Study of Genes Relating To Degradation of Aromatic Compounds and Carbon Metabolism in Mycobacterium Sp. Strain KMS

Chun Zhang
Utah State University

Follow this and additional works at: http://digitalcommons.usu.edu/etd

Part of the Biology Commons, Environmental Sciences Commons, and the Microbiology Commons

Recommended Citation
STUDY OF GENES RELATING TO DEGRADATION OF AROMATIC
COMPOUNDS AND CARBON METABOLISM IN MYCOBACTERIUM SP. STRAIN
KMS

by

Chun Zhang

A dissertation submitted in partial fulfillment
of the requirements for the degree
of
DOCTOR OF PHILOSOPHY
in
Biology

Approved:

____________________________  _____________________ _______
Anne J. Anderson, Ph.D.    Dennis L. Welker, Ph.D.
Major Professor      Committee Member

____________________________  _____________________ _______
Jeanette M. Norton, Ph.D.    Ronald C. Sims, Ph.D.
Committee Member     Committee Member

____________________________   ____________________ ________
Charles D. Miller, Ph.D.    Mark R. McLellan, Ph.D.
Committee Member     Vice President for Research and
Dean of the School of
Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2013
ABSTRACT

Study of Genes Relating to Degradation of Aromatic Compounds and Carbon Metabolism in *Mycobacterium sp.* Strain KMS

by

Chun Zhang, Doctor of Philosophy

Utah State University, 2013

Major Professor: Dr. Anne J. Anderson
Department: Biology

Polycyclic aromatic hydrocarbons, produced by anthropological and natural activities, are hazardous through formation of oxidative radicals and DNA adducts. Growth of *Mycobacterium sp.* strain KMS, isolated from a contaminated soil, on the model hydrocarbon pyrene induced specific proteins. My work extends the study of isolate KMS to the gene level to understand the pathways and regulation of pyrene utilization.

Genes encoding pyrene-induced proteins were clustered on a 72 kb section on the KMS chromosome but some also were duplicated on plasmids. Skewed GC content and presence of integrase and transposase genes suggested horizontal transfer of pyrene-degrading gene islands that also were found with high conservation in five other pyrene-degrading *Mycobacterium* isolates. Transcript analysis found both plasmid and chromosomal genes were induced by pyrene. These processes may enhance the survival of KMS in hydrocarbon-contaminated soils when other carbon sources are limited.
KMS also grew on benzoate, confirming the functionality of an operon containing genes distinct from those in other benzoate-degrading bacteria. Growth on benzoate but not on pyrene induced a gene, *benA*, encoding a benzoate dioxygenase α-subunit, but not the pyrene-induced *nidA* encoding a pyrene dioxygenase α-subunit; the differential induction correlated with differences in promoter sequences.

Diauxic growth occurred when pyrene cultures were amended with benzoate or acetate, succinate, or fructose, and paralleled delayed expression of *nidA*. Single phase growth and normal expression of *benA* was observed for benzoate single and mixed cultures. The *nidA* promoters had potential cAMP-CRP binding sites, suggesting that cAMP could be involved in carbon repression of pyrene metabolism.

Growth on benzoate and pyrene requires gluconeogenesis. Intermediary metabolism in isolate KMS involves expression from genes encoding a novel malate:quinone oxidoreductase and glyoxylate shunt enzymes. Generation of C3 structures involves transcription of genes encoding malic enzyme, phosphoenolpyruvate carboxykinase, and phosphoenolpyruvate synthase. Carbon source modified the transcription patterns for these genes.

My findings are the first to show duplication of pyrene-degrading genes on the chromosome and plasmids in *Mycobacterium* isolates and expression from a unique benzoate-degrading operon. I clarified the routes for intermediary metabolism leading to gluconeogenesis and established a potential role for cAMP-mediated catabolite repression of pyrene utilization.
Polycyclic aromatic hydrocarbons are produced from incomplete combustion of organic materials by human or natural activities. These polycyclic aromatic hydrocarbons are classified as pollutants because of their toxic, mutagenic, and carcinogenic characteristics. *Mycobacterium sp.* strain KMS, isolated from a contaminated soil, grows on the model polycyclic aromatic hydrocarbon, pyrene, with its degradation to water and carbon dioxide. This study locates genes on the chromosome and plasmids of isolate KMS relating to pyrene degradation, elucidates the influence of other carbon sources available in the habitats of isolate KMS on degradation of pyrene, and deduces possible metabolic pathways used by isolate KMS for its survival.

Pyrene-degrading genes are clustered on the KMS chromosome. Duplication of some of the genes may help the degradation of pyrene by KMS in contaminated soils. Pyrene-degrading genes are clustered and are in similar positions in the genome of six pyrene-degrading *Mycobacterium* isolates. Those genes were possibly acquired from a common ancestor.

Isolate KMS also grows on a simple aromatic compound, benzoate. The presence of benzoate does not help pyrene degradation by inducing the pyrene-degrading genes. Also the presence of carbon sources typical of those in the rhizosphere, such as sugars, organic acids, and benzoate, delayed utilization of pyrene until these carbon sources were utilized. Benzoate did not repress growth on the sugars or organic acids.

Analysis of gene transcription and enzyme activities revealed that isolate KMS had novel pathways for interconversion of carbon compounds required for effective cell growth, especially for the structures required for the thick cell wall of the mycobacteria.
ACKNOWLEDGMENTS

I would like to appreciate my major advisor, Dr. Anne J. Anderson, for your guidance, help, support, and inspiration on my study. When I began my study in Utah State University, I had few experience on how to make presentation and how to write a logical paper in English. Spending lots of efforts and time on my study, Dr. Anderson trained me how to think, how to organize the sentences in the writing and how to prove my ideas in scientific researches. I’ll memorize your kindly help.

I am appreciative to my committee members, Dr. Charles D. Miller, Dr. Dennis L. Welker, Dr. Jeanette M. Norton, and Dr. Ronald C. Sims, for discussing experiments, reviewing papers, enlightening ideas, supporting my study, and using the machines. I will remember your kindly help.

I will thank Dr. Jon Takemoto, Dr. Brad Kropp, Dr. Joseph Li, Dr. John Stark, Dr. Andy Anderson, Dr. Paul Cliften, Dr. Dane Hansen, Dr. Alan Savitzky, Dr. Edmund Brodie, Dr. Marie Walsh, Dr. Jeff Broadbent, Dr. Yajun Wu, Dr. Judith Sims, Dr. David Britt, Dr. Ningling Yin, Nancy Kay Harrison, Kami Mcneil, and Brian Joy for using the machines in my research and for your advice and help in my research.

I will also thank the people in my lab and students from the Department of Biology and Biological Engineering: Dr. Young Choel Kim, Dr. Christian Dimkpa, Dr. Songhee Han, Brian Pettee, Dr. Tian Yu, Dr. Sitaram Harihar, Dr. Yun Peng, Dr. Pin Liu, Dr. Junling Huo, Yukie Kawasaki, Sanjib Shrestha, Angela Cefalo, Han Xu, and Alyssa Calder for your help in my research. Thanks are also given to the Department of Biology, Utah State University, and the nice people in Logan, Utah.
I acknowledge the Utah Water Research Laboratory for partial financial support for my research.

Finally, I would like to thank my parents, Qingwei Zhang and Mengling Gao, for your encouragement and love as well as my friends Han Feng, Dong Wang, Hong Lv, Yi Ren, Dongliang Wang, Danning Huang, Jiahua Cao, Hui Wang, Yongjia Wang, Lei Quan, Wang Tian, Dongli Wang, Ji Qi, Hao Yu, Hao Feng, Yue Li, Lin Yang, Gary Chiu, Frankie Chiu, Fluorence Chiu, Wenbin Yu, Mingbo Li, Guanchiang Chen, Xin Dai, Chet Lu, Siewsun Wong, Mai Li, Lei Ye, Yongjun Tang, Dong Chen, Zhonghua Xu, and Ce Wang for your concern and friendship.

Chun Zhang
# CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>iii</td>
</tr>
<tr>
<td>PUBLIC ABSTRACT</td>
<td></td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td></td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td></td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td></td>
<td>xvii</td>
</tr>
</tbody>
</table>

## CHAPTER

### 1 INTRODUCTION

Outline of dissertation.................................................................................................1
Chemical structures and origins of aromatic hydrocarbons.........................................4
Biological risks from aromatic hydrocarbons...............................................................7
Biodegradation of PAHs by microbes in soils.................................................................8
*Mycobacterium sp.* strain KMS......................................................................................12
Pyrene-degradation mechanisms in mycobacteria..........................................................15

Summary of pyrene-degradation pathways in isolates KMS and PYR-1................................15
Components of ring-hydroxylating dioxygenases............................................................20
Organization of pyrene-degrading genes in *Mycobacterium* isolates..............................22
Pyrene-induced proteins in *Mycobacterium sp.* strain KMS.........................................26

Degradation of MAHs and MAH derivatives by *Mycobacterium* isolates.........................26
Key pathways for carbon metabolism in *Mycobacterium* isolates..................................30
Alternative habitats for environmental *Mycobacterium* isolates..................................37
Catabolite repression in *Mycobacterium*......................................................................40
Objectives to be studied.................................................................................................42
References......................................................................................................................43

### 2 POLYCYCLIC AROMATIC HYDROCARBON-DEGRADING GENE ISLANDS IN FIVE PYRENE-DEGRADING *Mycobacterium* ISOLATES FROM DIFFERENT GEOGRAPHIC LOCATIONS

Abstract.........................................................................................................................63
4 UTILIZATION OF PYRENE AND BENZOATE IN *MYCOBACTERIUM* ISOLATE KMS IS REGULATED DIFFERENTIALLY BY CATABOLIC REPRESSION

Introduction.............................................................................150
Materials and Methods.................................................................154
Bacterial strain and cultivation conditions.................................154
Measurement of fructose concentration when KMS grown on acetate or succinate plus fructose as carbon sources..................155
Primer design for PCR amplification.............................................155
RNA isolation and end-point reverse transcriptase PCR..............156
Bioinformatic prediction of potential cAMP-binding site and search of genes relating to cyclic AMP metabolism and genes encoding cyclic AMP receptor proteins..........................................................158
Results......................................................................................159

Growth of *Mycobacterium* *sp.* isolate KMS on different carbon sources..................................................................159
Expression from benA gene and nidA genes under different growth conditions.........................................................163
Potential cAMP-binding site with nidA genes and benA gene........167
Potential genes encoding adenylate cyclases, phosphodiesterases, and cyclic AMP receptor proteins............................168
Discussion..................................................................................170
References..................................................................................174

5 THE GLUCONEOGENIC PATHWAY IN A SOIL *MYCOBACTERIUM* ISOLATE WITH BIOREMEDITION ABILITY

Abstract.....................................................................................181
Introduction..................................................................................182
Materials and Methods................................................................186
Bacterial strain and cultivation conditions …………………………………… 186
Identification of target genes through bioinformatic analysis …………. 187
RNA isolation and end-point reverse transcriptase PCR ……………… 187
Preparation of cell membranes and cell extracts ……………………..…. 190
Assay of PEP synthase and malate:quinone oxidoreductase ……… 190

Results …………………………………………………………………………. 192

Growth of isolate KMS on defined carbon sources ……………….… 192
Genes involved in generation and transformation of malate, OAA, pyruvate, and PEP …………………………………………………… 192
Transcript analyses from selected genes ………………………… ……… 196
Activities of malate:quinone oxidoreductase (MQO) and PEP synthase (PEPS)………………………………………………………… 198

Discussion ……………………………………………………………………… 201
Reference …………………………………………………………………… 205

6 SUMMARY AND FUTURE DIRECTIONS ………………………………… 210

Summary ……………………………………………………………………… 210

Pyrene-degrading gene islands in Mycobacterium isolates:
gene duplication …………………………………………………………… 210
Expression from distinct operons containing divergent dioxygenase
genes occurs for the utilization of benzoate and pyrene in isolate KMS……………………………………………………………………… 213
Absence of other available carbon sources promotes pyrene
degradation by isolate KMS ……………………………………………… 220
Intermediary carbon shuffling in isolate KMS is efficient ………….… 224

Future studies ……………………………………………………………………… 228
References …………………………………………………………………… 230

APPENDICES …………………………………………………………………… 235

CURRICULUM VITAE …………………………………………………………… 264
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Summary of sources of PAHs-degrading <em>Mycobacterium</em> sp. strains and the PAHs degraded by each isolate.</td>
</tr>
<tr>
<td>1.2</td>
<td>Overview of the genomes of six PAH-degrading mycobacterium isolates and another environmental isolate <em>Mycobacterium smegmatis</em> MC² 155</td>
</tr>
<tr>
<td>1.3</td>
<td>Identification of seventeen proteins <em>de novo</em> or highly induced by pyrene in <em>Mycobacterium</em> isolate KMS.</td>
</tr>
<tr>
<td>1.4</td>
<td>Twelve key metabolites needed for biosynthesis of all cell components.</td>
</tr>
<tr>
<td>2.1</td>
<td>Specific primers for genes potentially encoding the pyrene-induced proteins and the 16S rRNA genes for strain KMS.</td>
</tr>
<tr>
<td>2.2</td>
<td>Nucleotide identity between locus tags of genes potentially encoding the proteins accumulating in <em>Mycobacterium sp.</em> strain KMS during pyrene degradation.</td>
</tr>
<tr>
<td>2.3</td>
<td>Transcript accumulation for 25 chromosomal and plasmid genes potentially encoding 17 pyrene-induced proteins from isolate KMS.</td>
</tr>
<tr>
<td>2.4</td>
<td>Synteny in gene order between five PAH-degrading <em>Mycobacterium</em> isolates from three different sites in the USA.</td>
</tr>
<tr>
<td>3.1</td>
<td>Specific primers for genes potentially encoding the subunits of dioxygenases and the 16S rRNA genes for strain KMS.</td>
</tr>
<tr>
<td>3.2</td>
<td>Distribution of genes encoding dioxygenase α- and β-subunits in the genome of KMS.</td>
</tr>
<tr>
<td>3.3</td>
<td>The homologies of all the chromosomal ring-hydroxylating dioxygenases β-subunit genes with chromosomal nidB gene in <em>Mycobacterium</em> isolate KMS.</td>
</tr>
<tr>
<td>3.4</td>
<td>The homologies of all the chromosomal ring-hydroxylating dioxygenases α-subunit genes with chromosomal nidA gene in <em>Mycobacterium</em> isolate KMS.</td>
</tr>
<tr>
<td>3.5</td>
<td>The homologies of the lone ring-hydroxylating dioxygenases α-subunit genes of <em>Mycobacterium</em> isolate KMS.</td>
</tr>
</tbody>
</table>
3.6 The homologies of the replicated ring-hydroxylating dioxygenases α- and β-subunit genes of *Mycobacterium* isolate KMS on the circular and linear plasmids to the chromosomal genes.................................129

3.7 The fold change in expression for genes encoding the β-subunit of dioxygenase normalized by the expression of 16S rRNA genes with growth of isolate KMS with or without pyrene.................................136

4.1 Specific primers for genes potentially encoding the duplicated pyrene-hydroxylating dioxygenase α-subunit genes *nidA*, benzoate dioxygenase α-subunit gene *benA* and the 16S rRNA genes for strain KMS.............157

4.2 Variable growth potential of *Mycobacterium* sp. KMS on acetate, succinate, fructose, benzoate, or pyrene as sole carbon sources in PBM medium..........161

4.3 Genes in KMS potentially encoding adenylate cyclase for cAMP generation, cAMP binding regulatory proteins and phosphodiesterase for cAMP degradation.................................................................170

5.1 Specific primers used for PCR in the transcript analysis for strain KMS during growth on different carbon sources.............................................................188

5.2 Variable growth potential of *Mycobacterium* sp. KMS on acetate, succinate, fructose, mannitol, benzoate, or pyrene as sole carbon sources in PBM medium.................................................................194

5.3 Comparison of chosen metabolism genes and gene products with the identified genes and their products in other bacteria.................................195

6.1 Potential cAMP-CRP binding sites found in the promoter regions of two operons containing *nidB*-nidA genes and two operons containing *nidB2* genes.........222
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Chemical structures of some common MAHs and MAH derivatives</td>
<td>5</td>
</tr>
<tr>
<td>1.2</td>
<td>Chemical structures of some common PAHs</td>
<td>6</td>
</tr>
<tr>
<td>1.3</td>
<td>Proposed pyrene degradation pathway for <em>Mycobacterium sp.</em> isolate KMS</td>
<td>16</td>
</tr>
<tr>
<td>1.4</td>
<td>Pathway of pyrene degradation deduced for <em>Mycobacterium vanbaalenii</em> PYR-1</td>
<td>17</td>
</tr>
<tr>
<td>1.5</td>
<td>Schematic electron transfer pathway from reducing equivalent NAD(P)H for oxidation of pyrene</td>
<td>21</td>
</tr>
<tr>
<td>1.6</td>
<td>Identified pyrene-degrading gene clusters in four <em>Mycobacterium</em> isolates</td>
<td>24</td>
</tr>
<tr>
<td>1.7</td>
<td>Clustering of genes involved in pyrene-degradation in the genome of isolate PYR-1</td>
<td>27</td>
</tr>
<tr>
<td>1.8</td>
<td>Two main benzoate degradation pathways utilized by bacteria and fungi</td>
<td>31</td>
</tr>
<tr>
<td>1.9</td>
<td>Schematic representation of the mycobacterial cell wall</td>
<td>34</td>
</tr>
<tr>
<td>1.10</td>
<td>Potential pathways for transformation of C4 TCA cycle intermediates malate and OAA to C3 pyruvate and PEP for gluconeogenesis</td>
<td>38</td>
</tr>
<tr>
<td>1.11</td>
<td>Proposed pathways for intermediary metabolism from carbon sources generated from complete pyrene and benzoate degradation in isolate KMS</td>
<td>39</td>
</tr>
<tr>
<td>1.12</td>
<td>cAMP-CRP regulated catabolite repression in <em>E. coli</em></td>
<td>41</td>
</tr>
<tr>
<td>2.1</td>
<td>Pyrene-degrading genes are clustered on both the chromosome and circular plasmid of isolate KMS</td>
<td>75</td>
</tr>
<tr>
<td>2.2</td>
<td>Synteny analysis by the Mauve program of pyrene-degrading gene islands in <em>Mycobacterium</em> isolates</td>
<td>95</td>
</tr>
<tr>
<td>2.3</td>
<td>Synteny analysis by the Mauve program of pyrene-degrading gene islands in <em>Mycobacterium</em> isolates</td>
<td>98</td>
</tr>
</tbody>
</table>
2.4 Identified pyrene-degrading gene clusters in four *Mycobacterium* isolate......102

3.1 Growth curves of isolate KMS grown on sodium benzoate or pyrene as sole carbon sources.................................118

3.2 The locations of the genes encoding ring-hydroxylating dioxygenase α- and β-subunit genes in the genome of *Mycobacterium sp.* isolate KMS..........................................................121

3.3 Phylogenetic trees generated by MEGA 4 for the predicted protein sequences of dioxygenase α- and β-subunits...................131

3.4 The genome of *Mycobacterium sp.* isolate KMS contains an array of dioxygenase α-subunits based on the classification method.........133

3.5 Alignments of the predicted promoter regions of all the ring-hydroxylating dioxygenase genes for isolate KMS.................................134

3.6 Transcriptional analysis of genes *benA* and *nidA* when isolate KMS was cultured on either sodium benzoate or pyrene as the sole carbon source........................................135

3.7 Alignment of BenC and BenD encoded by *Acinetobacter sp.* ADP1 and *Rhodococcus sp.* RHA1 with the peptide encoded by Mkms_1389 of *Mycobacterium sp.* KMS.................................................140

3.8 Comparison of benzate degrading gene cluster in *Mycobacterium* strains KMS, MCS, JLS, PYR-GCK, and Spyr1 as well as two other clusters from *Rhodococcus* strain RHA1 and *Acinetobacter* strain ADP1........141

4.1 Benzoate to catechol conversion in *Mycobacterium* isolate KMS.............152

4.2 Growth curve of *Mycobacterium sp.* KMS grown on acetate, succinate, fructose, benzoate, or pyrene as the sole carbon source in minimal medium.................160

4.3 Growth curve of isolate KMS on fructose plus acetate or succinate, and fructose alone as well as the corresponding concentration of fructose in the KMS cultures......................................................162

4.4 Growth curve of cells of *Mycobacterium sp.* KMS from an from Luria-Bertani medium agar plate or a minimal medium agar plate with fructose as the sole carbon source.........................................................163

4.5 Effect of additions of acetate, succinate, and fructose on growth of isolate KMS with benzoate and effect of additions of acetate, succinate, fructose, and benzoate on growth of isolate KMS with pyrene ...............165
4.6 Effect of growth of isolate KMS on mixed carbon sources on expression of gene benA and nidA genes.................................166

4.7 Sequences of the promoters for the chromosomal and plasmid nidA genes and benA gene of isolate KMS and the reported cAMP-CRP binding motif in Mycobacterium strain JC1 DSM 3803 cut gene cluster...........169

5.1 Genomic loci and compounds involved in benzoate degradation to Acetyl-CoA and succinyl-CoA for isolate KMS.............................................183

5.2 Proposed pathways for intermediary metabolism in isolate KMS.............184

5.3 Growth curves (OD₆₀₀ nm) of isolate KMS on different carbon sources with shaking at 220 rpm at 25 ºC.................................................................193

5.4 End-point RT PCR amplified with specific primers for genes encoding 16S rRNA, isocitrate lyase, malate synthase, malate:quinone oxidoreductase, PEP carboxykinase, malic enzyme, and PEP synthase.................................197

5.5 Specific activity of malate:quinone oxidoreductase and PEP synthase when KMS was grown on different carbon sources.................................199

6.1 Phylogenetic trees generated by MEGA 4 for the predicted protein sequences for the α-subunits of dioxygenases for Mycobacterium isolate KMS.........................................................215

6.2 Alignments of the predicted promoter regions of all the ring-hydroxylating dioxygenase genes for isolate KMS by ClustalX2........................................217

6.3 Benzoate degradation in isolate KMS, the operon and the deduced benzoate degradation pathway in the pyrene-degrading gene islands in isolate KMS.....219

6.4 The position of genes encoding potential regulators IclR and MarR in the pyrene-degrading gene islands in isolate KMS.............................................220

6.5 Proposed pathways for generation of phosphoenolpyruvate (PEP) from C2 or C4 carbon sources in isolate KMS.....................................................226
LIST OF ABBREVIATIONS

ABC  ATP binding cassette
AMP  Adenosine monophosphate
ATP  Adenosine triphosphate
BLAST Basic local alignment search tool
BSM  Basal salt medium
cAMP Cyclic adenosine monophosphate
cDNA Complementary deoxyribonucleic acid
CFU  Colony forming units
COG  Clusters of orthologous group
CRISPR Clustered regularly interspaced short palindromic repeats
CRP  cAMP receptor protein
DCPIP 2,6-dichlorophenol indophenol
DDW Distilled and deionized water
DNA  Deoxyribonucleic acid
DNase Deoxyribonuclease
dNTP 2-deoxynucleotide 5'-triphosphate
EDTA Ethylenediaminetetraacetic acid
EI enzyme I
EIIA enzyme IIA
ETC Electron transfer chain
FAD Flavin adenine dinucleotide
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTP</td>
<td>Guanosine-5'-triphosphate</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
<tr>
<td>HMW</td>
<td>High-molecule-weight</td>
</tr>
<tr>
<td>HPr</td>
<td>Histidine-containing phosphocarrier protein</td>
</tr>
<tr>
<td>IPR</td>
<td>Interpro protein sequence</td>
</tr>
<tr>
<td>KCN</td>
<td>Potassium cyanide</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>LMW</td>
<td>Low-molecule-weight</td>
</tr>
<tr>
<td>MAH</td>
<td>Monocyclic aromatic hydrocarbon</td>
</tr>
<tr>
<td>MDH</td>
<td>Malate dehydrogenase</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium chloride</td>
</tr>
<tr>
<td>MQO</td>
<td>Malate:quinone oxidoreductase</td>
</tr>
<tr>
<td>NAD⁺</td>
<td>Nicotinamide adenine dinucleotide</td>
</tr>
<tr>
<td>NADH</td>
<td>Nicotinamide adenine dinucleotide (reduced form)</td>
</tr>
<tr>
<td>NADP⁺</td>
<td>Nicotinamide adenine dinucleotide phosphate</td>
</tr>
<tr>
<td>NADPH</td>
<td>Nicotinamide adenine dinucleotide phosphate (reduced form)</td>
</tr>
<tr>
<td>OAA</td>
<td>Oxaloacetic acid</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PAH</td>
<td>Polycyclic aromatic hydrocarbon</td>
</tr>
<tr>
<td>PBM</td>
<td>Phosphate-based minimal medium</td>
</tr>
<tr>
<td>PEP</td>
<td>Phosphoenolpyruvate</td>
</tr>
<tr>
<td>Q PCR</td>
<td>Quantitative PCR</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>Q RT PCR</td>
<td>Quantitative reverse transcriptase PCR</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PTS</td>
<td>Phosphoenolpyruvate-dependent phosphotransferase transport system</td>
</tr>
<tr>
<td>RHD</td>
<td>Ring-hydroxylating dioxygenase</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive oxygen species</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>RT PCR</td>
<td>Reverse transcriptase PCR</td>
</tr>
<tr>
<td>TCA cycle</td>
<td>Tricarboxylic acid cycle</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(Hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

OUTLINE OF DISSERTATION

In the last ten years, significant advances have been made to broaden the knowledge of how polycyclic aromatic hydrocarbons (PAHs) are degraded. The higher molecular weight structures, such as pyrene with four aromatic rings and benzo[a]pyrene with six rings, constitute environmental risks. Metabolites and enzymes involved in pyrene degradation have been elucidated for several Mycobacterium isolates. Although by 2001 identification of pyrene-degrading genes was initiated for Mycobacterium isolates, the distribution of these genes within mycobacterium genomes and the extent of genes involved in PAH breakdown lacked full analysis prior to complete genome sequencing of five PAH-degrading mycobacterial isolates in 2008. Most of the PAH-degrading mycobacterial strains are isolated from PAH-contaminated soils and sludges. These strains have a broader environmental habitat through colonization of plant rhizospheres in soils and in water.

In the Introduction, I provide background on the structure and sources of aromatic hydrocarbons and discuss biodegradation of low and high molecular weight PAHs by microbes. I introduce background knowledge of isolate Mycobacterium sp. KMS that is the focus of my studies. I compare the predicted pathway for pyrene degradation in this isolate to that published for isolate PYR-1. I use the literature reports prior to genome sequencing to indicate that there was a likelihood of duplication and clustering for the genes involved in pyrene degradation in several Mycobacterium isolates. I introduce the
background information that the utilization of benzoate, a single ring oxidized aromatic, may vary between mycobacterial strains and bacterial strains from other genera. Study on benzoate degradation in *Mycobacterium* isolate is novel because no operons encoding genes for benzoate degradation had been characterized in mycobacteria prior to my work. Because the products of pyrene mineralization include C2, C3, and C4 intermediates, I discuss the intermediary metabolic pathways that connect their metabolism and other pathways to permit cell growth. I discuss how cAMP-dependent mechanisms may be involved in governing gene expression related to preference of carbon sources in bacteria including mycobacteria.

My research chapters, Chapters 2-5, show the use of bioinformatics and transcript analysis to reveal novel information on functional genomics for the environmental strain *Mycobacterium* sp. KMS, with respect to its soil habitat, its ability to perform bioremediation of aromatic hydrocarbons, and its ability to colonize plant roots.

In Chapter 2, the research is focused on determining in isolate *Mycobacterium* sp. KMS, the location of genes encoding proteins induced by pyrene; these proteins are postulated to function in pyrene degradation (90). Bioinformatic analysis shows that the genes cluster and some are duplicated on both chromosome and circular plasmid. Transcript analysis was performed to determine whether duplicated chromosomal and plasmid genes are induced by pyrene. Synteny alignment was used to compare the order of the pyrene-degrading gene clusters in isolate KMS to five other mycobacterial isolates with sequenced genomes: MCS, JLS isolated from the same site as KMS, isolates PYR-1
and PYR-GCK, which were isolated from geographically separate sites in the USA, as well as isolate Spyr1, which was isolated from Greece.

In Chapter 3, the research focused on the diversity of genes in the genome of isolate KMS with ring-hydroxylating dioxygenase activity. Phylogenetic trees based on alignment of genes encoding dioxygenase α- and β-subunits at both nucleic acid and amino acids levels as well as the base pairs in the promoter regions were constructed. Such analysis was performed to provide information on the extent of duplication and gene evolution. Genes relating to benzoate degradation were predicted in KMS because of high homology to genes encoding benzoate operons in other bacteria. An important target of this study was to find whether genes encoding the degradation of pyrene and benzoate were co-induced or not.

In Chapter 4, the research focused on whether the mycobacterium isolates have preferences for carbon utilization. This work compared the utilization of simple organic compounds, organic acids and sugars, to the PAH pyrene and MAH benzoate as sole carbon sources or as mixtures. Induction of genes encoding dioxygenase subunits responsible for the initial degradation steps of either pyrene or benzoate was studied. I raised the possibility that catabolite repression of pyrene utilization involves a cAMP-dependent mechanism.

In Chapter 5, the research focused on the possible pathways for intermediary metabolism in mycobacterium. Transcript analysis showed activation of genes encoding isocitrate lyase and malate synthase involved in glyoxylate shunt when KMS was grown on different carbon sources. A gene encoding a novel malate:quinone oxidoreductase was
suggested as a house-keeping gene involved in completing the C4 transformations in the TCA cycle. PEP for gluconeogenesis could be synthesized from both OAA and malate; PEP generation also is linked to utilization of fructose in the PST transport system.

In Chapter 6, conclusions on the significance of the findings are discussed along with future research directions.

Chemical structures and origins of aromatic hydrocarbons. Aromatic hydrocarbons range from monocyclic aromatic hydrocarbons (MAHs) with only one benzene ring (Fig. 1.1) (35, 75) to the polycyclic aromatic hydrocarbons (PAHs), where benzene rings are fused in different arrangements (Fig. 1.2) (91). Based on the number of fused benzene rings, PAHs are classified into low-molecular-weight (LMW) compounds, with two or three rings, and high-molecular-weight (HMW) compounds, with more than three rings (67).

Benzene is the simplest MAH and is isolated from petroleum and coal tar. It is used as a solvent in industry and constitutes the major component of gasoline. Toluene and xylene occur naturally in crude oil (8, 47); they also are present in gasoline and are common industrial solvents (114).

Many modified MAHs occur naturally, being released into soils due to decay of the plant cell wall polymer lignin by bacteria and fungi (20, 122, 134). Lignin has a complex oxidized aromatic structure (48) and its degradation products include benzoate (134). The oxidized phenolic salicylic acid is synthesized by plants as a signal compound for defense responses (97, 148). Soil microbes themselves also produce aromatic secondary metabolites that are involved in microbe-microbe and plant-microbe interactions (135).
Others, such as phthalate and protocatechuic acid, are part of degradation pathways, are derived from breakdown of pyrene as discussed in a following section. Such natural aromatic compounds are potential carbon sources for soil microbes. In addition to the naturally produced MAHs, phenol and phthalate esters are produced abundantly by industry for synthesis of commercial materials (7, 149).

**FIG 1.1** Chemical structures of some common MAHs and MAH derivatives
PAHs are generated from both natural activities during forest fires (121) and anthropological activities such as incomplete combustion of fossil fuels (3, 23). PAHs also are found in coal tar (73), sewage sludge (150), and oil sand (72). They exist in petroleum or are purposely produced (144).

Naphthalene is the simplest PAH which contains two fused benzene rings and is largely found in coal tar (60, 108). Naphthalene is naturally produced by fungi as an insect repellent (30). The LMW PAHs anthracene and phenanthrene contain three fused aromatic rings with different arrangements. They are components of coal tars (60, 108).
The four-ring structure, pyrene also is found in coal tar (108). Pyrene is used as a model high molecular weight PAH.

**Biological risks from aromatic hydrocarbons.** Benzene is a hazardous pollutant because during its metabolism reactive oxygen species are produced (4, 109). The generation of the reactive oxygen species and the metabolic reactivity of the oxidation products from benzene contribute to its behavior as a human carcinogen (62). Although toluene and xylene are not carcinogenic, exposure to toluene or xylene lead to neurological effects including headache, dizziness, fatigue, anxiety, and impaired short-term memory (5, 6). Phthalate esters, which vary in their side chains, affect female reproductive systems (93). Phenol is toxic to liver, kidney, and nervous system of animals (2). Thus, these MAHs threaten the health of human beings and methods for MAH degradation are required when soils and water become contaminated.

An increase in the number of benzene rings and the angularity of a PAH molecule, such as four or more fused rings, results in a concomitant increase in hydrophobicity and electrochemical stability (56, 154). The HMW PAHs have health risks because their degradation products are carcinogenic, mutagenic, and toxic to mammals (23, 94, 110, 115, 132, 142). Exposure to PAHs, such as benzo[a]pyrene, is especially a risk to young children (65). Toxicity is correlated with the formation of intermediate dihydrodiol epoxides and quinones as well as reactive oxygen species and other radicals as the PAHs are degraded. The PAH oxidation products form covalent adducts with DNA that have mutagenic potential (38, 95). Mutations also are generated from the depurination of DNA
by radical cations, formed by peroxidase or cytochrome P450 oxidation of the PAH degradation products (96).

Due to their hydrophobic, persistent, and recalcitrant character (103, 155), PAHs accumulate in the environment (141). Benzo[a]anthracene (BaA), benzo[a]pyrene (BaP), benzo[b]fluoranthene (BbF), benzo[k]-fluoranthene (BkF), chrysene (Chry), dibenzo[a,h]anthracene (DBA), and indeno[123-c,d]pyrene (IP) are classified as human carcinogens by the U.S. Environmental Protection Agency (1). These materials are among the 17 PAHs listed as priority pollutants by the U.S. Environmental Protection Agency (140). Indeed the HMW PAHs pyrene and benzo[a]pyrene with four- and five-ring structures, respectively, are monitored to represent the level of PAHs in different environments (12, 26, 128). Because of the risk to human health there is a strong interest in the mechanisms by which the HMW PAHs can be remediated from contaminated sites.

**Biodegradation of PAHs by microbes in soils.** PAHs are found associated with soil particles (152) and their remediation involves physical, chemical, and biological techniques (116). Physical and chemical techniques include soil venting, soil flushing, solvent extraction, and chemical oxidation (38, 49, 76). However, the effectiveness of these methods is limited for PAHs because they have low volatility for soil venting and low solubility for soil flushing so that high concentrations of solvent are needed for extraction.

In contrast, bioremediation is a natural inexpensive process that can function on-site or under confined conditions if the contaminated material is removed for processing. However, some problems need to be considered when bioremediation is applied, for
example: 1) accessibility of the PAHs to microbes, 2) control and maintenance of growth of the microbe in the contaminated sites, and 3) influence of other available nutrients in the contaminated area on the bioremediation potential of the organism. Consequently, basic knowledge concerning the microbe of interest is important when making decisions concerning bioremediation potential.

PAHs are mineralized by bacteria, fungi, and algae generating nontoxic CO$_2$ and H$_2$O or are partially degraded to less toxic chemicals (54). A broad range of bacteria have PAH-degrading activity: *Burkholderia*, *Pseudomonas*, *Mycobacterium*, *Rhodococcus*, and *Sphingomonas* (71, 77, 81, 119, 146). Most of the PAH-degrading bacteria are isolated from industrially-generated, PAH-contaminated soils (24, 99, 124), or sediments where pollutants have accumulated (31, 57), or ocean water contaminated by crude oil (69, 145).

The findings in this dissertation focus on PAH degradation by *Mycobacterium* isolates and as shown in Table 1.1 such isolates are found globally at sites contaminated with the different PAHs. PAH-degrading mycobacteria were detected initially in the early 1990s and continue to be documented from different global locations. Mycobacterium cells are Gram-positive and belong to the phylum *Actinobacteria* so that they have similar complex cell wall structures to *Corynebacterium* and *Nocardia* (29). The genus contains both pathogenic and environmental isolates; none of the pathogenic isolates possesses PAH-degrading ability. Some strains are reported to degrade only the LMW PAHs (51, 138), whereas others also degrade the more complex structures (27, 31, 57, 84, 99, 126). Pyrene is used as a model HMW PAH in many of the degradation studies...
because, although it is recalcitrant, it is less toxic than the carcinogenic five ring structure, benzo[a]pyrene (67).

**TABLE 1.1** Summary of sources of PAHs-degrading *Mycobacterium sp.* strains and the PAHs degraded by each isolate

<table>
<thead>
<tr>
<th>Isolate name</th>
<th>Source of Isolate</th>
<th>PAH substrate</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. sp.</em> strain</td>
<td>PAH polluted sediments, New Hampshire, USA</td>
<td>Phenanthrene</td>
<td>51</td>
</tr>
<tr>
<td>BG1</td>
<td>Petrogenic chemicals contaminated soil, Texas, USA</td>
<td>Phenanthrene, pyrene, fluoranthene, benzo[a]pyrene</td>
<td>57</td>
</tr>
<tr>
<td><em>M. vanbaalenii</em></td>
<td>Coal gasification plants soil, Illinois, USA</td>
<td>Pyrene, benzo[a]pyrene</td>
<td>50, 147</td>
</tr>
<tr>
<td>PAH135 (close to <em>M. aichiense</em>)</td>
<td>Air pollution</td>
<td>Pyrene, benzo[a]pyrene</td>
<td>50, 147</td>
</tr>
<tr>
<td><em>M. sp.</em> strain BB1</td>
<td>Coal gasification site, Germany</td>
<td>Phenanthrene, pyrene, fluoranthene, fluorene</td>
<td>15</td>
</tr>
<tr>
<td><em>M. sp.</em> isolate (close to <em>M. gilvum</em>)</td>
<td>Coal gasification site soil</td>
<td>Phenanthrene, pyrene, fluoranthene</td>
<td>41</td>
</tr>
<tr>
<td><em>M. sp.</em> strain</td>
<td>Coal gasification site, Germany</td>
<td>Fluorene, phenanthrene, fluoranthene, and pyrene</td>
<td>136</td>
</tr>
<tr>
<td>Organism</td>
<td>Location</td>
<td>Pollutants</td>
<td>References</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>M. sp. strain RJGII135</td>
<td>Coal gasification site, Illinois, USA</td>
<td>Pyrene, benz[a]anthracene, benzo[a]pyrene, chrysene</td>
<td>124</td>
</tr>
<tr>
<td>M. sp. strain</td>
<td>Soil, New Jersey, USA</td>
<td>Pyrene</td>
<td>64</td>
</tr>
<tr>
<td>M. hodleri sp. nov. EMI2T</td>
<td>Fluoranthene and benzo[a]pyrene-polluted soil</td>
<td>Fluoranthene</td>
<td>82</td>
</tr>
<tr>
<td>M. gilvum PYR-GCK</td>
<td>PAH polluted sediments, Indiana, USA</td>
<td>Phenanthrene, pyrene, fluoranthene</td>
<td>31</td>
</tr>
<tr>
<td>M. sp. strain</td>
<td>PAHs contaminated river sediments, New York, USA</td>
<td>Pyrene</td>
<td>102</td>
</tr>
<tr>
<td>M. sp. strain S1</td>
<td>Not mentioned</td>
<td>Anthracene only</td>
<td>138</td>
</tr>
<tr>
<td>M. sp. strain CH-1</td>
<td>PAH polluted sediments, Michigan, USA</td>
<td>Phenanthrene, pyrene, fluoranthene</td>
<td>27, 55</td>
</tr>
<tr>
<td>M. sp. strain AP1</td>
<td>Oil polluted marine beach, Spain</td>
<td>Phenanthrene, anthracene, fluoranthene, pyrene, benzo(a)anthracene</td>
<td>92, 145</td>
</tr>
<tr>
<td>M. sp. strain</td>
<td>Not mentioned</td>
<td>Pyrene, phenanthrene</td>
<td>84</td>
</tr>
</tbody>
</table>
Mycobacterium sp. strain KMS. Mycobacterium sp. strain KMS was isolated from soil contaminated with wood preservatives based on creosote and pentachlorophenol in a

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Environment Details</th>
<th>Contaminants</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>6PY1</td>
<td>Jet-fuel contaminated site, Canada</td>
<td>Pyrene, phenanthrene, fluoranthene</td>
<td>126</td>
</tr>
<tr>
<td>M. sp. strain S65</td>
<td>Creosote and pentachlorophenol (PCP) contaminated soil, Montana, USA</td>
<td>Phenanthrene, pyrene, benzo[a]pyrene, chrysene</td>
<td>99</td>
</tr>
<tr>
<td>M. sp. strain KMS</td>
<td>Creosote and pentachlorophenol (PCP) contaminated soil, Montana, USA</td>
<td>Phenanthrene, pyrene, benzo[a]pyrene, chrysene</td>
<td>99</td>
</tr>
<tr>
<td>M. sp. strain JLS</td>
<td>Creosote and pentachlorophenol (PCP) contaminated soil, Montana, USA</td>
<td>Phenanthrene, pyrene, benzo[a]pyrene, chrysene</td>
<td>99</td>
</tr>
<tr>
<td>M. sp. strain CH-2</td>
<td>PAH contaminated gas plant, Washington, USA</td>
<td>Phenanthrene, pyrene, fluoranthene</td>
<td>28, 55</td>
</tr>
<tr>
<td>M. sp. strain Spyr1</td>
<td>Creosote polluted soil, Greece</td>
<td>Phenanthrene, pyrene, fluoranthene</td>
<td>68</td>
</tr>
</tbody>
</table>
prepared land treatment-bed established for bioremediation at a wood-preserving plant at Libby, Montana, USA (99). These land treatment soils displayed rapid pyrene mineralization ability (59). Two other PAH-degrading isolates, MCS and JLS, also are isolated from these soils (99). The 16S rRNA genes sequence reveals all three strains are *Mycobacterium* isolates. However their catalase and superoxide dismutase (SOD) isozyme patterns differ between the strains and another PAH-degrading *Mycobacterium* isolate PYR-1 (99). The three Montana isolates also reveal different patterns of genomic DNA upon digestion with restriction enzymes and have differences in the composition of their long-chain fatty acids in their cell walls (100).

Genome sequencing of the Libby site isolates KMS, MCS, and JLS confirmed that each of these strain was unique. Their genomes were annotated by the DOE Joint Genome Institute with public access at http://img.jgi.doe.gov/cgi-bin/w/main.cgi. The genomic information shows that plasmids are differentially present in the Libby isolates: isolate KMS has one circular plasmid and one linear plasmid, isolate MCS has one linear plasmid, and isolate JLS lacks a plasmid (Table 1.2). The two other sequenced PAH-degrading isolates PYR-1 and PYR-GCK were isolated from Texas and Indiana, respectively. *Mycobacterium vanbaalenii* PYR-1 lacks plasmids whereas *Mycobacterium gilvum* PYR-GCK, which was previously named *M. flavescens* PYR-GCK (18), has three linear plasmids (Table 1.2). Complete sequencing of another PAH-degrading *Mycobacterium* isolate Spyrl, which was isolated in Greece (68), was finished recently in July, 2011. Each genome from these environmental isolates is large and has high GC content. These findings are similar to the environmental isolate *Mycobacterium*
*smegmatis* MC² 155 (Table 1.2) that like the PAH-degrading isolates is fast-growing and non-pathogenic. *M. smegmatis* is often studied as a surrogate for the pathogenic *Mycobacterium* species (125).

**TABLE 1.2** Overview of the genomes of six PAH-degrading mycobacterium isolates and another environmental isolate, *Mycobacterium smegmatis* MC² 155. Data from database at DOE Joint Genome Institute website (http://img.jgi.doe.gov/cgi-bin/w/main.cgi)

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Genome size</th>
<th>Plasmid presence</th>
<th>% GC content</th>
<th>Number of predicted ORFs</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. sp.</em> KMS</td>
<td>6.2 Mb</td>
<td>One circular plasmid and one linear plasmid</td>
<td>68%</td>
<td>5939</td>
</tr>
<tr>
<td><em>M. sp.</em> MCS</td>
<td>5.9 Mb</td>
<td>One linear plasmid</td>
<td>68%</td>
<td>5691</td>
</tr>
<tr>
<td><em>M. sp.</em> JLS</td>
<td>6.0 Mb</td>
<td>None</td>
<td>68%</td>
<td>5759</td>
</tr>
<tr>
<td><em>M. vanbaalenii</em></td>
<td>6.4 Mb</td>
<td>None</td>
<td>68%</td>
<td>6061</td>
</tr>
<tr>
<td>PYR-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. gilvum</em></td>
<td>5.9 Mb</td>
<td>Three linear plasmids</td>
<td>68%</td>
<td>5606</td>
</tr>
<tr>
<td>PYR-GCK</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. sp.</em> Spyr1</td>
<td>5.8 Mb</td>
<td>Two plasmids</td>
<td>68%</td>
<td>5434</td>
</tr>
<tr>
<td><em>M. smegmatis</em> MC² 155</td>
<td>6.9 Mb</td>
<td>None</td>
<td>67%</td>
<td>6829</td>
</tr>
</tbody>
</table>
Pyrene-degradation mechanisms in mycobacteria

Summary of pyrene-degradation pathways in isolates KMS and PYR-1. Many of the PAH-degrading *Mycobacterium* isolates utilize pyrene as well as the LMW PAHs as carbon sources (Table 1.1). Metabolites from pyrene degradation and the enzymes involved in pyrene degradation have been most studied in isolate KMS (Fig. 1.3) (90) and isolate PYR-1 (Fig. 1.4) (79). The pathways deduced by Liang et al. (90) and Kim et al. (79) basically use the same array of enzymes although differences between the two strains exist.

Initiation of pyrene degradation is catalyzed by pyrene ring-hydroxylating dioxygenases to generate *cis*-pyrene-dihydriodols (Fig 1.3 and Fig. 1.4) (79, 90). This oxidation is catalyzed by a dioxygenase that uses molecular oxygen as the oxidant. In isolate PYR-1 oxidation of fused benzene rings happens on different sites, for example at C1 to generate *cis*-1,2-pyrene-dihydriodiol and C2 or C4 and C5 to produce the major product *cis*-4,5-pyrene-dihydriodiol (Fig. 1.4, S1 and S22) (79). Generation of *cis*-1,2-pyrene-dihydriodiol leads to a dead end pathway for isolate PYR-1 because methylation of the hydroxyl-groups prevents further oxidation.

In PYR-1 the *cis*-4,5-pyrene-dihydriodiol is further oxidized to eventually produce a C4 and a C2 product that can enter intermediary metabolism (Fig. 1.4) (79). Studies by Kim et al. (79) found the first reaction to generate *cis*-4,5-pyrene-dihydriodiol is catalyzed by dioxygenases that also will oxidize the three ring PAH, phenanthrene, and another HMW PAH fluoranthene (Fig. 1.4), showing effectiveness of the enzyme on three- and four-ring PAHs.
In addition to initial ring oxidation by the ring-hydroxylating dioxygenase, Liang et al. (90) provide evidence for an alternative route of initial metabolism involving a monooxygenase; this alternative pathway is not reported for PYR-1 (Fig. 1.3) The novel metabolite pyrene-4,5-dione identified only in KMS culture (90) indicates that details of the pathways may differ between mycobacterium strains.

**FIG 1.3** Proposed pyrene degradation pathway for *Mycobacterium sp.* isolate KMS. Metabolites were identified by a combination of gas chromatography-mass spectral analysis. Enzymes were identified by nano-LC-tandem mass spectrometry to determine the size of pyrene-induced peptides separated by two dimension gel electrophoresis. Metabolites and enzymes in brackets are deduced and were not verified in this study (90). Novel steps and metabolites only found in isolate KMS are circled by blue boxes. The metabolites also identified in PYR-1 pyrene-grown cultures are circled in red boxes.
FIG 1.4 Pathway of pyrene degradation deduced for *Mycobacterium vanbaalenii* PYR-1. Metabolites were identified by combination of gas chromatography-mass spectroscopy and enzymes are identified from proteomic study (79).

Intermediates are labeled by P numbers as follows: P1, pyrene *cis*-4,5-dihydrodiol; P2, 4,5-dihydroxypyrene; P3, phenanthrene-4,5-dicarboxylate; P4, phenanthrene-4-carboxylate; P5, *cis*-3,4-dihydroxyphenanthrene-4-carboxylate; P6, 3,4-dihydroxyphenanthrene; P7, 2-hydroxy-2H-benzo[h]chromene-2-carboxylate; P8, 1-hydroxy-2-naphthaldehyde; P9, 1-hydroxy-2-naphthoate; P10, *trans*-2-carboxybenzalpyruvate; P11, 2-carboxybenzaldehyde; P12, phthalate; P13, phthalate
3,4-dihydrodiol; P14, 3,4-dihydroxyphthalate; P15, protocatechuate; P16, β-carboxy-
cis,cis-muconate; P17, γ-carboxymuconolactone; P18, β-ketoadipate enol-lactone; P19, 
β-ketoadipate; P20, β-ketoadipyl-CoA; P21, pyrene cis-1,2-dihydrodiol; P22, 1,2-
dihydroxypyrene; P23, 1-methoxy-2-hydroxypyrene; P24, 1-hydroxy-2-methoxypyrene; 
P25, 1,2-dimethoxypyrene. The metabolites also identified in KMS pyrene culture are 
circled in red boxes.

Each step (shown by S numbers) is catalyzed by enzymes as follows: S1, 
Pyrene/phenanthrene ring-hydroxylating dioxygenase, Fluoranthene/pyrene ring-
hydroxylating dioxygenase, Dioxygenase ferredoxin, and Dioxygenase reductase; S2, 
Dihydrodiol dehydrogenase; S3, Ring cleavage dioxygenase; S4, Decarboxylase; S5, 
Phenanthrene ring-hydroxylating dioxygenase, Dioxygenase ferredoxin, and 
Dioxygenase reductase; S6, Dihydrodiol dehydrogenase; S7, Ring cleavage dioxygenase; 
S8, Hydratase-aldolase; S9, Aldehyde dehydrogenase; S10, 1-Hydroxy-2-naphthoate 
dioxygenase; S11, trans-2'-Carboxybenzalpyruvate hydratase-aldolase; S12, 2-
Carboxylbenzaldehyde dehydrogenase; S13, Phthalate 3,4-dioxygenase, Dioxygenase 
ferredoxin, and Dioxygenase reductase; S14, Phthalate 3,4-dihydrodiol dehydrogenase; 
S15, Decarboxylase; S16, Protocatechuate 3,4-dioxygenase; S17, β-Carboxy-cis,cis-
muconate cycloisomerase; S18 and S19, γ-Carboxymuconolactone decarboxylase/β-
ketoadipate enol-lactone hydrolase; S20, β-Ketoadipate succinyl-CoA transferase; S21, 
β-Ketoadipyl-CoA thiolase; S22, Fluoranthene/pyrene ring-hydroxylating dioxygenase, 
Dioxygenase ferredoxin, and Dioxygenase reductase; S23, Dihydrodiol dehydrogenase; 
S24 and S25, Catechol O-methyltransferase. In the complete degradation pathway, steps 
(S1, S5, and S13) catalyzed by ring-hydroxylating dioxygenases are circled by blue 
boxes, steps (S2, S6, S9, S12, and S14) catalyzed by dehydrogenases by pink boxes, steps 
(S3, S7, S10, and S16) catalyzed by ring-cleavage dioxygenases by orange boxes, steps 
(S4 and S15) catalyzed by decarboxylases by green boxes, steps (S8 and S11) catalyzed 
by hydratase-aldolases by turquoise boxes.
Further degradation of pyrene in isolate PYR-1 includes aromatic ring oxidation catalyzed by other ring-hydroxylating dioxygenases (Fig 1.4, S5 and S13) oxidation catalyzed by dehydrogenases (Fig. 1.4, S2, S6, S9, S12, and S14), ring-cleavage, in which aromatic rings are broken, catalyzed by ring-cleavage dioxygenases (Fig. 1.4 S3, S7, S10, and S16), decarboxylation, through which CO$_2$ is released, catalyzed by decarboxylases (Fig. 1.4, S4 and S15), and side chain cleavage, through which pyruvate is released, catalyzed by hydratase-aldolases (Fig. 1.4, S8 and S11) (79). After these steps, the first nonaromatic compound, β-carboxy-cis,cis-muconate is generated (Fig. 1.4, P16) from protocatechuic acid (Fig. 1.4, P15). This compound is further degraded to final products, acetyl-CoA and succinyl-CoA, via an important intermediate β-ketoadipatyl-CoA (Fig. 1.4, P20) (79). The final products, acetyl-CoA and succinyl-CoA, from pyrene degradation would flow into intermediary metabolism with the tricarboxylic acid cycle (TCA) cycle (79) or the acetyl-CoA could be used for fatty acid synthesis, essential for cell wall formation (see section “Key pathways for carbon metabolism in Mycobacterium isolates” for discussion).

Similarly, enzymes including ring-hydroxylating dioxygenases, dehydrogenase, ring-cleavage dioxygenase, decarboxylase, and hydratase-aldolase are identified or hypothesized for isolate KMS (Fig. 1.3 and Table 1.3) (90). Common intermediates found for both pyrene degradation pathways in KMS and PYR-1 (Fig. 1.3 and Fig. 1.4) include PAHs: pyrene cis-4,5-dihydrodiol, 4,5-dihydroxypyrene, phenanthrene-4,5-dicarboxylate, and phenanthrene-4-carboxylate, MAHs: phthalate and protocatechuic acid, and CO$_2$ released from decarboxylation steps (79, 90).
TABLE 1.3 Identification of seventeen proteins de novo synthesized or highly induced by pyrene in Mycobacterium isolate KMS (90)

<table>
<thead>
<tr>
<th>Protein no.</th>
<th>Protein identification</th>
<th>NCBI accession no.</th>
<th>Spot density in indicated culture</th>
<th>Mascot score</th>
<th>No. of peptides matched</th>
<th>Coverage</th>
<th>Observed migration*</th>
<th>Theoretical migration*</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>4Fe-4S ferredoxin, iron-sulfur binding</td>
<td>ZP_01282568</td>
<td>0.06 0.25 0.26</td>
<td>153 2 46</td>
<td>16.3 4.3</td>
<td>10.5 4.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>Aromatic-ring-hydroxylating dioxygenase, β subunit</td>
<td>ZP_01282558</td>
<td>ND 0.15 0.17</td>
<td>569 8 59</td>
<td>16.7 4.4</td>
<td>18.9 4.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>Aromatic-ring-hydroxylating dioxygenase, β subunit</td>
<td>ZP_01286731</td>
<td>ND 0.17 0.13</td>
<td>49 2 14</td>
<td>18.3 4.4</td>
<td>18.8 4.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>Sterol binding protein</td>
<td>ZP_01286725</td>
<td>ND 0.26 0.22</td>
<td>471 12 59</td>
<td>19.5 4.8</td>
<td>19.5 4.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>Hypothetical protein Mm98DRAFT_0077</td>
<td>ZP_01286725</td>
<td>ND 0.26 0.22</td>
<td>471 12 59</td>
<td>19.5 4.8</td>
<td>19.5 4.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>Aromatic-ring-hydroxylating dioxygenase, β subunit</td>
<td>AAAT54177</td>
<td>ND 0.16 0.11</td>
<td>525 10 60</td>
<td>18.7 5.1</td>
<td>19.4 5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td>Aromatic-ring-hydroxylating dioxygenase, α subunit (α/β)</td>
<td>AAQ993917</td>
<td>ND 0.20 0.19</td>
<td>420 7 39</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P8</td>
<td>Aromatic-ring-hydroxylating dioxygenase, β subunit</td>
<td>ZP_01287262</td>
<td>ND 0.11 0.09</td>
<td>405 7 34</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P9</td>
<td>Phthalalate dehydrogenase</td>
<td>ZP_01286725</td>
<td>ND 0.26 0.22</td>
<td>471 12 59</td>
<td>19.5 4.8</td>
<td>19.5 4.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P10</td>
<td>Ring-hydroxylating dioxygenase, α subunit (Rieske [2Fe-2S] region)</td>
<td>ZP_01286725</td>
<td>ND 0.20 0.19</td>
<td>420 7 39</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P11</td>
<td>Glycosyl hydrolase, BNR</td>
<td>ZP_01286725</td>
<td>ND 0.11 0.09</td>
<td>405 7 34</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P12</td>
<td>Epoxydehydrogenase-like alpha beta hydrodrolase fold</td>
<td>ZP_01286725</td>
<td>ND 0.11 0.09</td>
<td>405 7 34</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P13</td>
<td>Ring-hydroxylating dioxygenase, α subunit (Rieske [2Fe-2S] region)</td>
<td>ZP_01286725</td>
<td>ND 0.20 0.19</td>
<td>420 7 39</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P14</td>
<td>Aromatic-ring-hydroxylating dioxygenase, α subunit (nAD)</td>
<td>ZP_01286725</td>
<td>ND 0.20 0.19</td>
<td>420 7 39</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P15</td>
<td>Flavoprotein-like fumarate reductase/nucleotide dehydrogenase (FAD-dependent oxidoreductase)</td>
<td>ZP_01286725</td>
<td>ND 0.20 0.19</td>
<td>420 7 39</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P16</td>
<td>PhdG (hydrazone-aldolase)</td>
<td>ZP_01286725</td>
<td>ND 0.20 0.19</td>
<td>420 7 39</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P17</td>
<td>Aldolase dehydrogenase (NAD+)</td>
<td>ZP_01286725</td>
<td>ND 0.20 0.19</td>
<td>420 7 39</td>
<td>24.9 6.1</td>
<td>22.4 5.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Components of ring-hydroxylating dioxygenases. The pyrene degradation pathway in both mycobacterium isolates uses ring-hydroxylating dioxygenases (RHDs) in the first transformation step and in subsequent transformations. Ring-hydroxylating dioxygenases are a group of unique enzymes containing Rieske non-heme iron (86). Most RHDs contain multiple components including a large α-subunit and a small β-subunit. The active site of an RHD is the 2Fe-2S Rieske center present in the α-subunit (21, 113). Proposed functions of the β-subunit are to maintain the spatial structure of RHDs (63) and to be involved in substrate recognition (53). Additionally for the dioxygenase
complex to function, a small electron transfer chain, ferredoxin, and a reductase are required (21, 86, 113). A schematic for an electron transfer pathway from NAD(P)H for oxidation of PAHs involving a ring dioxygenase is shown in Fig. 1.5. The involvement of the electron transfer chains means that the cells utilize electrons from reducing equivalents NAD(P)H along with molecular oxygen to oxidize the aromatic rings (86). Thus, intermediary metabolism is important to generate the initial reducing equivalent for PAH catabolism.

The electron transfer chains for the ring-hydroxylating dioxygenase may be complete as shown in Fig 1.5 or lack the ferredoxin component (86). The reductases differ in whether glutathione or ferredoxin is involved and in the arrangement of domains for binding ferredoxin, NAD⁺, or flavin (86). The ferredoxins of RHDs in bacteria other than Mycobacterium contain either 2Fe-2S or 3Fe-4S active centers However, the types of ferredoxin in the RHDs of Mycobacterium isolates PYR-1, S65, and 6PY1 are still unknown (86).

FIG 1.5 Schematic electron transfer pathway from reducing equivalent NAD(P)H for oxidation of pyrene modified from Habe et al. (52).
RHDs have specificity to certain aromatic hydrocarbon substrates due to the spatial structures of the enzyme and substrates (87). RHDs catalyze only the oxidation of the aromatic substrates whose structures spatially fit the active center of the RHDs (87). This restriction implies that to degrade a broad range of aromatic hydrocarbons, different types of RHDs need to be encoded. In isolate PYR-1, and we assume KMS, complete pyrene degradation requires four different kinds of RHDs including pyrene/phenanthrene dioxygenase, fluoranthene/pyrene dioxygenase, phenanthrene dioxygenase, and phthalate 3,4-dioxygenase (Fig. 1.4) (79). These would have differential specificity for pyrene, fluoranthene, phenanthrene, and phthalate.

**Organization of pyrene-degrading genes in *Mycobacterium* isolates.** Prior to genomic sequencing the importance of dioxygenase in PAH degradation lead to the identification of small gene clusters containing genes encoding the subunits of dioxygenases from several mycobacterium isolates: PYR-1, 6PY1, KMS, MCS, JLS, S65, and CH-2 (28, 74, 78, 84, 99, 126).

Genes in these mycobacterium strains were termed *nidA/B* for naphthalene inducible dioxygenase or *pdoA/B* for phenanthrene dioxygenase (Fig. 1.6) The function of genes encoding subunits NidA/NidB and NidA3/NidB3 in PYR-1 was proven by Khan et al. (74) and Kim et al. (78) through expression of the specific genes in *E. coli* and demonstrating the conversion of pyrene to pyrene dihydrodiol only when the genes were expressed. An additional copy of a gene, designated as *nidB2*, with 100% identity to gene *nidB* encoding the β-subunit of the naphthalene-inducible dioxygenase was identified in isolate PYR-1 at a short distance from the loci for the adjacent *nidB-nidA* genes (130).
These findings were the first to suggest gene duplication occurred for the dioxygenase \textit{nid} genes involved in the initiation of pyrene/naphthalene/phenanthrene degradation. Gene duplication further was supported by observations with isolate S65, where high homology of nucleotide sequence (more than 85\%) was found between corresponding genes in the \textit{nid} and \textit{pdo} clusters (126).

Stingley et al. (130) also showed by gene walking that the \textit{nidB2} cluster and the \textit{nidB-nidA} cluster in PYR-1 were contiguous with other genes, \textit{phtAa} and \textit{phtAb}, encoding \(\alpha\)- and \(\beta\)-subunits of phthalate dioxygenase, respectively, an enzyme involved in phenanthrene degradation, as well in the pyrene degradation pathway (79) (Fig. 1.4). These pathways merge at 3,4-dihydroxyphenanthrene (Fig. 1.4, P6).

Krivobok et al. (84) proposed the function of the \textit{pdoA1/pdoB1} and \textit{pdoA2/pdoB2} genes detected in isolate 6PY1 from the high identity of the deduced amino acid sequences of the genes to the \textit{nidA} and \textit{nidB} genes previously characterized in isolate PYR-1.

Sequencing in these isolates revealed the order of the genes and the nature of adjacent genes (Fig. 1.6). In isolate PYR-1 one sequence was \textit{nidD-nidB-nidA-nidC}, where \textit{nidB} and \textit{nidA} respectively encode the \(\beta\)- and \(\alpha\)-subunits of the dioxygenase (74). The gene order \textit{nidB-nidA} also was observed with the KMS, MCS and JLS isolates studied by our group (99), the \textit{pdoB1-pdoA1} genes of isolate 6PY1 (84), the \textit{pdoB-pdoA-pdoC} cluster of isolate S65 (126), and the \textit{nid} cluster (\textit{nidB-nidA-nidC}) of isolate CH-2 (28) (Fig. 1.6). However, the gene sequence \textit{nidA} followed by \textit{nidB} also was identified in other DNA regions from these strains; \textit{nidA3-nidB3} for isolate PYR-1 and \textit{pdoA2-pdoB2}
FIG 1.6 Identified pyrene-degrading gene clusters in four *Mycobacterium* isolates, modified from Khan et al. (74), Krivobok et al. (84), Sho et al. (126), Kim et al. (78), and Churchill et al. (28). Genes *nidA*, *nidA3*, *nidX*, *pdoA*, *pdoA1*, *pdoA2*, and *pdoX* encode dioxygenase α-subunits; genes *nidB*, *nidB3*, *pdoB*, *pdoB1*, *pdoB2* encode dioxygenase β-subunits; genes *nidC*, *nidD*, *orf4*, *nidH*, and *pdoH* encode dehydrogenases; genes *phdF* and *pdoF* encode extradiol dioxygenases; genes *orf1*, *orfN4*, *orfP6*, and *orf72* encode unknown proteins; genes *nidR* and *araC* encode regulatory proteins. Genes encoding ferredoxin and reductase are also labeled.
for isolates 6PY1 and CH2 (Fig 1.6). Thus, genes encoding similar ring-hydroxylating dioxygenases were found in more than one gene order in several PAH-degrading mycobacterial strains.

By 2007 the pyrene-degrading gene clusters in PYR-1 (Fig 1.6) were located to large contigs from the partially finished genome (genes colored in black in Fig. 1.7) (79). After the genome was completed these genes were shown to be clustered within a 102 kb region (Fig. 1.7) (80). This region contains the genes, \textit{nidB2, nidB/nidA} and \textit{nidA3/nidB3} encoding pyrene/fluoranthene/naphthalene/phenanthrene ring-hydroxylating dioxygenase \(\alpha\)- and \(\beta\)-subunits (74, 78, 130), as well as genes relating to phenanthrene degradation, designated as \textit{phd} genes, and genes relating to phthalate degradation, designated as \textit{pht} genes (Fig. 1.7) (79). Thus the cluster contains all the genes encoding the enzymes required for the initial as well as the downstream steps in PAH transformation. In 2007, Kim et al. (79) finalized the pyrene degradation pathway in PYR-1 by correlating the structure of chemical intermediates with the enzymes detected by proteomics and the genes on the chromosome.

Although a proteomic study showed that pyrene induced the expression of many of these proteins in isolate PYR-1 (79), no studies were performed with PYR-1 at the transcriptional level to show whether induction of pyrene-degrading enzymes is due to transcriptional or post-transcriptional regulation. Also the initial studies of the genome of KMS indicated that gene duplication in this strain was more complex than in PYR-1 because of the existence of pyrene-degrading genes on the plasmids harbored in KMS.
Consequently, in KMS the question existed whether both chromosomal and plasmid copies of a gene were expressed.

**Pyrene-induced proteins in *Mycobacterium sp. strain KMS***. Using proteomics as a tool to understand pyrene degradation in isolate KMS, Liang et al. (90) identified seventeen proteins induced during growth on pyrene. Liang et al. (90) correlated the identity of these proteins with enzymatic functions in the transformation of pyrene through different intermediates as shown in Fig. 1.3. As shown in Table 1.3, these proteins included ring-hydroxylating dioxygenase α- and β-subunits, the components of the electron transport chains for dioxygenase, dehydrogenases, and a hydratase-aldolase (90).

At the time of Liang’s publication no correlations could be made between the induced proteins and transcription from putative gene loci for isolate KMS. Consequently, I initiated my studies with the goal of understanding the location of the genes induced by pyrene within the KMS genome. The findings are presented in Chapter 2 of this thesis along with other studies pertinent to the bioinformatics analysis that I performed. These studies led to a comparative analysis of pyrene-degrading gene islands between five isolates for which the full genomes were available.

**Degradation of MAHs and MAH derivatives by *Mycobacterium* isolates.** Degradation of reduced MAHs benzene, toluene, and xylene by *Mycobacterium* isolates has been reported but not at the level of functional genomics (19, 37, 129, 133). In some Gram-negative bacteria, initiation of benzene oxidation is catalyzed by benzene dioxygenase (153), indicating ring oxidation is an initial step. Catechol is detected as an
FIG 1.7 Clustering of genes involved in pyrene-degradation in the genome of isolate PYR-1. Contigs containing the clusters of genes involved in pyrene degradation in PYR-1 genomic DNA are shown in black as published by Kim et al. (79). Positioning of these genes within the full genome sequence as published by Kim et al. (80) is shown with color annotation of important gene functions: red represents PAH catabolism genes; pink transcriptional regulator genes; yellow, DNA mobilization genes; green, membrane transporter genes; and white, genes encoding hypothetical proteins. Gene locus numbers (Mvan_) are shown above the gene sequence. Numbers provided in the genes denoted with a black color correlate to the enzymes corresponding to the steps presented in the pathway illustrated in Fig.1.4. The stars show genes corresponding to proteins detected as being pyrene-induced in the proteomic analysis of Kim et al. (79).
intermediate from oxidation of both phenol and benzene (36, 40). However, a *Mycobacterium* isolate transformed benzene into phenol and hydroquinone (19, 37) suggesting a different pathway but one in which oxygenases are used.

The oxidized MAH derivatives phthalate and protocatechuicate, as well as β-ketoadipate, are intermediates formed during pyrene degradation by isolate PYR-1 (Fig 1.4) (79, 90) leading finally to acetyl-CoA and succinyl-CoA (79). Protocatechuicate, catechol and β-keto adipyl-CoA also are demonstrated products in the degradation of the benzoate by fungi and bacteria (Fig. 1.8).

The MAH benzoate occurs naturally in plants (107) and is likely to be found in soils containing decaying plant material through its generation from the degradation of lignin by fungi (134). Complete benzoate degradation pathways are proposed for a broad range of bacteria and fungi that inhabit soils. The bacteria include *Pseudomonas, Burkholderia, Rhodococcus, Acinetobacter*, and *Azoarcus* (33, 39, 45, 89, 98, 112) (Fig. 1.8). One pathway (Fig. 1.8A) involves an initial ring oxidation step, catalyzed by a benzoate ring-hydroxylating dioxygenases. This pathway is found in both bacteria and fungi (33, 39, 45, 89, 98, 112). However, protocatechuicate is identified as an intermediate in fungi whereas catechol is found in bacteria (Fig. 1.8A) (33, 39, 45, 89, 98, 112). The second pathway, the benzoyl coenzyme A pathway, involves benzoyl-CoA formation as the initial step catalyzed by benzoate-CoA ligase (Fig. 1.8B). This pathway, which also generates β-ketoadipate, is found in *Burkholderia* and *Azoarcus* (34, 46). Acetyl-CoA and succinyl-CoA are the final products from each of these benzoate-degrading pathways (33, 39, 45, 89, 98, 112). Oxidation of benzoate by pathogenic *Mycobacterium* species is
reported with catechol as an intermediate (42, 43). A later study found that β-ketoadipate is an intermediate from the oxidation of catechol by *M. butyricum* (127). These studies suggest that mycobacteria probably use a catechol pathway dependent on initial dioxygenase-ring oxidation.

This possibility was strengthened by the funding for isolate 6PY1 that a dioxygenase was involved. Krivobok et al. (84) found induction of *pdoA1* and *pdoB1* encoding ring-hydroxylating dioxygenase subunits by benzoate as well as by pyrene. However, isolate PYR-1 does not degrade benzoate (80), although the chemical structure of benzoate is close to phthalate and protocatechuate that are intermediates in pyrene degradation. Kim et al. (80) deduced that the genome of *M. vanbaalenii* PYR-1 lacks genes encoding an effective benzoate dioxygenase. Consequently, I extended the studies of PAH/MAH degradation with isolate KMS to determine whether like PYR-1 it lacked ability to use benzoate or whether like 6PY1 it was able to metabolize this oxidized MAH. I continued these studies to examine the basis for differential induction of dioxygenase genes in isolate KMS.

**Key pathways for carbon metabolism in *Mycobacterium* isolates.** All cells when growing on C1-C5 compounds require intermediary metabolism to generate the structures, proteins, carbohydrates, lipids, and nucleic acids required for cell growth. Twelve key metabolites must be generated by cells to permit cell growth (22). Table 1.4 shows how these intermediates act as precursors for the cell components and names the pathways that are involved in their generation.
FIG 1.8 Two main benzoate degradation pathways, the ring-cleavage pathway (A) and benzoyl-CoA pathway (B), utilized by bacteria and fungi, modified from Gescher et al. (45). Names of key intermediates discussed in the text are provided with red lettering.

TABLE 1.4 Twelve key metabolites needed for biosynthesis of all cell components

<table>
<thead>
<tr>
<th>Metabolite and chemical structure</th>
<th>Product synthesized from the metabolite</th>
<th>Pathway involved in synthesis of product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetyl-CoA</td>
<td>Fatty acids</td>
<td>Glycolysis/pyrene</td>
</tr>
<tr>
<td>CH₃-CO-S-CoA</td>
<td></td>
<td>degradation in KMS</td>
</tr>
<tr>
<td>Succinyl-CoA</td>
<td>Heme and TCA cycle</td>
<td>TCA cycle</td>
</tr>
<tr>
<td>COOH-(CH₂)₂-CO-S-CoA</td>
<td>intermediates</td>
<td></td>
</tr>
<tr>
<td>Pyruvate</td>
<td>Pyruvate family amino</td>
<td>Glycolysis/pyrene</td>
</tr>
<tr>
<td>Chemical Formula</td>
<td>Description</td>
<td>Pathway</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>CH$_3$-CO-COOH</td>
<td>acids (Ala, Val, Leu, Ile, and Lys)</td>
<td>degradation in KMS</td>
</tr>
<tr>
<td>$\alpha$-ketoglutarate</td>
<td>Glutamate family amino acids</td>
<td>TCA cycle</td>
</tr>
<tr>
<td>COOH-(CH$_2$)$_2$-CO-COOH</td>
<td>acids (Glu, Gln, Arg, and Pro), glutathione, purine, pyrimidine, and TCA cycle intermediates</td>
<td></td>
</tr>
<tr>
<td>Oxaloacetate (OAA)</td>
<td>Aspartate family amino acids</td>
<td>TCA cycle</td>
</tr>
<tr>
<td>COOH-CH$_2$-CO-COOH</td>
<td>acids (Asp, Asn, Met, and Thr), purine, pyrimidine, and TCA cycle intermediates</td>
<td></td>
</tr>
<tr>
<td>Phosphoenolpyruvate (PEP)</td>
<td>Aromatic amino acids (Phe, Tyr, and Trp)</td>
<td>Glycolysis/Gluconeogenesis</td>
</tr>
<tr>
<td>CH$_2$=COP-COOH</td>
<td>Serine family amino acids (Ser, Gly, and Cys) and glycerol</td>
<td>Glycolysis/Gluconeogenesis</td>
</tr>
<tr>
<td>3-phosphoglycerate</td>
<td>N-acetylglucosamine, N-acetylmuramic acid, monosaccharides, disaccharides, and</td>
<td>Glycolysis/Gluconeogenesis</td>
</tr>
</tbody>
</table>
For *Mycobacterium* isolates there is a special need for C6- carbon sources because the complex thick cell walls require a high carbon input. A typical cell wall structure of *Mycobacterium* (Fig. 1.9) contains a layer of peptidoglycan beyond the cell membrane. The peptidoglycan backbone is covalently attached to arabinogalactan, a branched-chain polysaccharide consisting of a proximal chain of galactose units linked to a distal chain of arabinose subunits. The termini of arabinogalactan are esterified with mycolic acids (Fig. 1.9). The upper layer contains the lipoarabinomannan, which is not covalently linked to the arabinogalactan-peptidoglycan complex, but is attached to the cytoplasmic membrane by a phosphatidylinositol anchor (11, 16, 17). The components of the *Mycobacterium* cell wall such as mycolic acids and furan-rings of carbohydrates are carbon-rich so that

<table>
<thead>
<tr>
<th>Glucose-6-phosphate</th>
<th>Monosaccharides, disaccharides, and polysaccharides</th>
<th>Glycolysis/Gluconeogenesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHO-(CHOH)₄-CH₂O-P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ribose-5-phosphate</td>
<td>Purine and histidine (His)</td>
<td>Pentose phosphate pathway</td>
</tr>
<tr>
<td>CHO-(CHOH)₃-CH₂O-P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erythrose-4-phosphate</td>
<td>Aromatic amino acids</td>
<td>Pentose phosphate pathway</td>
</tr>
<tr>
<td>CHO-(CHOH)₂-CH₂O-P</td>
<td>(Phe,Tyr,and Trp), NAD,</td>
<td></td>
</tr>
<tr>
<td>Triose-3-phosphate</td>
<td>Dihydroxyacetone and</td>
<td>Pentose phosphate pathway</td>
</tr>
<tr>
<td>CHO-CHOH-CH₂O-P</td>
<td>glycerol-3-phosphate</td>
<td></td>
</tr>
</tbody>
</table>
species in *Mycobacterium* must have an efficient system for shuttling carbon into these wall structures.

**FIG 1.9** Schematic representation of the mycobacterial cell wall (11).

Inspection of the complete pathway for pyrene degradation in mycobacterium (79, 80) predicts the generation of CO$_2$, acetyl-CoA, pyruvate and succinyl-CoA which then must be utilized for growth (see discussion relative to Fig 1.4). Analysis of $^{13}$C-labeled pyrene degradation by isolate KMS using nuclear magnetic resonance (NMR) implies that pyrene degradation happens inside the cells and also reveals that the majority of pyrene metabolites are incorporated into cellular material (105). The C2 compound acetyl-CoA and the C4 compound succinyl-CoA are generated from complete benzoate degradation, independent of the pathway used (33, 39, 45, 89, 112).

Acetyl-CoA and succinyl-CoA metabolism are both intermediates of the TCA cycle. Acetyl-CoA and succinate also participate in the glyoxylate shunt through the enzymes
isocitrate lyase and malate synthase (80). Early findings with *E. coli* show the essential role of the glyoxylate shunt for assimilation of acetate (83). However, analysis of the expression of genes encoding isocitrate lyase and malate synthase in pathogenic *Mycobacterium* species shows the shunt is not only for growth on acetate and fatty acids but also is involved for growth on glucose (58). These findings suggest that in mycobacteria the glyoxylate shunt has a broader role than just metabolism when C2 compounds are provided for growth. In pathogenic *M. avium* and *M. tuberculosis*, two genes, *icl* and *aceA*, encoding proteins with isocitrate lyase activity are both expressed during growth on glucose (58) but both are repressed by the C4 compound succinate (58). The *aceB* gene encoding malate synthase would complete the pathway and is present in the mycobacterial pathogens’ genomes. For the PAH-degrading strain PYR-1, AceB and AceA but not Icl are expressed with growth on sorbitol (80), suggesting that expression of the pathway is modified from that in the pathogenic isolates.

Operation of the TCA cycle requires the regeneration of OAA and this step normally involves malate transformation catalyzed by a cytoplasmic NAD$^+$-linked malate dehydrogenase (143). In isolate PYR-1, malate dehydrogenase is expressed when the cells are grown on sorbitol (80). However, when pyrene is added with sorbitol, no malate dehydrogenase was detected (80) suggesting that an alternative enzyme must catalyze the malate-OAA interconversion. Other microbes possess an additional enzyme malate:quinone oxidoreductase, which is membrane bound and uses a quinone involved in the electron transport chain as the cofactor. This enzyme only catalyzes the one-directional reaction from malate to OAA, whereas the reaction catalyzed by NAD$^+$-linked
malate dehydrogenase is determined by the ratio of $\text{NAD}^+/\text{NADH}$ (101). Malate:quinone oxidoreductase appears to function alone in *Helicobacter pylori* (70) or together with malate dehydrogenase in *E. coli* and *Corynebacterium glutamicum* (101, 143).

For growth on carbon sources other than C6 compounds, gluconeogenesis is required to supply the C3 compound PEP from which C5 and C6 sugar skeletons can be built. The pathways used for the conversion of the C4, TCA cycle intermediates malate and OAA into the C3 compounds, pyruvate and PEP required for gluconeogenesis have not been reported for *Mycobacterium* isolates.

Two pathways through which C3 compounds are generated are found in both Gram-positive and Gram-negative bacteria, such as *E. coli*, *Bacillus subtilis* and *C. glutamicum* (123). In one pathway, PEP is generated directly from OAA catalyzed by PEP carboxykinase; and in the second pathway, PEP is generated from malate via pyruvate catalyzed by malic enzyme followed by PEP synthase, as illustrated in Fig. 1.10.

I focused on the pathways in isolate KMS for intermediary metabolism involved in the interconversion of C4 compounds and transformations between C4 and C3 compounds. I analyzed genes involved in the crucial metabolic steps relating to generation of PEP for gluconeogenesis from acetyl-CoA and succinyl-CoA, both final products from degradation of pyrene and benzoate (Fig. 1.11). I asked the following questions: 1) Is there differential activation of genes for the glyoxylate shunt enzymes during growth on different carbon sources in isolate KMS? 2) What genes encode enzymes involved in the malate-OAA interconversion? and 3) What genes are involved in the generation of C3 compounds for gluconeogenesis from C4 TCA intermediates?
These questions are not only important for the growth of KMS on pyrene but also during growth on other substrates that the bacterium may encounter in the soil habitat.

**Alternative habitats for environmental mycobacterium isolates.** Previous studies with isolate KMS showed, unexpectedly, that it is a competent colonizer of plant rhizospheres (25). The bacterium colonized plant roots competitively when co-inoculated with a pseudomonad isolate (25). Growth on plant root exudates is observed with utilization of the sugars present (25). Recently *M. gilvum* isolates also are shown to inhabit plant root rhizospheres (139). Interaction between root exudates and bacteria is found suggesting that root exudates are utilized by the *M. gilvum* isolates (139).

In addition to sugars such as fructose and glucose (25), organic acids such as acetate and succinate are common components in the root exudates (13). Another carbohydrate, mannitol, is produced by plant and fungi for storage and as stress response metabolite (88, 151). Thus, mannitol together with benzoate, which is generated from lignin degradation by fungi (88, 134), may be included in the many metabolites presented to mycobacteria as carbon sources in the rhizosphere.

Transport of these varied compounds into the KMS cells differs. For the soil isolate *M. smegmatis*, fructose and glucose are transported through a phosphoenolpyruvate dependent phosphotransferase transport system (PTS) whereas sugar alcohols, such as mannitol, are transported through an ATP-binding cassette (ABC) system (137). Consequently growth on the sugars may place more of a demand for PEP on the cell than growth on mannitol.
Pyrene mineralization is detected from plant rhizospheres colonized by isolate KMS and during culture on barley root washes (25). However, more pyrene is mineralized by KMS grown on 1/10 diluted barley root wash than on full strength root wash (25). This finding suggested that catabolite repression may operate in isolate KMS. With other microbes such as *Pseudomonas*, utilization of the PAH phenanthrene is repressed by substrates in root exudates such as organic acids, sugars, and amino acids (118). However Toyama et al. (139) suggest that the accelerated degradation of pyrene by rhizosphere-inhabiting *M. gilvum* isolates could be due to induction of pyrene-degrading enzymes. An alternative view is that the root exudates increase the populations of the microbes so that more degradation could take place.

**FIG 1.10** Potential pathways for transformation of C4 TCA cycle intermediates malate and OAA to C3 pyruvate and PEP for gluconeogenesis modified from Sauer et al. (123). Enzymes are shown are: MDH: malate dehydrogenase, MQO: malate:quinone oxidoreductase, MAE: malic enzyme, PEPck: PEP carboxykinase, PPS: PEP synthase.
FIG 1.11 Proposed pathways for intermediary metabolism from carbon sources generated from complete pyrene and benzoate degradation in isolate KMS. The products, acetyl-CoA and succinyl-CoA, from degradation of both pyrene and benzoate are noted by stippled boxes, and pyruvate also released during pyrene degradation, is denoted by a grey box. Key metabolites are underlined. The locus tags for isolate KMS for genes encoding enzymes in the glyoxylate shunt (isocitrate lyase and malate synthase) and those involved in completion of the TCA cycle and C4-C3 exchanges are shown. OAA, oxaloacetate; PEP, phosphoenolpyruvate; 3-PG, 3-phosphoglycerate; F-6-P, fructose-6-phosphate; G-6-P, glucose-6-phosphate. Not all intermediates are shown.
Catabolite repression in *Mycobacterium*. When more than one carbon source is available to a bacterium, these sources are either co-catabolized or there is a preferred carbon source (32). Catabolite repression happens when utilization of the preferred carbon source inhibits the utilization of the other carbon sources. Catabolite repression dependent on cyclic AMP (cAMP) has been studied in *Enterobacteriaceae* such as *E. coli* and *Serratia marcescens* (66, 104).

In *E. coli* the preferred carbon source, glucose, is transported into the cell through a phosphoenolpyruvate-dependent phosphotransferase transport system (PTS) (Fig. 1.12) (85). A series of cytoplasmic and membrane proteins transfer the phosphate group from phosphoenolpyruvate (PEP) to the sugar including enzyme I (EI), histidine-containing phosphocarrier protein (HPr), and the glucose-specific enzyme IIA (EIIA\(^{Glc}\)). In *E. coli*, a low glucose leads in a high level of the phosphorylated PTS protein EIIA\(^{Glc}\) that activates the adenylate cyclase for generation of cyclic AMP (cAMP) from ATP (14). The cAMP interacts with a cAMP receptor protein (CRP) to form a complex (111). Gene regulation then involves binding of the cAMP-CRP complex to discrete promoter sites to influence RNA polymerase activity (117). Sequences, such as TGTGA-\(N_6\)-TCACA, are detected for binding cAMP-CRP in the promoters of the cAMP-regulated genes and account for both up- and down-regulation for genes required for utilization of carbohydrates other than glucose (61, 131). The crux of the cAMP-regulated system in *E. coli* is the mechanism permitting activation of adenylate cyclase activity.
Genes encoding adenylate cyclase, which catalyzes the generation of cAMP from ATP, are abundant in pathogenic *Mycobacterium* species as well as genes encoding phosphodiesterase for degradation of cAMP. Regulation of gene expression involving cAMP and a cAMP receptor protein (CRP) complex is suggested to function in *Mycobacterium* (9, 106). Genes encoding proteins with the potential to bind cAMP also are identified in mycobacteria (10) consistent with the functioning of cAMP as a signaling compound. Gene regulation by cAMP-CRP is linked with the pathogenicity of disease-causing mycobacteria (10). One of the cAMP-induced genes under low oxygen conditions encodes a malate dehydrogenase (44), which is involved in the TCA cycle indicating that cAMP levels may modify this important pathway in intermediary metabolism in a mycobacterium.
Analysis of an environmental *Mycobacterium* isolate JC1 DSM3803 suggests that glucose repression of a carbon monoxide dehydrogenase gene cluster (*cut*) involves cAMP regulation (106). A typical palindromic cAMP-CRP binding motif is found between the transcriptional start site and translational start codon of the *cut* gene cluster. This motif is conserved in several soil *Mycobacterium* isolates, including isolate KMS (106), suggesting that cAMP-CRP is a possible regulator of metabolism in KMS.

With this background I raised the questions of whether: 1) There is catabolite repression of the degradation of the PAH pyrene by components of root exudates such as organic acids and sugars; 2) Utilization of the MAH benzoate from degradation of phenolic polymers in the rhizosphere influences pyrene degradation; and 3) cAMP-CRP is possibly involved in gene regulation involved in benzoate and pyrene utilization.

**OBJECTIVES TO BE STUDIED**

Based on the background described in the Introduction, four objectives were proposed for my PhD research project:

Objective 1. To identify in the genome of isolate KMS the distribution of the gene loci that encode the seventeen pyrene-induced proteins identified in a previous proteomic study by Liang et al. (90). To use analysis of transcript abundance to determine the extent to which transcriptional control was involved in expression of the pyrene-induced proteins in KMS. To use synteny analysis to reveal the relationships between the pyrene-inducing islands in isolate KMS to those determined by Cerniglia’s group in isolate PYR-1 and three other PAH-degrading strains, PYR-GCK, JLS and KMS. My findings are discussed in Chapter 2.
Objective 2. To examine the array of dioxygenase genes in isolate KMS and by phylogenetic analysis determine their relationships based on nucleic acid and amino acid sequences as well as to look at the diversity of the promoter sequences for the dioxygenase genes. To find whether isolate KMS can grow on benzoate as the sole carbon source and to predict the genomic basis for benzoate utilization. I determined whether pyrene dioxygenase genes are induced by both benzoate and pyrene. My findings are discussed in Chapter 3.

Objective 3. To determine whether isolate KMS has a preferred carbon source among compounds present in the soil and plant rhizosphere and to examine its use with respect to utilization of benzoate or pyrene. To determine the correlation between growth and expression of key dioxygenase genes for pyrene and benzoate utilization. To determine whether cAMP-regulation was involved in differential expression of dioxygenase genes for benzoate and pyrene. This work is presented in Chapter 4.

Objective 4. To explore intermediary metabolism in isolate KMS. To determine the status of expression for genes involved in the glyoxylate shunt and for malate-OAA interconversion when isolate KMS is cultured on different carbon sources. To deduce possible conversion pathways for generation of the C3 compounds pyruvate and PEP when isolate KMS is grown on pyrene or benzoate or other carbon sources requiring gluconeogenesis. This work is presented in Chapter 5.

REFERENCES


31:3291-3307.


60. **Hyun S, Jafvert CT, Lee LS, Rao PS.** 2006. Laboratory studies to characterize the efficacy of sand capping a coal tar-contaminated sediment. Chemosphere 63:1621-


82. **Kleespies M, Kroppenstedt RM, Rainey FA, Webb LE, Stackebrandt E.** 1996. *Mycobacterium hodleri* sp. nov., a new member of the fast-growing mycobacteria


98. **Middelhoven WJ.** 1993. Catabolism of benzene compounds by ascomycetous and basidiomycetous yeasts and yeastlike fungi. A literature review and an experimental


116. **Rehmann L, Prpich GP, Daugulis AJ.** 2008. Remediation of PAH contaminated soils:


J. Environ. Monit. 7:1378-1387.


CHAPTER 2
POLYCYCLIC AROMATIC HYDROCARBON-DEGRADING GENE ISLANDS IN FIVE PYRENE-DEGRADING MYCOBACTERIUM ISOLATES FROM DIFFERENT GEOGRAPHIC LOCATIONS\textsuperscript{1}

Abstract: Mycobacterium sp. strain KMS utilizes pyrene, a high-molecular weight polycyclic aromatic hydrocarbon (PAH), as a sole carbon source. Bioinformatic analysis of the genome of isolate KMS predicted 25 genes with the potential to encode seventeen pyrene-induced proteins identified by proteomics; these genes were clustered on both the chromosome and a circular plasmid. RT-PCR analysis of total RNA isolated from KMS cells grown with or without pyrene showed that the presence of pyrene increased the transcript accumulation of 20 of the predicted chromosomal and plasmid located genes encoding pyrene-induced proteins. The transcribed genes from both chromosome and circular plasmid were within larger regions containing genes required for PAH degradation constituting PAH-degrading gene islands. Genes encoding integrases and transposases were found within and outside the PAH-degrading gene islands. The lower GC content of the genes within the gene island (61-63 %) compared with the average

\textsuperscript{1}Coauthored by Dr. Anderson AJ., Polycyclic aromatic hydrocarbon-degrading gene islands in five pyrene-degrading Mycobacterium isolates from different geographic locations, © 2012 with permission from NRC Research Press.
genome content (68 %) suggested that these mycobacteria initially acquired these genes by horizontal gene transfer. Synteny was detected for the PAH-degrading gene island in the genomes of two additional Mycobacterium isolates from the same PAH-polluted site and of two other pyrene-degrading Mycobacterium from different sites in the United States of America. Consequently the gene islands have been conserved from a common ancestral strain.

Introduction

Polycyclic aromatic hydrocarbons (PAHs) are formed naturally during the incomplete burning of organic substances and also purposely for industrial needs (U.S. Department of Health and Human Services 1995). The recalcitrance of PAHs increases with the number of rings and angularity of the structure (Zander 1983; Harvey 1997). High molecular weight PAHs are listed by the EPA as being among the most persistent, bioaccumulative, and toxic pollutants (Kanaly and Harayzma 2000). However, mineralization of complex PAHs is catalyzed by microbes (Cerniglia 1992). For instance, environmental mycobacterial isolates degrade the four-aromatic ring PAH, pyrene, and the five-ring benzo[a]pyrene. Pathways for pyrene degradation have been resolved for Mycobacterium vanbaalenii strain PYR-1, obtained from a PAH-polluted site in Texas, USA (Kim et al. 2007; Kim et al. 2008). Analysis of the PYR-1 genome reveals 194 chromosomal genes associated with PAH utilization with most genes being clustered in
two regions, one of 150 kb and another 31 kb (Kim et al. 2008). Key genes involved in
the initial stages of pyrene degradation are those encoding α- and β-subunits of aromatic
ring-hydroxylating dioxygenases (Kim et al. 2006; Kweon et al. 2010). The genes,
termed nidA/nidB, nidB2 and nidA3/nidB3, are found in the larger 150 kb cluster (Kim et
al. 2008). The α-subunits of aromatic ring-hydroxylating dioxygenases NidA and NidA3
have highly related amino acid sequences (Kweon et al. 2010).

*Mycobacterium sp.* strain KMS from another PAH-contaminated site also, like strain
PYR-1, mineralizes the higher ring PAHs, pyrene and benzo[a]pyrene (Miller et al.
2004). Unlike PYR-1 that lacks plasmids, KMS possesses both a circular (293 genes) and
a linear plasmid (235 genes) in addition to the chromosome (5561 genes)
(http://img.jgi.doe.gov/cgi-bin/w/main.cgi). Chemical identification of intermediates of
pyrene degradation by isolate KMS and proteomic analysis allowed the prediction of a
potential degradative pathway (Liang et al. 2006). Pathway enzymes included
dioxygenases, along with the proteins required for electron transfer, a reductase and iron-
sulfur protein, dihydrodiol dehydrogenase, a quinone reductase, intradiol dioxygenase
and a quinone oxidase (Liang et al. 2006).

In this study, we performed bioinformatic analysis to deduce the likely locations of
genes encoding pyrene-induced proteins in isolate KMS. We report that the pyrene-
induced genes were clustered on the chromosome but a subset also was duplicated on the
circular plasmid in isolate KMS. We performed reverse transcriptase-PCR (RT-PCR)
using primers for specified genes to explore the issue of whether genes from both the chromosome and plasmid of isolate KMS were induced during pyrene degradation to correlate with the proteomic analysis (Liang et al. 2006). We extended the bioinformatic analysis to four other fully-sequenced genomes of PAH-degrading mycobacteria including two strains from the same contaminated site as KMS and two isolates from other USA locations, *M. gilvum* PYR-GCK and *M. vanbaalennii* PYR-1. This analysis for the first time reveals the extensive synteny in regions of genes concerned with PAH degradation suggesting conservation from a common ancestor.

**Materials and methods**

**Bacterial strain and growth conditions**

*Mycobacterium* *sp.* strain KMS cells from 0.5 mL of a 15% glycerol stock were transferred to 200 mL of a modified basal salts medium (BSM⁺) (Liang et al. 2006) in a 500 mL flask to generate a cell concentration of 10⁶ cfu/mL. Cultures were grown at 25 °C at 220 rpm for 7 days when the OD 600 nm exceeded 0.2 but cells were still in logarithmic growth phase. The cells were pelleted by centrifugation at 3,200 g, and washed once with sterile, distilled, and deionized water (DDW) before suspension in 6 ml sterile DDW. This suspension was used as inoculum for 1 liter BSM⁺ medium with or without 20 mg pyrene (Sigma-Aldrich Corporation, MO, USA) at a final concentration 100 µmol/L. Pyrene was dissolved in methanol at 4 mg/mL and 5 mL were added to the
flask. The same volume 5 mL methanol was added to the control flask. All of the methanol was evaporated in a sterile transfer hood before sterile BSM medium was added and the media inoculated.

**Identification of the potential gene loci for pyrene-induced proteins**

We deduced by using BLAST (http://www.ncbi.nlm.nih.gov/) the potential genomic loci most likely to encode for the pyrene-induced proteins in our earlier proteomic study (Liang et al. 2006). When these predictions showed more than one potential gene, we used the CLUSTALW program on the San Diego supercomputer website (http://workbench.sdsc.edu) to identify unique nucleotides for each potential gene. The position of these genes on plasmids or on the chromosome in isolate KMS was determined based on the information at Joint Genome Institute, Integrated Microbial Genomes website (http://img.jgi.doe.gov/cgi-bin/w/main.cgi). This database was used to obtain the order of genes in the genomes for two other PAH-degrading isolates from the same site, isolates MCS and JLS, and two from other sites, *M. vanbaalenii*, isolate PYR-1, and *M. gilvum*, isolate PYR-GCK.

**Authenticity of primer sets**

To examine expression from discrete loci, specific primers were designed (Table 2.1), using the Primer3 software (http://frodo.wi.mit.edu/primer3/). Each primer was examined using BLAST to ensure a single match in the KMS genome. The unique nucleotide sequences specific for genes with high homology were used at the 3′ end of
primers. Primer authenticity was evaluated by performing PCR with genomic DNA from isolate KMS to confirm product with the anticipated size and sequence. PCR reaction mixtures contained in 50 µL: 30.8 µL H₂O, 8 µL MgCl₂ (25 mM), 5 µL 10× reaction buffer, 2 µL dNTP mix (10 mM each), 1 µL forward primer (final concentration was 1 µM), 1 µL reverse primer (final concentration was 1 µM), 2 µL KMS genomic DNA (0.2 µg), 0.2 µL Taq polymerase (1 unit) (Fermentas Inc., MD, USA). The PCR program used was: 94 ºC for 2 min; 30 cycles at 94 ºC for 30 sec, 58 ºC for 40 sec, 72 ºC for 1 min; 72 ºC for 10 min and holding at 4 ºC.

Table 2.1. Specific primers for genes potentially encoding the pyrene-induced proteins and the 16S rRNA genes for strain KMS.

<table>
<thead>
<tr>
<th>Gene locus</th>
<th>Corresponding protein</th>
<th>Primer sequences</th>
<th>Product length</th>
<th>Annealing temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>For ATCGAAAGCATCGACGAAAC</td>
<td>160 bp</td>
<td>58 ºC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rev ACAACCGACAGGAATGACAGT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_1662</td>
<td>P1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rev AGCGCGGTTTCATCGTCAAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_5623</td>
<td>P2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rev AGCGCGGTTTCATCGTCAAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_1673</td>
<td>P3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rev AGCGCGGTTTCACCGTGAA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_5631</td>
<td>P3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Primer</td>
<td>For Sequence</td>
<td>Length (bp)</td>
<td>Temperature (°C)</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>--------------</td>
<td>-------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>kms_1674</td>
<td>P4</td>
<td>GATCACCGATGGTGAGGCCG</td>
<td>76</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev GGTTGCAGTCAGAAGCTGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_5632</td>
<td>P5</td>
<td>GATCACCGATGGTGAGGCCA</td>
<td>72</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev GCAGTCAGAACGAAGCCAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_5625</td>
<td>P6</td>
<td>TGTTCCGGGAGGCGGAGCTC</td>
<td>93</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev TGGTCGACCAGGAGACATAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_1660</td>
<td>P7</td>
<td>TCGATGGGGGGCAGTGTCGA</td>
<td>339</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev CGTCGACCCGTTCGGCGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_1667</td>
<td>P7</td>
<td>GGGTGAGGTGCGATCGGCGAT</td>
<td>147</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev CCGGTCCGGCAACCCTGCGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_1667</td>
<td>P7</td>
<td>GGAACGACGTACACGGGTTG</td>
<td>177</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev CCGGCTGCGGAAAGACAGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_1648</td>
<td>P8</td>
<td>CGTTACCGTTGCGACTCGGTCT</td>
<td>221</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev CGCGGAATGACTGACTTGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_1650</td>
<td>P9</td>
<td>GAGTACGACACCGTGATCGGGA</td>
<td>184</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev CCTGAATAACCGGTACCAAGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_1672</td>
<td>P10</td>
<td>CTGCATGGTGCGATCGGGGGA</td>
<td>176</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev CATAGGGCGGTGCTGACCAAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>P</td>
<td>For</td>
<td>Rev</td>
<td>bp</td>
</tr>
<tr>
<td>--------</td>
<td>----</td>
<td>------------</td>
<td>------------</td>
<td>-----</td>
</tr>
<tr>
<td>kms_5630</td>
<td>10</td>
<td>CTGCATGGTGCATGACGGGG</td>
<td>CATAGGGCTCGCTGACCAGC</td>
<td>176</td>
</tr>
<tr>
<td>kms_1671</td>
<td>11</td>
<td>GCAGGTCATGGCCATGGCTT</td>
<td>AGTCCAGTGGTCTCGCCCA</td>
<td>97</td>
</tr>
<tr>
<td>kms_5629</td>
<td>11</td>
<td>GGCATGGTTTCTCGCCACCGC</td>
<td>GACGACGTCAGCTTCTATGGA</td>
<td>279</td>
</tr>
<tr>
<td>kms_1686</td>
<td>12</td>
<td>TTTAGCATTCCACGTGC</td>
<td>ACTGATCGAATCGGTTTGG</td>
<td>196</td>
</tr>
<tr>
<td>kms_5635</td>
<td>12</td>
<td>TTTTGCATCGACGTCGCT</td>
<td>ACTGATCGAATCGGTTTGG</td>
<td>196</td>
</tr>
<tr>
<td>kms_1668</td>
<td>13</td>
<td>ACCGCGCACTTCCAATGCT</td>
<td>GCCGCCGGGATAGGCCTCC</td>
<td>96</td>
</tr>
<tr>
<td>kms_5626</td>
<td>14</td>
<td>CACCGCGCACTTCCAATGCC</td>
<td>GCCGCCGGGATAGGCCTC</td>
<td>97</td>
</tr>
<tr>
<td>kms_1661</td>
<td>15</td>
<td>GGCATTCGCAATGTATAGTC</td>
<td>ACTTCAGCCACTGATATTG</td>
<td>248</td>
</tr>
<tr>
<td>kms_5622</td>
<td></td>
<td>ACTTCAGCCACTGATATTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kms_1658</td>
<td>16</td>
<td>AACGCGGTCCAGTACTACAA</td>
<td>AGACGAGGACCGACATCTTT</td>
<td>220</td>
</tr>
</tbody>
</table>
Isolation and purification of total RNA

After 7 days of growth, cultures of *Mycobacterium* sp. strain KMS were centrifuged at 3,200 g and the cells were washed once with DEPC-treated water. Cells, about 0.05 g wet weight, were transferred to a 2 mL screw-cap tube (BioSpec Products, Inc. Bartlesville, OK, USA) with 0.5 mL 0.1 mm Zirconia/Silica Beads (BioSpec Products, Inc. Bartlesville, OK, USA) and suspended in 1 mL TRI-reagent (Molecular Research Center, Inc. OH, USA). The cell walls were broken using a bead-beater for 5 min (BioSpec Products, Inc. Bartlesville, OK, USA). After centrifugation at 5,000 g for 30 sec, the supernatant was transferred to a new 1.5 mL tube. Chloroform (0.2 ml) was added and vortexed vigorously, and the mixture was maintained on ice for 5 min. The mixture was centrifuged at 12,000 g for 15 min. The supernatant was transferred to a new tube and 0.5 mL cold isopropanol was added with mixing. After 10 min at 22 °C, the
mixture was centrifuged at 12,000 g for 8 min. The supernatant was discarded and the white pellet of RNA was washed twice by 80% ethanol and centrifuged at 7,500 g for 5 min at room temperature. After brief air drying, the pelleted RNA fraction was dissolved in DEPC-treated water and kept at -70 °C for future use.

**End-point reverse-transcriptase (RT) PCR**

Total RNA was treated with RNase-free DNase (Promega Corporation, Madison WI, USA) to remove the contaminating DNA before the RT-PCR. The concentration of purified RNA was measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, DE, USA). An end-point PCR was performed using 1 µL of the DNase-treated RNA with the 16S rRNA primers to ensure no DNA contamination. First-strand cDNA was synthesized by following the protocol for M-MuLV reverse transcriptase (Fermentas Inc., MD, USA). The same amount (1 µg) of non-induced and pyrene-induced RNA samples was treated in final reaction volumes of 20 µL. The synthesized cDNA was used directly in the RT-PCR analyses.

The PCR reaction mixtures contained in 20 µL: 9.8 µL H₂O, 4 µL MgCl₂ (25 mM), 2 µL 10x reaction buffer, 1 µL dNTP mix (10 mM each), 1 µL forward primer (final concentration 1 µM), 1 µL reverse primer (final concentration 1 µM) (Table 2.1), 1 µL cDNA (0.05 µg), and 0.2 µL Taq polymerase (Fermentas Inc., MD, USA). The PCR program was: 94 °C for 2 min; 30 cycles at 94 °C for 30 sec, 55 °C or 58 °C depending on primer set for 40 sec (Table 2.1), 72 °C for 1 min; 72 °C for 10 min and holding at 4 °C.
An Eppendorf Mastercycler gradient machine was used (Certified GeneTool, Inc, CA, USA). The PCR products (20 μL) were loaded to a 2% agarose gel. Each gel was examined by using a UV transilluminator (UVP Inc., CA, USA) and their images were recorded photographically (AEase™ version 3.25, A Innotech Corporation, CA, USA). The intensities of the bands were read by Quantity-One software from BIORAD.

**Synteny analysis of the genomes of five pyrene-degrading isolates**

The Mauve software from the University of Wisconsin-Madison, (ftp://ftp.ncbi.nih.gov/genomes/), was used for synteny alignment of chromosomes and plasmids. The genome databases were obtained from the NCBI GenBank with the accession numbers as follows: isolate KMS chromosome, NC_008705; isolate KMS circular plasmid, NC_008703; isolate KMS linear plasmid, NC_008704; isolate MCS chromosome, NC_008146; isolate MCS linear plasmid, NC_008147; isolate JLS chromosome, NC_009077; isolate PYR-1 chromosome, NC_008726; isolate PYR-GCK chromosome, NC_009338.

**GC content of the pyrene-degrading islands on the chromosomes and plasmids**

Nucleotide sequences of all the pyrene-degrading islands on the chromosomes and plasmids of five mycobacterium isolates were generated by Artemis (Rutherford et al. 2000) from Sanger Institute (http://www.sanger.ac.uk/resources/software/artemis/). The GC content was obtained using an oligo calculator (http://www.sciencelauncher.com/oligo_calc.html).
Results

Pyrene-degrading genes are duplicated on both the chromosome and circular plasmid of isolate KMS

Analysis of pyrene-induced proteins in isolate KMS, designated as P1-17 by Liang et al. (2006), showed that for P1 - P8, P10, P11, P13, P15, and P17 there was a 100 % match to predicted products from genes within the KMS genome (Table 2.2 and Fig. 2.1). P12 and P14 had 99 % and P9 and P16 had 98 % identity at the amino acid level (Table 2.2 and Fig. 2.1). Proteins P8, P9, and P16 could only have been expressed by a single chromosomal gene, whereas others could be encoded by either chromosomal or genes present on the circular plasmid of isolate KMS (Table 2.2 and Fig. 2.1). For instance, three genes (chromosomal genes Mkms_1660 and Mkms_1667 and circular plasmid gene Mkms_5621) could encode P7, an aromatic ring-hydroxylating dioxygenase β-subunit.

In isolate KMS, the genes predicted to be activated by pyrene were clustered on the chromosome, between loci Mkms_1648 and Mkms_1686 (Fig. 2.1). Part of this sequence was found duplicated on a circular plasmid in isolate KMS, between loci Mkms_5615 and Mkms_5635 (Fig. 2.1). Chromosomal and plasmid genes Mkms_1667 and Mkms_5625 were homologous to the nidB gene (Mvan_0487) designated in isolate PYR-1. Genes Mkms_1660 and Mkms_5621 are homologous to the nidB2 gene (Mvan_0483) in the PYR-1 isolate (Khan et al. 2001; Kim et al. 2007). Thus, we used the same designation for the homologous genes in isolate KMS.
Fig. 2.1. Pyrene-degrading genes are clustered on both the chromosome and circular plasmid of isolate KMS. These genes were identified based on similarity of predicted gene products with the amino acid sequences of seventeen pyrene-induced proteins from the proteomic study by Liang et al. (2006). Gene locus designations were acquired from the JGI-IMG website (http://img.jgi.doe.gov/cgi-bin/w/main.cgi). The correlation between locus tag and the P designation of the proteins (P1-P17) identified by Liang et al. (2006) are: P1, Fe-4S ferredoxin; P2, aromatic-ring-hydroxylating dioxygenase β-subunit; P3, aromatic-ring-hydroxylating dioxygenase β-subunit; P4, sterol binding protein; P5, hypothetical protein; P6, aromatic-ring-hydroxylating dioxygenase β-subunit; P7, aromatic-ring-hydroxylating dioxygenase β-subunit; P8, aromatic-ring-hydroxylating dioxygenase β-subunit; P9, phthalate dihydrodiol dehydrogenase; P10, aromatic-ring-hydroxylating dioxygenase α-subunit; P11, glycosyl hydrolase; P12, epoxide hydrolase; P13, aromatic-ring-hydroxylating dioxygenase α-subunit; P14, aromatic-ring-hydroxylating dioxygenase α-subunit; P15, Fe-4S ferredoxin reductase; P16, hydratase-aldolase; P17, aldehyde dehydrogenase.
Table 2.2. Nucleotide identity between locus tags of genes potentially encoding the proteins accumulating in *Mycobacterium sp.* strain KMS during pyrene degradation.

<table>
<thead>
<tr>
<th>Protein number and predicted function</th>
<th>Locus tag and match at amino acid level</th>
<th>Nucleotide identity between genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chromosome</td>
<td>Circular plasmid</td>
</tr>
<tr>
<td>P1 Fe-4S ferredoxin</td>
<td>Mkms_1662</td>
<td>Mkms_5623</td>
</tr>
<tr>
<td></td>
<td>(100 %)</td>
<td>(100 %)</td>
</tr>
<tr>
<td>P2 Aromatic ring-hydroxylating β-subunit</td>
<td>Mkms_1673</td>
<td>Mkms_5631</td>
</tr>
<tr>
<td></td>
<td>(100 %)</td>
<td>(96 %)</td>
</tr>
<tr>
<td>P3 Aromatic ring-hydroxylating β-subunit</td>
<td>Mkms_1673</td>
<td>Mkms_5631</td>
</tr>
<tr>
<td></td>
<td>(96 %)</td>
<td>(100 %)</td>
</tr>
<tr>
<td>P4 Sterol binding protein</td>
<td>Mkms_1674</td>
<td>Mkms_5632</td>
</tr>
<tr>
<td></td>
<td>(100 %)</td>
<td>(97 %)</td>
</tr>
<tr>
<td>P5 Sterol binding protein</td>
<td>Mkms_1674</td>
<td>Mkms_5632</td>
</tr>
<tr>
<td></td>
<td>(97 %)</td>
<td>(100 %)</td>
</tr>
<tr>
<td>P6</td>
<td>Mkms_1660</td>
<td>Mkms_5621</td>
</tr>
<tr>
<td>P7</td>
<td>Aromatic ring-hydroxylating dioxygenase β-subunit</td>
<td>Mkms_1667</td>
</tr>
<tr>
<td>----</td>
<td>------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(97 %)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mkms_1660</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(100 %)</td>
</tr>
<tr>
<td>P8</td>
<td>Aromatic ring-hydroxylating dioxygenase β-subunit</td>
<td>Mkms_1667</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(100 %)</td>
</tr>
<tr>
<td>P9</td>
<td>Phthalate dihydrodiol dehydrogenase</td>
<td>Mkms_1650</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(98 %)</td>
</tr>
<tr>
<td>P10</td>
<td>Aromatic ring-hydroxylating dioxygenase α-subunit</td>
<td>Mkms_1672</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(100 %)</td>
</tr>
<tr>
<td>P11</td>
<td></td>
<td>Mkms_1671</td>
</tr>
</tbody>
</table>
Glycosyl hydrolase  

P12  
Mkms_1686  Mkms_5635  99 %

Epoxide hydrolase  

P13  
Mkms_1668  Mkms_5626  99 %

Aromatic ring-hydroxylating dioxygenase α-subunit  

P14  
Mkms_1668  Mkms_5626  99 %

Fe-4S ferredoxin reductase  

P15  
Mkms_1661  Mkms_5622  100 %

Hydratase-aldolase  

P16  
Mkms_1658

Aldehyde dehydrogenase  

P17  
Mkms_1663  Mkms_5624  99 %

\(^a\) Proteins induced by growth of isolate KMS on pyrene and identified by LC-MS-MS by 2D gel separation (Liang et al. 2006) are designated by number (P1, P2 etc). Corresponding gene locus numbers are from annotation of the genome of isolate KMS available at the Joint Genome Institute website.
Pyrene-induced expression from loci in KMS

To study gene expression, gene-specific PCR primers were designed (Table 2.1) based on divergent positions in the DNA sequences of the genes. Because the genes, Mkms_1662 and Mkms_5623, encoding the 4S-4Fe ferredoxin (P1) and genes, Mkms_1661 and Mkms_5622, encoding the ferredoxin reductase (P15) were 100% identical for the plasmid and chromosomal copies, differential primers could not be designed. Pyrene growth caused enhanced transcript accumulation from both chromosomal and corresponding circular plasmid copies of genes (Table 2.3). Pyrene enhanced transcript accumulation from 20 of the 25 predicted genes: Mkms_1668, Mkms_5626, Mkms_1672/5630 encoding ring-hydroxylating dioxygenase α-subunits identified as peptides P13, P14 and P10; Mkms_1648, Mkms_5625, Mkms_1660/1667/5621, Mkms_1673, and Mkms_5631 encoding the β-subunits P8, P6, P7, P2 and P3; Mkms_1650 and Mkms_1663/5624 encoding dehydrogenases P9 and P17; Mkms_1658 encoding hydratase aldolase P16; and Mkms_1671 encoding glycosyl hydrolase P11. De novo induction was detected for genes (Mkms_1662 and Mkms_5623) encoding 4S-4Fe ferredoxin (P1) and genes (Mkms_1661 and Mkms_5622) encoding a ferredoxin reductase (P15).

Two genes were expressed before exposure to pyrene: Mkms_1674 and Mkms_5632 encoding sterol binding proteins P4 and P5, respectively. Increased transcript accumulation also was detected for these two genes. The RT-PCR analysis did not detect any RNA accumulation in both control and pyrene-induced samples from gene, Mkms_5629, encoding
a glycosyl hydrolase (P11) and genes Mkms_1686 and Mkms_5635, encoding epoxide hydrolases (99% and 98% homologous to P12, respectively).

Table 2.3. Transcript accumulation for 25 chromosomal and plasmid genes potentially encoding 17 pyrene-induced proteins (Liang et al. 2006) from isolate KMS.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Circular plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene locus</td>
<td>Gene locus</td>
</tr>
<tr>
<td>(Protein encoded and P number from Liang et al. (2006))</td>
<td>Product a</td>
</tr>
<tr>
<td></td>
<td>Chromosome</td>
</tr>
<tr>
<td>Mkms_1648</td>
<td>Mkms_1648</td>
</tr>
<tr>
<td>(dioxygenase β-subunit, P8)</td>
<td>0</td>
</tr>
<tr>
<td>Mkms_1650</td>
<td>Mkms_1650</td>
</tr>
<tr>
<td>(phthalate dihydrodiol dehydrogenase, P9)</td>
<td>0</td>
</tr>
<tr>
<td>Mkms_1658</td>
<td>Mkms_1658</td>
</tr>
<tr>
<td>(hydratase aldolase, P16)</td>
<td>0</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Mkms_1660</td>
<td>dioxygenase β-subunit nidB2, P7</td>
</tr>
<tr>
<td>Mkms_1661*</td>
<td>4Fe-4S ferredoxin reductase, P15</td>
</tr>
<tr>
<td>Mkms_1662#</td>
<td>4Fe-4S ferredoxin, P1</td>
</tr>
<tr>
<td>Mkms_1663</td>
<td>aldehyde dehydrogenase, P17</td>
</tr>
<tr>
<td>Mkms_1667</td>
<td>dioxygenase β-subunit nidB, P7</td>
</tr>
<tr>
<td>Mkms_1668</td>
<td>dioxygenase α-subunit nidA, P13</td>
</tr>
<tr>
<td>Mkms_1671</td>
<td>glycosyl hydrolase, P11</td>
</tr>
<tr>
<td>Mkms_5621</td>
<td>dioxygenase β-subunit nidB2, P7</td>
</tr>
<tr>
<td>Mkms_5622*</td>
<td>4Fe-4S ferredoxin reductase, P15</td>
</tr>
<tr>
<td>Mkms_5623#</td>
<td>4Fe-4S ferredoxin, P1</td>
</tr>
<tr>
<td>Mkms_5624</td>
<td>aldehyde dehydrogenase, P17</td>
</tr>
<tr>
<td>Mkms_5625</td>
<td>dioxygenase β-subunit nidB, P6</td>
</tr>
<tr>
<td>Mkms_5626</td>
<td>dioxygenase α-subunit nidA, P14</td>
</tr>
<tr>
<td>Mkms_5629</td>
<td>glycosyl hydrolase, P11</td>
</tr>
<tr>
<td>Genes</td>
<td>Relative Intensity</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Mkms_1672 (dioxygenase α-subunit, P10)</td>
<td>0 51.9±1.9</td>
</tr>
<tr>
<td>Mkms_1673 (dioxygenase β-subunit, P2)</td>
<td>0 50.2±0.9</td>
</tr>
<tr>
<td>Mkms_1674 (sterol-binding protein, P4)</td>
<td>9.3±2.6 20.1±3.3</td>
</tr>
<tr>
<td>Mkms_1686 (epoxide hydrolase, P12)</td>
<td>0 0</td>
</tr>
<tr>
<td>Mkms_5630 (dioxygenase α-subunit, P10)</td>
<td></td>
</tr>
<tr>
<td>Mkms_5631 (dioxygenase β-subunit, P3)</td>
<td></td>
</tr>
<tr>
<td>Mkms_5632 (sterol-binding protein, P5)</td>
<td></td>
</tr>
<tr>
<td>Mkms_5635 (epoxide hydrolase, P12)</td>
<td></td>
</tr>
</tbody>
</table>

*The numbers represent the relative intensities of PCR product bands read by Quantity-One software from BIORAD. The results are typical from three independent studies for KMS cells grown with (Pyrene) or without pyrene (Control) in the BSM* medium. The symbols * and # indicate that the products from the chromosomal and plasmid copies of the genes cannot be distinguished from one another because their nucleotide sequences are 100% identical.

**Synteny of pyrene-degrading gene islands of five pyrene-degrading isolates**

In addition to gene duplication between the chromosome and circular plasmid in isolate KMS, other blocks of genes potentially significant in pyrene degradation were
present on the linear plasmid. This gene amplification is demonstrated in Fig. 2.2A and shows block duplication with retention of gene sequence but with some block rearrangements between the chromosome, the circular plasmid and the linear plasmid in isolate KMS (Fig. 2.2A): these regions were from Mkms_1647 to Mkms_1718 on the chromosome, from Mkms_5609 to Mkms_5651 on the circular plasmid and from Mkms_5837 to Mkms_5855 on the linear plasmid. A survey of other genes in these regions showed the presence of transposases and integrases as well as omission of gene blocks, notably for genes for mammalian cell entry that are found only on the chromosome (Table 2.4 and Fig. 2.2A). The island region in the KMS chromosome was flanked at loci Mkms_1653 and 1720 by insertion elements of the IS3/IS911 family and at loci Mkms_1652 and 1721 by integrases. Similarly the clusters associated with pyrene degradation on the linear and circular plasmids also possessed integrase and insertion elements. For the circular plasmid, these included Mkms_5607, transposase gene family IS3/IS911, integrase gene Mkms_5605, Mkms_5653 and 5655 plus additional IS116/110/902 type transposase genes at loci Mkms_5506, 5513, 5520, 5526, 5538 and integrase genes at loci Mkms_5548 and 5552. For the linear plasmid, genes for mutator type transposases were predominant although one gene for an IS3/IS911 family transposase was present at locus Mkms_5829. The average GC content of the genes within these regions, as shown in Fig. 2.2A, was lower (61-63 %) than the average for the whole genome (68 %).
A single copy of this PAH-degrading island, as defined by loci from Mkms_1647 to Mkms_1718 from KMS, was present in two other mycobacterium isolates, MCS and JLS, from the same Montana contaminated soil site as isolate KMS (Fig. 2.3A). Both isolates MCS and JMS also degraded pyrene (Miller et al. 2004). The homologous pyrene-degrading gene island in isolate MCS extended from chromosomal gene Mmcs_1622 to gene Mmcs_1691 and in isolate JLS from gene Mjls_1595 to gene Mjls_1662 (Fig. 2.3A). A linear plasmid in isolate MCS had high identity with the linear plasmid of KMS as shown by Mauve synteny analysis in Fig. 2.3B. Both linear plasmids had identity in the region of the PAH-degrading cluster. Isolate JLS did not harbor plasmids.

PAH-degrading gene islands were also present in Mycobacterium isolates (PYR-1 and PYR-GCK) from two other geographic regions, although rearrangements and differences in gene content are apparent (Fig. 2.2B). The regions in isolate PYR-1 extended from gene Mvan_0463 to gene Mvan_0568 and in isolate PYR-GCK from gene Mflv_0536 to gene Mflv_0685 (Fig. 2.2B). Although *M. gilvum* PYR-GCK possesses three small linear plasmids (321.2 Kb, 25.3 Kb, and 16.7 Kb), these plasmids lack any homologues of the pyrene-inducible genes. However, in *M. gilvum* PYR-GCK, genes with high identity to those in the PAH-induced island were present at two chromosomal locations that were closely situated (from Mflv_0546 to Mflv_0588 and from Mflv_0686 to Mflv_0628) (Fig. 2.2B). Minor differences in these alignments were caused by the absence and insertion of genes as well as genes designated as encoding transposases and
integrases (Table 2.4, Fig. 2.2A, 2.2B, and Fig. 2.3B). One notable feature was the block of mammalian cell entry genes present in the islands within all of the chromosomes but not in the plasmids. Again the GC contents of the designated pyrene-degrading gene islands in isolates PYR-1, and PYR-GCK were lower at 63-64 % than the 68 % for their chromosomes.

**Discussion**

The pyrene-induced proteins in the PAH-degrading *Mycobacterium* sp. strain KMS observed in the proteomic study by Liang et al. (2006) were encoded by genes clustered both on the chromosome and a circular plasmid. These clusters were rich in genes encoding dioxygenases that are key enzymes in initiating pyrene degradation (Kim et al. 2007; Kim et al. 2008) as well as other genes encoding proposed functions in ring oxidation and opening. In the genome of isolate KMS, the differences in GC usage and the presence of transposase/integrase genes near and within these pyrene-induced gene clusters are consistent with the genes being part of a genomic island promoting the utilization of PAHs. The concept of genomic islands being involved in metabolism of aromatic compounds has been previously suggested (Toussaint et al. 2003). Here we have shown that in isolate KMS, block gene duplication resulted in their duplicated locations on the chromosome as well as a linear and a circular plasmid. These regions possess genes encoding transposases and integrases that are proposed to drive genome modifications (Tillier and Collins 2000; Wang et al. 2008). The transposases drive
excision of mobile elements and the integrases drive recombination events (Duval-Valentin et al. 2004). However, the positioning of this genomic island in strain KMS was not at tRNA sites as has been demonstrated for other mycobacterial genomic islands (Becq et al. 2007; Danelishvili et al. 2007).

The genes in the PAH-degrading islands in strain KMS encode proteins involved in the early steps of pyrene degradation for ring oxidation and ring opening. Some of the transcribed products have functions that are unresolved such as the proteins with a sterol-binding domain (COG3255) encoded by Mkms_1674 and 5632. In this paper we demonstrated that the pyrene-induced protein accumulation observed by Liang et al. (2006) generally was correlated with enhanced transcript accumulations. The genes most highly expressed encoded the 4Fe-4S ferredoxin reductase (P15), NidB2 (P7), and NidA (P14). Our work failed to detect transcripts corresponding to P12 for an epoxide hydrolase and only transcripts from the chromosomal gene for a glycosyl hydrolase, P11, were detected. Only the genes for the sterol-binding protein P4 and P5 were expressed in the absence of pyrene, although transcript abundance was enhanced in its presence. Thus although Liang et al. (2006) found low levels of proteins P1, P10, P11, P14, P15, P16 in extracts from cells grown without pyrene, we did not correlate their presence with detectable mRNA levels. Differences between cell growth conditions between the proteomic and transcript analysis could be responsible or the level of transcript could have been below detection in the RT-PCR studies.
Synteny was observed with the pyrene-degrading island in the chromosomes of all five PAH-degrading isolates examined irrespective of their diverse sites of origin within the United States. The genes in the pyrene-degrading island of isolate KMS had synteny with the genes in the PAH-degrading region, termed region A by Kim et al. (2008) but the region designated as region B in isolate PYR-1, correlated with biphenyl usage was not fully duplicated in isolate KMS. Indeed, KMS does not grow on biphenyl (Zhang et al. unpublished). Duplication of some of the blocks containing pyrene-related genes was observed between chromosome and plasmids for the Montana isolates KMS and MCS but not for the plasmids of _M. gilvum_. Rather, in this isolate there was a partial sequence duplication of the pyrene-degrading gene island on the chromosome itself. These gene islands had sizes from 66 kb in isolate JLS (1.2 % of the genome) to 140 kb (2.7 % of the genome) in isolate PYR-GCK; gene numbers in the pyrene-degrading islands, based on the sequences in isolate KMS, were 72 in isolate KMS, 70 in isolate MCS, 68 in isolate JLS, 86 in isolate PYR-1, and 150 in isolate PYR-GCK. These gene islands are comparable in size to the pathogenicity islands of 70 kb and 190 kb found in _Escherichia coli_ (Blum et al. 1994).
Table 2.4. Synteny in gene order between five PAH-degrading Mycobacterium isolates from three different sites in the U.S.A.

<table>
<thead>
<tr>
<th>Protein encoded</th>
<th>Locus on M. sp. strain KMS chromosome</th>
<th>Locus on M. sp. strain KMS circular plasmid</th>
<th>Locus on M. sp. strain KMS linear plasmid</th>
<th>Locus on M. sp. strain MCS chromosome</th>
<th>Locus on M. sp. strain MCS linear plasmid</th>
<th>Locus on M. vanbaalenii PYR-1 chromosome</th>
<th>Locus on M. gilvum chromosome (region 1)</th>
<th>Locus on M. gilvum chromosome (region 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/β hydrolase fold</td>
<td>Mkms_5609</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, α subunit</td>
<td>Mkms_5610</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, β subunit</td>
<td>Mkms_5611</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxidoreductase domain protein</td>
<td>Mkms_5612</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extradiol ring-cleavage dioxygenase</td>
<td>Mkms_5613</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class II aldolase/adducin family protein</td>
<td>Mkms_5614</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-chain dehydrogenase/reductase</td>
<td>Mkms_5615</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regulatoy protein</td>
<td>Mkms_1646</td>
<td>Mmcs_1621</td>
<td>Mjls_1594</td>
<td>Mvan_0462</td>
<td>Mflv_0686</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IclR type</td>
<td>Mkms_1647</td>
<td>Mmcs_1622</td>
<td>Mjls_1595</td>
<td>Mvan_0463</td>
<td>Mflv_0685</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, α subunit</td>
<td>Mkms_1648</td>
<td>Mmcs_1623</td>
<td>Mjls_1596</td>
<td>Mvan_0464</td>
<td>Mflv_0684</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Type</td>
<td>Genomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothetical protein Mkms_1649</td>
<td>Mjls_1597 Mvan_0465 MfLV_0683</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-chain dehydrogenase Mkms_1650</td>
<td>Mmcs_1624 Mjls_1598 Mvan_0466 MfLV_0682</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAD-dependent pyridine nucleotide-disulphide oxidoreductase Mkms_1651</td>
<td>Mmcs_1625 Mjls_1599 Mvan_0467</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrase Mkms_1652</td>
<td>Mmcs_1626</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transposase Mkms_1653</td>
<td>Mmcs_1627</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cupin 2, conserved barrel domain protein Mkms_1654</td>
<td>Mmcs_1628 Mjls_1601 Mvan_0468 MfLV_0677</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dihydrodipicolinate synthetase Mkms_1655</td>
<td>Mmcs_1629 Mjls_1602 Mvan_0469 MfLV_0676 MfLV_0588</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyoxalase/bleomycin resistance protein/dioxygenase Mkms_1656 Mkms_5616</td>
<td>Mmcs_1630 Mjls_1603 Mvan_0470 MfLV_0675 MfLV_0587</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothetical protein Mkms_1657 Mkms_5617</td>
<td>Mmcs_1631 Mjls_1604 Mvan_0471 MfLV_0674 MfLV_0586</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mkms_5618 Unknown function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydratase aldolase, PhdG Mkms_1658</td>
<td>Mmcs_1632 Mjls_1605 Mvan_0472 MfLV_0673 MfLV_0585</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown function Mkms_1659 Mkms_5620</td>
<td>Mmcs_1633 Mjls_1606</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Family</td>
<td>Gene ID 1</td>
<td>Gene ID 2</td>
<td>Gene ID 3</td>
<td>Gene ID 4</td>
<td>Mvan</td>
<td>Mflv 1</td>
<td>Mflv 2</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>------</td>
<td>--------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, β subunit</td>
<td>Mkms 1660</td>
<td>Mkms 5621</td>
<td>Mmcs 1634</td>
<td>Mjls 1607</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succinate dehydrogenase</td>
<td>Mkms 1661</td>
<td>Mkms 5622</td>
<td>Mmcs 1635</td>
<td>Mjls 1608</td>
<td>Mvan 0483</td>
<td>Mflv 0670</td>
<td>Mflv 0572</td>
<td></td>
</tr>
<tr>
<td>flavoprotein domain protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4Fe-4S ferredoxin, iron-sulfur binding domain protein</td>
<td>Mkms 1662</td>
<td>Mkms 5623</td>
<td>Mmcs 1636</td>
<td>Mjls 1609</td>
<td>Mvan 0484</td>
<td>Mflv 0669</td>
<td>Mflv 0571</td>
<td></td>
</tr>
<tr>
<td>Aldehyde dehydrogenase</td>
<td>Mkms 1663</td>
<td>Mkms 5624</td>
<td>Mmcs 1637</td>
<td>Mjls 1610</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown function</td>
<td>Mkms 1664</td>
<td></td>
<td>Mmcs 1638</td>
<td>Mjls 1611</td>
<td>Mvan 0486</td>
<td>Mflv 0667</td>
<td>Mflv 0569</td>
<td></td>
</tr>
<tr>
<td>Unknown function</td>
<td>Mkms 1665</td>
<td></td>
<td>Mmcs 1639</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown function</td>
<td>Mkms 1666</td>
<td></td>
<td>Mmcs 1640</td>
<td>Mjls 1613</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown function</td>
<td></td>
<td></td>
<td></td>
<td>Mjls 1614</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Integrase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mjls 1615</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Transposase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, β subunit</td>
<td>Mkms 1667</td>
<td>Mkms 5625</td>
<td>Mmcs 1641</td>
<td>Mjls 1616</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, α subunit</td>
<td>Mkms 1668</td>
<td>Mkms 5626</td>
<td>Mmcs 1642</td>
<td>Mjls 1617</td>
<td>Mvan 0487</td>
<td>Mflv 0663</td>
<td>Mflv 0568</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>nidB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

90
<table>
<thead>
<tr>
<th><strong>Short-chain dehydrogenase/reductase SDR</strong></th>
<th>Mkms_1669</th>
<th>Mkms_5627</th>
<th>Mmcs_1643</th>
<th>Mjls_1618</th>
<th>Mvan_0489</th>
<th>Mflv_0661</th>
<th>Mflv_0566</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alcohol dehydrogenase, zinc-binding domain protein</strong></td>
<td>Mkms_1670</td>
<td>Mkms_5628</td>
<td>Mmcs_1644</td>
<td>Mjls_1619</td>
<td>Mvan_0490</td>
<td>Mflv_0660</td>
<td>Mflv_0565</td>
</tr>
<tr>
<td><strong>Hypothetical protein</strong></td>
<td>Mkms_1671</td>
<td>Mkms_5629</td>
<td>Mmcs_1645</td>
<td>Mjls_1620</td>
<td>Mvan_0491</td>
<td>Mflv_0659</td>
<td>Mflv_0564</td>
</tr>
<tr>
<td><strong>Dioxygenases, α subunit</strong></td>
<td>Mkms_1672</td>
<td>Mkms_5630</td>
<td>Mmcs_1646</td>
<td>Mjls_1621</td>
<td>Mvan_0492</td>
<td>Mflv_0658</td>
<td>Mflv_0563</td>
</tr>
<tr>
<td><strong>Dioxygenase, β subunit</strong></td>
<td>Mkms_1673</td>
<td>Mkms_5631</td>
<td>Mmcs_1647</td>
<td>Mjls_1622</td>
<td>Mvan_0493</td>
<td>Mflv_0657</td>
<td>Mflv_0562</td>
</tr>
<tr>
<td><strong>Sterol-binding domain protein</strong></td>
<td>Mkms_1674</td>
<td>Mkms_5632</td>
<td>Mmcs_1648</td>
<td>Mjls_1623</td>
<td>Mvan_0494</td>
<td>Mflv_0656</td>
<td></td>
</tr>
<tr>
<td><strong>Intradiol ring-cleavage dioxygenase</strong></td>
<td>Mkms_1675</td>
<td>Mkms_5633</td>
<td>Mmcs_1649</td>
<td>Mjls_1624</td>
<td>Mvan_0495</td>
<td>Mflv_0655</td>
<td></td>
</tr>
<tr>
<td><strong>Unknown function</strong></td>
<td>Mkms_1676</td>
<td>Mkms_5634</td>
<td>Mmcs_1650</td>
<td>Mjls_1625</td>
<td>Mvan_0498</td>
<td>Mflv_0654</td>
<td></td>
</tr>
<tr>
<td><strong>Unknown function</strong></td>
<td>Mkms_1677</td>
<td>Mkms_1678</td>
<td>Mmcs_1651</td>
<td>Mjls_1626</td>
<td>Mvan_0499</td>
<td>Mflv_0653</td>
<td></td>
</tr>
<tr>
<td><strong>Virulence factor Mce family protein</strong></td>
<td>Mkms_1683</td>
<td>Mkms_1653</td>
<td>Mmcs_1652</td>
<td>Mjls_1627</td>
<td>Mvan_0500-</td>
<td>Mflv_0652</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mvan_0501,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mflv_0652</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mvan_0515-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mflv_0647</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mvan_0518</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Type</td>
<td>ID 1</td>
<td>ID 2</td>
<td>ID 3</td>
<td>ID 4</td>
<td>ID 5</td>
<td>ID 6</td>
<td>ID 7</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>Mkms_1684</td>
<td>Mkcs_1654-</td>
<td>Mkcs_1659</td>
<td>Mjls_1633</td>
<td>Mvan_0519</td>
<td>Mflv_0646</td>
<td></td>
</tr>
<tr>
<td>Hypothetical protein</td>
<td>Mkms_1685</td>
<td>Mkcs_1660</td>
<td>Mjls_1634</td>
<td>Mvan_0520</td>
<td>Mflv_0645</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epoxide hydrolase domain protein</td>
<td>Mkms_1686</td>
<td>Mkms_1662</td>
<td>Mjls_1635</td>
<td>Mvan_0521</td>
<td>Mflv_0644</td>
<td>Mflv_0561</td>
<td></td>
</tr>
<tr>
<td>Aldehyde dehydrogenase</td>
<td>Mkms_1687</td>
<td>Mkms_1663</td>
<td>Mjls_1636</td>
<td>Mvan_0522</td>
<td>Mflv_0643</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drug resistance transporter</td>
<td>Mkms_1688</td>
<td>Mkms_1664</td>
<td>Mjls_1637</td>
<td>Mvan_0523</td>
<td>Mflv_0642</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcriptional regulator</td>
<td>Mkms_1689</td>
<td>Mkms_1665</td>
<td>Mjls_1638</td>
<td>Mvan_0524</td>
<td>Mflv_0641</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MarR type</td>
<td>Mkms_1690</td>
<td>Mkms_1666</td>
<td>Mjls_1639</td>
<td>Mflv_0525</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, α subunit</td>
<td>Mkms_1691</td>
<td>Mkms_1667</td>
<td>Mjls_1640</td>
<td>Mflv_0526</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, β subunit</td>
<td>Mkms_1692</td>
<td>Mkms_1668</td>
<td>Mjls_1641</td>
<td>Mflv_0527</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol dehydrogenase, zinc-binding</td>
<td>Mkms_1693</td>
<td>Mkms_1669</td>
<td>Mjls_1642</td>
<td>Mflv_0528</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-carnitine dehydratase/bile acid-</td>
<td>Mkms_1694</td>
<td>Mkms_1670</td>
<td>Mjls_1643</td>
<td>Mflv_0529</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>inducible protein F</td>
<td>Mkms_1695</td>
<td>Mkms_1671</td>
<td>Mjls_1644</td>
<td>Mflv_0530</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein Type</td>
<td>Gene IDs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>---------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Putative nonspecific lipid-transfer protein</td>
<td>Mkms_1694, Mkms_5643, Mkms_5851, Mmcs_1670, Mmcs_5454, Mjls_1643, Mvan_0529, Mflv_0636, Mflv_0554</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein of unknown function</td>
<td>Mkms_1695, Mkms_5644, Mkms_5850, Mmcs_1671, Mmcs_5453, Mjls_1644, Mvan_0530, Mflv_0635, Mflv_0553</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conserved hypothetical protein</td>
<td>Mkms_1696, Mkms_5645, Mkms_5849, Mmcs_1672, Mmcs_5452, Mjls_1645, Mvan_0531, Mflv_0634, Mflv_0552</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conserved hypothetical protein</td>
<td>Mkms_1697, Mkms_5646, Mkms_5848, Mmcs_5451, Mjls_1646, Mvan_0532, Mflv_0633, Mflv_0551</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, α subunit</td>
<td>Mkms_1698, Mkms_5647, Mkms_5847, Mmcs_1673, Mmcs_5450, Mjls_1647, Mvan_0533, Mflv_0632, Mflv_0550</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, β subunit</td>
<td>Mkms_1699, Mkms_5648, Mkms_5846, Mmcs_1674, Mmcs_5449, Mjls_1648, Mvan_0534, Mflv_0631, Mflv_0549</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cupin 2, conserved barrel domain protein</td>
<td>Mkms_1700, Mkms_5649, Mkms_5845, Mmcs_1675, Mmcs_5448, Mjls_1649, Mvan_0535, Mflv_0630, Mflv_0548</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-phosphogluconate dehydrogenase, NAD-binding</td>
<td>Mkms_1701, Mkms_5650, Mkms_5844, Mmcs_1676, Mmcs_5447, Mjls_1650, Mvan_0536, Mflv_0629, Mflv_0547</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcriptional regulator MarR type</td>
<td>Mkms_1702, Mkms_5651, Mkms_5843, Mmcs_1677, Mmcs_5446, Mjls_1651, Mvan_0537, Mflv_0628, Mflv_0546</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transposase</td>
<td>Mkms_1703</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrase</td>
<td>Mkms_1704</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrase</td>
<td>Mkms_1705</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transposase</td>
<td>Mkms_1706</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transposase</td>
<td>Mkms_1707</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Family</td>
<td>Accession Numbers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/β hydrolase fold</td>
<td>Mkms_1708 Mkms_5837 Mkms_5837 Mkms_5838 Mkms_1710 Mkms_1711 Mkms_1712 Mkms_1713 Mkms_5842</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, α subunit</td>
<td>Mkms_1709 Mkms_5838 Mkms_5838 Mkms_5839 Mkms_1710 Mkms_5840 Mkms_5841 Mkms_5842</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, β subunit</td>
<td>Mkms_1709 Mkms_5838 Mkms_5838 Mkms_5839 Mkms_1710 Mkms_5840 Mkms_5841 Mkms_5842</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxidoreductase domain protein</td>
<td>Mkms_1709 Mkms_5838 Mkms_5838 Mkms_5839 Mkms_1710 Mkms_5840 Mkms_5841 Mkms_5842</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extradiol ring-cleavage dioxygenase</td>
<td>Mkms_1713 Mkms_5842 Mkms_5842 Mkms_5843 Mkms_1714 Mkms_1684 Mkms_5445 Mmc_1652 Mmc_1655</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class II aldolase/adducin family protein</td>
<td>Mkms_1713 Mkms_5842 Mkms_5842 Mkms_5444 Mkms_1652 Mmc_1653 Mmc_1654 Mmc_1655 Mmc_1656</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-chain dehydrogenase/reductase</td>
<td>Mkms_1714 Mkms_1685 Mkms_1685 Mmc_1658 Mmc_1658 Mmc_1659 Mmc_1659 Mmc_1659 Mmc_1659</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyoxalase/bleomycin resistance protein/dioxygenase</td>
<td>Mkms_1715 Mkms_1686 Mkms_1686 Mmc_1659 Mmc_1659 Mmc_1659 Mmc_1659 Mmc_1659 Mmc_1659</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, α subunit</td>
<td>Mkms_1716 Mkms_1689 Mkms_1689 Mmc_1660 Mmc_1660 Mmc_1660 Mmc_1660 Mmc_1660 Mmc_1660</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dioxygenase, β subunit</td>
<td>Mkms_1717 Mkms_1690 Mkms_1690 Mmc_1661 Mmc_1661 Mmc_1661 Mmc_1661 Mmc_1661 Mmc_1661</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conserved hypothetical protein</td>
<td>Mkms_1718 Mkms_1691 Mkms_1691 Mmc_1662 Mmc_1662 Mmc_1662 Mmc_1662 Mmc_1662 Mmc_1662</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The bold underlined genes are those identified as being involved in pyrene degradation by Kim et al. 2007 for isolate PYR-1 and in this study for strain KMS. The names of proteins were from the annotations from JGI-IMG website (http://img.jgi.doe.gov/cgi-bin/w/main.cgi).
Fig. 2.2. Synteny analysis by the Mauve program of pyrene-degrading gene islands in *Mycobacterium* isolates. (A) Duplication of blocks from the pyrene-degrading gene island from the chromosome, the circular plasmid, and the linear plasmid of isolate KMS. (B) Synteny of pyrene-degrading islands in the chromosomes of isolates KMS, PYR-1, and PYR-GC. The color blocks generated by the Mauve indicate the homology of the genes within the blocks. The locus numbers of the first and last genes of the pyrene-degrading islands are labeled. The *nidA*, *nidB*, *nidB2*, *nidA3*, and *nidB3* genes were designated by Cerniglia’s group for isolate PYR-1 (Kim et al. 2007). Single genes are represented by rectangles. The ring-hydroxylating dioxygenase α- and β-subunit genes are shown in red and blue respectively. The genes annotated as encoding integrases and transposases are shown in green and pink, respectively. The genes annotated as encoding virulence factor Mce family proteins are shown in yellow.
Pyrene-degrading gene island on Mycobacterium sp. isolate KMS chromosome (GC content 63%)

Duplicated pyrene-degrading genes on Mycobacterium sp. isolate KMS circular plasmid (GC content 63%)

Duplicated pyrene-degrading genes on Mycobacterium sp. isolate KMS linear plasmid (GC content 61%)
Pyrene-degrading gene island on *Mycobacterium* sp. isolate KMS chromosome (GC content 63%)

Pyrene-degrading gene island on *Mycobacterium vanbaalenii* PYR-1 chromosome (GC content 64%)

Pyrene-degrading gene island on *Mycobacterium gilvum* PYR-GCK chromosome (GC content 63%)

(B)
**Fig. 2.3.** Synteny analysis by the Mauve program of pyrene-degrading gene islands in *Mycobacterium* isolates. (A) Synteny of the pyrene-degrading gene islands in the chromosomes of isolates KMS, MCS, and JLS. (B) Synteny of blocks of pyrene-degrading genes on the linear plasmids of isolates KMS and MCS. The color blocks generated by the Mauve indicate the homology of the genes within the blocks. The locus numbers of the first and last genes of the pyrene-degrading islands are labeled. The *nidA, nidB, nidB2, nidA3,* and *nidB3* genes were designated by Cerniglia’s group for isolate PYR-1 (Kim et al. 2007). Single genes are represented by rectangles. The ring-hydroxylating dioxygenase α- and β-subunit genes are shown in red and blue respectively. The genes annotated as encoding integrases and transposases are shown in green and pink, respectively. The genes annotated as encoding virulence factor Mce family proteins are shown in yellow.
Duplicated pyrene-degrading genes on *Mycobacterium* sp.

isolate KMS linear plasmid (GC content 61%)

![Diagram of KMS linear plasmid with genes and GC content]

Duplicated pyrene-degrading genes on *Mycobacterium* sp.

isolate MCS linear plasmid (GC content 61%)

![Diagram of MCS linear plasmid with genes and GC content]
Evidence from PAH-degrading pseudomonads demonstrated clustering and duplication of genes involved in PAH degradation on the chromosome (Takizawa et al. 1994) as well as plasmids (Eaton 1994). Previous work with isolate PYR-1 from 2001 to 2006 (Khan et al. 2001; Stingley et al. 2004; Kim et al. 2006) showed clustering of genes for PAHs degradation on the mycobacterial chromosome with the presence of two types of arrangements of *nid* genes, one where the gene for the α-subunit preceded that for the β-subunit and a second where gene order was reversed as shown in Fig. 2.4. Also the sequences of genes termed *nidB* and *nidB2* were 99% identical suggesting that single gene duplication had occurred in PYR-1 (Stingley et al. 2004). These findings were confirmed upon complete sequencing of the PYR-1 genome (Kim et al. 2007; Kim et al. 2008). Krivobok et al. (2003) also found two clusters one with gene order *pdoA1/pdoB1* and a second with gene order *pdoA2/pdoB2* in *Mycobacterium* isolate 6PY1. These genes had high identity to the deduced amino acid sequences of *nidA* and *nidB* genes previously characterized in isolate PYR-1. Similarly Churchill et al. (2008) found two different clusters, one with *nidB-nidA* order and one with *pdoA2-pdoB2* order in *Mycobacterium sp.* CH-2. Gene duplication in mycobacteria also was supported by observations with *Mycobacterium sp.* isolate S65, where high identity of nucleotide sequence (more than 85%) was found between corresponding genes in the *nid* and *pdo* clusters identified by Sho et al. (2004). Only limited knowledge is available on genes adjacent to the dioxygenase genes as shown in Fig. 2.4 but the identity of these genes also is consistent to gene order in the KMS and PYR-genomes. Whether any of the genes in these other mycobacterial isolates are plasmid borne is not reported. However, this limited
knowledge supports the conservation of sequences for pyrene-degrading genes between isolates of varied origins.

In conclusion, there was a high degree of conservation in both gene sequence and gene order for the genes involved in PAH degradation in environmental mycobacterium isolates from three diverse PAH-contaminated sites in the USA. Duplication of clusters of these genes was apparent both on the chromosome and also on plasmids for three isolates from one geographic site. Regulation of expression of these genes during growth on pyrene was conserved between two isolates KMS and PYR-1 as determined by proteomics (Liang et al. 2006; Kim et al. 2007) and transcriptional studies of this paper. Transcripts were detected from duplicated copies on the chromosomal and plasmid of several genes involved transformation of the PAH ring structures.

**Fig. 2.4.** Identified pyrene-degrading gene clusters in four Mycobacterium isolates: PYR-1, modified from Khan et al. (2001); Kim et al. (2006), 6PY1 modified from Krivobok et al. (2003), S65 modified from Sho et al. (2004), and CH-2 modified from Churchill et al. (2008) showing the gene arrangements in these clusters. Genes nidA, nidA3, nidX, pdoA, pdoA1, pdoA2, and pdoX encode dioxygenase α-subunits; genes nidB, nidB3, pdoB, pdoB1, pdoB2 encode dioxygenase β-subunits; genes nidC, nidD, orf4, nidH, and pdoH encode dehydrogenases; genes phdF and pdoF encode extradiol dioxygenases; genes orf1, orfN4, orfP6, and orf72 encode unknown proteins; genes nidR and araC encodes regulatory proteins. Genes encoding ferredoxin and reductase deduced from using the BLAST program based on their sequences are labeled accordingly.
References


CHAPTER 3
MULTIPLICITY OF GENES FOR AROMATIC RING-HYDROXYLATING DIOXYGENASES IN *MYCOBACTERIUM* ISOLATE KMS AND THEIR REGULATION

Abstract *Mycobacterium* sp. strain KMS has bioremediation potential for polycyclic aromatic hydrocarbons (PAHs), such as pyrene, and smaller ring aromatics, such as benzoate. Degradation of these aromatics involves oxidation catalyzed by aromatic ring-hydroxylating dioxygenases. Multiple genes encoding dioxygenases exist in KMS: ten genes encode large-subunits with homology to phenylpropionate dioxygenase genes, sixteen pairs of adjacent genes encode α- and β-subunits of dioxygenase and two genes encode β-subunits. These genes include orthologs of *nid* genes essential for degradation of multi-ring PAHs in *M. vanbaalenii* isolate PYR-1. The multiplicity of genes in part is explained by block duplication that results in two or three copies of certain genes on the chromosome, a linear plasmid, and a circular plasmid within the KMS genome. Quantitative real-time PCR showed that four dioxygenase β-subunit *nid* genes from operons with almost identical promoter sequences otherwise unique in the genome were induced by pyrene to similar extents. No induction occurred with benzoate. Unlike isolate PYR-1, isolate KMS has an operon specifying benzoate catabolism and the expression of the α-subunit dioxygenase gene was activated by benzoate but not pyrene. These studies showed that isolate KMS had a genome well adapted to utilization of different aromatic compounds.

---

1Coauthored by Dr. Anderson AJ., Multiplicity of genes for aromatic ring-hydroxylating dioxygenases in *Mycobacterium* isolate KMS and their regulation, © 2012 with permission from Springer.
Introduction

Aromatics range from single ring, such as benzoate, to multiple fused-ring polycyclic aromatic hydrocarbons (PAHs). Benzoate naturally occurs in plants and is utilized as an antimicrobial agent in the food industry (Orlova et al. 2006; Zengin et al. 2010). PAHs are generated from incomplete combustion of organics, and are notable for their toxic, mutagenic, and carcinogenic properties (U.S. Department of Health and Human Services, 1995; Tannheimer et al. 1997). Aerobic degradation of aromatic hydrocarbons frequently is initiated by oxidation catalyzed by ring-hydroxylating dioxygenases (Butler and Mason 1997; Peng et al. 2010). The dioxygenases work in conjunction with small electron transport chains so that a multi-protein structure is involved in the oxygenation process (Butler and Mason 1997; Correll et al. 1993; Jiang et al. 1999; Kweon et al. 2008; Peng et al. 2010). Such dioxygenases generally are complexes of a large α-subunit and a small β-subunit (Butler and Mason 1997; Parales and Resnick 2006), although a subset only has one α-subunit (Parales and Resnick 2006). Kweon et al. (2008) derived a new classification scheme based on homology of the oxygenases and the composition of the electron transport chains. The benzoate dioxygenase is found in Type III which has a simpler electron transport chain than the Type V dioxygenases where the pyrene-degrading dioxygenases are found (Kweon et al. 2008).

Working with *Mycobacterium vanbaalenii* PYR-1, Cerniglia’s group demonstrates that the initial ring oxidation of high molecular weight PAHs, such as pyrene and fluoranthene, involves the dioxygenases encoded by chromosomal genes, *nidB/nidA*, *nidB2* and *nidA3/nidB3* (Khan et al. 2001; Kim et al. 2006, 2007; Kweon et al. 2010). Specificity for the ring structures that are oxidized is conditioned by the α-subunits more
than the β-subunits (Parales and Resnick 2006), although contributions from the β-
elegantly demonstrate that the preferred activity for a substrate correlates with the size of
the substrate pocket in the α-subunit. However, they also comment that the substrate
plays a role in regulating the production of the enzyme. Proteomic analysis reveals that
the enzyme that uses pyrene as the preferred substrate NidA/B is induced in pyrene but
not fluoranthene cultures and that NidA3/B3, which better degrades fluoranthene than
pyrene, is produced to a greater extent when fluoranthene is the carbon source (Kweon et
al. 2010).

Our group found that the environmental *Mycobacterium sp.* strain KMS from PAH-
contaminated soils in Montana (Miller et al. 2004) has genes encoding orthologs of *nidB*,
*nidB2* and *nidA* (Zhang and Anderson 2012). These genes are embedded in segments of
the chromosome that have synteny with other PAH-degrading mycobacterium isolates
(Zhang and Anderson 2012). In this paper we determined the full complement of
dioxygenase genes present in the KMS genome. Previous studies with isolate PYR-1
focus only on the Nid proteins (Kweon et al. 2008, 2010). We provide evidence that
block duplication events have caused certain of these genes to be present as two and three
copies with locations on the chromosome as well as on circular and linear plasmids.
Bioinformatics predicts that one set of paired α- and β-dioxygenase genes could be
involved in benzoate degradation in isolate KMS; isolate PYR-1 lacks such genes and
does not grow on benzoate (Kim et al. 2008). We verified growth of isolate KMS on
benzoate and examined the influence of the growth substrate, benzoate versus pyrene, on
dioxygenase gene expression. We compared growth substrate regulation of the *benA*
gene, encoding the benzoate dioxygenase α-subunit, with the expression of the chromosomal copy of the \textit{nidA} gene. We also examined expression from the homologous \textit{nidB} and \textit{nidB2} genes located on the chromosome and the circular plasmid in isolate KMS. These findings were correlated with identity in the sequences of the predicted promoters for each of the dioxygenase genes. Comparison of the benzoate-degrading gene clusters of isolate KMS as well as those in four other mycobacterium isolates to those in other benzoate-degrading microbes, the Gram-positive \textit{Rhodococcus} (Kitagawa et al. 2001) and Gram-negative \textit{Acinetobacter} isolates (Collier et al. 1997) was performed. The mycobacterium isolates were from the same USA site as isolate KMS, isolates JLS and MCS, and from two other locations in the USA, isolates PYR-1 and PYR-GCK, as well as an isolate, Spyr1, from Greece (Karabika et al. 2009). These studies showed that the arrangement of benzoate catabolism genes in the mycobacterium isolates differs from those of the other microbes but is conserved among these isolates from different geographic areas.

\textbf{Materials and Methods}

Growth of isolate KMS on pyrene or sodium benzoate as the sole carbon source

Cells of \textit{Mycobacterium sp.} strain KMS were stored at -80 °C in 15 % glycerol. The cells were cultured on modified minimal medium (Kim et al. 2003) plates amended with 2.5 mM pyrene or 5 mM sodium benzoate (Sigma Aldrich, St. Louis, MO, USA). Cells from the plates were transferred to liquid medium with defined concentrations of pyrene or benzoate to determine growth rate. The pyrene was dissolved in methanol at 5 mg ml\textsuperscript{-1} and 5 ml were transferred to each flask. The methanol was evaporated before adding 50 ml medium. The sodium benzoate was sterilized by filtration through a 0.2 μm pore size
Whatman filter (Whatman Inc., NJ, USA) and added separately to the minimal medium. The cells in liquid medium were shaken at 220 rpm at 25 °C. Cell growth was monitored at OD 600 nm, and colony forming units per ml (CFU ml⁻¹) were assessed by dilution plating on Luria-Bertani (LB) agar medium; colonies were counted after 7-10 days growth at 25 °C.

Bioinformatic analysis of genes encoding ring-hydroxylating dioxygenases in isolate KMS

MicrobesOnline (http://www.microbesonline.org) was used to identify genes encoding the α- and β-subunits of aromatic ring dioxygenases in the genome of *Mycobacterium* isolate KMS. These loci were confirmed using the JGI-IMG website (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi). The predicted amino acid sequences were used in BLASTp to determine matches to proteins in other microbes. The phylogenetic relationships at the amino acid level were examined using the software program MEGA4 (http://www.megasoftware.net/) using neighbor joining analysis (Saitou and Nei 1987; Tamura et al. 2007). This methodology has been used in other studies of bacterial gene duplication (Gu et al. 2009). The analysis was performed with 1000 replications.

The potential promoter regions for the forty-four genes encoding aromatic ring-hydroxylating dioxygenase α- and β-subunits were predicted by the online software Neural Network Promoter Prediction (http://www.fruitfly.org/seq_tools/promoter.html) (Reese 2001). The minimum promoter score was set as 0.4, from the range of 0 to 1, where 1 represents the highest possibility of being a promoter region. The position of the promoters also was based on the predicted operons in MicrobesOnline and by searching through Artimis (Rutherford et al. 2000). The sequences totaling 60 bp from 50 bp
upstream from the predicted transcriptional start sites for all the ring-hydroxylating
dioxygenase α- and β-subunits genes were aligned by ClustalX2 acquired from
http://www.clustal.org/.

Induction of different dioxygenase genes by pyrene or benzoate

Total RNA was isolated by bead beating using TRI-reagent from middle-log phase
cells (OD 600 nm at 0.15-0.20 for pyrene-induced cultures grown on 2.5 mM pyrene and
OD 600 nm at 0.25-0.30 for cultures grown with 5 mM sodium benzoate (Molecular
Research Center, Inc. OH, USA). The RNA was treated by RNase-free DNase (Promega
Corporation, Madison, WI, USA) and purified by extraction with
phenol:chloroform:isoamyl alcohol (25:24:1) to remove the DNase. The concentration of
the purified RNA was determined by using a NanoDrop ND-1000 spectrophotometer
(NanoDrop Technologies, DE, USA). PCR was performed with the DNase-treated RNA
as the template to ensure the complete digestion of DNA. First-strand cDNA was
synthesized by using specific reverse primers for the benA gene Mkms_1391, the
chromosomal nidA gene (Mkms_1668), or the 16S rRNA genes with 1 µg purified total
RNA. The primer sequences are listed in Table 3.1. The specificity of the primers was
verified using genomic DNA and sequencing of the product prior to transcript analysis.
The protocol for M-MuLV reverse transcriptase (Promega Corporation, Madison WI,
USA) was followed. The synthesized first strand cDNA was used in a 20 µl volume end-
point PCR reaction containing: 11.3 µl H₂O, 2 µl MgCl₂ (25 mM), 2 µl 10× Taq DNA
polymerase buffer, 1 µl dNTP mix (10 mM each), 1 µl forward primer (final
ccentration 1 µM), 1 µl reverse primer (final concentration 1 µM), 1 µl cDNA, 0.3 µl
Taq DNA polymerase (1.5 unit) (Fermentas Inc., MD, USA). The PCR was performed in
an Eppendorf Mastercycler gradient PCR machine (Certified GeneTool, Inc., CA, USA). The reactions were heated at 94 °C for 2 min followed by 35 cycles of 94 °C 30 sec, 58 °C 40 sec, and 72 °C 40 sec, and an additional 72 °C 10 min was performed before the end of the PCR. The PCR products were checked by running a 0.8% agarose gel under 80 V in 1 × TAE buffer. The gel was examined by a UV transilluminator (UVP Inc., CA, USA) and images were recorded photographically (AEase™ version 3.25, A Innotech Corporation, CA, USA).

Quantification of change of transcriptional accumulation of four dioxygenase β-subunit genes with homologous promoters

*Mycobacterium sp.* strain KMS was cultured in a modified basal salts medium (BSM⁺) based on Liang et al. (2006) with or without pyrene at a final concentration 100 µM. Methods described above were used to generate first strand cDNA from isolated total RNA.

DNA products amplified from KMS genomic DNA using primers for two *nidB* genes (Mkms_1667 and Mkms_5625), two *nidB2* genes (Mkms_1660 and Mkms_5621), and 16S rRNA genes (Table 3.1) were used as controls and to standardize the quantitative real time PCR (Q-PCR). The primers were designed to be specific for each of the genes through using the differences in base pair sequences that exist between the homologs as described in Zhang and Anderson (2012). As shown in Table 3.1 the products generated were of different sizes for each of these genes. The specificity was confirmed by sequencing the DNA products derived from PCR with genomic DNA. Six 10-fold dilutions were made for each of the DNA products to make a concentration gradient used in the Q-PCR to create standards.
Table 3.1 Specific primers for genes potentially encoding the subunits of dioxygenases and the 16S rRNA genes for strain KMS

<table>
<thead>
<tr>
<th>Gene locus</th>
<th>Product encoded</th>
<th>Primer sequences</th>
<th>Product length</th>
<th>Annealing temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkms_1660</td>
<td>Dioxygenase β-subunit</td>
<td>For TCGATGCGGGCGCAGTTCCGA, Rev CGTCGACCGGTTCCGGCAG</td>
<td>339 bp</td>
<td>58 °C</td>
</tr>
<tr>
<td>Mkms_1667</td>
<td>Dioxygenase β-subunit</td>
<td>For GGGTGGGTGCGATCGCGCAT, Rev CCGGTCCGGGAAACCGTGCGA</td>
<td>147 bp</td>
<td>58 °C</td>
</tr>
<tr>
<td>Mkms_5621</td>
<td>Dioxygenase β-subunit</td>
<td>For GGAACGACGACTACACGGGT, Rev CGCGGCTGCGGAAGAACAAG</td>
<td>177 bp</td>
<td>58 °C</td>
</tr>
<tr>
<td>Mkms_5625</td>
<td>Dioxygenase β-subunit</td>
<td>For TGTTCCGGGAGGCGGAGCTC, Rev TGCACCGGAAACACATAG</td>
<td>93 bp</td>
<td>58 °C</td>
</tr>
<tr>
<td>Mkms_1391</td>
<td>Dioxygenase α-subunit</td>
<td>For ATCGTCGACCTACACCTACGAC, Rev GTGAACCCTTCACCATGCTC</td>
<td>320 bp</td>
<td>58 °C</td>
</tr>
<tr>
<td>Mkms_1668</td>
<td>Dioxygenase α-subunit</td>
<td>For ACCCGCGACTTCCAATGCCT, Rev GCCCGCCGGATAGGCCCTCC</td>
<td>96 bp</td>
<td>55 °C</td>
</tr>
<tr>
<td>Mkms_R0021</td>
<td>16S rRNA genes</td>
<td>For TGTCGTGAGATGTTGGGATAG, Rev ATCTGCATAGCGACTCC</td>
<td>288 bp</td>
<td>58 °C</td>
</tr>
</tbody>
</table>
Each Q-RT-PCR reaction contained 1 µl DNA, 0.4 µl of primers (final concentration 1 µM) for the relevant gene, 10 µl 2 × master mix from DyNAmo™ HS SYBR® Green Q-PCR Kit (Invitrogen, Carlsbad, CA, USA), and 8.6 µl H₂O. The PCR was performed in a DNA Engine Opticon 2 with continuous fluorescence detection system (MJ Research, Incorporated, MA, USA). The samples were heated at 95 °C for 15 min before cycling for 40 cycles of 94 °C 30 sec, 58 °C 40 sec, 72 °C 1 min, 73 °C 1 sec. For each cDNA sample, triplicate reactions were performed. The average Ct value was plotted against the log₂DNA mass and regression analysis was performed. Changes in transcript accumulation of the two nidB genes and two nidB2 genes were calculated based on the method mentioned in Stratagene PCR manual (2007) by using the transcriptional accumulation of 16S rRNA genes as control.

**Results**

Isolate KMS utilizes both pyrene and benzoate as the sole carbon source

Isolate KMS grew on minimal medium agar plates with either 2.5 mM pyrene or 5.0 mM sodium benzoate as the sole carbon sources. Growth in liquid medium from plate-grown inocula (Fig. 3.1) showed the lag phase was more extended as the concentration of benzoate increased from 5 to 25 mM benzoate and no growth was observed at 50 mM. The final culture density increased as the substrate concentration was raised from 5 to 25 mM benzoate. Growth on pyrene as the sole carbon source was achieved after a longer lag phase (160 h) and then the growth rate was slower than on benzoate (Fig. 3.1). From CFU ml⁻¹ data, the doubling time of KMS grown with 5-25 mM sodium benzoate was 5.8 ± 0.6 h compared with the doubling time on 2.5 mM pyrene of 57.3 ± 11.5 h.
Distribution of genes encoding $\alpha$- and $\beta$-subunits of dioxygenases within the genome of *Mycobacterium* isolate KMS

Forty-four genes encoding $\alpha$- and $\beta$-subunits of dioxygenase were detected on the chromosome, circular plasmid, and linear plasmid of isolate KMS (Fig. 3.2 and Table 3.2). We have used the designations *adi* and *bdi* followed by a distinguishing letter to specify the genes encoding dioxygenase $\alpha$- and $\beta$-subunits respectively. We examined the annotations from MicrobesOnline where domains are predicted based on the clusters of orthologous group (COG) and interpro protein sequence (IPR) characterization (Tatusov et al. 2000; Mulder et al. 2005). The $\alpha$-subunits were recognized by COG4638 (large subunits of phenylpropionate dioxygenase and ring-hydroxylating dioxygenases), or IPR001663 and the $\beta$-subunits by COG5517 (small subunits of ring hydroxylating dioxygenases). Genes encoding $\alpha$-subunits outnumbered the genes encoding the $\beta$-subunits; ten large subunit genes were lone $\alpha$-subunit genes without an adjacent $\beta$-subunit gene. Domain recognition by MicrobesOnline revealed all of the lone $\alpha$-subunit genes as being associated with the Rieske [2Fe-2S] iron-sulphur domain (IPR017941). Five of the genes, *adiH, adiI, adiJ, adiK* and *adiL* had COG4638 as the other major conserved domain characteristic of phenylpropionate dioxygenases; a second group, *adiM, adiN, adiO* and *adoP* had the IPR001663 domain characteristic of an $\alpha$-subunit dioxygenase.

The gene *adiG* contained the domains COG4638, IPR001663, and IPR015879, characteristic of the C terminus of an aromatic ring dioxygenase $\alpha$-subunit. Orthologs to the *nid* genes (*nidA, nidB, nidA3*, and *nidB3*) already characterized in isolate PYR-1 by Cerniglia’s group for their proven functionality in pyrene degradation (Khan et al. 2001; Kim et al. 2006) were detected in the genome of KMS. These genes were among the
eight pairs of dioxygenase genes clustered on the chromosome, from Mkms_1390/1391 to Mkms_1717/1718 (Fig. 3.2). Only two genes at loci Mkms_1660 and 5621 encoding the β-subunit, NidB2, were unpaired with an α-subunit gene. Five sets of chromosomal genes were duplicated on the linear and/or circular plasmid (Fig. 3.2 and Table 3.2); this group included orthologs nidA/B, nidB2 and nidA3/B3.

**Fig. 3.1** Growth curves of isolate KMS grown on 5, 10, 15, 20, 25, or 50 mM sodium benzoate or 2.5 mM pyrene as sole carbon sources in liquid minimal medium (Kim et al. 2003) with shaking at 220 rpm at 25 °C. Data and the standard error are the means from two replicates.
Table 3.2 Distribution of genes encoding dioxygenase α- and β-subunits in the genome of KMS

<table>
<thead>
<tr>
<th></th>
<th>Single copy <em>a</em></th>
<th>Double copy <em>a</em></th>
<th>Triple copy <em>a</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lone genes</td>
<td>Mkms_2506 (<em>adiG</em>), ch;</td>
<td>Mkms_1660 (<em>nidB2</em>), ch;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mkms_3775 (<em>adiH</em>), ch;</td>
<td>Mkms_5621 (<em>nidB2</em>), cp.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mkms_3783 (<em>adiI</em>), ch;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mkms_3790 (<em>adiJ</em>), ch;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mkms_3820 (<em>adiK</em>), ch;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mkms_4493 (<em>adiL</em>), ch;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mkms_0306 (<em>adiM</em>), ch;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mkms_2305 (<em>adiN</em>), ch;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mkms_3642 (<em>adiO</em>), ch;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mkms_4728 (<em>adiP</em>), ch.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired genes</td>
<td>Mkms_1390/1391 (<em>benB/benA</em>), ch;</td>
<td>Mkms_1667/1668 (<em>nidB/nidA</em>), ch;</td>
<td>Mkms_1690/1691 (<em>nidA3/nidB3</em>), ch;</td>
</tr>
<tr>
<td></td>
<td>Mkms_1647/1648</td>
<td>Mkms_5625/5626</td>
<td>Mkms 5639/5640</td>
</tr>
</tbody>
</table>
Analysis using ALIGN of the nucleotide sequences revealed the *nidB* and *nidB2* were homologs (96.9 %) possibly only recently duplicated according to criteria discussed by Gevers et al. (2004). All of the other β-subunit genes have homology with *nidB* greater than 30% for over 150 amino acids characterizing them as homologs according to

\[ \text{ch means chromosome, cp, circular plasmid, and lp, linear plasmid} \]
Fig. 3.2 The locations of the genes encoding lone ring-hydroxylating dioxygenase α- and β-subunit genes and the paired ring-hydroxylating genes in the genome of *Mycobacterium sp.* isolate KMS. Designation of genes *nidA*, *nidB*, *nidB2*, *nidA3*, and *nidB3*
followed those used for *M. vanbaalenii* PYR-1 (Khan et al. 2001; Kim et al. 2006; Kim et al. 2007) and designation of genes *benA* and *benB* followed those identified in *Rhodococcus jostii* RHA1 (Kitagawa et al. 2001). Other genes are designated as *adi* (α-subunits of dioxygenases) or *bdi* (β-subunits of dioxygenases) followed by a specifying letter. The locus numbers of the gene is shown. The red capitalized P shows the position of the predicted promoter and the red line predicts the operon under control of that promoter. Block duplication between the chromosome and plasmids is shown by the blue lines.
Gevers et al. (2004): identity of nidB to bdiW was 57.9 %, and nidB3 55.8 %; other β-subunit genes had between 48.6 to 52.1 % homology to nidB (Table 3.3). Similarly all α-subunit genes were homologs: compared with the nidA sequence all genes had between 58 % for nidA3 and adiW to 47.1 % homology (Table 3.4). Homology among the lone α-subunit genes was between 47.0% (adiH and adiO) and 70.5% (adiM and adiP) (Table 3.5).

BLASTp analysis predicts that other dioxygenase genes encode proteins related to utilization of an array of substrates. Genes at loci Mkms_1391 and Mkms_1390 encoded α- and β-subunits of a benzoate dioxygenase respectively and, consequently, we use the designation of benA and benB. The BenA protein had 71 % identity with the α-subunit of benzoate dioxygenase (YP_702347.1) and BenB 74% with the β-subunit of benzoate dioxygenase of *Rhodococcus jostii* RHA1 (YP_702348.1). The protein AdiV had 95% identity to phthalate 3,4-dioxygenase α-subunit (YP_707370.1) from *Rhodococcus jostii* RHA1. AdiW had 90 % identity with NidA or NidA3 (ADH94655.1 and ADH94645.1) encoded by *Mycobacterium sp.* strains and other naphthalene-degrading ring-hydroxylating dioxygenases in other bacteria. AdiX had 97 % identity to a Rieske (2Fe-2S) domain protein (YP_001261401.1) in *Sphingomonas* and 99 % identity to a Rieske (2Fe-2S) domain protein (ZP_06846381.1) in *Burkholderia* strains. AdiY had 96 % identity to the α-subunit of dibenzofuran dioxygenase (ZP_06829928.1) in a *Terrabacter* strain and 93 % identity to a type IV biphenyl 2,3-dioxygenase (YP_002777102.1) in a *Rhodococcus* strain. There was 99 % identity between AdiZ and the α-subunits of a biphenyl 2,3-dioxygenase α-subunit (YP_707265.1) and a benzene dioxygenase (YP_002777102.1) in *Rhodococcus* strains. Thus, this characterization indicates that the
putative dioxygenases encoded by the genes could metabolize an array of aromatic ring structures.

**Table 3.3** The homologies of all the chromosomal ring-hydroxylating dioxygenase β-subunit genes with the chromosomal $nidB$ gene (locus Mkms_1667) in *Mycobacterium* isolate KMS as analyzed by program ALIGN (http://workbench.sdsc.edu/)

<table>
<thead>
<tr>
<th>Locus number and designation</th>
<th>Homology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkms_1390, $benB$</td>
<td>50.3%</td>
</tr>
<tr>
<td>Mkms_1648, $bdiV$</td>
<td>49.1%</td>
</tr>
<tr>
<td>Mkms_1660, $nidB2$</td>
<td>96.9%</td>
</tr>
<tr>
<td>Mkms_1667, $nidB$</td>
<td>100%</td>
</tr>
<tr>
<td>Mkms_1673, $bdiX$</td>
<td>48.6%</td>
</tr>
<tr>
<td>Mkms_1691, $nidB3$</td>
<td>55.8%</td>
</tr>
<tr>
<td>Mkms_1699, $bdiY$</td>
<td>51.2%</td>
</tr>
<tr>
<td>Mkms_1710, $bdiZ$</td>
<td>52.1%</td>
</tr>
<tr>
<td>Mkms_1717, $bdiW$</td>
<td>57.9%</td>
</tr>
</tbody>
</table>

Block duplication

Block duplication of segments containing dioxygenase genes in isolate KMS is apparent (Fig. 3.2). The genes pairs designated as $a/bdiZ$ and $a/bdiY$ and $nidA3B3$ were
present in three copies, located on the chromosome, circular plasmid and linear plasmid. The two lone \textit{nidB2} genes (one chromosomal and one plasmid) were grouped in potential operons with the \textit{nidB/nidA} genes showing block duplication on the chromosome and the circular plasmid (Fig. 3.2). However, rearrangement of the block duplications is noted between the chromosome and plasmids as denoted by using blue lines in Fig. 3.2. Block duplication was supported by BLAST analysis showing that chromosomal and plasmid duplicated genes had 96 \% to 100 \% homology, with the exception of lower (90 \%) homology between the \textit{bdiY} genes at loci Mkms_1699 (chromosome) and 5486 (linear plasmid) (Table 3.6).

Phylogenetic analysis

The phylogenetic relationship between \(\alpha\)- or \(\beta\)-subunit protein sequences was determined to better understand the divergence of these genes within the single microbe (Fig. 3.3). The tree for the \(\alpha\)-subunits (Fig. 3.3a and 3.3b) showed the clustering of sequences from unpaired genes encoding \(\alpha\)-subunits (AdiG-P) away from the \(\alpha\)-subunits that function with \(\beta\)-subunits. Further analysis showed that the \(\alpha\)-subunits in isolate KMS could be classified into the different types based on the system published by Kweon et al. (2008). This finding indicates that the proteins are associated with different electron transport systems and have distinct substrate specificities. Fig. 3.4 shows the tree obtained for the KMS \(\alpha\)-subunit proteins using the standard proteins employed in the Type designation. There are no proteins that correspond to Type III\(\alpha\beta\) or IV designations within the KMS genome. Also a subset of the proteins, the clustered lone \(\alpha\) subunits AdiG/H/I/J/K/L and the clustered AdiZ and AdiY proteins have no homology to the
standard proteins used for the type classification by Kweon et al. 2010. The phylogenetic relationships for the β-subunits are simpler and support that NidB2 and NidB subunits likely arose from duplication of a single gene.

**Table 3.4** The homologies of all the chromosomal ring-hydroxylating dioxygenase α-subunit genes with the chromosomal nidA gene (locus Mkms_1668) in *Mycobacterium* isolate KMS as analyzed by program ALIGN (http://workbench.sdsc.edu/)

<table>
<thead>
<tr>
<th>Locus number and designation</th>
<th>Homology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkms_0306, adiM</td>
<td>48.6%</td>
</tr>
<tr>
<td>Mkms_2305, adiN</td>
<td>47.6%</td>
</tr>
<tr>
<td>Mkms_3642, adiO</td>
<td>47.1%</td>
</tr>
<tr>
<td>Mkms_4728, adiP</td>
<td>48.4%</td>
</tr>
<tr>
<td>Mkms_2506, adiG</td>
<td>50.1%</td>
</tr>
<tr>
<td>Mkms_3775, adiH</td>
<td>51.5%</td>
</tr>
<tr>
<td>Mkms_3783, adiI</td>
<td>51.7%</td>
</tr>
<tr>
<td>Mkms_3790, adiJ</td>
<td>51.1%</td>
</tr>
<tr>
<td>Mkms_3820, adiK</td>
<td>51.4%</td>
</tr>
<tr>
<td>Mkms_4493, adiL</td>
<td>51.2%</td>
</tr>
<tr>
<td>Mkms_1391, benA</td>
<td>53.2%</td>
</tr>
</tbody>
</table>
Alignment of the predicted promoters

Sequences encompassing 60 bp from 50 bp upstream from the predicted transcriptional start sites for all dioxygenase genes were aligned by ClustalX2 (Fig. 3.5). The genes nidA, nidB, adiX, and bdiX were predicted to be in the same operon and, thus, have the same promoter. The alignment showed that certain promoters had high identity supporting duplication of blocks of genes. There was 75-99% identity for predicted promoters of four operons containing nidB/A and nidB2 genes, 92.5% to 97.5% identity for the triple copies of nidA3/B3 genes, 97.5% to 37.5% for the copies of the a/bdiZ genes, and 100% homology for the promoters of the triple copies of a/bdiY genes. The promoters of other paired genes benA/B, a/bdiV, and a/bdiW or the lone α-subunit genes adiG/H/I/J/K/L/M/N/O/P lacked alignments.

<table>
<thead>
<tr>
<th>Gene Combination</th>
<th>Identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkms_1647, adiV</td>
<td>51.2%</td>
</tr>
<tr>
<td>Mkms_1668, nidA</td>
<td>100%</td>
</tr>
<tr>
<td>Mkms_1672, adiX</td>
<td>50.1%</td>
</tr>
<tr>
<td>Mkms_1690, nidA3</td>
<td>58.4%</td>
</tr>
<tr>
<td>Mkms_1698, adiY</td>
<td>49.7%</td>
</tr>
<tr>
<td>Mkms_1709, adiZ</td>
<td>51.4%</td>
</tr>
<tr>
<td>Mkms_1716, adiW</td>
<td>58.9%</td>
</tr>
</tbody>
</table>
Table 3.5 The homologies of the lone ring-hydroxylating dioxygenase α-subunit genes of *Mycobacterium* isolate KMS as analyzed by program ALIGN (http://workbench.sdsc.edu/)

<table>
<thead>
<tr>
<th>Homology</th>
<th>adiG</th>
<th>adiH</th>
<th>adiI</th>
<th>adiJ</th>
<th>adiK</th>
<th>adiL</th>
<th>adiM</th>
<th>adiN</th>
<th>adiO</th>
<th>adiP</th>
</tr>
</thead>
<tbody>
<tr>
<td>adiG</td>
<td>50.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adiH</td>
<td>51.4%</td>
<td>51.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adiI</td>
<td>52.9%</td>
<td>51.5%</td>
<td>66.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adiJ</td>
<td>50.6%</td>
<td>53.0%</td>
<td>61.1%</td>
<td>67.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adiK</td>
<td>49.7%</td>
<td>54.2%</td>
<td>51.8%</td>
<td>52.4%</td>
<td>52.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adiL</td>
<td>50.7%</td>
<td>47.7%</td>
<td>50.4%</td>
<td>50.7%</td>
<td>49.4%</td>
<td>49.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adiM</td>
<td>51.5%</td>
<td>47.0%</td>
<td>48.6%</td>
<td>48.7%</td>
<td>49.3%</td>
<td>46.4%</td>
<td>52.1%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adiN</td>
<td>51.3%</td>
<td>47.7%</td>
<td>50.1%</td>
<td>50.6%</td>
<td>48.5%</td>
<td>46.3%</td>
<td>51.5%</td>
<td>52.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adiO</td>
<td>50.0%</td>
<td>49.6%</td>
<td>49.5%</td>
<td>49.2%</td>
<td>49.4%</td>
<td>49.1%</td>
<td>70.5%</td>
<td>52.6%</td>
<td>51.7%</td>
<td></td>
</tr>
</tbody>
</table>

Benzoate and pyrene differentially induced ring-hydroxylating dioxygenase genes

The differences between the promoter sequences suggested that they could dictate different regulatory conditions. Consequently, we examined accumulation of transcripts from the *benA* gene, at locus Mkms_1391, predicted to encode the large subunit for a benzoate dioxygenase and the *nidA* gene, encoding the large subunit of a dioxygenase.
involved in transformation of pyrene. Reverse transcriptase (RT) PCR analysis revealed accumulation of transcripts from *benA* occurred in cells cultured on benzoate but not pyrene. In contrast the *nidA* gene, from locus Mkms_1668, was expressed during pyrene growth but not during growth on benzoate. The PCR products from these analyses are shown along with PCR products from the 16S rRNA genes used as a control for equal loading of samples in Fig. 3.6.

**Table 3.6** The homologies of the replicated ring-hydroxylating dioxygenases α- and β-subunit genes of *Mycobacterium* isolate KMS on the circular and linear plasmids to the chromosomal genes as analyzed by the program ALIGN (http://workbench.sdsc.edu/)

<table>
<thead>
<tr>
<th>Homology to the chromosomal gene</th>
<th>Locus number (Mkms_)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chromosomal gene</td>
</tr>
<tr>
<td>Designation</td>
<td></td>
</tr>
<tr>
<td>Ring-hydroxylating dioxygenases</td>
<td>nidA</td>
</tr>
<tr>
<td>α-subunit genes</td>
<td>nidA3</td>
</tr>
<tr>
<td>adiX</td>
<td>1672</td>
</tr>
<tr>
<td>adiY</td>
<td>1698</td>
</tr>
<tr>
<td>adiZ</td>
<td>1709</td>
</tr>
<tr>
<td>Ring-hydroxylating</td>
<td>nidB</td>
</tr>
</tbody>
</table>
Dioxygenase genes with homologous promoters were similarly regulated

The homology between predicted promoters for both the chromosomal and plasmid copies of the *nidB* and *nidB2* genes suggested that these genes would be coordinately expressed. This possibility was tested using Q-RT-PCR to assess the RNA accumulating from these genes during growth on pyrene. The fold increases in expression for each gene, normalized based upon expression of the 16S rRNA genes, had a mean fold increase of 3-fold and, thus, were of similar magnitudes (Table 3.7).

<table>
<thead>
<tr>
<th>Dioxygenases β-subunit genes</th>
<th>nidB2</th>
<th>1660</th>
<th>5621 (98%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nidB3</td>
<td>1691</td>
<td>5640 (98%)</td>
</tr>
<tr>
<td></td>
<td>bdiX</td>
<td>1673</td>
<td>5631 (98%)</td>
</tr>
<tr>
<td></td>
<td>bdiY</td>
<td>1699</td>
<td>5648 (99%)</td>
</tr>
<tr>
<td></td>
<td>bdiZ</td>
<td>1710</td>
<td>5611 (99%)</td>
</tr>
</tbody>
</table>
Fig. 3.3 Phylogenetic trees generated by MEGA 4 for the predicted protein sequences for the α-subunits of dioxygenases (a) and the β-subunits of dioxygenase (b) for *Mycobacterium* isolate KMS. The tree was obtained by neighbor-joining and represents 1000 replications. Overlayed on this tree are balloons indicating the classification of the dioxygenases into the group (Type I to Type V) based on the work of Kweon et al. (2008) as determined by analysis shown in Fig. 3.4. Designation of proteins NidA, NidB, NidB2, NidA3, and NidB3 followed those identified in *M. vanbaalenii* PYR-1 (Khan et al. 2001;
Kim et al. 2006; Kim et al. 2007). Designation of proteins BenA and BenB was because of homology to proteins in *Rhodococcus jostii* RHA1 (Kitagawa et al. 2001). Other as yet uncharacterized aromatic ring-hydroxylating dioxygenase α-subunits are designated as Adi, and the β-subunits as Bdi. The term ch indicates a chromosomal gene, cp a circular plasmid gene, and lp a linear plasmid genes. The locus numbers of the genes encoding the proteins are shown. The scale bar means 0.2 changes per amino acid position.
Fig. 3.4 The genome of *Mycobacterium sp.* isolate KMS contains an array of dioxygenase α-subunits based on the classification method of Kweon et al. (2008, 2010). This phylogenetic tree generated by MEGA 4 using neighbor-joining with 1000 replications for all the ring-hydroxylating dioxygenase α-subunits potentially encoded by isolate KMS as well as ring-hydroxylating dioxygenase α-subunits from other bacteria used as standards in the classification method of Kweon et al. (2008, 2010). The scale bar shows 0.2 amino acid changes per amino acid site. Ring-hydroxylating dioxygenases were classified into five groups (Type I to Type V) (Kweon et al. 2008). α-subunits
belonging to type Iα are noted by green, type Iαβ by red, type II by orange, type IIIα by yellow, type IIIαβ by turquoise, type IV by blue, and type V by pink. Designation of proteins NidA and NidA3 followed those identified in *Mycobacterium vanbaalenii* PYR-1 (Khan et al. 2001; Kim et al. 2006; Kim et al. 2007). Designation of protein BenA was because of homology to the benzoate dioxygenase in *Rhodococcus jostii* RHA1 (Kitagawa et al. 2001). Other as yet uncharacterized aromatic ring-hydroxylating dioxygenase α-subunits are designated as Adi. The term ch indicates a chromosomal gene, cp a circular plasmid gene, and lp a linear plasmid genes. The locus tags for the genes encoding the proteins are shown

**Fig. 3.5** Alignments of the predicted promoter regions of all the ring-hydroxylating dioxygenase genes for isolate KMS. The alignment by ClustalX2 was of 60 bp total showing 50 bp upstream from the predicted transcriptional start site. The symbol ch means chromosomal gene, cp circular plasmid gene, and lp linear plasmid gene. The promoters with homology are shown in groups
Fig. 3.6 Transcriptional analysis of genes benA (Mkms_1391) and nidA (Mkms_1668) when isolate KMS was cultured on either 5 mM sodium benzoate or 2.5 mM pyrene as the sole carbon source. The expression of 16S rRNA genes was used as control. Lane 1, 100 bp gene ladder; lane 2 product from cDNA of corresponding genes from cells grown on 5 mM sodium benzoate; lane 3 product from cDNA of corresponding genes from cells grown on 2.5 mM pyrene. Genes used in this study and their products (in brackets) are shown. The study is from one of two studies with the same results.
Table 3.7 The fold change in expression for genes encoding the β-subunit of dioxygenase normalized to the expression of 16S rRNA genes during growth of isolate KMS on BSM⁺ medium amended with pyrene relative to growth on BSM⁺ medium without pyrene

<table>
<thead>
<tr>
<th>Selected gene and location</th>
<th>Fold change based on the mass of cDNA ¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkms_1660 (nidiB2)</td>
<td>3.3±1.5</td>
</tr>
<tr>
<td>Chromosome</td>
<td></td>
</tr>
<tr>
<td>Mkms_1667 (nidiB)</td>
<td>3.0±1.0</td>
</tr>
<tr>
<td>Chromosome</td>
<td></td>
</tr>
<tr>
<td>Mkms_5621 (nidiB2)</td>
<td>2.8±1.1</td>
</tr>
<tr>
<td>Plasmid</td>
<td></td>
</tr>
<tr>
<td>Mkms_5625 (nidiB)</td>
<td>3.2±0.9</td>
</tr>
</tbody>
</table>

¹ Data were normalized based on the expression of the 16S rRNA genes and standard deviations were calculated based on results from four separate Q-RT-PCR studies
Discussion

The pyrene-degrading *Mycobacterium* isolate KMS had a plethora of genes encoding ring-hydroxylating dioxygenase subunits that were clustered on the chromosome as well as on the circular and linear plasmids. There is evidence of block duplication for the operons containing the genes that based on observations of Cerniglia’s group initiated ring oxidation in the high molecular weight PAHs (Khan et al. 2001; Kim et al. 2006; Kweon et al. 2010). Duplication of genes leading to the existence of families with related function is common in bacteria (Gevers et al. 2004; Serres et al. 2009). These events are postulated to permit adaption to different environments. Andersson and Hughes (2009) posited “Whenever cellular growth is restricted escape from these growth restrictions often occurs by gene duplication-amplification events that resolve the selective problem.” They provide examples with *Alcaligenes*, *Pseudomonas* and *Acinetobacter* isolates where gene duplication relates to adapted utilization of contaminants such as alkanes and benzoate. Similarly, Gevers et al. (2004) stated that “gene duplication is an evolutionary response in bacteria exposed to different selection pressures such as starvation.” The ability of certain environmental *Mycobacterium* isolates, such as KMS which came from PAH-contaminated soils, to develop the ability to utilize the recalcitrant PAH, pyrene, fits well within this concept of gene duplication to overcome starvation and promote adaption to novel carbon sources. Supporting evidence comes from comparison of the KMS genome with the genomes of pathogenic, *M. tuberculosis*. Dioxygenase genes involved in house-keeping pathways (taurine catabolism, breakdown of nitro-organics) exist in both environmental and pathogenic isolates (Eichhorn et al. 2000; Stiens et al. 2006). However the pathogenic isolates possess limited genes encoding enzymes for
modification of complex aromatics (e.g. two genes for large subunit phenylpropionate dioxygenase in \textit{M. tuberculosis} H37Ra among the 25 predicted dioxygenase genes with general metabolic activities). In contrast there is an extensive ring-hydroxylating dioxygenase gene family with specificity for the high molecular PAHs in KMS. This finding is seen also in two other \textit{Mycobacterium} isolates, MCS and JCS, from the same Montana site as KMS, as well as two isolates from other PAH-contaminated sources, \textit{M. vanbaalennii} PYR-1 and \textit{M. gilvum} PYR-GCK (Kim et al. 2008; Zhang and Anderson 2012). Additional genes with the COG4638 signature for phenylpropionate \( \alpha \)-subunits are found in other \textit{Mycobacterium} isolates, such as the pathogenic \textit{M. marinum}, \textit{M. avium} and environmental \textit{M. smegmatis} according analysis with MicrobesOnline Treebrowser. Thus there is circumstantial evidence that the expanded gene family offers an advantage when the microbes are forced to utilize the multi-aromatic ring PAHs.

In many cases, members of the duplicated genes are modified further (e.g. by mutation or acquisition of novel domains) such that there is clear divergence of function (Iwabe et al. 1989; Notebaart et al. 2005). Indeed extensive superfamilies have been annotated within many bacterial genomes (Serres et al. 2009). For example the genes in the crotonase family in \textit{Escherichia coli} encode for enzymes exhibiting a range of distinct activities (Serres et al. 2009). Clearly this family of dioxygenase genes has not diversified to such a large extent as they retained high homology.

The existence of some of the amplified genes on plasmids in isolate KMS is consistent with similar observations in other microbes. Andersson and Hughes (2009) cited the extensively studied cases of amplification of genes for drug resistance on plasmids. In KMS, the circular and the linear plasmids have few genes with predicted
functions and lack genes predicted to be associated with drug resistance. The circular plasmid has genes involved in resistance towards cadmium and copper, suggesting that genes involved in other needed environmental adaptions are present; these metal-resistance genes were not duplicated on the chromosome.

BLASTp analysis of the ring hydroxylating dioxygenase genes revealed homology with genes encoding proteins with assayed functionality. For instance, the additional paired genes we designated as V, W, X, Y, and Z genes in isolate KMS were predicted to encode the $\alpha$- and $\beta$-subunits of dioxygenase acting on small ring structures (e.g. V, phthalate; W, naphthalene; Y and Z, biphenyl). In this paper we correlated expression from the $\text{benA}$ gene at locus Mkms_1391 encoding a dioxygenase $\alpha$-subunit with benzoate degradation because transcripts from this locus accumulated when isolate KMS was grown on benzoate. The genes adjacent to this locus are predicted to encode other proteins with functions in benzoate degradation: Mkms_1388 encoded a regulator protein with 97% identity to a LuxP family transcriptional regulator, Mkms_1390 $\text{benB}$ encoding the $\beta$ subunit of the dioxygenase and Mkms_1389 encoded a chimeric protein of a benzoate dioxygenase reductase (BenC) and a diol dehydrogenase (BenD) as shown in Fig. 3.7. This benzoate-degrading gene cluster in KMS differed from two other types of clusters endowing benzoate catabolism (Fig. 3.7); the operons in strains of a pseudomonad and a Rhodococcus species contained five genes (Cowles et al. 2000; Kitagawa et al. 2001) whereas in Acinetobacter there were eight genes (Collier et al. 1997). Identical sequences to that in strain KMS were found in four other PAH-degrading mycobacterium strains MCS, JLS, $M. \text{gilvum}$ PYR-GCK, and Spyr1 (Fig. 3.8) but $M. \text{vanbaalenii}$ PYR-1 lacked such genes for benzoate catabolism as indicated in Kim et al.
Alignment of BenC (benzoate dioxygenase reductase) and BenD (diol dehydrogenase) encoded by *Acinetobacter sp.* ADP1 and *Rhodococcus sp.* RHA1 with the peptide encoded by Mkms_1389 of *Mycobacterium sp.* KMS by BLASTp on the NCBI website. The numbers indicate the amino acids in the predicted proteins from Mkms_1389. Amino acids homologous to the benzoate dioxygenase reductase are shown in light blue whereas amino acids homologous to the diol dehydrogenase are shown in darker blue.
Fig. 3.8 Comparison of benzoate degrading gene cluster in *Mycobacterium* strains KMS, MCS, JLS, PYR-GCK, and Spyr1 as well as two other clusters endowing benzoate catabolism from *Rhodococcus* strain RHA1 and *Acinetobacter* strain ADP1. The gene cluster in *Rhodococcus* strain RHA1 contains five genes (*benA, benB, benC, benD,* and *benK*) encoding α- and β-subunits of benzoate dioxygenase, benzoate dioxygenase reductase, benzoate diol dehydrogenase, and a benzoate transporter protein respectively (Kitagawa et al. 2001). The gene cluster in *Acinetobacter* strain ADP1 has eight genes (*benP, benK, benR, benA, benB, benC, benD,* and *benE*) encoding a putative porin protein for benzoate transport, a benzoate transporter protein, a regulator protein, α- and β-subunits of benzoate dioxygenase, benzoate dioxygenase reductase, benzoate diol dehydrogenase, and a benzoate membrane transport protein respectively (Collier et al. 1997). The gene cluster in strain KMS contains gene *benA* encoding a predicted benzoate dioxygenase α-subunit, *benB* the dioxygenase β-subunit, locus Mkms_1389 encoding a protein with two domains one with homology to a benzoate reductase and a second with benzoate diol dehydrogenase, and a gene with homology to *luxR* encoding a regulator. The gene clusters in strains MCS, JLS, and Spyr1 are identical to that in strain KMS whereas the gene cluster in PYR-GCK differs with the presence of two genes annotated as LuxR family regulators (Mflv_1352 and Mflv_1351)
Benzoate degrading gene cluster in *Rhodococcus* sp. RHA1:

```
benA ---- benB ---- benC ---- benD ---- benK
```

Benzoate degrading gene cluster in *Acinetobacter* sp. ADP1:

```
benP ---- benK ---- benR ---- benA ---- benB ---- benC ---- benD ---- benE
```

Benzoate degrading gene cluster in *Mycobacterium* sp. KMS:

```
benA ---- benB ---- \text{Mmns\textsubscript{1388} (benC benD)} ---- \text{Mmns\textsubscript{1388} (benR)}
```

Benzoate degrading gene cluster in *Mycobacterium* sp. MCS:

```
Mmns\textsubscript{1373} ---- Mmns\textsubscript{1372} ---- Mmns\textsubscript{1371} ---- Mmns\textsubscript{1370}
```

Benzoate degrading gene cluster in *Mycobacterium* sp. JLS:

```
Mjls\textsubscript{1407} ---- Mjls\textsubscript{1406} ---- Mjls\textsubscript{1405} ---- Mjls\textsubscript{1404}
```

Benzoate degrading gene cluster in *Mycobacterium gilvum* PYR-GCK:

```
Mfgy\textsubscript{1355} ---- Mfgy\textsubscript{1354} ---- Mfgy\textsubscript{1353} ---- Mfgy\textsubscript{1352} ---- Mfgy\textsubscript{1351}
```

Benzoate degrading gene cluster in *Mycobacterium* sp. Spyrl1:

```
Mspyrl\textsubscript{148230} ---- Mspyrl\textsubscript{148240} ---- Mspyrl\textsubscript{148250} ---- Mspyrl\textsubscript{148260}
```
(2008). These strains are from diverse geographical sites showing conservation of function.

Our studies illustrated regulated expression from the *benA* gene by benzoate versus pyrene. Similarly the *nidA* gene was not expressed during growth on benzoate but was induced by pyrene. Because the clusters containing *nidA* and *benA* have distinct promoters, these findings are consistent with differential activation of the promoters by aromatic ring compounds. This concept was supported by observing during pyrene growth the very similar levels of activation from the promoters that have high sequence identity of four *nidB* genes, whether these genes were located on the chromosome or plasmid.

Our findings illustrate the adaption of *Mycobacterium sp.* KMS to maximize its effectiveness in using aromatic ring structures. Adaption appears to have evolved gene duplication with divergence of the subunits of the dioxygenases towards different substrate specificity as well as the development of distinct promoter sequences. We demonstrated expression simultaneously from plasmid and chromosomal copies of a duplicated gene and we show functionality of two distinct promoter sequences.

**References**


Jiang H, Parales RE, Gibson DT (1999) The α subunit of toluene dioxygenase from *Pseudomonas putida* F1 can accept electrons from reduced ferredoxin$_{TOL}$ but is


Stratagene (2007) Introduction to quantitative PCR Methods and application guide, Stratagene La Jolla, San Diego


CHAPTER 4

UTILIZATION OF PYRENE AND BENZOATE IN MYCOBACTERIUM ISOLATE KMS IS REGULATED DIFFERENTIALLY BY CATABOLIC REPRESSION

The soil isolate, *Mycobacterium* *sp.* strain KMS, utilizes an array of carbon compounds including the aromatics benzoate and pyrene as sole carbon sources. Growth on pyrene induced both chromosomal and plasmid *nidA* genes encoding pyrene ring-hydroxylating dioxygenase α-subunits for pyrene oxidation. Diauxic growth occurred when KMS was cultured with pyrene plus either acetate, succinate, fructose, or benzoate and *nidA* expression only was detected in the second slower log-phase period. Potential cAMP-CRP binding sites exist within the promoter region of both *nidA* genes, indicating that cAMP-CRP may be involved in catabolite repression of pyrene utilization. When cultured with benzoate plus either acetate, succinate, or fructose, there was no diauxic growth. Also there was no diauxic growth on fructose plus succinate or acetate. Expression of a *benA* gene, encoding a benzoate dioxygenase α-subunit involved in the initiation of benzoate oxidation, was detected in log-phase cells from the benzoate-mixed substrate cultures at the same level as when the cells were cultured on benzoate alone. These findings suggested that catabolite repression of pyrene but not benzoate occurred in isolate KMS. These differences may help the microbe exploit the varied carbon sources available in the soil and rhizosphere environments.

---

1Coauthored by Dr. Anderson AJ., Utilization of pyrene pyrene and benzoate in *Mycobacterium* *isolate* KMS is regulated differentially by catabolic repression, © 2012 with permission from John Wiley and Sons.
Introduction

When different carbon sources are supplied simultaneously to a bacterium, they may be co-catabolized or used sequentially causing diauxic growth patterns [1]. Diauxic growth often is explained as a consequence of catabolite repression, which is a complex regulatory process. The mechanism of catabolite repression varies with the microbe [2]. Frequently the use of the preferred carbon source relates to inactivation of the transport system required for uptake of the other carbon sources [3] or regulation of genes involved in their catabolism [4, 5]. Such regulatory mechanisms are important when the bacterium is faced with multiple choices of carbon sources as well as under conditions of starvation. A soil bacterium would encounter patchy nutrition depending on plant root exudates and other soil carbohydrate sources.

In *Escherichia coli* the preferred carbon source, glucose, is transported into the cell through a phosphoenolpyruvate-dependent phosphotransferase transport system (PTS) [6]. A series of cytoplasmic and membrane proteins transfer the phosphate group from phosphoenolpyruvate (PEP) to the sugar. In *E. coli*, low glucose results in a high level of the phosphorylated PTS protein EIIA\(^{\text{Glc}}\) that activates the adenylate cyclase for generation of cyclic AMP (cAMP) from ATP [7]. The cAMP interacts with binding proteins such as the cAMP receptor protein (CRP) [8]. Gene regulation involves binding of the cAMP-CRP complex to discrete sites to influence RNA polymerase activity [9]. Sequences for binding cAMP-CRP are detected in the promoters of the regulated genes and account for both up- and down-regulation; regulated genes include those required for utilization of carbohydrates other than glucose [10, 11]. It is likely mycobacteria also use
the PTS system for uptake of glucose and fructose [12] but any role of the intermediates in this system in catabolite repression is not established.

This paper focuses on a better understanding of the metabolism of an environmental *Mycobacterium* sp. strain KMS, isolated from a polycyclic aromatic hydrocarbon (PAH)-contaminated soil [13]. KMS utilizes both the PAH, pyrene [14], and the one ring aromatic compound, benzoate, as sole carbon sources [15]. Initiation of pyrene degradation by isolate KMS is catalyzed by ring-hydroxylating dioxygenase [14] bearing an active site, a Rieske center, within the $\alpha$-subunit NidA [16]. The gene, *nidA*, is duplicated on the chromosome and the circular plasmid in isolate KMS [17]. At a different chromosomal location, a four-gene operon (Mkms_1391 to Mkms_1388) accounts for the degradation of benzoate with catechol as an intermediate. These genes encode both $\alpha$- and $\beta$-subunits of a benzoate dioxygenase, a specific dioxygenase reductase and diol dehydrogenase, and a LuxR-family regulator protein. The likely pathway (Fig. 4.1) for benzoate degradation is similar to that in other soil microbes, *Pseudomonas*, *Burkholderia*, *Rhodococcus*, *Acinetobacter*, and *Azoarcus* [18-24]. The first step in benzoate degradation also is ring oxidation catalyzed by a dioxygenase. The gene *benA*, the first gene of the operon encoding the $\alpha$-subunit of the dioxygenase, is induced by benzoate but not pyrene [15]. Benzoate occurs naturally in plants [25] and is present in soils through fungal degradation of lignin [26].
Figure 4.1. Benzoate to catechol conversion in *Mycobacterium* isolate KMS showing a key role of benzoate dioxygenase in initial ring oxidation. The pathway is based on gene annotation from the JGI-IMG database (http://img.jgi.doe.gov/cgi-bin/w/main.cgi). Genes *benA* and *benB* encode the benzoate dioxygenase subunits.

Mineralization of the PAH pyrene by isolate KMS is promoted by the presence of plants in the soil [27]. A similar observation occurs with isolates related to *M. gilvum* as they colonize the roots of aquatic reeds [28]. Several factors may be involved. The solubility of the PAH may be increased by factors in the plant root exudate [29]. Also the nutrients in root exudates, including sugars and organic acids, would support higher cell populations; KMS grows effectively on barley root exudates with demonstrated utilization of fructose [30]. However mineralization studies in barley root exudate show
that pyrene degradation is reduced as the concentration of the root wash is increased [30].

This finding raises the question of whether carbon repression of PAH utilization was occurring in KMS by the components in the root wash. Catabolite repression of the degradation of simple aromatic compounds such as benzoate by sugars and organic acid is observed in *Acinetobacter baylyi* [23]. Similarly organic acids, sugars, and amino acids in plant root exudates [31, 32], are cited as causing catabolite repression of degradation of the PAH, phenanthrene, by *Pseudomonas putida* [33].

Carbon repression by glucose in an environmental *Mycobacterium* isolate JC1 is demonstrated in utilization of carbon monoxide [34]. The observed glucose repression of expression of an operon required for growth on carbon monoxide is linked to a cAMP-based mechanism [34]. A palindrome region (TGTGA-N$_6$-TCACA) predicted to bind a cAMP-CRP complex exists between the transcriptional and translational start sites for the promoter of this operon [34]. Gene regulation by cAMP also is documented for pathogenic isolates of *M. tuberculosis* for cells grown under oxygen limitation, mimicking the conditions in human macrophages [35, 36] Regulation of gene expression by cAMP in mycobacteria is feasible because genes involved in synthesis and degradation and encoding cAMP binding proteins are present the genomes of pathogenic isolates [36, 37]. Thus cAMP-CRP is suggested to function as regulator for catabolite repression in the pathogenic isolates. However, other roles of cAMP also are invoked including a role of the secreted nucleotide in manipulating the defense responses of the pathogen’s host cells [36].

In this paper we examined catabolite repression in the *Mycobacterium* isolate KMS by determining first whether diauxic growth occurred when cells exposed to the
aromatics, benzoate and pyrene, also were provided with fructose or the organic acids, acetate and succinate. To explore whether components in the uptake system for fructose were involved in catabolite repression, the occurrence of diauxic growth was examined on fructose plus acetate or succinate. Expression from both the chromosomal and plasmid \( nidA \) genes and the chromosomal \( benA \) genes, required for the initiation of pyrene and benzoate degradation respectively, was assessed at defined growth phases to understand the relationship with preferred carbon source. The promoters for the operons containing the dioxygenase \( \alpha \)-subunit genes were examined to determine whether control of expression could be connected to a cAMP-based mechanism. The genome of isolate KMS was explored for the existence of genes encoding enzymes involved in cAMP synthesis and degradation and CRP-like proteins. This is the first report relating to catabolite repression, caused by simple carbon sources, available in plant root exudates and the soil, in metabolism of complex aromatics in an environmental \( Mycobacterium \) isolate.

**Materials and methods**

**Bacterial strain and cultivation conditions**

Cells of \( Mycobacterium \) sp. strain KMS were stored at -80 °C in 15 % glycerol stock. The cells were resuscitated by streaking on a Luria-Bertani (LB) medium agar plate. Sodium acetate, succinate, fructose, sodium benzoate, and pyrene were from Sigma Aldrich, St. Louis, MO, USA. Seven-day-old KMS cells from the LB plate were scraped and suspended in modified phosphate-based minimal (PBM) medium [38] and the OD \( 600 \) nm of the cell suspension was adjusted to 0.2. Aliquots of 0.5 ml of cell suspension were inoculated into 50 ml PBM medium in 100 ml flasks supplemented with one of the
following substrates as the sole carbon source with carbon (C) availability of about 50 mM: sodium acetate (25 mM), succinate (12.5 mM), fructose (8.3 mM), glucose (8.3 mM), sodium benzoate (7 mM), and pyrene (3 mM). When mixed carbon sources were used, the carbon supplied from each carbon source was 25 mM. The media contained: (1) sodium benzoate (3.5 mM) plus sodium acetate (12.5 mM) or succinate (6.25 mM), or fructose (4.15 mM), sodium or (2) pyrene (1.5 mM) plus acetate (12.5 mM) or succinate (6.25 mM) or fructose (4.15 mM) or sodium benzoate (3.5 mM) or (3) fructose (4.15 mM) plus sodium acetate (12.5 mM) or succinate (6.25 mM).

Sodium salts of acetate, succinate or benzoate and fructose were dissolved in distilled and deionized water (DDW) and the pH was adjusted to 7.0. All the carbon sources were sterilized by filtration through a 0.2 µm pore size Whatman filter (Whatman Inc., NJ, USA). Pyrene was dissolved in methanol and was added to sterile flasks in a sterile hood. The methanol was evaporated before sterile PBM medium was added. The cells in PBM medium were shaken at 220 rpm at 25 ºC. Cell growth was monitored at OD \(_{600\text{ nm}}\), and colony forming units per ml (cfu ml\(^{-1}\)) were assessed by dilution plating on LB agar medium; colonies were counted after 7-10 days growth at 25 ºC.

**Measurement of fructose concentration when KMS grown on acetate or succinate plus fructose as carbon sources**

The utilization of fructose by the KMS cells was followed by sampling at 4 h intervals and centrifugation at 12,000 g to remove cells. The fructose concentration of the supernatant was determined by the colorimetric phenol-sulfuric acid method with authentic fructose as the standard [39].
Primer design for PCR amplification

Nucleotide sequences of the chromosomal and plasmid *nidA* genes, the benzoate dioxygenase α-subunit gene *benA*, and the 16S rRNA genes of *Mycobacterium* isolate KMS were acquired from Joint Genome Institute (JGI) website (http://img.jgi.doe.gov/cgi-bin/w/main.cgi). Specific primers for these genes were designed using Primer3 (http://frodo.wi.mit.edu/primer3/) (Table 4.1). Each primer was used in the BLAST program to ensure only one match in the KMS genome. The authenticity of the primers was evaluated by performing PCR with genomic DNA from isolate KMS to observe the anticipated product size; the correct sequence was confirmed by sequencing the PCR product.

RNA isolation and end-point reverse transcriptase PCR

KMS cells cultured on different carbon sources were harvested at middle-log phase according to cell density measured at OD$_{600\text{nm}}$. Thus, cells were harvested at OD$_{600\text{nm}} =$ 0.30 for cells grown on acetate, succinate, fructose, benzoate, or pyrene as sole carbon source. In cultures with diauxic growth, cells were harvested in the two growth phases: at OD$_{600\text{nm}} =$ 0.25 and 0.40 for cells grown on sodium acetate plus pyrene; at OD$_{600\text{nm}} =$ 0.30 and 0.60 for cells grown on succinate plus pyrene or fructose plus pyrene; at OD$_{600\text{nm}} =$ 0.30 and 0.55 for cells grown on sodium benzoate plus pyrene.

Total RNA was isolated by bead beating cells with TRI-reagent (Molecular Research Center, Inc. OH, USA) following instruction from the manufacturer. The RNA was treated with RNase-free DNase (Promega Corporation, Madison, WI, USA) and purified by extraction with phenol:chloroform:isoamyl alcohol (25:24:1) to remove the DNase. The concentration of the purified RNA was determined by using a NanoDrop ND-1000
s spectrophotometer (NanoDrop Technologies, DE, USA). First-strand cDNA was synthesized by using random primers with 1 µg purified total RNA to a final volume 20 µl following the procedure (Fermentas Inc., MD, USA) of Maxima Universal first strand cDNA synthesis kit (Fermentas Inc., MD, USA).

**Table 4.1.** Specific primers for genes potentially encoding the duplicated pyrene ring-hydroxylating dioxygenase α-subunit genes *nidA*, benzoate dioxygenase α-subunit gene *benA* and the 16S rRNA genes for strain KMS.

<table>
<thead>
<tr>
<th>Gene locus</th>
<th>Protein encoded</th>
<th>Primer sequences (5’ → 3’)</th>
<th>Product length</th>
<th>Annealing temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkms_1668 (chromosomal <em>nidA</em>)</td>
<td>pyrene ring-hydroxylating dioxygenase α-subunit</td>
<td>For ACCCGCGCATTCCAATGCCT</td>
<td>96 bp</td>
<td>55 °C</td>
</tr>
<tr>
<td>Mkms_5626 (plasmid <em>nidA</em>)</td>
<td>pyrene ring-hydroxylating dioxygenase α-subunit</td>
<td>For CACCACGTGACTTCCAATGCC</td>
<td>97 bp</td>
<td>55 °C</td>
</tr>
<tr>
<td>Mkms_1391 (chromosomal <em>benA</em>)</td>
<td>benzoate dioxygenase α-subunit</td>
<td>For ATCGTCGACCTACCTACCTAGC</td>
<td>320 bp</td>
<td>58 °C</td>
</tr>
<tr>
<td>Mkms_R0021 (16S rRNA)</td>
<td>For TGTCGTGAGATGTGTTGGGTAAG</td>
<td>288 bp</td>
<td>58 °C</td>
<td>58 °C</td>
</tr>
<tr>
<td>Mkms_R0038</td>
<td>Rev ATCTGCATTAGCGACTCC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The synthesized first strand cDNA was used in a 20 µl volume end-point PCR reaction containing: 11 µl H₂O, 2 µl MgCl₂ (25 mM), 2 µl 10× Taq DNA polymerase buffer, 1 µl dNTP mix (10 mM each), 1 µl forward primer (final concentration 1 µM), 1 µl reverse primer (final concentration 1 µM), 1 µl cDNA (transcribed from 50 ng total RNA), 1 µl Taq DNA polymerase (1 unit) (Fermentas Inc., MD, USA). The genomic DNA of strain KMS and the DNase-treated RNA were used separately as templates using the same PCR condition as positive and negative control. The PCR was performed in an Eppendorf Mastercycler gradient PCR machine (Certified GeneTool, Inc, CA, USA). The reactions were heated at 94 °C for 2 min followed by 25 cycles of 94 °C 30 sec, 58 °C 40 sec, and 72 °C 40 sec, and an additional 72 °C 10 min was performed before the end of the PCR. This cycle number for the end-point PCR was determined by experimentation with each gene to show that the product intensity was not at saturation. The studies were initially run with 20, 30 and 40 cycles to determine the optimum for distinguishing expression levels. The PCR products were checked by running a 0.8 % agarose gel under 80 V in 1 × TAE buffer. The gel was examined by a UV transilluminator (UVP Inc., CA, USA) and images were recorded photographically (AEase™ version 3.25, A Innotech Corporation, CA, USA).

**Bioinformatic prediction of potential cAMP-binding site and search of genes relating to cyclic AMP metabolism and genes encoding cyclic AMP receptor proteins**

The potential promoter regions and the transcriptional start sites for chromosomal and plasmid nidB/nidA gene clusters and benA gene were predicted by the software Neural Network Promoter Prediction (http://www.fruitfly.org/seq_tools/promoter.html)
The -35 and -10 regions in the promoter were predicted based on the potential transcriptional start site. Potential cAMP-CRP binding sites close to the palindrome sequence TGTGA-N₆-TCACA within the promoter regions were searched through Artemis [41].

Genes potentially encoding adenylate cyclases, phosphodiesterases and CRP proteins [36] in the KMS genome were searched based on key words using the JGI-IMG website. Additionally, BLASTp analysis was used to detect proteins with homology to CRP encoded by Rv3676 in M. tuberculosis [36].

Results

Growth of Mycobacterium sp. isolate KMS on different carbon sources

Isolate KMS utilized sodium acetate, succinate, fructose, benzoate, or pyrene as sole carbon sources (Fig. 4.2). However, no growth was observed on glucose as the sole carbon source. Under the conditions of the assay, different lag phases were observed: for cultures on benzoate and succinate, 36 h, acetate, 48 h, fructose, 84 h, and pyrene, 120 h. The doubling times of KMS grown on different carbon sources calculated from the measurements of cfu ml⁻¹ are shown in Table 4.2. The doubling times were similar, between 6.7 h to 9 h for growth on acetate, succinate, fructose, and benzoate, whereas growth on pyrene was very slow, with a doubling time greater than 60 h (Table 4.2). Growth of KMS on sodium acetate or pyrene supported lower final cell densities than growth on succinate, fructose and benzoate (Table 4.2). These findings suggested that incorporation of carbon into cellular components had different efficiencies for each of the carbon sources.
Figure 4.2. Growth curves of *Mycobacterium* sp. KMS on sodium acetate, succinate, fructose, sodium benzoate, or pyrene as the sole carbon source in PBM medium. Cultures were shaken at 220 rpm at 25 °C and growth determined by OD$_{600\text{ nm}}$. The data are the means with standard errors of OD$_{600\text{ nm}}$ from three independent replicates.

When isolate KMS was cultured on fructose plus acetate or succinate as carbon sources, a single growth phase was observed (Fig. 4.3). The lag phase lasted 60 h for acetate plus fructose (Fig. 4.3a) and 48 h for succinate plus fructose (Fig. 4.3b) compared with about 96 h for fructose alone (Fig. 4.3c). Subculturing the cells from a culture grown on to the fructose liquid medium reduced the lag time only to 72 h; a lag was not eliminated (Fig. 4.4). The timing of fructose utilization by the KMS cells was similar whether or not acetate or succinate was present (Fig. 4.3a, 4.3b, 4.3c). These findings suggested that there was no catabolic repression involved with these substrates.
Table 4.2. Variable growth potential of *Mycobacterium sp.* KMS on acetate, succinate, fructose, benzoate, or pyrene as sole carbon sources in PBM medium. The media were inoculated at an initial concentration of $2.5 \times 10^5 \pm 1.4 \times 10^5$ cfu ml$^{-1}$ and shaken at 220 rpm at 25 ºC. The data are the means and standard errors from three independent replications.

<table>
<thead>
<tr>
<th>Carbon source</th>
<th>Doubling time (hour)</th>
<th>Cell concentration at stationary phase (cfu ml$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 mM acetate</td>
<td>7.9 ± 1.0</td>
<td>$7.2 \times 10^7 \pm 7 \times 10^6$</td>
</tr>
<tr>
<td>12.5 mM succinate</td>
<td>9.0 ± 1.2</td>
<td>$1.7 \times 10^8 \pm 3 \times 10^7$</td>
</tr>
<tr>
<td>8.3 mM fructose</td>
<td>8.9 ± 1.9</td>
<td>$1.4 \times 10^8 \pm 2 \times 10^7$</td>
</tr>
<tr>
<td>7 mM benzoate</td>
<td>6.7 ± 0.8</td>
<td>$3.6 \times 10^8 \pm 6 \times 10^7$</td>
</tr>
<tr>
<td>3 mM pyrene</td>
<td>60.2 ± 9.8</td>
<td>$4.8 \times 10^7 \pm 9 \times 10^6$</td>
</tr>
</tbody>
</table>

When isolate KMS was cultured on benzoate amended with either acetate, succinate, or fructose, only a single growth phase was observed (Fig. 4.5a). The growth rate of KMS on the mixed carbon sources was close to the rates when the microbe was grown on benzoate, acetate, succinate, or fructose as the sole carbon source. However, diauxic growth was observed when isolate KMS was cultured on mixture of pyrene plus acetate, or succinate, or fructose, or benzoate (Fig. 4.5b). In these cultures with mixed carbon sources, the growth rates in the first log phase were close to those grown on the single carbon sources: acetate, succinate, fructose, or benzoate (Fig. 4.5b). The growth rate in the second log phase when KMS was cultured on succinate plus pyrene, fructose plus pyrene, and benzoate plus pyrene (Fig. 4.5b) was very slow and similar to that for growth on pyrene alone: with acetate plus pyrene growth rate was faster (Fig. 4.5b).
Figure 4.3. Growth curve of isolate KMS on fructose plus acetate (a) or succinate (b) and fructose alone (c) as well as the corresponding concentration of fructose in the KMS cultures. KMS was cultured in PBM medium amended with the designated carbon sources with shaking at 220 rpm at 25 °C. Cell growth was measured by \( \text{OD}_{600} \). The data are the means with standard errors from three independent studies.
Figure 4.4. Growth curve of cells of Mycobacterium sp. KMS from a Luria-Bertani (LB) medium agar plate or a minimal medium agar plate with fructose as the sole carbon source. Cultures were grown on PBM liquid medium amended with 8.3 mM fructose as the sole carbon source with shaking at 220 rpm at 25 °C. Cell growth was measured by OD$_{600}$ nm. The data are the means with standard errors from three independent replicates.

Expression from *benA* gene and *nidA* genes under different growth conditions

Expression of the 16S rRNA genes was used as a control to indicate the quality and quantity of the total cDNA. As shown in Fig. 4.6, a product of the anticipated size (Table 4.1) was amplified using specific primers for the 16S rRNA genes from KMS cells grown on sole carbon sources: acetate, succinate, fructose, benzoate, and pyrene or mixtures of benzoate plus either acetate, succinate, fructose, or pyrene. The similar band intensity indicated that the expression level of 16S rRNA genes from log-phase cells was independent of growth substrate. Transcripts of *benA* accumulated when benzoate was the sole carbon source: no transcripts from *benA* were detected from acetate-, succinate-, fructose-, or pyrene-cultured cells (Fig. 4.6a). Transcript accumulation from the *benA*
gene was detected in both growth phases when isolate KMS was cultured on a mixture of benzoate and pyrene as the carbon sources (Fig. 4.6a). However, benA transcript accumulation in the first growth phase was much stronger than the accumulation in the second growth phase (Fig. 4.6a). These findings indicated that the benA gene was specifically induced by benzoate and/or benzoate derivatives. The presence of carbon sources (acetate, succinate, fructose, and pyrene) in addition to benzoate in the culture medium did not reduce benA transcript levels (Fig. 4.6a) suggesting that there was no catabolite repression for expression from this gene.

Because there was diauxic growth when KMS was cultured on pyrene plus acetate, succinate, fructose, or benzoate (Fig. 4.5b), total RNA was isolated from the first and the second log-phases. Expression of the 16S rRNA genes from all cells was similar (Fig. 4.6b). Transcripts from both the chromosome and plasmid copies of nidA were detected when cells were grown on pyrene as the sole carbon source (Fig. 4.6b). No expression was detected when acetate, succinate, fructose, and benzoate were used as sole carbon sources. Similarly no expression was observed in the first phase of growth for the cells from mixed substrate cultures (Fig. 4.6b). Expression from both copies of nidA was detected during the second growth phase in the mixed carbon cultures (Fig. 4.6b).

End-point RT PCR products were obtained using specific primers for the 16S rRNA genes and the α-subunit genes after 25 cycles. Panel A: G PCR product from genomic DNA as template; PCR product from cDNA of RNA extracted from cells grown on: A acetate; S succinate, F fructose, B benzoate, P on pyrene, A+B on acetate plus benzoate, S+B on succinate plus benzoate, F+B fructose plus benzoate, P+B pyrene plus benzoate. Panel B: G PCR product from genomic DNA as template; PCR product from cDNA of
Figure 4.5. (a) Effect of additions of acetate, succinate, and fructose on growth of isolate KMS with benzoate. (b) Effect of additions of acetate, succinate, fructose, and benzoate on growth of isolate KMS with pyrene. The carbon sources were used in PBM medium shaken at 220 rpm at 25 ºC and growth was measured as OD$_{600}$ nm. The data show the means with standard errors of OD$_{600}$ nm from three independent replicates.
RNA extracted from cells grown on: A sodium acetate, S succinate, F fructose, B sodium benzoate, P pyrene, A+P₁ and A+P₂ from the first and second log-phase cells on sodium acetate plus pyrene, S+P₁ and S+P₂ from the first and second log-phase cells on succinate plus pyrene, F+P₁ and F+P₂ from the first and second log-phase cells on fructose plus pyrene, B+P₁ and B+P₂ from the first and second log-phase cells on sodium benzoate plus. The results are typical of those from two independent experiments.

(a)

Figure 4.6. Effect of growth of isolate KMS on mixed carbon sources on expression of (a) the benzoate dioxygenase α-subunit gene benA (Mkms_1391) and (b) the chromosomal (Mkms_1668), and plasmid (Mkms_5626) copies of the nidA gene encoding the pyrene dioxygenase α-subunit.
Potential cAMP-binding site with nidA genes and benA gene

Promoters and transcriptional start sites for the chromosomal and plasmid nidA genes and benA gene were predicted. Promoters of nidA genes were located upstream of an operon in which nidA follows after the first gene, nidB. Both of the nidA genes had the same predicted -35 and -10 regions, TTGACG and TACATT (Fig. 4.7). The benA gene is the first gene in the benzoate-degrading operon. The promoter of the benzoate operon at the -35 and -10 regions differed from those of the nidA genes at the projected -35 and -10 region sequences, TTCACA and TCGAAT (Fig. 4.7). Additionally in the nidA promoters there were potential cAMP-CRP binding sequences, TGTGA-N₆-TTACA in the chromosome and sequence TGA₆AGA-N₆-TTACA for the plasmid. These sequences are one or two nucleotides different from the reported cAMP-CRP binding site TGTGA-N₆-
TCACA in the carbon monoxide dehydrogenase gene cluster of *Mycobacterium* isolate JC1 DSM 3803 [34]. Sequences TGTGA and TCACA were found downstream from the transcriptional start site predicted for the *benA* gene, with a location close to the translational start codon of *benA* gene, but there was a very long intervening sequence (1045 bp) (Fig. 4.7c).

**Potential genes encoding adenylyl cyclases, phosphodiesterases, and cyclic AMP receptor proteins**

Further bioinformatics analysis with the KMS genome confirmed the existence of enzymes associated with cAMP. There are 15 genes encoding putative adenylyl cyclases (Table 4.3) similar to the multiplicity, 16-17 genes, in pathogenic mycobacterial isolates [37]. Twelve of these genes in isolate KMS were chromosomal while three were on the circular plasmid; none was on the linear plasmid. The potential chromosomal adenylyl cyclase genes were far from the chromosomal pyrene-degrading gene island located between Mkms_1647 to Mkms_1718 [17], whereas the potential plasmid adenylyl cyclase genes Mkms_5675 and Mkms_5688 were close to the circular plasmid pyrene-degrading gene island from Mkms_5609 to Mkms_5651 [17]. The genes encoding potential adenylyl cyclases on the circular plasmid and chromosome have limited homology with each other. Four genes potentially encoding cyclic nucleotide-binding proteins were annotated in the KMS chromosome (Table 4.3) and another gene, annotated by JGI as a putative transcriptional regulator at chromosomal locus Mkms_4904, had 89% identity with the proposed CRP protein RV3676 in *M. tuberculosis* [37]. Several genes potentially encoding phosphodiesterases catalyzing cAMP degradation also were present in the KMS genome (Table 4.3). Thus, as shown in
Table 4.3, genes involved in synthesis, function and degradation of cAMP are present in the KMS genome.

Figure 4.7. Sequences of the promoters for the chromosomal (a) and plasmid (b) nidA genes and (c) benA gene of isolate KMS and (d) the reported cAMP-CRP binding motif in Mycobacterium strain JC1 DSM 3803 for the cut gene cluster [34]. Promoters and the transcriptional start sites of genes in KMS were predicted by Neural Network Promoter Prediction. Translational start codons were identified based on the annotation from the JGI-IMG website. Potential cAMP-CRP binding regions are shown in bold with underlining, the -35 and -10 regions are shown by gray boxes, the transcriptional start site are underlined, and translation start codons are shown by boxes.
Table 4.3. Genes in KMS potentially encoding adenylate cyclase for cAMP generation, cAMP binding regulatory proteins and phosphodiesterase for cAMP degradation.

<table>
<thead>
<tr>
<th>Gene product</th>
<th>Locus number of potential genes&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenylate cyclase</td>
<td>Mkms_0810/1253/1254/1255/1566/3019/3084/3904/3942/4023/4168/4881 (ch)</td>
</tr>
<tr>
<td></td>
<td>Mkms_5675/5688/5737 (cp)</td>
</tr>
<tr>
<td>cAMP receptor protein (CRP)</td>
<td>Mkms_1085/2510/2535/4370 (ch), Mkms_4904 (ch)</td>
</tr>
<tr>
<td>Phosphodiesterase</td>
<td>Mkms_1604/1795/2261/4833 (ch)</td>
</tr>
</tbody>
</table>

<sup>a</sup>ch, chromosomal gene; cp, circular plasmid gene.

Discussion

The findings from this paper support that *Mycobacterium sp.* strain KMS has preferences in the use of different carbon sources. When the PAH pyrene was present with other carbon sources, diauxic growth occurred and the sugar fructose, or the organic acids succinate or acetate, were used preferentially over this PAH. Fructose, succinate, and acetate are among compounds secreted by plant roots [30-32] and their catabolism agreed with the maintenance of populations of KMS in a plant rhizosphere [30]. The finding that fructose, succinate and acetate delayed expression of both the chromosomal and plasmid copies of the *nidA* genes, encoding subunits of the dioxygenase required for the first stage of pyrene degradation, suggested that the presence of alternative carbon sources would influence pyrene mineralization. This finding may explain why pyrene mineralization was higher when KMS was cultured in 1/10 diluted barley root exudates.
than when non-diluted exudates was provided [30]. Preferential use of carbon substrates has previously been implicated in repression of PAH catabolism for other microbes [33].

Diauxic growth and delayed *nidA* expression was observed when the single ring aromatic compound benzoate was added into the pyrene culture. These findings for regulation of *nidA* expression suggested that the bioremediation potential of KMS would be affected in soils where degradation of lignin by other microbes would supply benzoate [26]. Whether other simple aromatic compounds would hinder remediation requires further study. Not all plant- or soil-derived phenolics are used as growth substrates by isolate KMS; we have not observed growth of KMS on salicylic acid, coumarin, lignin, or humic acids (Zhang et al., unpublished). We propose that the relative lack of ready carbon sources in the land treatment soils [42] from which KMS was isolated was a key factor in the remediation of pyrene and other PAHs [13]. Under such carbon-starved conditions expression of *nidA* and other genes required for PAH degradation [14] should be expressed.

The addition of fructose, acetate or succinate to benzoate cultures did not produce diauxic growth suggesting simultaneous metabolism of these carbon sources occurred. Indeed expression of the *benA* gene, encoding the α-subunit of the dioxygenase involved in the initial step of benzoate degradation, was observed with log phase cells from the mixed carbon cultures at a level similar to that from cells grown on benzoate alone. Co-metabolism would be valuable for maximizing growth potential in the soil and plant rhizospheres where limited but variable carbon sources would be available. In another soil bacterium *A. baylyi*, the catabolism of benzoate was strongly repressed by acetate or succinate [23]. It is speculated that repression was due feedback inhibition of enzymes
involved in the degradative $\beta$-ketoadipate pathway by accumulation of acetyl-CoA or succinyl-CoA [23].

Our bioinformatics study of promoter sequences suggested that cAMP may be involved in the repression of *nidA* gene expression. Potential cAMP-CRP binding sites TGTGA-$N_6$-TTACAGA-$N_6$-TTACA were found within the promoter regions of both chromosomal and plasmid *nidA* genes suggesting that cAMP serves as a signaling molecule governing gene expression. These sequences differed in only one or two bases from the reported cAMP-CRP binding site TGTGA-$N_6$-TCACA for the operon required for carbon monoxide utilization in *Mycobacterium* strain JC1 DSM 3803 [34]. Their bioinformatic analysis showed that this cAMP-CRP binding site in the promoters for the carbon monoxide dehydrogenase operon was conserved in other soil mycobacterial isolates KMS, MCS, JLS and *M. smegmatis* [34]. However the motif was not found in pathogenic species *M. tuberculosis*, *M. bovis*, and *M. ulcerans* suggesting that cAMP-CRP associated catabolite repression of carbon monoxide utilization was associated with bacterial habitat. Findings with other bacteria show the base pairs in cAMP-CRP binding sequences are varied, for example, in *Sinorhizobium meliloti*, TGTGA-$N_6$-TCCGT; in *Agrobacterium radiobacter* TGTCA-$N_6$-TCACT; in *E. coli* TGTGA-$N_6$-TCACA; and both TGTGA-$N_7$-TTGCT and TGTGA-$N_7$-TCGTA for *Rhizobium leguminosarum* bv. *Trifolii* [43]. Variability of the placement of the cAMP-CRP binding site, within the promoter or between the transcriptional start site and translational start codon, also is found [34, 43]; cAMP-CRP binding motifs are not found after the translational start site. We find with additional bioinformatic analysis that potential cAMP-CRP binding motifs with high identity were present within the predicted promoters of the *nidB-nidA* and
*nidB2* operons of four other PAH-degrading mycobacteria, MCS, JLS, PYR-1, and PYR-GCK [13, 44, 45], all isolated from sources within the USA as well as in strain Spyr1, recently isolated in Greece [46]. Thus this sequence is conserved between isolates of different geographic origin.

In KMS, palindromic sequences TGTGA and TCACA were present in the promoter of the benzoate operon although they were separated by 35 rather than 6-7 nucleotides. A similar finding was observed in the upstream regions of the benzoate operons in the isolates MCS, JLS, PYR-GCK and Spyr1 (data not shown). This abnormal spacing might not allow the binding of a regulatory cAMP-CRP complex and, consequently, we correlated this finding with the lack of carbon repression for benzoate utilization. Additionally a gene with homology to a *luxR* gene is embedded within the benzoate operon suggesting that it may serve a regulatory function for this operon.

The genome of isolate KMS has genes encoding enzymes that will generate and degrade cAMP and there are genes encoding CRP-like proteins. The mechanisms by which cAMP is regulated in the mycobacteria are not resolved. In other bacteria, intracellular cAMP levels have been associated with acetyl-CoA regulation of adenylate cyclase activity [47], the presence of succinate [48], or the utilization of the PTS system to import fructose [49]. The findings that the KMS isolate grew without diauxic growth on fructose plus succinate or acetate suggested to us that the PTS intermediates were not involved in catabolite repression by fructose of the utilization of these organic acids. Of interest is the finding that for isolate KMS failure to grow on glucose correlated with the lack of genes encoding components of PTS transport system homologous to those found in the soil isolate, *M. smegmatis* [12].
The results of this study advance our understanding of the process of PAH-bioremediation by mycobacterium isolate KMS. We demonstrated that the microbe uses other carbon sources in preference to the PAH, pyrene. A potential soil metabolite, benzoate, as well as components in plant root exudates, organic acids and a sugar, repressed growth of KMS on pyrene and inhibited expression from the *nidA* genes, required for PAH degradation. KMS co-utilized other carbon sources with benzoate without repression of the *benA* gene required for catabolism of this single aromatic ring structure. We propose that repression of growth and gene expression for pyrene could be related to mechanisms involving cAMP levels in the cell. In the field it would appear that pyrene remediation by KMS would be enhanced by the absence of other preferred carbon sources whereas growth on plant and soil metabolites could occur concurrently.

References


Abstract *Mycobacterium* sp. strain KMS was isolated from soils where remediation of polycyclic aromatic hydrocarbons was active. This isolate is a competent plant root colonizer through utilization of an array of carbon substrates available in the root exudates. Bioinformatic analyses based on the KMS genome propose pathways for C4- and C3- intermediate conversions during growth of the isolate on substrates requiring gluconeogenesis. Expression of candidate genes for these pathways was compared using semi-quantitative RT-PCR from cells grown on acetate, succinate, benzoate, or pyrene as sole carbon sources requiring gluconeogenesis during growth. Expression was examined for cells grown on fructose and mannitol where gluconeogenesis would not be essential. Transcript accumulation in cells grown on all the carbon sources confirmed expression from genes involved in the glyoxylate shunt and a gene encoding a novel enzyme to complete the tricarboxylic acid cycle, a membrane-associated malate:quinone oxidoreductase. Transcript accumulations for genes encoding phosphoenolpyruvate carboxykinase, malic enzyme, and phosphoenolpyruvate (PEP) synthase were weak for mannitol growth but were detected for the other carbon sources. Activities for PEP
synthase and the membrane–associated malate:quinone oxidoreductase were confirmed in cell extracts at different levels indicating feasibility of their function in production of PEP for gluconeogenesis in this soil mycobacterium.

Introduction

*Mycobacterium* sp. strain KMS was isolated from creosote- and pentachlorophenol-contaminated soils where bioremediation of these toxic compounds was active [25]. Mycobacterial degradation of pyrene generates carbon dioxide (CO$_2$), the C2 compound acetyl-CoA, the C3 intermediate pyruvate, and the C4 tricarboxylic acid cycle (TCA) intermediate succinyl-CoA in the ratio 4:1:2:1 [16, 22]. KMS grows on the simple aromatic benzoate as the sole carbon source [35] with the likely generation of succinyl-CoA as well as acetyl-CoA as the final products based on presence of genes detected in the KMS genome; this pathway is identical to that used by *Rhodococcus jostii* RHA1 [18]. The initial step in benzoate breakdown is catalyzed by a benzoate dioxygenase encoded by genes *benA* and *benB*, which are part of an operon containing additional genes encoding a benzoate dioxygenase reductase, a diol dehydrogenase and a LuxR-like regulatory protein (Fig. 5.1). We find that benzoate induces the expression from *benA* in KMS [35].

The survival of KMS in soils correlates with abilities to degrade a different array of substrates, such as sugars and amino acids [29], present in root exudates [6]. For growth on compounds with carbon skeletons of less than C6 units, gluconeogenesis would be required. The specialized thick walls of mycobacterial cells impose a heavy requirement for both this pathway and fatty acid metabolism because there is a need to build several types of sugar skeletons and very long chain fatty acids [3, 4]. The C2 acetyl-CoA
building block for fatty acid synthesis and transformation of C4 metabolites to the C3 metabolites, pyruvate and phosphoenolpyruvate (PEP), are required for cell proliferation, as illustrated in Fig. 5.2. The TCA intermediates, α-ketoglutarate and oxaloacetic acid (OAA), as well as pyruvate, are used in the generation of amino acids for cell protein production. The formation of succinyl-CoA from succinate recently was demonstrated to play a key role in biomass production for *Mycobacterium tuberculosis* as succinyl-CoA is the precursor of the methyl-branched lipids found in the bacterial cell walls [2].

**Fig. 5.1** Genomic loci and compounds involved in benzoate degradation to acetyl-CoA and succinyl-CoA for isolate KMS. Locus tags are provided as Mkms_. The names of the proteins encoded by the genes and the intermediates in a benzoate degradation pathway are shown.
Fig. 5.2 Proposed pathways for intermediary metabolism in isolate KMS. The products, acetyl-CoA and succinyl-CoA, from degradation of both pyrene and benzoate are noted by stippled boxes, and pyruvate, released during pyrene degradation, is denoted by a grey box. Other intermediates, important because they are the precursors of amino acids, are underlined. The locus tags for isolate KMS for genes encoding enzymes in the glyoxylate shunt (isocitrate lyase and malate synthase) and those involved in completion of the TCA cycle and C4-C3 exchanges are shown. OAA, oxaloacetic acid; PEP, phosphoenolpyruvate; 3-PG, 3-phosphoglycerate; F-6-P, fructose-6-phosphate; G-6-P, glucose-6-phosphate. Not all intermediates are shown.

Proteomic studies with the PAH-degrading isolate *Mycobacterium vanbaalenii* PYR-1 demonstrate a functional, complete TCA cycle and operation of the glyoxylate shunt
when cells are grown on pyrene [17]. Estimations of the abundance of most of the TCA cycle and the glyoxylate shunt enzymes show little difference for PYR-1 cells grown on sorbitol versus sorbitol amended with pyrene [17]. The glyoxylate shunt is active in pathogenic mycobacteria [2, 9, 11]. Thus it seems that the glyoxylate shunt would be active in isolate KMS.

We extended our studies with isolate KMS to probe hitherto unknown functions of the enzymes involved in conversion of the C4 TCA intermediates into the C3 compounds needed for gluconeogenesis. Two pathways illustrated in Fig. 5.2 are feasible based on genome annotation of isolate KMS. One pathway involves decarboxylation of OAA catalyzed by PEP carboxykinase (OAA + GTP $\rightarrow$ PEP + CO$_2$ + GDP) to generate PEP directly [30]. In pathogenic *Mycobacterium* species, the PEP carboxykinase is implicated in cellular growth on fatty acids [23, 24]. A second pathway first involves the decarboxylation of malate catalyzed by the malic enzyme to generate pyruvate (malate + NAD(P)$^+$ $\rightarrow$ pyruvate + CO$_2$ + NAD(P)H) [10] followed by conversion to PEP. Several bacteria possess the enzyme PEP synthase that would generate the PEP directly from pyruvate thus initiating carbon flow into gluconeogenesis [5, 8, 28]. The operation of this enzyme in *Mycobacterium* isolates is not published.

Our transcript analyses involved cells grown on different carbon sources. Pyrene and benzoate were used to compare metabolism of the poly- and mono- aromatic ring structures. Because KMS is a competitive colonizer of plant roots, growth on acetate, succinate and fructose was used to explore gene expression with substrates present in plant root exudates [1]. Additionally, we examined KMS for growth and gene expression on mannitol that is produced by plants and fungi [21, 34] and, thus, could be present in
soils. Mannitol was used to test the hypothesis that cells would require more PEP when grown on fructose than on mannitol because of differences in transport mechanisms. Fructose in *M. smegmatis* is imported by the PEP-dependent phosphotransferase system (PTS) whereas import of mannitol involves an ATP-binding cassette (ABC) [32]. The genome of isolate KMS possesses the genes encoding components of PTS for fructose import [36]. To maintain an intact TCA cycle, transformation from malate to OAA is required yet bioinformatic analysis shows that KMS lacks a gene encoding malate dehydrogenase. In this work we explored expression from an alternative gene encoding a membrane-associated enzyme malate:quinone oxidoreductase (MQO) and we assayed the level of this activity in cell membrane extracts. Additionally we confirmed by activity assay the production of PEP synthase in cells expressing the PEP synthase gene.

**Materials and Methods**

**Bacterial Strain and Cultivation Conditions**

Cells of *Mycobacterium sp.* strain KMS, stored at -80 °C in 15 % glycerol stock, were resuscitated by streaking on a Luria-Bertani (LB) medium agar plate. Sodium acetate (25 mM), sodium succinate (12.5 mM), fructose (8.3 mM), mannitol (8.3 mM), sodium benzoate (7 mM), and pyrene (3 mM) (Sigma Aldrich, St. Louis, MO, USA) were used in a modified minimal medium (PBM) at pH 7 [15]; these concentrations supplied the same potential levels of carbon to the cell. The carbon sources were sterilized by filtration through a 0.2 µm pore size Whatman filter (Whatman Inc., NJ, USA). Pyrene was dissolved in methanol and was added to sterile flasks in the sterile hood. The methanol was evaporated before the PBM medium was added. These media were inoculated with a suspension of KMS cells from 7-day old LB plate (OD$_{600 \text{ nm}} = 0.2$)
using 0.5 ml per 50 ml in 100 ml flasks. The cells in PBM medium were shaken at 220 rpm at 25 °C. Cell growth was monitored at OD\textsubscript{600 nm} and colony forming units per ml (cfu/ml) were assessed by dilution plating on LB agar medium; colonies were counted after 7-10 days growth at 25 °C.

### Identification of Target Genes Through Bioinformatic Analysis

Nucleotide sequences of genes encoding two isocitrate lyases (\textit{aceA} and \textit{icl}), malate synthase (\textit{aceB}), PEP carboxykinase (\textit{pck}), malic enzyme (\textit{mez}), malate:quinone oxidoreductase (\textit{mqo}), PEP synthase (\textit{pps}), and the amino acid sequences of these proteins were extracted from the annotated genome of isolate KMS generated by the Joint Genome Institute (http://img.jgi.doe.gov/cgi-bin/w/main.cgi). Sequences were compared with genes in other microbes using the ALIGN program on the San Diego Super Computer workbench (http://workbench.sdsc.edu/). Specific primers for these genes were designed by using Primer3 (http://frodo.wi.mit.edu/primer3/) (Table 5.1). The specificity of the primers only to the gene of interest was evaluated by performing PCR with genomic DNA from isolate KMS to observe the anticipated product size and the correct sequence was confirmed by sequencing these PCR products.

### RNA Isolation and End-Point Reverse Transcriptase PCR

Growth curves were determined for isolate KMS on the different carbon sources. For transcript analysis, cells were harvested at log phase (OD\textsubscript{600 nm}= 0.30). Total RNA was isolated by bead beating cells with TRI-reagent (Molecular Research Center, Inc. OH, USA). The RNA was treated with RNase-free DNase (Promega Corporation, Madison, WI, USA) and purified by extraction with phenol:chloroform:isoamyl alcohol (25:24:1) to remove the DNase. The concentration of the purified RNA was determined by using a
NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, DE, USA). First-strand cDNA was synthesized by using random primers with 1 µg purified total RNA to a final volume 20 µl following the protocols of the Maxima Universal first strand cDNA synthesis kit (Fermentas Inc., MD, USA).

**Table 5.1** Specific primers used for PCR in the transcript analysis for strain KMS during growth on different carbon sources

<table>
<thead>
<tr>
<th>Gene locus</th>
<th>Protein encoded</th>
<th>Primer sequences</th>
<th>Product length</th>
<th>Cycles in PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkms_4051</td>
<td>Malic enzyme</td>
<td>For GTGGCTGAAACTGTTGTTGGTT</td>
<td>435 bp</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev CGAGATGTCCTCGAGGTGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mkms_0185</td>
<td>PEP carboxykinase</td>
<td>For GAACTCATACCTGGCACTCTCC</td>
<td>415 bp</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev GGTGATGTACTTGGTGTCGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mkms_3469</td>
<td>PEP synthase</td>
<td>For CCAGCAGGACTCTACTACCTCAAC</td>
<td>318 bp</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev TAGGTGGCTTGGTGACATAGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mkms_2861</td>
<td>Malate synthase</td>
<td>For GACATGTCGAACAGAGATCG</td>
<td>248 bp</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev AGGATCGACTGCGTAGGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mkms_0196</td>
<td>Isocitrate lyase</td>
<td>For GATGTGGAGCGAGCAGTACAG</td>
<td>207 bp</td>
<td>25</td>
</tr>
<tr>
<td>(aceA)</td>
<td></td>
<td>Rev TTCTCGGCATAAGTGTTCCTGT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The synthesized first strand cDNA was used in a 20 µl volume end-point PCR reaction containing: 11 µl H₂O, 2 µl MgCl₂ (25 mM), 2 µl 10× Taq DNA polymerase buffer, 1 µl dNTP mix (10 mM each), 1 µl forward primer (final concentration 1 µM), 1 µl reverse primer (final concentration 1 µM), 1 µl cDNA (transcribed from 50 ng total RNA), 1 µl Taq DNA polymerase (1 unit) (Fermentas Inc., MD, USA). The genomic DNA of strain KMS and the DNase-treated RNA were also used as template in the same PCR reaction as positive and negative control. The PCR was performed in an Eppendorf Mastercycler gradient PCR machine (Certified GeneTool, Inc, CA, USA). The reactions were heated at 94 °C for 2 min followed by appropriate cycles (from 24 cycles to 35 cycles depending on the intensity of the product) (Table 5.2) of 94 °C 30 sec, 58 °C 40 sec, and 72 °C 40 sec, and an additional 72 °C 10 min was performed before the end of the PCR. PCR products were visualized with ethidium bromide stained gels after electrophoresis. Product formation using primers specific for the 16S rRNA genes were used as the control for equal loading.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>Length bp</th>
<th>Tm °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkms_0649</td>
<td>For GGACCAAGGAAGGCTTCTACC</td>
<td>Rev GCCAGATCGAACATCGAGTAGT</td>
<td>328</td>
<td>25</td>
</tr>
<tr>
<td>Mkms_2115</td>
<td>For GCTGTGCGAGCTCAACTACAC</td>
<td>Rev AGGTACTCCACGTCTTGCAC</td>
<td>204</td>
<td>30</td>
</tr>
<tr>
<td>Mkms_R0021</td>
<td>For TGTCGTGAGATGTTGGTTAAG</td>
<td>Rev ATCTCGATTACAGCGACTCC</td>
<td>288</td>
<td>24</td>
</tr>
</tbody>
</table>

**Mkms_0649** (icl) Isocitrate lyase

**Mkms_2115** Malate-quinone oxidoreductase

**Mkms_R0021** 16S rRNA

---

**Mkms_R0038**
Preparation of Cell Membranes and Cell Extracts

Cells (500 ml) from middle log phase (OD_{600nm} 0.3) of growth on the different carbon sources were collected by centrifugation at 10,000g for 20 min at 4 °C and were washed twice with 2 ml of 30 mM Tris-HCl containing 10 mM EDTA (pH 7.5). Cells were resuspended in 2 ml of the same Tris/EDTA buffer containing a bacterial protease inhibitor cocktail from the Protease Inhibitor Set #1206893 (Roche Diagnostics Corporation, IN, USA) and were disrupted by treatments in a French Pressure cell (Dayton 12-Ton Hydraulic Jack model 2Z194) (Dayton Electric MFG. Co., Chicago, IL, USA) at 4 °C; extracts were passed three times through the pressure cell to maximize lysis. The lysate was centrifuged at 12,000 g for 20 min at 4 °C to remove unbroken cells and cell debris. The supernatants were retained and cell membranes were recovered from by centrifugation at 100,000 g for 90 min at 4 °C. The pellets were resuspended in 150 µl of 30 mM Tris-HCl (pH 7.5). Both the cell membrane suspension and the supernatant after ultracentrifugation were stored at -80 °C until enzyme activities were determined.

To determine protein concentrations, the preparations were treated with the Compat-Able™ Protein Assay Preparation Reagent (Thermo Fisher Scientific Inc., IL, USA) to remove possible interfering materials. Protein was determined using the Pierce® BCA Protein Assay Kit (Thermo Fisher Scientific Inc., IL, USA) by measuring absorbance at 562 nm. The assay was standardized by using bovine serum albumen (0-2 mg/ml) (Thermo Fisher Scientific Inc., IL, USA).

Assay of PEP Synthase and Malate:Quinone Oxidoreductase

Activity of PEP synthase was determined based on the methods of Hutchins et al. [12] and Tjaden et al. [31]. Reactions were performed in a total volume of 100 µl
containing 30 mM Tris-HCl (pH 7.5), 20 mM β-mercaptoethanol, 0.3 mM pyruvate, 10 mM ATP, 10 mM MgCl₂, and 50 µl cell extract supernatant. Remaining pyruvate was determined through a second reaction. The solution (total volume 1 ml) contained in 30 mM Tris-HCl (pH 7.5) 20 mM MgCl₂, 40 µM NADH, 10 units lactic acid dehydrogenase (Catalog # L-2500, Sigma Aldrich, St. Louis, MO, USA), and 100 µl of the first reaction mixture. Absorption at 340 nm was measured immediately in a Genesys 10 Bio spectrometer (Thermo Electron Scientific Instruments LLC, Madison, WI, USA). The controls used boiled cell preparations and running the first reaction without pyruvate as well as without cell extract supernatant. One unit of PEPS activity was defined the amount that transformed 0.1 nmol pyruvate in 1 min under the assay conditions described.

Activity of malate:quinone oxidoreductase (MQO) was determined based on the methods of Kretzschmar et al. [20]. The reaction was performed in a total volume of 1 ml in 30 mM Tris-HCl (pH 7.5) containing 0.05 mM 2,6-dichlorophenol indophenol (DCPIP) (Catalog # D1878-5G, Sigma Aldrich, St. Louis, MO, USA), 0.1 mM phenazine methosulfate, 5 mM KCN, 10 µM FAD, 1 mM L-malate (Sigma Aldrich, St. Louis, MO, USA), and 50 µl cell membrane suspension. The activity of MQO was determined by measuring the decrease in absorbance at 600 nm of DCPIP. The controls used boiled cell membrane suspension at 95 ºC for 15 min as well as reactions without adding cell preparations. One unit MQO activity was defined as the amount that caused a decreased absorbance of 2,6-dichlorophenol indophenol of 0.01 at 600 nm in 1 min.
Results

Growth of Isolate KMS on Defined Carbon Sources

Isolate KMS grew on acetate, succinate, fructose, mannitol, benzoate, and pyrene as sole carbon sources (Fig. 5.3). The shortest lag phase, about 36 h, was on benzoate and succinate with increasingly longer lag phases for acetate (48 h), fructose (72 h), mannitol (96 h), and pyrene (120 h). Benzoate supported the most rapid growth, followed by acetate, succinate and fructose (Table 5.2). Growth on mannitol was about three-fold slower and growth on pyrene was very slow (Table 5.2). The original cell concentration after inoculation in minimal medium was at a level of $10^5$ cfu/ml. Although the same level of carbon was supplied with each of the carbon sources, the least cell concentration was supported by growth on acetate and pyrene (final concentration $10^7$ cfu/ml) compared with $10^8$ cfu/ml for the other carbon sources. These findings suggested that efficiencies of carbon incorporation by isolate KMS using these carbon sources were different from each other.

Genes Involved in Generation and Transformation of Malate, OAA, Pyruvate, and PEP

A search of the published annotated KMS genome showed two isocitrate lyase genes ($aceA$ and $icl$ at loci Mkms_0196 and Mkms_0649) and a malate synthase gene ($aceB$ at Mkms_2861). No gene for malate dehydrogenase corresponding to the enzyme expressed in PYR-1 [17] was detected. However, a gene encoding a membrane-associated oxidoreductase malate:quinone oxidoreductase (malate + quinone $\rightarrow$ OAA + quinol) was observed ($mqo$ at Mkms_2115). Gene loci for PEP carboxykinase, malic enzyme, and PEP synthase were found (Table 5.1). Identity of the annotated genes was confirmed using the ALIGN program; all the genes had high similarity to genes in $M.\ vanbaalenii$. 
PYR-1, as well as isolates of *Corynebacterium glutamicum* and *Ralstonia eutropha*, where functionality of the genes had been published (Table 5.3). The genes were present on the chromosome of isolate KMS but without clustering. PCR products of anticipated size and sequence were generated using primer pairs designed for each gene using KMS genomic DNA (Fig. 5.4).

**Fig. 5.3** Growth curves (OD<sub>600 nm</sub>) of isolate KMS on different carbon sources with shaking at 220 rpm at 25 ºC. Data are the averages with standard deviations from three independent replicates.
Table 5.2 Variable growth potential of *Mycobacterium sp.* KMS on acetate, succinate, fructose, mannitol, benzoate, or pyrene as sole carbon sources in PBM medium. The media were inoculated at an initial concentration of $2.5 \times 10^5 \pm 1.4 \times 10^5$ cfu/ml and shaken at 220 rpm at 25 °C. The data are the means and standard errors from three independent replications.

<table>
<thead>
<tr>
<th>Carbon source</th>
<th>Doubling time (hour)</th>
<th>Cell concentration at stationary phase (cfu/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 mmol l⁻¹ acetate</td>
<td>7.9 ± 1.0</td>
<td>$7.2 \times 10^7 \pm 7 \times 10^6$</td>
</tr>
<tr>
<td>12.5 mmol l⁻¹ succinate</td>
<td>9.0 ± 1.2</td>
<td>$1.7 \times 10^8 \pm 3 \times 10^7$</td>
</tr>
<tr>
<td>8.3 mmol l⁻¹ fructose</td>
<td>8.9 ± 1.9</td>
<td>$1.4 \times 10^8 \pm 2 \times 10^7$</td>
</tr>
<tr>
<td>8.3 mmol l⁻¹ mannitol</td>
<td>30.5 ± 5.1</td>
<td>$9.5 \times 10^7 \pm 1.5 \times 10^7$</td>
</tr>
<tr>
<td>7 mmol l⁻¹ benzoate</td>
<td>6.7 ± 0.8</td>
<td>$3.6 \times 10^8 \pm 6 \times 10^7$</td>
</tr>
<tr>
<td>3 mmol l⁻¹ pyrene</td>
<td>60.2 ± 9.8</td>
<td>$4.8 \times 10^7 \pm 9 \times 10^6$</td>
</tr>
</tbody>
</table>
### Table 5.3 Comparison of chosen metabolism genes and gene products with the identified genes and their products in other bacteria by program ALIGN on San Diego Super Computer workbench

<table>
<thead>
<tr>
<th>Enzyme name/gene symbol</th>
<th>Gene locus in KMS</th>
<th>Identity nucleotides</th>
<th>Reference gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEP carboxykinase / pck</td>
<td>Mkms_0185</td>
<td>68.8% 65.6%</td>
<td>C. glutamicum pck gene (cg3169) (NCBI accession no. 3345234) [30]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>88.4% 90.6%</td>
<td>M. vanbaalenii PYR-1 Mvan_0201</td>
</tr>
<tr>
<td>PEP synthase / pps</td>
<td>Mkms_3469</td>
<td>64.1% 53.7%</td>
<td>R. eutropha pps gene (NCBI accession no. YP_726506) [5]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>84.0% 89.0%</td>
<td>M. vanbaalenii PYR-1 Mvan_4106</td>
</tr>
<tr>
<td>Malic enzyme / mez</td>
<td>Mkms_4051</td>
<td>62.8% 59.3%</td>
<td>C. glutamicum mez gene (cg3335) (NCBI accession no. 3344434) [10]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82.0% 87.7%</td>
<td>M. vanbaalenii PYR-1 Mvan_4481</td>
</tr>
</tbody>
</table>
Transcript Analyses from Selected Genes

Transcripts were observed for all the genes examined but each had a unique expression pattern depending on the carbon source used to grow the cells (Fig. 5.4). The *icl* gene encoding one of the two isocitrate lyases was most induced in acetate-grown cells as well as cells grown on benzoate or pyrene. Least induction of *icl* occurred with...
Fig. 5.4 End-point RT PCR amplified with specific primers for genes encoding 16S rRNA, isocitrate lyase from locus tags Mkms_0196 (aceA) and Mkms_0649 (icl), malate synthase, malate:quinone oxidoreductase, PEP carboxykinase, malic enzyme, and PEP synthase. PCR was performed as described in Materials and Methods and the results were identical for three independent experiments. Product from genomic DNA was generated as a control for each gene.

The second isocitrate lyase gene, aceA, was expressed almost equally for cells grown on all carbon sources except mannitol. The gene for malate synthase was expressed to similar levels independent of carbon source. Likewise, transcripts from the gene, mqo, encoding malate:quinone oxidoreductase, were similar in
level when KMS was grown on all the six carbon sources, suggesting that the reductase served as a house-keeping gene to complete the function of the TCA cycle.

Isolate KMS has the potential to use two pathways to generate PEP during gluconeogenesis. Expression of the gene for PEP carboxykinase was induced most strongly from KMS cells grown on benzoate, induced to a lesser extent with growth on succinate, pyrene, and fructose, and induced least when mannitol was used as the sole carbon source (Fig. 5.4). For the gene encoding PEP synthase expression was highest for KMS cells grown on fructose, intermediate for acetate, succinate, benzoate, lower for pyrene and least for mannitol (Fig. 5.4).

Activities of Malate:Quinone Oxidoreductase (MQO) and PEP Synthase (PEPS)

Enzyme activities for the key enzymes MQO and PEPS were detected in extracts from cells grown on each carbon source. The specific activities are shown in Fig. 5.5. MQO activity was detected in the membrane preparations from KMS cells grown on each carbon source indicating that MQO was expressed constitutively (Fig. 5.5a). The finding correlated with the detection of transcripts from the \( \text{mqo} \) gene. ANOVA analysis showed that specific activity of MQO was at different levels based on the carbon source used by KMS. Growth on aromatics, benzoate and pyrene, resulted in the highest level of specific activity of MQO (Fig. 5.5a) indicating MQO activity was crucial for the utilization of aromatic compounds as carbon sources. Specific activity of MQO was at the same level when KMS was grown on succinate and fructose whereas it was lower for growth on acetate and mannitol (Fig. 5.5a). This finding may be due to reduced flux of carbon through the TCS cycle because some acetyl CoA would be used directly for fatty acid
Fig. 5.5 Specific activity of malate:quinone oxidoreductase (MQO) (a) and PEP synthase (PEPS) (b) when KMS was grown on different carbon sources. Cells were harvested at OD$_{600\,nm}$ 0.3 in the middle log phase and disrupted by French pressure. One unit of MQO was defined as 0.01 decrease of absorbance of 2,6-dichlorophenol indophenols at 600 nm in 1 min. Specific activity of MQO was shown as activity per mg total protein in cell membrane. One unit of PEPS was defined as transform of 0.1 nmol pyruvate in 1 min. Specific activity of PEPS was shown as activity per mg total protein in cell extracts. Levels of specific activity of MQO and PEPS from cells grown on different carbon sources were analyzed by ANOVA ($p=0.05$) from http://www.xlstat.com/en/. The results were typical of two independent studies.
synthesis. The slow rate of growth on mannitol may explain why a low MQO activity would still allow for growth.

PEP synthase from expression from the gene at locus Mkms_3469 was detected in all extracts at similar specific activities except for growth on mannitol (Fig. 5.5b). Extracts from cells grown on mannitol had the least PEPS activity possibly due to slow growth rate and reduced need for gluconeogenesis. High activity of PEPS was detected from KMS cells grown on organic acids and aromatic carbon sources, where gluconeogenesis was needed. High PEPS activity also was detected from KMS cells grown on fructose. We speculate that this high level of activity was due to the requirement for PEP in fructose uptake by the PTS although PEP also should be supplied during the C6 to C3 breakdown.
Discussion

Our findings suggested that *Mycobacterium* sp. isolate KMS used a combination of the TCA cycle and the glyoxylate shunt for C2- and C4-interconversions and PEP carboxykinase, malic enzyme, and/or PEP synthase to initiate the flow of carbon as C3 structures into gluconeogenesis. The pathways are illustrated in Fig. 5.2. The transcript abundance from genes encoding enzymes differed between carbon sources used for growth. Enzyme assays in KMS cell extracts confirmed production of two novel enzymes for mycobacterial metabolism, PEP synthase and a membrane-associated malate: quinone oxidoreductase. These findings substantiated the ability of KMS to utilize a variety of carbon sources for survival dependent on its habitat, such as plant root surfaces and the PAH-contaminated soils [7, 25].

Our transcript analysis in isolate KMS for genes in the glyoxylate shunt confirmed the findings which were observed at the proteomic level with mycobacterial isolate PYR-1 for isocitrate lyase and malate synthase [17]. However, KMS showed transcripts from both genes for isocitrate lyase, whereas with PYR-1, only *aceA* expression was observed by the proteomic approach when cells were grown on sorbitol or sorbitol plus pyrene [17]. This gene shared with no homology with the *icl* gene. Notable in our findings was the higher expression of the *icl* gene than the *aceA* gene for growth of KMS on acetate. This may occur because of the essential role of the glyoxylate pathway as an anaplerotic pathway as first proposed by Kornberg [19] but recently substantiated for *M. tuberculosis* [2]. In pathogenic *Mycobacterium* isolates, both AceA and Icl proteins were induced by acetate, glucose, and the fatty acid palmitate but were repressed by succinate even in the presence of palmitate [11]. Thus these findings with the pathogens were not consistent...
with that in isolate KMS where similar expression was measured for both the aceA and icl during growth on succinate. For KMS, expression of the malate synthase gene was constitutive on all growth substrates indicating the potential for the functionality of the glyoxylate shunt during metabolism of an array of substrates.

Our transcript and enzymatic analysis confirmed the use in isolate KMS of a membrane-bound malate:quinone oxidoreductase to complete the TCA cycle [14]. Although transcript accumulation of the gene was detected from KMS cells grown on all carbon sources (Fig. 5.4), enzyme activities in cell extracts varied between carbon sources (Fig. 5.5a). We speculate that the activity of MQO increased when the cells were grown on single ring and complex aromatic compounds due to the importance of the TCA cycle activity in maximizing the use of the C2 and C4 compounds generated from their breakdown. Variation of activity of MQO has been demonstrated with growth substrate for other bacteria [27, 33] although no aromatic compounds were used in these studies.

Transcript analysis suggested that both the direct conversion of OAA to PEP by PEP carboxykinase as well as the shuttling of malate to PEP, requiring the malic enzyme and then PEP synthase, could be functional during gluconeogenesis in the KMS cells (Fig. 5.2). PEP carboxykinase is unidirectional [30]. Our detection of transcripts for the malic enzyme differed from the results of the proteomic study with M. vanbaalenii PYR-1 grown on pyrene and sorbitol where the protein was not reported [17]. However, the detection of malic enzyme activity in mycobacterium has been problematical as discussed by Beste et al. [2]. A positive role for the malic enzyme in gluconeogenesis has been proposed for growth of pathogenic M. bovis BCG and M. tuberculosis on acetate or fatty
acids [23, 24]. Transcript accumulation of the gene encoding PEP carboxykinase, Mkms_0185, was detected from KMS cells grown on organic acids and aromatic compounds when gluconeogenesis was needed. Transcript accumulation from the gene encoding PEP carboxykinase reinforced that KMS needs to keep a high level of PEP pool for fructose import when fructose was used as carbon source but not mannitol.

To the best of our knowledge our findings for both transcript and activity for PEP synthase are the first report to evaluate their functions in mycobacterium. Interestingly, bioinformatic analyses showed the presence of the gene for PEP synthase to vary: it is present in the PAH-degrading isolates (*M. vanbaalenii* PYR-1, *M. gilvum* isolates Spyr1 and PYR-GCK, and *Mycobacterium* sp. MCS and JLS from the same PAH-contaminated site) and in the nonpathogenic soil isolate *M. smegmatis* but no genes encoding a PEP synthase are annotated in genomes of pathogenic mycobacterium. Instead, the genomes for the pathogenic bacteria (isolates of *M. tuberculosis*, *M. avium*, and *M. bovis*) have a gene encoding an alternative enzyme, pyruvate phosphate dikinase. Perhaps unexpectedly the functioning of this enzyme was not discussed in the two recent metabolomics papers for carbon flow in pathogenic mycobacteria [2, 9].

The detection of transcripts encoding PEP synthase and activity of the enzyme suggested that it has the potential to function in gluconeogenesis in isolate KMS, although the enzyme could be bidirectional [13]. Unexpectedly although transcript accumulation for this gene was highest with growth on fructose, specific activity was similar for growth on acetate and benzoate, suggesting that factors more than transcriptional control occur. PEP is not only involved in gluconeogenic flow for isolate KMS, but also is essential for the import of fructose. Genes in KMS at the locus tags,
Mkms_0088 to Mkms_0092, have high identity to those from Msmeg_0084 to Msmeg_0088 encoding the components of the fructose-PTS transport system in *M. smegmatis* [32]. The dependence of fructose transport on PEP may require the transformation of pools of C4 malate and OAA to PEP even with fructose as the carbon source. In contrast, the low level of transcripts and activity for PEP synthase of mannitol-grown cells may reflect a reduced need for gluconeogenesis due to slow growth rate and the use of an import system dependent on ATP for mannitol. Genes at locus tags Mkms_4440 to Mkms_4443 in KMS have identity to genes at locus tags Msmeg_5571 to Msmeg_5574 encoding an ATP-binding cassette (ABC) involved in alditol transport in *M. smegmatis* [32] suggesting that growth on mannitol does not involve PEP-dependent import.

In conclusion, we provided data at the transcript and activity levels to illustrate potential pathways of intermediary metabolism utilized by isolate KMS when growing on different carbon sources relevant to its survival on plant roots and in PAH-contaminated soils. Our findings illustrate that even microbes of the same genus and with similar functions demonstrate plasticity in the enzymes employed in basic metabolic flow. One novelty is the reliance of this strain on MQO in isolate KMS to complete the TCA cycle rather than a malate dehydrogenase as exemplified in isolate PYR-1. Transcript abundance studies revealed that expression of the genes for isocitrate lyase in KMS was more regulated than the gene for malate synthesis so that interchange of C2 and C4 intermediates from the TCA cycle could be affected. Our data support roles for PEP formation by PEP carboxykinase and the combination of malic enzyme and PEP synthase for initiation of gluconeogenesis. The operation of these metabolic pathways may help
the bacterium to competitively colonize plant root surfaces as well as to grow under conditions of relative starvation when recalcitrant pollutants, such as the PAH pyrene, are the only available carbon sources in soil.

References
*Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. Chemistry and Biology 17:1122-1131


hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. Biodegradation 19:859-881


and cytoplasmic malate dehydrogenases in the citric acid cycle of *Escherichia coli*. J Bacteriol 182:6892-6899


CHAPTER 6
SUMMARY AND FUTURE DIRECTIONS

SUMMARY

My findings extend the knowledge relating to aromatic hydrocarbon utilization in Mycobacterium sp. strain KMS in four novel ways: 1) Some of the expressed clustered genes for pyrene-degradation are replicated on plasmids within the isolate; the sequences for these regions in KMS are conserved between PAH-degrading mycobacteria of different geographical origin with evidence of their origin from horizontal transfer; 2) Expression from distinct operons containing divergent dioxygenase genes occurs for the utilization of benzoate and pyrene in isolate KMS; 3) There is preference in the use of carbon sources by the isolate, possibly involving cAMP-dependent mechanisms, the PAH pyrene is used only when other carbon sources are depleted; 4) Intermediary carbon shuffling involves a constitutive glyoxylate shunt, a novel malate:OAA quinone oxidoreductase, and two potential pathways to transform C4 structures from the TCA cycle to C3 intermediates for gluconeogenesis.

1) Pyrene-degrading gene islands in Mycobacterium isolates: gene duplication.

Through BLAST analysis, gene loci potentially encoding seventeen pyrene-induced proteins identified for isolate KMS (21) were found clustered on the chromosome in a 72 kb region including 71 annotated ORFs. This work confirms findings from the Cerniglia group (16) that the pyrene-degrading genes are clustered in PAH-degrading mycobacteria. However, isolate KMS differs from isolate PYR-1 because it contains plasmids that also possess duplicated copies of dioxygenase subunit genes, including the \( \text{nidA} \), \( \text{nidB} \), and \( \text{nidB2} \) genes encoding subunits for the dioxygenases responsible for
initial oxidation of the HMW-PAH pyrene (21). Block duplication involving intact operons is evident. Both the chromosomal and plasmid genes in the pyrene-degrading gene islands were expressed.

Block duplication may enhance the PAH-degrading ability of the microbe by producing a higher level of protein expression. Also minor changes to the sequence of the open reading frames and the promoters for the operons containing these genes may provide variations to the conditions under which the genes are expressed and the specificity of the products that can be attacked. Additionally, the duplicated gene copies increase the likelihood that the isolate will retain PAH-degrading ability even if one of the gene copies becomes mutated. Most PAHs are highly mutagenic through the formation of DNA adducts and the formation of reactive oxygen species (ROS) as the PAH compounds are transformed (9, 22). I suspect that when the microbes are in a PAH-contaminated soil that the potential for mutation is enhanced because of the heightened damage by ROS to DNA as the PAHs are metabolized.

Because of the similarity between the sequences of the genes for pyrene-degradation in KMS to that published for PYR-1 (16), I extended the genome analysis to the other four sequenced mycobacteria: isolates MCS, JLS, PYR-GCK, and Spyr1. Synteny analysis showed that the pyrene-degrading gene islands were essentially conserved between KMS and two other PAH-degrading mycobacterial isolates from the same site, isolates MCS and JLS. This was especially true for the linear plasmids of MCS and JLS. Synteny also was observed for isolate *M. gilvum* PYR-GCK from a third geographic location in the USA. The pyrene-degrading gene island also was found in the chromosome of isolate Spyr1 from Greece. These findings indicate strong conservation
of the pyrene degrading gene islands between isolates of diverse geographic origin. The similarity in the pyrene-degrading gene islands may be explained by an initial horizontal gene transfer event of genes into a common mycobacterium ancestor followed by limited gene duplication and divergence. Evidence for horizontal transfer include the finding that the GC content of the genes in the pyrene-degrading islands (61-63% for those on the plasmid and 63-64% for those on the chromosome) is lower than the average GC content (68%) of other genes. Also there are genes encoding transposases and integrases within or close to the pyrene-degrading gene islands as remarked by Stingley et al. (29) for isolate PYR-1. These activities may be involved in duplication and rearrangement of the genes within the pyrene-degrading gene islands (32, 35).

Mechanisms for horizontal transfer include transfer of plasmids through conjugation, uptake of naked DNA, or transduction by phage. I speculate that conjugation is the most likely method used for horizontal gene transfer with the mycobacterium; plasmid transfer through conjugation is reported between the soil isolate *M. smegmatis* and *E. coli* (11). Evidence for plasmid sharing between the isolates from the PAH-Libby site is provided by the high identity between the genes in the linear plasmid in KMS to that in MCS. Uptake of naked DNA is suggested by a broad range of bacteria, but *Mycobacterium* is not referenced as a recipient (31). The thick cell wall possessed by KMS may make it difficult to acquire foreign naked DNA without serious trauma to the cell. Transduction through phages is another possible method for horizontal gene transfer in KMS. Phages with the potential to attack *Mycobacterium* isolates (12) are present in soils and composts (13), habitats similar to that from which KMS was isolated. Evidence of phage attack on a bacterial cell is correlated in some bacteria with the existence of clustered regularly
interspaced short palindromic repeats (CRISPR) in the genomes; these regions are thought to relate to resistance to bacteriophages (2). To date CRISPR regions have not been detected in most sequenced *Mycobacterium* isolates including isolate KMS (12).

2) Expression from distinct operons containing divergent dioxygenase genes occurs for the utilization of benzoate and pyrene in isolate KMS. A total of 44 paralogous genes encoding ring-hydroxylating dioxygenase subunits are found in the genome of isolate KMS including ten lone α-subunits genes, two lone β-subunits genes, and sixteen paired α- and β-subunits genes. In contrast, annotation from MicrobesOnline (http://www.microbesonline.org/) revealed that the pathogenic isolate, *Mycobacterium tuberculosis* H37Ra, only possesses two genes both encoding ring-hydroxylating dioxygenase α-subunits characterized by the COG number 4638 corresponding to the presence of a Rieske center in a phenylpropionate dioxygenase. This designation is the same as for the lone α-subunit genes found in KMS. No genes encoding β-subunits to generate the complexes involved in ring opening of the HMW PAHs are annotated as being present in the pathogenic isolate. I speculate that the difference between the pathogen and the PAH-degrading environmental isolate is an indication that the soil isolate KMS has acquired and maintained this extensive array of dioxygenase genes to degrade a variety of aromatic compounds to maximize its flexibility in the environment.

Bioinformatic analysis of the predicted amino acid homology of the dioxygenase genes in KMS suggests that this isolate has a potential to degrade a broad range of aromatic hydrocarbons and derivatives in addition to pyrene and benzoate. Cerniglia’s group (19, 20) classified the ring-hydroxylating dioxygenases into five types according to the protein structure and the types of electron transfer chains with which the terminal
oxidase functions. Presumably changes in the open reading frames would permit enzymatic attack on different aromatic compounds because the spatial sizes of the substrate-binding pocket in the active center of dioxygenases were determined by the amino acid sequence (20). As discussed above, the lone genes encoding α-subunits contain Rieske centers characteristic of phenylpropionate dioxygenases that belong to either Type Iα or Type IIIα (Fig. 6.1). BLASTp analysis of the paired genes suggests that genes adiV and bdiV encode phthalate dioxygenase and are orthologs of phtAa and phtAb in PYR-1 with a Type V classification (20) (Fig. 6.1). Predicted products AdiW and BdiW have identity to Type V naphthalene dioxygenase whereas AdiX and BdiX are identical to Type Iα undefined aromatic dioxygenases (Fig. 6.1). However, based on standard proteins used in Cerniglia’s classification, there are no type homologs for AdiY and AdiZ. BLASTp analysis shows that AdiY and BdiY have identity to dibenzofuran and biphenyl dioxygenases from other microbes and AdiZ and BdiZ have identity to biphenyl and benzene dioxygenases (Fig. 6.1). Homologs of genes encoding A/BdiX/Y/Z are also found in M. vanbaalenii PYR-1 but the function of the gene products have not been identified yet.

Phylogenetic analysis for the ring-hydroxylating dioxygenase α-subunits at the amino acid levels shows obvious separation between the lone α-subunit genes and the paired dioxygenase α-subunit genes suggesting the two types of α-subunit genes diverged in an earlier time. Alignment analysis confirmed the duplication of several α- and β-subunit genes in the genome of isolate KMS. From the phylogenetic analysis, I speculate that the common ancestors of dioxygenase α- and β-subunit genes were acquired from a foreign source and diverged into paralogs early in the ancestral strain.
FIG 6.1 Phylogenetic trees generated by MEGA 4 for the predicted protein sequences for the α-subunits of dioxygenases for *Mycobacterium* isolate KMS. The tree was obtained by neighbor-joining and represents 1000 replications. Overlayered on this tree are balloons indicating the classification of the dioxygenases into the group (Type I to Type V) based on the work of Kweon et al. (19). Designation of proteins NidA and NidA3 followed those identified in *M. vanbaalenii* PYR-1 (16). Designation of protein BenA was because of homology to proteins in *Rhodococcus jostii* RHA1 (18). Other as yet uncharacterized aromatic ring-hydroxylating dioxygenase α-subunits are designated as
Adi. The term ch indicates a chromosomal gene, cp a circular plasmid gene, and lp a linear plasmid genes. The locus numbers of the genes encoding the proteins are shown. The scale bar means 0.2 changes per amino acid position.

Alignment of 60 bp segments for the potential promoters for operons containing the dioxygenase subunit genes showed the promoters for duplicated segments were highly homologous (Fig. 6.2). Outside of the promoters for duplicated genes the base pairs sequences lacked similarity. Such variation in sequence between promoters would permit regulation by different mechanisms enabling expression under different environmental pressures.

Transcript analysis by PCR showed similar expression of two *nidB* and two *nidB2* genes encoding pyrene dioxygenase β-subunits that are regulated by promoters that for 60 bp upstream of the predicted transcriptional start site have sequences with high identity (Fig. 6.2). My work with expression from two operons containing dioxygenase genes shows that the promoters for the operons have the potential to be differentially regulated. Evidence includes the induction of the pyrene dioxygenase α-subunit gene *nidA* by pyrene but not benzoate and the induction of the benzoate dioxygenase α-subunit gene *benA* by benzoate but not pyrene. These findings indicate that not all aromatic ring compounds will induce the same operons.
FIG 6.2 Alignments of the predicted promoter regions of all the ring-hydroxylating dioxygenase genes for isolate KMS. The alignment by ClustalX2 was of 60 bp total showing 50 bp upstream from the predicted transcriptional start site. The symbol ch means chromosomal gene, cp circular plasmid gene, and lp linear plasmid gene. The promoters with homology are shown in groups.

An operon containing genes potentially involved in benzoate degradation (Fig. 6.3), encoding the ring-hydroxylating dioxygenase subunits, a dioxygenase reductase, a diol dehydrogenase, and a LuxR family-regulatory protein, differed from clusters found in other benzoate-degrading microbes: no gene encoding a benzoate transporter was present in the operon in KMS. However, potential benzoate transporter genes were found distant from the benzoate-degrading gene operon. The gene at locus Mkms_2800 encoding a drug resistant transporter was homologous to the benzoate transport protein, BenK, identified in \textit{Rhodococcus jostii} RHA1 (18). Also the gene at locus Mkms_4242 had
homology to the dctA gene in *Pseudomonas chlororaphis* O6, which encodes a transporter active with benzoate as well as with dicarboxylic acids (24). Consequently, I suggest that genes encoding a benzoate transporter may be located elsewhere on the KMS genome.

A single gene at locus Mkms_1389 is predicted to encode the benzoate dioxygenase reductase at the N-terminus and the diol dehydrogenase at the C-terminus. In other benzoate-degrading bacteria, benzoate dioxygenase reductase and diol dehydrogenase are encoded by two separate genes (8, 18). Whether products of gene Mkms_1389 are functional awaits additional study. However, the rapid growth of KMS on benzoate correlated with demonstrated accumulation of transcripts from the *benA* gene encoding the α-subunit of the benzoate dioxygenase. These findings suggest that a functional pathway for benzoate utilization is present in KMS. I have detected in the KMS genome other genes encoding additional enzymes required for full degradation of benzoate to acetyl-CoA and succinyl-CoA; these genes are found adjacent to the benzoate operon as illustrated in Fig. 6.3. Because benzoate should be available in soils from lignin degradation (30) the presence of this pathway may aid in the survival of KMS in soils.

The mechanism underlying differential regulation of the promoters awaits determination. Bioinformatics suggest that other regulatory proteins could function to control the expression of the additional operons in the pyrene-degrading gene islands. The chromosomal single copy operon containing *adiV/bdiV* is possibly regulated by an IclR family regulator encoded by gene Mkms_1646 (Fig. 6.4). MarR family regulators encoded by chromosomal genes Mkms_1689 and Mkms_1702 as well as genes at the loci Mkms_5638 and Mkms_5651 on the circular plasmid and Mkms_5857 and Mkms_5843
on linear plasmid may function in control of the operons containing \( nidA3/nidb3 \) and \( adiYlbdY \). The block duplications for these regions duplicate the genes for substrate modification and those for potential regulators. The presence of different types of potential regulators would enhance the plasticity of isolate KMS to respond to complex PAH-contaminated soils.

**FIG 6.3** Benzoate degradation in isolate KMS, the operon and the deduced benzoate degradation pathway. Locus tags are provided as Mkms_. The names of the proteins encoded by the genes and the intermediates in benzoate degradation are shown.
FIG 6.4 The position of genes encoding potential regulators IclR and MarR in the pyrene-degrading gene islands in isolate KMS. Ring-hydroxylating dioxygenase subunit genes are shown by designation *nid* or *adil/bdi*. Genes encoding regulator proteins are labeled by locus number. Duplicated operons are linked by red lines.

3) Absence of other available carbon sources promotes pyrene degradation by **isolate KMS**. Child et al. (7) demonstrated that in addition to surviving in soils (15), isolate KMS was a competent colonizer of barley roots. This concept of *Mycobacterium* as a root colonizer was supported by the detection of PAH remediation when isolates of *M. gilvum* colonized aquatic reeds (34). In both cases growth of the microbes on root exudates permitted augmentation of cell numbers and PAH mineralization. Acetate, succinate, and fructose are common components of root exudates (3, 7) whereas benzoate is present in the rhizosphere from lignin degradation (30).

A survey of different carbon sources, including the root exudate components of acetate, succinate, fructose and benzoate, as well as pyrene, revealed that each is utilized by KMS as a sole carbon source. Isolate KMS had preferences in utilizing these carbon sources: benzoate is used most readily, followed by the organic acids (acetate and succinate) and, after an extensive lag period, the sugar (fructose). Growth on pyrene involved a long lag phase and protracted generation time.
Diauxic growth of KMS occurred in pyrene cultures amended with the other carbon sources. However, no diauxic growth of KMS occurs when benzoate is amended with acetate, succinate or fructose. No diauxic growth of KMS was found when fructose is amended with organic acids acetate and succinate either. Diauxic growth on pyrene cultures mixed with other carbon sources is correlated with delayed expression of the nid genes encoding pyrene dioxygenase α-subunits. This finding indicates that acetate, succinate, fructose, and benzoate are preferentially utilized by KMS over pyrene. However, expression from the benA gene encoding benzoate dioxygenase α-subunit was not delayed by the presence of acetate, succinate, and fructose. As discussed above and as shown in Fig. 6.2, the promoters for the benzoate and pyrene-degrading operons are quite different.

Bioinformatic analysis of the promoter sequences of the nidB-nidA gene operons suggested that cAMP may be involved in the repression of expression of the nidA gene but is without effect on the promoter of the benzoate operon. Potential cAMP-CRP binding sites TG(T/A)GA-N₆-T(T/A)(A/G)CA are found in the promoter regions of the operons containing nidB-nidA genes and two operons containing nidB2 genes (Table 6.1). Genes encoding the enzymes required for synthesis and degradation of cAMP and protein binding are present in isolate KMS. KMS possesses fifteen potential genes encoding adenylate cyclases for cAMP synthesis, four potential genes encoding phosphodiesterase for cAMP degradation, and five potential genes encoding CRP for binding with cAMP.

My suggestion that expression of certain pyrene-degrading genes is under cAMP control is the second example of the potential use of cAMP as a regulator in environmental mycobacteria. Previously a cAMP-CRP binding motif, TGTGA-N₆-
TCACA, was found upstream of the translation start site for the carbon monoxide dehydrogenase gene operon in mycobacterial isolate JC1 DSM 3803 (26). Thus, I speculate that cAMP may be involved in the regulation of expression of genes important for niche habitation by isolate KMS. Studies with pathogenic mycobacterium show that these cells secrete cAMP, a process that thought to trigger remodeling in the metabolism of the challenged host cells (1). Because cAMP is a general regulator in bacteria and fungi (5, 28), if secretion occurs by the environmental mycobacteria, they may use this as a mechanism to alter the metabolism of other microbes within their environmental habitats.

**TABLE 6.1** Potential cAMP-CRP binding sites found in the promoter regions of two operons containing *nidB-nidA* genes and two operons containing *nidB2* genes

<table>
<thead>
<tr>
<th>Operons</th>
<th>Potential cAMP-CRP binding motif (bold and underlined)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chromosomal</strong></td>
<td></td>
</tr>
<tr>
<td><em>nidB-nidA</em> operon</td>
<td>GTTGTTTTGGACGGCCGCTGTGA CGTGACTTTACATTCTCTA</td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
</tr>
<tr>
<td><em>nidB-nidA</em> operon</td>
<td>TTTGTCTTGGACGGCCGCTGTGA CGTGACTTTACATTCTCTA</td>
</tr>
<tr>
<td><strong>Chromosomal</strong></td>
<td></td>
</tr>
<tr>
<td><em>nidB2</em> operon</td>
<td>GTTGTTTTGGACGGCCGCTGTGA CGTATCTAGCATTTCTCT</td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
</tr>
<tr>
<td><em>nidB2</em> operon</td>
<td>TTTGTCTTGGACCGCCGCTGTGA GTGAATTACATTCTCTTA</td>
</tr>
</tbody>
</table>

Potential cAMP-CRP binding sites were not found in promoters of the other operons containing ring-hydroxylating dioxygenase subunits in the pyrene-degrading islands.
suggesting that these operons have other modes of regulation. Consequently it is possible that only the genes in the *nidB* and *nidB2* operons are regulated by cAMP-CRP. However, I cannot rule out that my bioinformatics analysis was not sufficiently robust.

Surprisingly, during the course of this study, I found that isolate KMS could not grow on glucose. Failure to utilize glucose is possibly due to the lack of an efficient glucose transport system; no genes corresponding to the PEP-linked uptake system operative in *M. smegmatis* (33) are found in KMS. The potential for a PEP-dependent uptake system for fructose is present in isolate KMS and I discuss the significance of this mechanism of transport on intermediary metabolism in the next section.

These findings on nutrient preference suggest that presence of other available nutrients in a contaminated soil possibly decreases the potential remediation of pyrene by isolate KMS but may have no influence on remediation of benzoate. Indeed utilization of benzoate can occur concurrently with the use of other carbon sources, an advantage for the bacterium when living in plant rhizospheres or in soil. Further studies are needed to determine whether or not there is diauxic growth on other PAHs and MAHs catabolized by KMS. Removal of other soluble carbon sources by water leaching may help in pyrene remediation. My findings fit with the isolation of KMS from an open land treatment unit without plants at the Libby Montana site (15); these soils would be periodically leached by rain and snowfall and very few carbon sources other than the pollutants would be available. Although carbon limitation in soil helps utilization of PAHs, microbe-assisted phytoremediation could still be a constructive approach to augment remediation. Growth of microbes on the plant roots would enhance the population size of the PAH-degrading mycobacteria to enhance the overall PAH remediation.
4) Intermediary carbon shuffling in isolate KMS is efficient. I hypothesized that isolate KMS must possess an efficient carbon retention system for growth especially for under nutrient limiting conditions, such as PAH-contaminated soils or rhizospheres. Studies with KMS suggested that the metabolites from pyrene catabolism were incorporated into cellular materials (25). Consequently growth on pyrene and benzoate would require intermediary metabolism to derive the C3 compounds essential for gluconeogenesis. *Mycobacterium* isolates have a high requirement for fructose-6-phosphate and fatty acids to produce the structures in their unique, thick, carbon-rich cell walls (6).

Utilization in isolate PYR-1 of the products from pyrene degradation, acetyl-CoA and succinyl-CoA, by the TCA cycle and the glyoxylate shunt is suggested from their proteomics studies (17). Their proteomic analysis suggests a malate dehydrogenase converts malate to OAA in the TCA cycle and expression of one of the two isocitrate lyases, AceA, and malate synthase was observed suggesting a functional glyoxylate shunt (17).

The findings with isolate KMS differ from that in PYR-1. I find expression of both genes for isocitrate lyase for isolate KMS, whereas with isolate PYR-1 no accumulation of the isocitrate lyase protein Icl was found during growth on sorbitol (17). Also the operation of the glyoxylate shunt seems to differ from that in pathogenic mycobacteria, in which succinate repressed expression of both isocitrate lyase genes *aceA* and *icl* (14), whereas expression of genes for the glyoxylate shunt was detected when KMS grew on succinate. I speculate that this difference may be necessary to maximize the utilization of
all carbon sources when a low level of mixed substrates is available to this bacterium as would occur in the soil environment.

A novel enzyme for conversion of malate to OAA, malate:quinone oxidoreductase, completes the TCA cycle in KMS. Expression from the gene encoding this enzyme is confirmed for isolate KMS; the genome of this strain apparently lacks a gene encoding the normal malate dehydrogenase. This membrane-linked enzyme uses the cytoplasmic quinone as the cofactor and thus catalyzes a one-directional reaction from malate to OAA (23). This reaction would favor cycling of carbon through the TCA cycle to build cellular components, to generate reducing power, NADH, needed for dioxygenase activity and to produce ATP and GTP required as enzyme substrates. Enzyme assay shows activity of MQO when KMS cells are cultured on several carbon sources suggesting constitutive expression from the gene. Activity of MQO is highest when aromatic compounds are used as the sole carbon source possibly because transport of aromatic compounds has influence on the membrane-linked enzyme.

The requirement for PEP or pyruvate for gluconeogenesis when cells are provided with the C2 or C4 carbon sources appears to be met in isolate KMS by two potential pathways diagramed as part of Fig. 6.5. In one pathway, PEP for gluconeogenesis is generated from OAA directly by PEP carboxykinase. The second pathway uses the malic enzyme to convert malate to pyruvate and then to PEP by PEP synthase. My evidence that these pathways could be functional is the detection of transcripts from the genes that encode the necessary enzymes. The further detection of PEP synthase activity demonstrates the possibility of direct generation of PEP from pyruvate in isolate KMS. I speculate that the malate pool is possibly depleted by its conversion to OAA due to the
function of malate:quinone oxidoreductase. The other reason may be that the GTP for the carboxykinase is generated from the TCA cycle during the conversion of succinyl-CoA to succinate.

**FIG 6.5** Proposed pathways for generation of phosphoenolpyruvate (PEP) from C2 or C4 carbon sources in isolate KMS. The products, acetyl-CoA and succinyl-CoA, from degradation of both pyrene and benzoate are noted by stippled boxes, and pyruvate, released during pyrene degradation or generated as intermediate, is denoted by a grey box. Other intermediates, important because they are the precursors of amino acids, are underlined. The locus tags for isolate KMS for genes encoding enzymes in the glyoxylate shunt (isocitrate lyase and malate synthase) and those involved in completion of the TCA cycle and C4-C3 exchanges are shown. OAA, oxaloacetic acid; PEP, phosphoenolpyruvate. Not all intermediates are shown.

I found that when the C6-sugar fructose was used as the sole carbon source, transcripts abundance from the genes encoding PEP carboxykinase, malic enzyme, and
especially PEP synthase, were detected. Expression levels were lower for these transcripts with growth on mannitol. I deduced that this pattern of expression may be related to the demand for PEP for transport of fructose. Fructose probably is transported into KMS through a PEP-dependent phosphotransferase system. These genes are present in the KMS genome and a functional study of this uptake system was demonstrated for the soil isolate *M. smegmatis* (33). In contrast, in *M. smegmatis*, mannitol is transported by an ABC type (ATP-binding cassette) system in which PEP is not needed (33). A lower expression of PEP synthase as well as the relative low PEP synthase activity when KMS was grown on this alditol would correlate with PEP-independent but ATP-dependent transport.

Overall, in this study, I expanded the knowledge of utilization of different carbon sources, with a focus on benzoate and fructose, by *Mycobacterium sp.* KMS. By using bioinformatics and transcriptional analysis I was able to establish that gene transcription occurred from discrete loci on the KMS genome for carbohydrate, organic acid and aromatic growth. I demonstrated differential regulation of both plasmid and chromosomal genes during growth on different carbon sources was relevant to environmental conditions typically encountered in a contaminated soil or the rhizosphere. I conclude that isolate KMS is well adapted to survive in carbon-restricted environments. In particular, I determined that degradation of pyrene by this microbe is promoted by carbon-limited conditions. My work also confirms the feasibility for phytoremediation where the function of the plant would be to enhance the populations of the PAH-degrading microbe as well as to transport the microbe on the root surface into new soil locations where undegraded PAHs absorbed to soil particles could be present.
FUTURE STUDIES

I propose to better understand the survival of isolate KMS in the soil and rhizosphere habitats. I would like to expand this study to determine whether expression of potential cAMP-regulated pyrene-degrading nidB and nidB2 operons is dependent on the intracellular cAMP level and whether cAMP is secreted outside KMS cells to alter the metabolism of other microbes. Although cAMP functions as a signaling molecule in both *E. coli* and mycobacteria, aspects of the mechanisms involved may differ. In *E. coli*, an increased cAMP level activates the operons repressed by the preferred carbon source glucose (4). However, no significant change of cAMP level is documented for pathogenic *Mycobacterium* growing on different carbon sources (1). Thus, it seems significant to determine whether cAMP-regulated catabolite repression in the environmental *Mycobacterium* isolate KMS relates to intracellular cAMP levels.

To answer this question, I plan to determine whether a reporter gene fused with the promoters containing the putative cAMP-CRP binding motif can be regulated in *E. coli* where the mechanism of cAMP action is resolved (27). Promoter regions of pyrene-degrading nidB and nidB2 operons containing the potential cAMP-CRP binding motif are shown in Table 6.1. DNA fragments containing the promoter regions will be amplified by PCR using the high fidelity DNA polymerase pfu (Fermentas Inc., MD, USA) to make sure no mismatch happens. The amplified DNA fragment will be ligated into plasmid pCR2.1 as a fusion with a promoterless lux gene cassette. These lux genes encode a luciferase, which when supplied with substrate will generate light (10). The plasmid will be transformed into *E. coli* in which the cAMP level will be regulated by culturing on different carbon sources (4). It is anticipated that the cAMP level will be low when
glucose is used as the carbon source and the cAMP level will increase with growth on acetate due to activation of the adenylate cyclase. I will measure the light from the transformed E. coli grown on glucose or acetate. If cAMP regulation is observed, then it means that this sequence actively recognizes cAMP-CRP from E. coli. Consequently it supports the concept that the operons containing the nidB-nidA or nidB2 genes are cAMP regulated in KMS.

A second objective in this study is to check on the intracellular and secreted levels of cAMP when KMS is grown under different conditions. Secretion of cAMP is a process proposed to occur with pathogenic Mycobacterium species (1). I will assay levels of cAMP during the two diauxic phases of growth in cultures with pyrene amended with fructose. The cAMP will be assayed using a cAMP ELISA kit (Assay Designs, Ann Arbor, MI, USA) (36). Cytoplasmic extracts will be obtained by pelleting the cells and performing cell fracture in a French pressure cell. The debris will be removed by centrifugation at 10,000 g for 15 min to leave the supernatant to be assayed. To assay for secreted product the ELISA assay will be performed with culture filtrates from which all cells will be removed by filtration through a 0.2 micron filter. From this assay I will determine whether catabolite repression of the nidB-nidA or nidB2 operon genes caused by the presence of fructose is related to the levels of cAMP either accumulated within the cell or secreted into the environment.

I will attempt to determine the role of de novo synthesis of cAMP in this process as was observed with adenylate cyclase activation in E. coli (4). I will design primer sets for all the potential adenylate cyclase genes and assay whether I can detect different expression levels during the diauxic phases of growth in pyrene-fructose mixed cultures.
From this research, I can answer 1) whether cAMP-regulated gene expression in isolate KMS is dependent on the cAMP level as in *E. coli* or is independent on the cAMP level as in pathogenic *Mycobacterium* species and 2) whether KMS secretes cAMP to the environment to potentially affect other microbes in its habitat within the soil and rhizosphere.

REFERENCES


APPENDICES
APPENDIX A

Growth of Isolate KMS on Different Compounds as Sole Carbon Source

The purpose of this growth study was to identify carbon sources including simple compounds from C1 to C6, aromatic compounds, and naturally complex compounds that could support the growth of KMS. The information will allow a better understanding of the survival of KMS in soil and rhizospheres. The extent of growth on each carbon source is shown in Table A1.

Environmental isolate KMS inhabits soils and rhizospheres (4, 9) where carbon sources are versatile. The simplest C1 compounds present in soils arise from degradation of pectin and lignin (7, 8), which are components of plant cell walls or are part of the root exudates or are formed by other soil organisms (5). Other environmental Mycobacterium isolates could grow on methanol and carbon monoxide (10, 11), but under the conditions tested KMS did not grow on any of the C1 compounds provided e.g. methanol, formaldehyde, formate, and formamide. Growth was tested by both OD$_{600\text{nm}}$ measurement and by counting the colony forming units (cfu) on Luria-Bertani (LB) plates inoculated with serial dilutions of the cultures.

Other C2 to C6 simple carbon sources including sugars, organic acids, amides and amino acids are found in plant root exudates (2) or produced by other soil microbes (1, 3). The results in Table A1 show that KMS can grow on acetamide and some amino acids, which contain nitrogen in addition to carbon. Unexpectedly from the annotation of the KMS genome for the function of a complete TCA cycle, the glyoxylate shunt, and gluconeogenesis pathway, KMS did not utilize citrate as a sole carbon source. I deduced that KMS lacks a transport protein for citrate. Also, KMS could not utilize the C2
compound glycine and the C3 compound serine but did grow on other C2 compounds and glycerol. I deduced that the lack of transport proteins may also underlie the KMS growth patterns.

While KMS could use benzoate and pyrene as sole carbon sources, it was unable to use the other aromatic compounds or phenolic polymers used in this study as sole carbon sources. Thus it seems that KMS could not function as a sole remediation agent for benzene, biphenyl, and toluene. Although salicylic acid has a structure close to phthalate, benzoate, and protocatechuate, which were degraded by KMS, the salicylic acid did not support growth. KMS may not have the transport protein for salicylic acid or the ring-hydroxylating dioxygenases possessed by KMS could not recognize salicylic acid as a substrate.

Growth of KMS on grape seed oil was detected, indicating a β-oxidation system to degrade fatty acids into acetyl-CoA to build cell structures.

**TABLE A1** Carbon-containing compounds and growth of isolate KMS

<table>
<thead>
<tr>
<th>Carbon number</th>
<th>Carbon source</th>
<th>Growth condition</th>
<th>Medium type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No growth (-)</td>
<td>Weak growth (-/+)</td>
</tr>
<tr>
<td>C1</td>
<td>methanol</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>formaldehyde</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>sodium formate</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>formamide</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>ethanol</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---</td>
<td>------------</td>
<td>---</td>
</tr>
<tr>
<td>sodium acetate</td>
<td>×</td>
<td>liquid &amp;</td>
<td></td>
</tr>
<tr>
<td>acetamide</td>
<td>×</td>
<td>liquid</td>
<td></td>
</tr>
<tr>
<td>glycine</td>
<td>×</td>
<td>liquid</td>
<td></td>
</tr>
<tr>
<td>glycerol</td>
<td>×</td>
<td>liquid</td>
<td></td>
</tr>
<tr>
<td>serine</td>
<td>×</td>
<td>liquid</td>
<td></td>
</tr>
<tr>
<td>succinate</td>
<td>×</td>
<td>liquid &amp;</td>
<td></td>
</tr>
<tr>
<td>sodium glutamate</td>
<td>×</td>
<td>liquid</td>
<td></td>
</tr>
<tr>
<td>glutamine</td>
<td>×</td>
<td>liquid</td>
<td></td>
</tr>
<tr>
<td>fructose</td>
<td>×</td>
<td>liquid</td>
<td></td>
</tr>
<tr>
<td>mannitol</td>
<td>×</td>
<td>liquid &amp;</td>
<td></td>
</tr>
<tr>
<td>glucose</td>
<td>×</td>
<td>liquid &amp;</td>
<td></td>
</tr>
<tr>
<td>inositol</td>
<td>×</td>
<td>liquid &amp;</td>
<td></td>
</tr>
<tr>
<td>citrate</td>
<td>×</td>
<td>liquid &amp;</td>
<td></td>
</tr>
<tr>
<td>sorbitol</td>
<td>×</td>
<td>liquid &amp;</td>
<td></td>
</tr>
<tr>
<td>Compound</td>
<td>Interaction</td>
<td>Phase</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Hexane</td>
<td>×</td>
<td>Liquid</td>
<td></td>
</tr>
<tr>
<td>Benzene (aromatic)</td>
<td>×</td>
<td>Agar plate</td>
<td></td>
</tr>
<tr>
<td>Sodium benzoate (aromatic)</td>
<td>×</td>
<td>Liquid</td>
<td></td>
</tr>
<tr>
<td>Salicylic acid (aromatic)</td>
<td>×</td>
<td>Liquid &amp; agar plate</td>
<td></td>
</tr>
<tr>
<td>Toluene (aromatic)</td>
<td>×</td>
<td>Liquid</td>
<td></td>
</tr>
<tr>
<td>Coumarin (aromatic)</td>
<td>×</td>
<td>Liquid &amp; agar plate</td>
<td></td>
</tr>
<tr>
<td>Biphenyl (aromatic)</td>
<td>×</td>
<td>Liquid &amp; agar plate</td>
<td></td>
</tr>
<tr>
<td>Pyrene (aromatic)</td>
<td>×</td>
<td>Liquid &amp; agar plate</td>
<td></td>
</tr>
<tr>
<td>Lignin (phenolic polymer)</td>
<td>×</td>
<td>Liquid &amp; agar plate</td>
<td></td>
</tr>
<tr>
<td>Humic acid salts (aromatic complex)</td>
<td>×</td>
<td>Liquid &amp; agar plate</td>
<td></td>
</tr>
<tr>
<td>Grape seed oil (mixture of fatty acids)</td>
<td>×</td>
<td>Liquid</td>
<td></td>
</tr>
</tbody>
</table>
The assays were performed in liquid phosphate-based minimal medium (PBM) (6) or 1.5% agar PBM medium plates amended with listed materials as a sole carbon source at a concentration of 0.1%. The starting pH of the medium with each carbon source was adjusted to 7.0 before sterilization. The growth temperature was 25 °C for both liquid cultures and agar plate growth. The data represents the results from at least two independent incubations. The growth status (growth, weak growth, or no growth) of KMS on each carbon source is designated by “×” in the corresponding column.

REFERENCES


APPENDIX B

Bioinformatic Tools Used in This Study

**Synteny analysis of pyrene-degrading gene islands by Mauve.** In this study, comparison of pyrene-degrading gene islands among five pyrene-degrading *Mycobacterium* isolates: KMS, MCS, JLS, PYR-1, and PYR-GCK was analyzed using the software Mauve.

Mauve was developed by the Genome Center of Wisconsin, at the University of Wisconsin, Madison. The function of Mauve is to align genomes and determine block gene rearrangements caused by recombination, horizontal gene transfer, and gene deletions (1). Mauve is installed by downloading the program from website http://asap.ahabs.wisc.edu/software/mauve/overview.html. Instructions on how to use Mauve are available at http://asap.ahabs.wisc.edu/mauve/mauve-user-guide/.

For my analyses, the genome sequences of chromosomes and plasmids of five pyrene-degrading *Mycobacterium* isolates were saved as files suffixed with .gbk in the NCBI (National Center for Biotechnology Information) database. These files were downloaded from the website ftp://ftp.ncbi.nih.gov/genomes/Bacteria/.

To illustrate the method I will use as an example the analysis for synteny of pyrene-degrading gene islands on KMS chromosome and the plasmids. The Mauve analysis was used in the research discussed in Chapter 2.

1) Open Mauve program, choose File → Align sequences (Fig. B1), and a window comes out.

2) Add the sequences of KMS chromosome and the plasmids from the hard drive, create an output path, and begin the alignment by clicking Align (Fig. B2).
3) The output was on alignment of the sequences at the whole genome level with the view style set as shown in Fig. B3. To check gene locations in detail, the genome should was zoomed in and shifted using the buttons in the panel for a scan (Fig. B3).

4) The same-color blocks on the chromosome and the plasmids represent the syntenic gene blocks (Fig. B4). The rectangular boxes represent the ORFs (open reading frames) (Fig. B4). Double clicking on one color block (e.g. the green one) in one replicon
(e.g. the chromosome of KMS) will put the corresponding syntenic gene blocks in the middle of the Mauve screen (Fig. B4).

**FIG B3** Set the view style and adjust the view size by using the panels.

**FIG B4** The view of the aligned gene blocks represented by the same color.

5) Zoom in the genome scale and check the information for any gene (locus number and gene product sequence) by single clicking the rectangular box representing that gene (e.g. Mkms_1686) (Fig. B5).

6) Export the screen containing the target gene blocks from Tools → Export → Export Image to a JPEG format figure to the hard drive (Fig. B6).
FIG B5 Enlarged view of genome to show the syntenic regions and gene details.

FIG B6 Save the current view by exporting the image.

**Gene and protein alignment and phylogenetic analysis by ClustalX2 and MEGA4.** In this study, alignment of ring-hydroxylation dioxygenase α- and β-subunits of the nucleotide and amino acid levels as well as the promoters of operons containing ring-hydroxylation dioxygenase subunit genes was conducted using ClustalX2. ClustalX2 performs a multiple sequence alignment sequentially by aligning the most closely related sequences first, and then adding additional sequences and groups. ClustalX2 was
downloaded from http://www.clustal.org/ and installed in the Windows operating system on an Acer computer.

Alignment by ClustalX2 can be used to produce a phylogenetic tree using software MEGA4 that uses the neighbor-joining method (2). MEGA 4 is a molecular evolutionary genetics analysis tool developed by the Center for Evolutionary Medicine and Informatics. It can be downloaded from http://www.megasoftware.net/mega4/index.html.

Use of ClustalX2 and MEGA4 to generate phylogenetic trees is discussed in Chapter 3. I would take construction of ring-hydroxylating dioxygenase β-subunit genes as an example. Its use is as follows:

1) The nucleotide sequences of all the ring-hydroxylating dioxygenase β-subunit genes were obtained from the JGI-IMG database and were put into one fasta format file, in which a “>” symbol is added ahead the name of each sequence (Fig. B7).

2) Open the ClustalX2 program and choose File → Load sequences to read the fasta file created in step 1, containing nucleotide sequences of ring-hydroxylating dioxygenase β-subunit genes (Fig. B8) and choose Alignment → Do complete alignment → Align to do the alignment (Fig. B8).

3) Genes were aligned as shown in Fig. B9 and an .aln format file was generated containing the aligned sequence, which includes the gaps shown by hyphens (-).

4) Open MEGA 4 program, choose File → convert to MEGA format → Data format choose .aln (clustal) → OK (the .aln format file was generated by ClustalX2 for complete alignment) (Fig. B10). After the .aln file was converted to MEGA4 format file, which can be read by MEGA4, save the file (Fig. B10).
FIG B7 Create a fasta format file containing the sequences to be aligned.

FIG B8 Load sequences and do a complete alignment using ClustalX2.

FIG B9 The aligned sequences shown by ClustalX2 and saved as .aln format file.
5) Open the generated MEGA4 file by choosing File → Open data, and choose the sequences as nucleotide or protein (Fig. B11), and then choose whether the sequence is protein-coding, followed by setting the genetic code as standard and clicking OK (Fig. B11).

6) Generate the phylogenetic tree by choosing Phylogeny → Bootstrap test of phylogeny → Neighbor-joining (Fig. B12). Set 1000 replicates and compute (Fig. B13).

7) The generated phylogenetic tree was saved by choosing Image → Save as TIFF file (Fig. B14).

FIG B11 Sequence analysis by MEGA4.
FIG B12 Set calculation method as Bootstrap Neighbor-Joining for generation of the phylogenetic tree.

FIG B13 Set replication times for Bootstrap analysis.
FIG B14 Save generated phylogenetic tree as a TIFF format figure.

REFERENCES


APPENDIX C
Permission Letters to Reprint Published Content

NRC RESEARCH PRESS LICENSE
TERMS AND CONDITIONS

Nov 20, 2012

This is a License Agreement between Chun Zhang ("You") and NRC Research Press ("NRC Research Press") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by NRC Research Press, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 3032510460652
License date Nov 19, 2012
Licensed content publisher NRC Research Press
Licensed content publication Canadian Journal of Microbiology
Licensed content title Polycyclic aromatic hydrocarbon degrading gene islands in five pyrene-degrading Mycobacterium isolates from different geographic locations
Licensed content author Chun Zhang, Anne J. Anderson
Licensed content date Jan 1, 2012
Volume number 58
Issue number 1
Type of Use Thesis/Dissertation
Requestor type Author (original work)
Format Print and electronic
Portion Full article
Order reference number
Title of your thesis / dissertation STUDY OF GENES RELATING TO DEGRADATION OF AROMATIC COMPOUNDS AND CARBON METABOLISM IN MYCOBACTERIUM SP. STRAIN
Expected completion date Dec 2012
Estimated size(pages) 300
Total 0.00 USD
Terms and Conditions

General Terms & Conditions

Permission is granted upon the requester's compliance with the following terms and conditions:

1. A credit line will be prominently placed in your product(s) and include: for books the author, book title, editor, copyright holder, year of publication; for journals the author, title of article, title of journal, volume number, issue number, and the inclusive pages. The credit line must include the
following wording: "© 2008 Canadian Science Publishing or its licensors. Reproduced with permission," except when an author of an original article published in 2009 or later is reproducing his/her own work.

2. The requester warrants that the material shall not be used in any manner that may be derogatory to the title, content, or authors of the material or to Canadian Science Publishing, including but not limited to an association with conduct that is fraudulent or otherwise illegal.

3. Permission is granted for the term (for Books/CDs-Shelf Life; for Internet/Intranet-In perpetuity; for all other forms of print-the life of the title) and purpose specified in your request. Once term has expired, permission to renew must be made in writing.

4. Permission granted is nonexclusive, and is valid throughout the world in English and the languages specified in your original request. A new permission must be requested for revisions of the publication under current consideration.

5. Canadian Science Publishing cannot supply the requester with the original artwork or a "clean copy."

6. If the Canadian Science Publishing material is to be translated, the following lines must be included: The authors, editors, and Canadian Science Publishing are not responsible for errors or omissions in translations.

v1.4

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500900134.

Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.
SPRINGER LICENSE
TERMS AND CONDITIONS

Nov 20, 2012

This is a License Agreement between Chun Zhang ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number: 3032481009655
License date: Nov 19, 2012
Licensed content publisher: Springer
Licensed content publication: Biodegradation
Licensed content title: Multiplicity of genes for aromatic ring-hydroxylating dioxygenases in Mycobacterium isolate KMS and their regulation
Licensed content author: Chun Zhang
Licensed content date: Jan 1, 2012
Volume number: 23
Issue number: 4
Type of Use: Thesis/Dissertation
Portion: Full text
Number of copies: 6
Author of this Springer article: Yes and you are a contributor of the new work
Order reference number:
Title of your thesis / dissertation: STUDY OF GENES RELATING TO DEGRADATION OF AROMATIC COMPOUNDS AND CARBON METABOLISM IN MYCOBACTERIUM SP. STRAIN
Expected completion date: Dec 2012
Estimated size(pages): 300
Total: 0.00 USD

Terms and Conditions
Introduction
The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

Limited License
With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry.
Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the licensing process.

This License includes use in an electronic form, provided its password protected or on the university’s intranet or repository, including UMI (according to the definition at the Sherpa website: http://www.sherpa.ac.uk/romeo/). For any other electronic use, please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com).

The material can only be used for the purpose of defending your thesis, and with a maximum of 100 extra copies in paper.

Although Springer holds copyright to the material and is entitled to negotiate on rights, this license is only valid, provided permission is also obtained from the (co) author (address is given with the article/chapter) and provided it concerns original material which does not carry references to other sources (if material in question appears with credit to another source, authorization from that source is required as well).

Permission free of charge on this occasion does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

Altering/Modifying Material: Not Permitted
You may not alter or modify the material in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s) and/or Springer Science + Business Media. (Please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com)

Reservation of Rights
Springer Science + Business Media reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

Copyright Notice:Disclaimer
You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Springer and the original publisher/journal title, volume, year of publication, page, chapter/article title, name(s) of author(s), figure number(s), original copyright notice) is given to the publication in which the material was originally published, by adding; with kind permission from Springer Science and Business Media"

Warranties: None

Example 1: Springer Science + Business Media makes no representations or warranties with respect to the licensed material.

Example 2: Springer Science + Business Media makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

Indemnity
You hereby indemnify and agree to hold harmless Springer Science + Business Media and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License
This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person.
without Springer Science + Business Media's written permission.

No Amendment Except in Writing
This license may not be amended except in a writing signed by both parties (or, in the case of Springer Science + Business Media, by CCC on Springer Science + Business Media's behalf).

Objection to Contrary Terms
Springer Science + Business Media hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC’s Billing and Payment terms and conditions. These terms and conditions, together with CCC’s Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer Science + Business Media (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC’s Billing and Payment terms and conditions, these terms and conditions shall control.

Jurisdiction
All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in The Netherlands, in accordance with Dutch law, and to be conducted under the Rules of the 'Netherlands Arbitrage Instituut' (Netherlands Institute of Arbitration).

OR:

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.

Other terms and conditions:

v1.3

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500900097.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.
This is a License Agreement between Chun Zhang ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

<table>
<thead>
<tr>
<th>License Number</th>
<th>3032521114539</th>
</tr>
</thead>
<tbody>
<tr>
<td>License date</td>
<td>Nov 19, 2012</td>
</tr>
<tr>
<td>Licensed content publisher</td>
<td>John Wiley and Sons</td>
</tr>
<tr>
<td>Licensed content publication</td>
<td>Journal of Basic Microbiology</td>
</tr>
<tr>
<td>Book title</td>
<td></td>
</tr>
<tr>
<td>Licensed content author</td>
<td>Chun Zhang, Anne J. Anderson</td>
</tr>
<tr>
<td>Licensed content date</td>
<td>Jun 26, 2012</td>
</tr>
<tr>
<td>Start page</td>
<td>n/a</td>
</tr>
<tr>
<td>End page</td>
<td>n/a</td>
</tr>
<tr>
<td>Type of use</td>
<td>Dissertation/Thesis</td>
</tr>
<tr>
<td>Requestor type</td>
<td>Author of this Wiley article</td>
</tr>
<tr>
<td>Format</td>
<td>Print and electronic</td>
</tr>
<tr>
<td>Portion</td>
<td>Full article</td>
</tr>
<tr>
<td>Will you be translating?</td>
<td>No</td>
</tr>
<tr>
<td>Order reference number</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.00 USD</td>
</tr>
</tbody>
</table>

 TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or a society for whom a Wiley Company has exclusive publishing rights in relation to a particular journal (collectively WILEY). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your Rightslink account (these are available at any time at http://myaccount.copyright.com)

Terms and Conditions

1. The materials you have requested permission to reproduce (the "Materials") are protected by copyright.

2. You are hereby granted a personal, non-exclusive, non-sublicensable, non-transferable, worldwide, limited license to reproduce the Materials for the purpose specified in the licensing process. This license is for a one-time use only with a maximum distribution equal to the number that you identified in the licensing process.
Any form of republication granted by this licence must be completed within two years of the date of the grant of this licence (although copies prepared before may be distributed thereafter). The Materials shall not be used in any other manner or for any other purpose. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Material. Any third party material is expressly excluded from this permission.

3. With respect to the Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Materials, or any of the rights granted to you hereunder to any other person.

4. The Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc or one of its related companies (WILEY) or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding (“Marks”) of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

5. NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

7. You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this
Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. Any fee required for this permission shall be non-refundable after thirty (30) days from receipt.

13. These terms and conditions together with CCC’s Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

14. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC’s Billing and Payment terms and conditions, these terms and conditions shall prevail.

15. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC’s Billing and Payment terms and conditions.

16. This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

17. This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state’s conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

**Wiley Open Access Terms and Conditions**

All research articles published in Wiley Open Access journals are fully open access: immediately freely available to read, download and share. Articles are published under the terms of the Creative Commons Attribution Non Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. The license is subject to the Wiley Open Access terms and conditions:

Wiley Open Access articles are protected by copyright and are posted to repositories and websites in accordance with the terms of the Creative Commons Attribution Non Commercial License. At the time of deposit, Wiley Open Access articles include all changes made during peer review, copyediting, and publishing. Repositories and websites that host the article are responsible for incorporating any publisher-supplied amendments or retractions issued subsequently.

Wiley Open Access articles are also available without charge on Wiley's publishing platform, Wiley Online Library or any successor sites.

**Use by non-commercial users**
For non-commercial and non-promotional purposes individual users may access, download, copy, display and redistribute to colleagues Wiley Open Access articles, as well as adapt, translate, text- and data-mine the content subject to the following conditions:

The authors' moral rights are not compromised. These rights include the right of "paternity" (also known as "attribution" - the right for the author to be identified as such) and "integrity" (the right for the author not to have the work altered in such a way that the author's reputation or integrity may be impugned).

Where content in the article is identified as belonging to a third party, it is the obligation of the user to ensure that any reuse complies with the copyright policies of the owner of that content.

If article content is copied, downloaded or otherwise reused for non-commercial research and education purposes, a link to the appropriate bibliographic citation (authors, journal, article title, volume, issue, page numbers, DOI and the link to the definitive published version on Wiley Online Library) should be maintained. Copyright notices and disclaimers must not be deleted.

Any translations, for which a prior translation agreement with Wiley has not been agreed, must prominently display the statement: "This is an unofficial translation of an article that appeared in a Wiley publication. The publisher has not endorsed this translation."

Use by commercial "for-profit" organisations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee. Commercial purposes include:

- Copying or downloading of articles, or linking to such articles for further redistribution, sale or licensing;
- Copying, downloading or posting by a site or service that incorporates advertising with such content;

The inclusion or incorporation of article content in other works or services (other than normal quotations with an appropriate citation) that is then available for sale or licensing, for a fee (for example, a compilation produced for marketing purposes, inclusion in a sales pack)

- Use of article content (other than normal quotations with appropriate citation) by for-profit organisations for promotional purposes
- Linking to article content in e-mails redistributed for promotional, marketing or educational purposes;
- Use for the purposes of monetary reward by means of sale, resale, licence, loan, transfer or other form of commercial exploitation such as marketing products

Print reprints of Wiley Open Access articles can be purchased from: corporatesales@wiley.com

Other Terms and Conditions:

BY CLICKING ON THE "I AGREE..." BOX, YOU ACKNOWLEDGE THAT YOU HAVE READ AND FULLY UNDERSTAND EACH OF THE SECTIONS OF AND PROVISIONS SET FORTH IN THIS AGREEMENT AND THAT YOU ARE IN AGREEMENT WITH AND ARE WILLING TO ACCEPT ALL OF YOUR OBLIGATIONS AS SET FORTH IN THIS AGREEMENT.

v1.7
If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500900156. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.
This is a License Agreement between Chun Zhang ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 3032490213794
License date Nov 19, 2012
Licensed content publisher Springer
Licensed content publication Current Microbiology
Licensed content title The Gluconeogenic Pathway in a Soil Mycobacterium Isolate with Bioremediation Ability
Licensed content author Chun Zhang
Licensed content date Jan 1, 2012
Type of Use Thesis/Dissertation
Portion Full text
Number of copies 6
Author of this Springer article Yes and you are a contributor of the new work
Order reference number
Title of your thesis / dissertation STUDY OF GENES RELATING TO DEGRADATION OF AROMATIC COMPOUNDS AND CARBON METABOLISM IN MYCOBACTERIUM SP. STRAIN
Expected completion date Dec 2012
Estimated size(pages) 300
Total 0.00 USD

Terms and Conditions

Introduction
The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

Limited License
With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry.

Licenses are for one-time use only with a maximum distribution equal to the number that you identified in the
licensing process.

This License includes use in an electronic form, provided its password protected or on the university’s intranet or repository, including UMI (according to the definition at the Sherpa website: http://www.sherpa.ac.uk/romeo/). For any other electronic use, please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com).

The material can only be used for the purpose of defending your thesis, and with a maximum of 100 extra copies in paper.

Although Springer holds copyright to the material and is entitled to negotiate on rights, this license is only valid, provided permission is also obtained from the (co) author (address is given with the article/chapter) and provided it concerns original material which does not carry references to other sources (if material in question appears with credit to another source, authorization from that source is required as well).

Permission free of charge on this occasion does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

Altering/Modifying Material: Not Permitted
You may not alter or modify the material in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s) and/or Springer Science + Business Media. (Please contact Springer at (permissions.dordrecht@springer.com or permissions.heidelberg@springer.com)

Reservation of Rights
Springer Science + Business Media reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

Copyright Notice:Disclaimer
You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Springer and the original publisher/journal title, volume, year of publication, page, chapter/article title, name(s) of author(s), figure number(s), original copyright notice) is given to the publication in which the material was originally published, by adding: with kind permission from Springer Science and Business Media”

Warranties: None

Example 1: Springer Science + Business Media makes no representations or warranties with respect to the licensed material.

Example 2: Springer Science + Business Media makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

Indemnity
You hereby indemnify and agree to hold harmless Springer Science + Business Media and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License
This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person
without Springer Science + Business Media's written permission.

No Amendment Except in Writing
This license may not be amended except in a writing signed by both parties (or, in the case of Springer Science + Business Media, by CCC on Springer Science + Business Media's behalf).

Objection to Contrary Terms
Springer Science + Business Media hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer Science + Business Media (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

Jurisdiction
All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in The Netherlands, in accordance with Dutch law, and to be conducted under the Rules of the 'Netherlands Arbitrage Instituut' (Netherlands Institute of Arbitration). OR:

All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.

Other terms and conditions:

v1.3

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500900105. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.
CURRICULUM VITAE

CHUN ZHANG
(March 2012)

Contact method

E-Mail: zhangch0918@gmail.com

Education

B.S. 2001 Microbiology Nankai University; China
M.S. 2004 Microbiology Nankai University; China
Ph.D. 2011 Biology Utah State University; USA

Research experience

1. Dissertation: Study of genes relating to degradation of aromatic compounds and carbon metabolism in *Mycobacterium sp.* strain KMS.
2. Identification of unknown bacteria from fire-burnt forest soils by sequencing their 16S rRNA genes.
3. Construction of biosensor *Pseudomonas putida* KT2440 strains containing a plasmid with *lux* gene fusion to Cu and Cd responsive promoters.
4. Construction of gene library of O-antigen gene clusters from *E. coli* and identification of the conserved genes in the O-antigen gene cluster.

Teaching experience

Teaching assistant for Department of Biology, Utah State University
2009-2010: Elementary Microbiology BIOL 2060
2010-2011: Microbial Physiology BIOL 5300

2011: Plant Pathology BIOL 4430

**Technical skills**

1. Bacterial DNA extraction and RNA extraction.
2. End-point PCR, long fragment PCR, reverse transcriptase PCR, real time PCR.
3. Isolation and identification of unknown bacteria.
4. Transformation and electroporation.
5. Setting up shot-gun gene bank.

**Publications**


