation steps were highly expressed (total spectral counts) when grown on hydrocarbons compared to lactate-grown cells. Benzene and toluene were degraded *via* the same set of enzymes. These included the multicomponent phenol hydroxylase and enzymes of catechol *meta*-cleavage pathway. Analysis revealed that 4-HBA degradation proceeds *via* protocatechuate pathway and this pathway has been shown to present in many non-halophiles. Protocatechuate is a known central intermediate in degradation of many man-made chemicals as well as lignin related compounds. Analysis also indicated that PAA degradation is initiated by a hybrid pathway (anaerobic and aerobic) where PAA undergoes initial activation by CoA ligase to form PA-CoA which is further undergoes ring-hydroxylation and hydrolytic ring-cleavage to form aliphatic compounds that are channeled through β-oxidation like steps to form acetyl-CoA and succinyl-CoA. Overall, the work described in Chapter II and III displays the metabolic versatility of *Arhodomonas* spp. and provides a comprehensive picture of the various hydrocarbon degradation pathways in high salt conditions.

Our data clearly suggests that the pathways and enzymes of hydrocarbon degradation in *Arhodomonas* spp. were similar to those described in non-halophiles. Although halophilic proteins perform identical functions as their non-halophilic homologs, they differ from their non-halophilic homologs by maintaining stability and activity at high salinity. Halophilic proteins bind significant amounts of salt and water in contrast to non-halophilic proteins that bind only

water and not salt (1, 2). At protein level, most halophiles are characterized by low hydrophobicity of the protein core, excess negative charge on the protein surface, and low lysine content (3). Excess negative charge on halophilic proteins is reflected in the low pI values of their proteins which are mostly due to high content of acidic residues on the protein surface (1). A similar observation was made in the hydrocarbon-degrading proteins identified in *Arhodomonas* sp. strain Seminole. These proteins displayed low-pI values (theoretical) compared to the non-halophilic homologs. Low pI is mostly due to high content of acidic residues like glutamate and aspartate on the surface of the proteins that forms a hydration shell to protect the enzyme from aggregation under high salinity (3, 4). These unique features enable halophilic proteins to function efficiently under extremely saline conditions where non-halophilic proteins lose their activity (3).

This work also focuses on the potential of archaeal enrichment to degrade aromatic compounds at 4 M NaCl (Chapter IV). Most halophilic archaea show optimum degradation of hydrocarbons at extreme salinity (>26% NaCl) when compared to halophilic bacteria (5). Similar observation was made in this work, where rate of benzoate degradation increased with increase in salt concentration and highest rates were measured at 4 and 5 M NaCl. Pyrosequencing was used to determine the diversity and community structure of the benzoate degrading community. Analysis revealed that the archaeal community was entirely composed of *Halobacteriaceae* members with majority (>90%) belonging to *Halopenitus* genus thus suggesting development of a highly enriched culture. In addition, the community also consisted of a significant proportion of unidentified putatively novel archaeal genera. Degenerate primers revealed the presence of ring-hydroxylating monooxygenase; 4-hydroxybenzoate 3-monooxygenase and ring-cleaving dioxygenase; protocatechuate 3,4-dioxygenase suggesting that perhaps the archaeal enrichment degrades benzoate *via* protocatechuate pathway.

Further research using deletion mutant experiments as well as detection of key pathway intermediates using GC-MS will help provide decisive proof of the proposed hydrocarbon degradation pathways in *Arhodomonas* spp. Also, the potential of *Halopenitus*, the most dominant genus found in the benzoate degrading archaeal community needs to be evaluated by isolation and characterization of a *Halopenitus* sp because to date there are no reports on the ability of these organisms to metabolize hydrocarbons in high salinity environments.

In summary (see Chapter I), several studies report on the ability of halophilic bacteria and archaea to degrade a variety of petroleum compounds under aerobic conditions. On the other hand, not much is known about anaerobic degradation of hydrocarbons under hypersaline conditions. This is important because, solubility of oxygen decreases as salinity increases thereby limiting the availability of oxygen in hypersaline brines (6, 7). Hence it is critical to explore the fate of anaerobic hydrocarbon degradation under hypersaline conditions. Isolation of pure cultures of anaerobic halophiles will help in understanding their roles during the hydrocarbon degradation process. Genomic, proteomic, GC-MS analyses will assist in identifying novel genes, pathways and intermediates of anaerobic degradation under hypersaline conditions. The knowledge generated by such studies will accelerate the development of bioremediation technologies for clean-up of petroleum-impacted hypersaline areas.

REFERENCES:

1. **Fukuchi S, Yoshimune K, Wakayama M, Moriguchi M, and Nishikawa K.** 2003. Unique amino acid composition of proteins in halophilic bacteria. J Mol Biol. **327:**347-357.

2. **Siglioccolo A, Paiardini A, Piscitelli M, and Pascarella S.** 2011. Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. BMC Structural Biol. **11:**50.

3. **Mevarech M, Frolow F, and Gloss LM.** 2000. Halophilic enzymes: proteins with a grain of salt. Biophys Chem. **86:**155-164.

4. **Reed CJ, Lewis H, Trejo E, Winston V, and Evilia C.** 2013. Protein adaptations in archaeal extremophiles. Archaea. **2013:**373275.

5. **Fathepure BZ.** 2014. Recent studies in microbial degradation of petroleum hydrocarbons in hypersaline environments. Front Microbiol. **5**:173.Doi:10.3389/fmicb.2014.00173.

6. **Martins LF, and Peixoto RS.** 2012. Biodegradation of petroleum hydrocarbons in hypersaline environments. Braz J Microbiol. **43:**865-872.

7. **McGenity TJ.** 2010. Halophilic Hydrocarbon Degraders. p. 1939-1951. *In* K. N. Timmis (ed.), Handbook of Hydrocarbon and Lipid Microbiology. Springer-Verlag, Berlin Heidelberg.

VITA

Sonal Dalvi

Candidate for the Degree of

Doctor of Philosophy

Thesis: MOLECULAR CHARACTERIZATION OF AROMATIC HYDROCARBON DEGRADATION IN EXTREME HALOPHILES

Major Field: Microbiology and Molecular Genetics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy/Education in Microbiology and Molecular Genetics at Oklahoma State University, Stillwater, Oklahoma in July, 2014.

Completed the requirements for the Master of Science in Botany at University of Mumbai, Mumbai, Maharashtra, India in 2006.

Completed the requirements for the Bachelor of Science in Botany at University of Mumbai, Mumbai, Maharashtra, India in 2004.

Experience: Teaching and Research Assistant at Department of Microbiology and Molecular Genetics at Oklahoma State University, Stillwater, OK.

Professional Memberships: American Society of Microbiology.