SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of Biotechnology and Biosciences

PhD in Biology and Biotechnology
XXXI cycle

Genomic and functional analysis of *Rhodococcus* strains to identify genes and degradative functions for soil quality evaluation

Di Canito Alessandra

705596

Tutor: Dr. Patrizia Di Gennaro

Coordinator: Prof. Paola Branduardi

ACADEMIC YEAR 2017/2018

Abstract **8**

Riassunto **10**

Chapter 1. **Introduction** **12**

1.1 Preface **13**

1.2 Soil quality **14**

1.3 Soil biodiversity **16**

1.3.1 Soil Microorganisms 16

1.4 *Rhodococcus* genus **19**

1.4.1 *Rhodococcus* genus genomes 21

1.4.1.1 *Rhodococcus opacus R7* 23

1.4.1.2 *Rhodococcus aetherivorans BCP1* *24*

1.4.1.3 *Rhodococcus erythropolis MI2* *24*

1.4.2 *R. opacus* R7 and *R. aetherivorans* BCP1 comparative analysis 25

1.5 Scope of the thesis work **27**

1.6 *References* **28**

Chapter 2. **Biodegradative potential of *R. opacus* R7 and *R. aetherivorans* BCP1: genome-based analysis for the identification of pathways and gene clusters** **35**

2.1 Introduction **36**

2.2 Material and methods **37**

2.2.1 Identification of genetic aspects for xenobiotic degradation pathways 37

2.3 Results **38**

2.3.1 Genetic aspects related to the great potential for organic/xenobiotic degradation in *R. opacus* R7 and *R. aetherivorans* BCP1 38

2.3.2 Aliphatic hydrocarbons and cycloalkanes degradation 38

2.3.3 Aromatic hydrocarbons degradation 39

2.3.4 Polycyclic aromatic hydrocarbons degradation 41

2.3.5 Naphthenic acids degradation 43

2.3.6 Genetic aspects related to aromatic peripheral pathways in *Rhodococcus opacus* R7 and *Rhodococcus aetherivorans* BCP1 45

2.4 Discussion **48**

2.5 *References* **49**

Chapter 3. **Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7 .. 52**

3.1 Introduction 54

3.2 Materials and methods 55

3.2.1 Statistical analysis of pm data 55

3.2.2 Identification of genetic aspects 55

3.3 Results 56

3.3.1 Tolerance to osmotic stress 56

3.3.2 Tolerance to acid and alkaline stress 58

3.3.3. Resistance to antibiotics, metals and other toxic compounds 59

3.4 Discussion 62

3.5 *References* 75

Chapter 4. **Genome-based analysis for the identification of genes involved in *o*-xylene degradation in *Rhodococcus opacus* R7 ..... 79**

4.1 Introduction 81

4.2 Materials and methods 83

4.2.1 Bacterial strain and growth conditions 83

4.2.2 Bioinformatic analysis: nucleotide sequence determination and protein sequence analysis 84

4.2.3 Preparation, analysis, and DNA manipulation 85

4.2.4 RNA extraction and RT-PCR, quantitative real-time RT-PCR (qPCR) 86

4.2.5 Mutant preparation 87

Transposon-induced mutagenesis in R. opacus R7 using IS1415 (pTNR-TA vector) ... 87

Analysis of pTNR-TA insertion sites 87

4.2.6 Construction of the recombinant strain *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) . 88

4.2.7 Bioconversion experiments of the recombinant strain *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) in presence of *o*-xylene 88

4.2.8 Analytical methods 88

4.3 Results 89

4.3.1 R7 genome sequence analysis 89

Analysis and clusterization of dioxygenases 89

Analysis and clusterization of monooxygenases/hydroxylases 91

4.3.2 Involvement of the *akb* genes in *o*-xylene degradation by RT-PCR experiments 93

4.3.3 Involvement of the *akb* genes in *o*-xylene degradation by the identification of *R. opacus* R7 mutants in this cluster 96

4.3.4 Identification of the involvement of the *phe* genes and the *prm* genes in *o*-xylene degradation by RT-PCR experiments 97

4.3.5 Quantitative real-time RT-PCR (qPCR) analysis 99
4.3.6 Involvement of the *prm* genes by cloning and expression of the activity in *R. erythropolis* AP..... 100
4.4 Discussion..... 101
4.6 *References* 112

***Chapter 5.* Soil quality assessment by the identified sequences and degradative functions of *R. opacus* R7 ..... 117**
5.1 Introduction..... 118
5.2 Materials and Methods..... 120
5.2.1 Characterization of sand sample 120
5.2.2 Preparation of slurry microcosms 121
5.2.3 Extraction and analysis of CHCA..... 121
5.2.4 Experimental plan 122
5.2.5 CHPCA cluster identification in R7 genome..... 122
5.2.6 RNA extraction, RT-PCR and RT-qPCR experiments..... 122
5.3 Results 124
5.3.1 Growth kinetic experiments of *R. opacus* R7 on CHCA 124
5.3.2 Identification of the *aliA1* and *pobA1* gene clusters and evaluation of their involvement in CHCA degradation in *R. opacus* R7 125
5.3.3 Characterization of sand samples..... 126
5.3.4 Growth kinetic experiments in microcosms in presence of CHCA 126
5.3.5 Gene expression of the marker sequence in microcosms by RT-PCR and RT-qPCR experiments..... 129
5.4 Discussion..... 131
5.5 *References* 133

***Chapter 6.* Investigation of 4,4'-dithiodibutyric acid (DTDB) utilization of *Rhodococcus erythropolis* MI2 by generation of marker-free deletion mutants**
(in collaboration with the Institut für Molekulare Mikrobiologie und Biotechnologie of the Westfälische Wilhelms-Universität of Münster, during the PhD period abroad)..... **137**
6.1 Introduction..... 138
6.1.1 PTEs production..... 138
6.1.2 DTDB degradation in *R. erythropolis* MI2 140
6.1.3 Removal of the sulfur group 142
6.2 Materials and Methods..... 143
6.2.1 Bacterial strains and culture conditions 143

6.2.2 Deletion constructs preparation 143
6.2.3 Isolation of the target gene flanking regions of *R. erythropolis* MI2 145
6.2.4 Recombinant pJET1.2/blunt vector constructions 146
6.2.5 *R. erythropolis* MI2 deletion mutant generations 146
6.2.6 Recombinant pJQ200mp18Tc suicide vector constructions 146
6.2.7 *R. erythropolis* MI2 competent cells preparation 148
6.2.8 *R. erythropolis* MI2 transformation by the deletion constructs 148
6.3 Results 149
6.3.1 *Rhodococcus erythropolis* MI2 characterization 149
6.3.2 Deletion mutants 150
6.3.3 Deletion of genes proposed to be involved in metabolism of DTDB 150
6.3.4 *Rhodococcus erythropolis* MI2 mutant generation 152
6.4 Discussion 154
6.5 *References* 155

***Chapter 7.* Conclusions and Future Perspectives 156**

***Scientific Contributions* 160**
Research papers 161
Conference contributions 161
Annex 162

Abstract

Soil quality has been one of the major issues of the last decades, because of the increase of anthropogenic pollution. Soil contains organisms that play vital functions such as nutrient and hydrological cycles and degradation of toxic compounds. Under stress conditions caused by contamination, soil microorganisms undergo several alterations. Molecular technologies use microbial communities as an ecological parameter in monitoring polluted sites, detecting community shifts in response to pollution.

Bacteria belonging to *Rhodococcus* genus have an important role in the degradation of recalcitrant compounds. It is a metabolically versatile genus, widely distributed in nature. *Rhodococcus* spp. can degrade a wide range of organic compounds (aliphatic and aromatic hydrocarbons, halogenated, polychlorinated biphenyls, nitroaromatics, heterocyclic, nitriles, sulfuric, steroids, herbicides) and to survive in presence of toxic compounds, under desiccation conditions, carbon starvation, wide range of temperatures, UV irradiation and osmotic stress. In line with their catabolic diversity, they possess large and complex genomes, containing a multiplicity of catabolic genes, high genetic redundancy and a sophisticated regulatory network.

The aim of this project is to obtain molecular tools in *Rhodococcus* strains to use as "marker" sequences for the assessment of soils, through genomic and functional analysis of metabolic pathways and catabolic gene clusters involved in the degradation of the most diffused environmental contaminants.

In particular, this work focused the attention on three *Rhodococcus* strain genomes: *R. opacus* R7, *R. aetherivorans* BCP1 and *R. erythropolis* MI2.

A Phenotype Microarray approach was used to evaluate R7 and BCP1 strains metabolic potential and their stress response. Also, their capability to utilize contaminants belonging to four categories (aliphatic hydrocarbons and cycloalkanes, aromatic compounds, polycyclic aromatic compounds (PAHs), naphthenic acids and other carboxylic acids) and to persist under various stress conditions (high osmolarity, pH stress, toxic compounds, antibiotics) was tested. A genome-based approach was used to relate their abilities to genetic determinants putatively involved in the analysed metabolisms (such as naphthalene, *o*-xylene, *n*-alkanes, naphthenic acids, phenols, phthalate) and in their environmental persistence.

In particular, *o*-xylene and naphthenic acids degradations were investigated in *R. opacus* R7. Computational and molecular analyses (RT and RT-qPCR, gene cloning, random mutants) revealed the putative involvement of several genes in these degradation pathways.

o-xylene is a pollutant that has great relevance for its toxic properties and persistence in environment. R7 can degrade *o*-xylene by the induction of the *akb* genes (by a deoxygenation) producing the corresponding dihydrodiol. Likewise, the redundancy of sequences encoding for monooxygenases and hydroxylases (*prmA* and *pheA1A2A3*), supports the involvement of other genes that induce the formation of phenols, converging to the phenol oxidation path. The activation of converging oxygenase systems represents a strategy in *Rhodococcus* genus to degrade recalcitrant compounds and to persist in contaminated environments.

Naphthenic acids (NAs) are a relevant group of new emerging contaminants, because of their toxicity and the uncountable environmental impact. NAs degradation pathway is not fully clear but two main routes have been proposed: i) aromatization of the cyclohexane ring ii) activation as CoA thioester. RT and RT-qPCR results showed that *R. opacus* R7 degrade cyclohexanecarboxylic acid (CHCA) molecule (chosen as a model compound) by a cyclohexane carboxylate CoA ligase (*aliA*).

The potential application of this work was demonstrated by a microcosm approach, simulating a bioaugmentation process with R7 strain. Autochthone bacteria and R7 capabilities to degrade CHCA were evaluated and compared; results indicated that R7 can degrade the contaminant faster than the microbial community and that its contribute increased CHCA consumption. The degradation was followed by RT and RT-qPCR, monitoring the expression of the *aliA* gene.

Moreover, a potential biotechnological application was investigated in *R. erythropolis* MI2, studying the disulfide 4,4-dithiodibutyric acid (DTDB) degradation pathway. DTDB is a promising substrate for polythioester (PTE) synthesis; indeed, its degradation produces the PTE building block 4-mercaptobutyric acid. The aim was pursued generating *R. erythropolis* MI2 marker-free deletion mutants for genes involved in the final steps of the pathway.

Riassunto

La qualità del suolo è una delle principali problematiche ambientali degli ultimi decenni a causa dell'aumento dell'inquinamento antropico. Il suolo è un sistema dinamico e gli organismi autoctoni giocano un ruolo fondamentale per il mantenimento delle sue funzioni (cicli idrologici e dei nutrienti e degradazione dei composti tossici). In condizioni di stress ambientali, i microrganismi del suolo subiscono alterazioni; esse possono essere usate come parametro ecologico per il monitoraggio dei siti contaminati mediante tecnologie molecolari.

I batteri appartenenti al genere *Rhodococcus* hanno un ruolo importante nella degradazione dei composti più recalcitranti. Sono versatili ed ampiamente distribuiti in natura; essi possono degradare diversi composti organici, tra cui idrocarburi alifatici ed aromatici, alogenati, nitroaromatici, eterociclici, nitrili, sulfuri, steroidi ed erbicidi. Inoltre, essi possono sopravvivere in presenza di composti tossici, in condizioni di essiccazione, carenza di carbonio, irradiazione UV e stress osmotico. Questa versatilità metabolica è correlata con la complessità dei loro genomi, i quali contengono molteplici geni catabolici, ridondanza genica e un sofisticato network regolatorio.

L'obiettivo di questo progetto è ottenere nuovi tools molecolari da ceppi di *Rhodococcus* da usare come marcatori per la valutazione della qualità dei suoli, attraverso analisi genomiche e funzionali dei pathway metabolici e dei cluster genici coinvolti nella degradazione dei contaminanti ambientali. In questo lavoro, l'attenzione è stata rivolta verso i genomi dei ceppi: *R. opacus* R7, *R. aetherivorans* BCP1 e *R. erythropolis* MI2.

Un'analisi fenotipica mediante microarray ha permesso di valutare il potenziale metabolico e la risposta allo stress dei ceppi R7 e BCP1; sono stati testati contaminanti appartenenti a quattro categorie (idrocarburi alifatici e cicloalcani, aromatici, policiclici aromatici, acidi naftenici ed altri acidi carbossilici) e varie condizioni di stress (alta osmolarità, differenti valori di pH, composti tossici, antibiotici). Quindi, è stato usato un approccio genomico per correlare le abilità metaboliche identificate a determinanti genici, putativamente coinvolti nei metabolismi analizzati (degradazione di naftalene, *o*-xilene, *n*-alcani, acidi naftenici, fenoli, ftalato) e nella persistenza ambientale.

In particolare, sono stati esaminati i pathway degradativi dell'o-xilene e degli acidi naftenici di *R. opacus* R7. Analisi bioinformatiche e molecolari (RT e RT-qPCR, clonaggio, mutanti) hanno permesso di valutare il coinvolgimento di diversi geni nei pathway degradativi.

L'*o*-xilene è un contaminante tossico e persistente in ambiente; R7 è in grado di degradarlo inducendo la trascrizione dei geni *akb* (sistema diossigenasico) formando il diidrodiolo. Tuttavia, la ridondanza di monossigenasi e idrossilasi (*prmA* and *pheA1A2A3*), ha suggerito il coinvolgimento di altri geni

convergenti nella via di ossidazione dei fenoli. L'attivazione di sistemi convergenti rappresenta una strategia dei rhodococci per degradare composti recalcitranti e persistere in ambienti contaminati.

Gli acidi naftenici (NAs) sono una classe di contaminanti emergenti, tossici e dall'enorme impatto ambientale. I loro pathway degradativi non sono ancora noti ma sono state proposte due possibili vie: i) aromatizzazione dell'anello del cicloesano ii) attivazione come CoA tioestere. I risultati delle RT e RT-qPCR hanno mostrato che R7 degrada l'acido cicloesanocarbossilico (CHCA) (scelto come composto modello), attraverso una cicloesano carbossilato-CoA ligasi (*aliA*).

L'applicazione di questo lavoro è stata dimostrata in esperimenti di microcosmo simulando condizioni reali con sabbia bioaugmentata con R7. Le capacità dei batteri autoctoni e di R7 di degradare il CHCA sono state comparate e i risultati mostrano che R7 degrada il contaminante più velocemente rispetto alla comunità microbica e che il suo contributo aumenta la velocità di degradazione del CHCA. La velocità di degradazione è stata seguita monitorando l'espressione del gene *aliA* mediante esperimenti di RT e RT-qPCR.

Una potenziale applicazione biotecnologica di questo lavoro è stata valutata in *R. erythropolis* MI2, studiando il pathway di degradazione dell'acido 4,4'-ditiobutirrico (DTDB). Esso è un promettente substrato per la sintesi dei politioesteri (PTE) poiché un suo intermedio metabolico, l'acido 4-mercaptobutirrico ne è un precursore. L'obiettivo di questo studio è stato perseguito generando mutanti di delezione del ceppo MI2 per i geni coinvolti nelle reazioni finali del pathway di degradazione.

Chapter 1.

Introduction

1.1 Preface

Soil quality has been one of the major issues of the last decades because of the increase of anthropogenic pollution, along with difficulties in the maintenance of its complex multi-functional aspects. Soil functions are fundamental representing a resource for agriculture, livestock and industry, and the soil conservation is essential for the life maintenance. Soil also plays an important role in supporting animal biodiversity according to the seasons; these species are directly connected to the water system, the region climate, the soil structure and the plant residues content (Zilli et al. 2003; Larigauderie, Prieur-Richard and Mace 2012; Pulleman et al. 2012). Moreover, the diverse soil organisms play many vital functions including decomposition of plant debris, acquire components from the atmosphere, aerate the soil as well as regulation of carbon, nutrient and hydrological cycles and degradation of toxic compounds. For all these reasons soil degradation implies changes in soil health status resulting in a reduced capacity of the ecosystem to provide goods and services.

In this context, the great number of facilities and processes have enhanced the widespread of environmental contamination affecting human health, water resources, ecosystems and other receptors (Burgess, 2013). The continual discovery of contaminated land over recent years has led to international efforts to remedy many of these sites for a dual purpose, either as a response to the risk of adverse health or environmental effects caused by contamination or to enable the site to be redeveloped for use (Vidali, 2001).

The necessity to monitor and avoid the negative effects of soil degradation has formed the basis of the dissertation and responsiveness on soil governance and has also seen the emergence of science and technology as mutual links between soil management and soil governance (http://www.fao.org/home/en/).

In response to public and government concern and because of the intriguing research problems presented, environmental scientists, biologists and chemists have been giving increased attention to the identification and the determination of the behavior and fate of organic compounds spreading in natural ecosystems (Alexander, 1981).

1.2 Soil quality

Environmental quality includes three main components: soil quality, water and air quality (Andrews et al., 2002). Soil quality can be defined as the capacity of a living soil to function, within natural or managed ecosystem boundaries, to sustain plant and animal productivity, maintain or enhance water and air quality, and promote plant and animal health (Doran, 2002). Water and air quality are defined mainly by their degree of pollution that impacts directly on human and animal consumption and health, or on natural ecosystems (Carter et al., 1997; Davidson, 2000). Soil health is not limited to the degree of soil pollution but is commonly defined much more broadly as "the capacity of a soil to function within ecosystem and land-use boundaries to sustain biological productivity, maintain environmental quality, and promote plant and animal health" (Doran and Parkin, 1994; 1996).
This definition reflects the complexity and site-specificity of the belowground part of terrestrial ecosystems as well as the many linkages between soil functions and soil-based ecosystem services (Bünemann et al., 2018) (**Figure 1.1**).

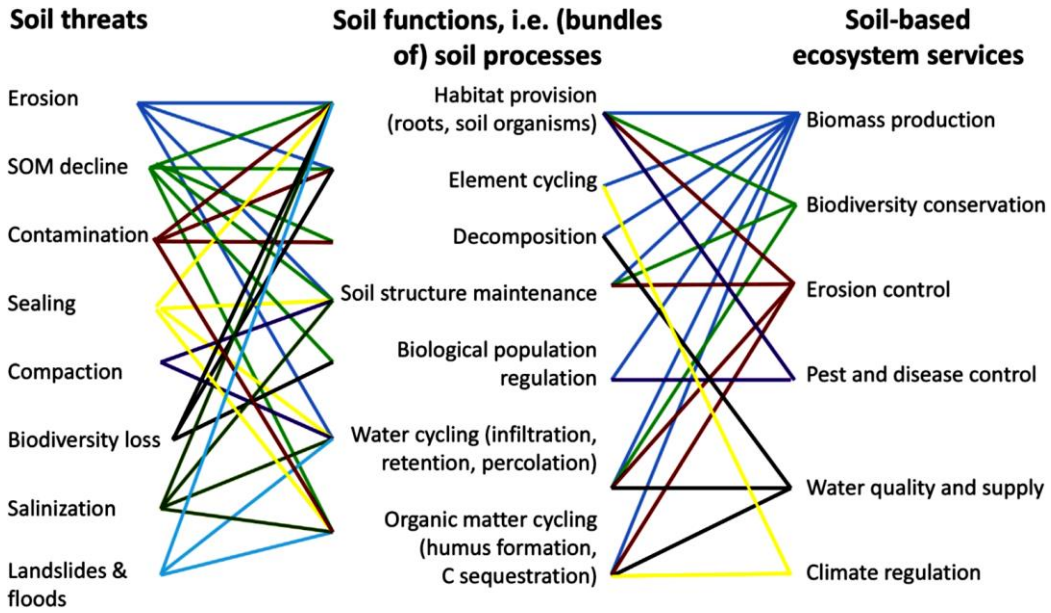


Figure 1.1 Linkages between soil threats, soil functions and soil-based ecosystem services (scheme presented by Kibblewhite et al. (2008a) and modified by Brussaard (2012)). (Bünemann et al., 2018)

The assessment of biological indicators of soil quality is required to connect abiotic soil properties to soil functions in terms of biochemical and biophysical transformations (Lehman et al., 2015). Consequently, adding biological and biochemical indicators can greatly improve soil quality assessments (Barrios, 2007). Nevertheless, soil biological indicators are still underrepresented in soil quality assessments, probably because they require specific knowledge and skills. This situation is

unfortunate because soil biota is considered the most sensitive indicators of soil quality due to their high responsiveness to changes in management and environmental conditions (Bünemann et al., 2018). The presence of a diverse soil microbial community is crucial to the productivity of any ecosystem, since microorganism affect all levels within the ecosystem (Kennedy and Stubbs, 2006). Its composition influences the rate of residue decomposition and nutrient cycling (Beare et al., 1993). Microorganisms are soil decomposer able to break down complex organic compounds into simple and possibly recalcitrant compounds; together with fungi, bacteria are crucial for the mineralization of nutrients, making them available to plants and other organisms (Henriksen and Breland, 1999). They also play a major role in soil structure and consequently will influence soil quality (Lynch and Bragg, 1985). Indeed, soil bacteria aid in weathering soil minerals, contribute to soil formation, and secrete polysaccharides to hold soil particles together and promote aggregate stability. These actions reduce erosion, allow for increased water infiltration and maintain adequate aeration of the soil.

Diverse soil microbial communities and dynamics within and among these communities can be used to evaluate changes in soil health. Their increase or decrease in activity and diversity indicate degradative or aggregative processes of the soil (Kennedy and Stubbs, 2006). Moreover, the diverse microorganisms have complex interactions among each other and have both redundant and specific functions. Therefore, microbial diversity and composition are key determinants of the ecological functions (Nannipieri et al., 2003; Philippot et al., 2013). For these reasons, microbial diversity and the related process level responses such as biomass, respiration rates and enzyme activities should be examined simultaneously under the context of global changes (Wang et al., 2018).

Recent advances in soil biology have encouraged the employment of indicators based on genotypic and phenotypic community diversity. Over the past few decades the advent of molecular technologies has led environmental microbiologists to recognize microbial communities as crucial ecological parameter in monitoring polluted sites either by detecting community shifts in response to pollution or their resilience towards anthropogenic disturbances (**Figure 1.2**). Molecular methods, especially omics approaches, hold great potential to perform faster, cheaper and more informative measurements of soil biota and soil processes than conventional methods (Bouchez et al., 2016). These methods include the determination of microbial diversity and activity, the screening for particular gene diversity, the gene quantification, and also the whole-genome sequencing of bacterial isolates and metagenome sequencing. Consequently, these novel approaches may yield new indicators that could substitute or complement existing biological and biochemical soil quality indicators in regular monitoring programs. The rapid evolution of these techniques and the decreasing costs associated with them will facilitate this development (Bünemann et al., 2018).

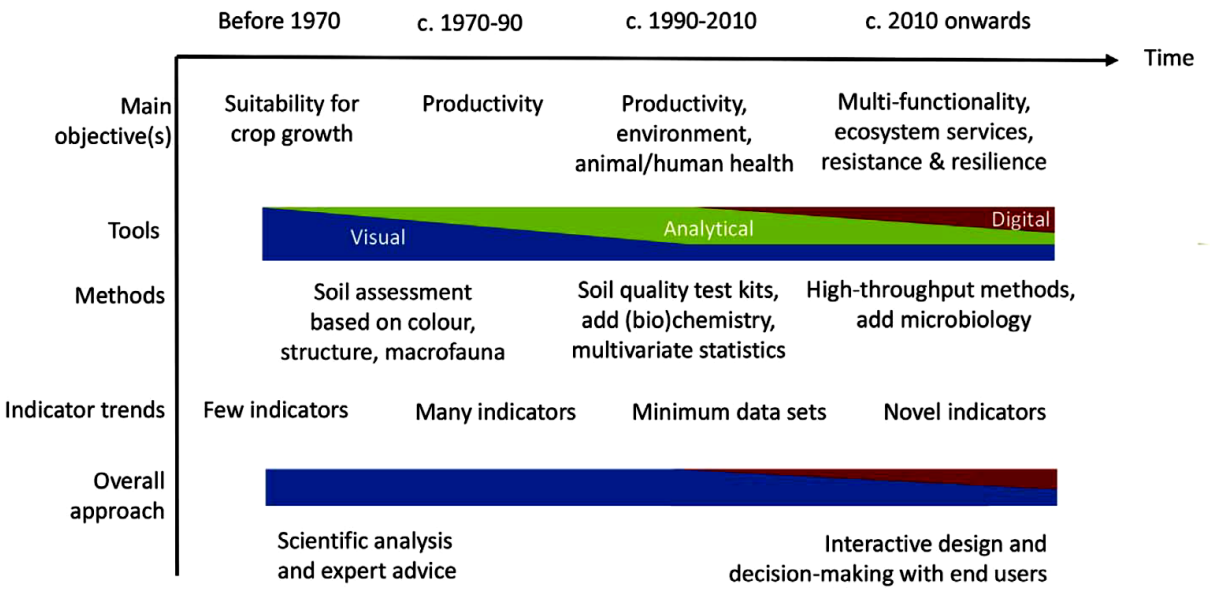


Figure 1.2. Main objectives, tools and approaches of soil quality assessment through history. (Bünemann et al., 2018)

1.3 Soil biodiversity

Soil biodiversity addresses the variability among living organisms including microorganisms (e.g. bacteria, fungi, protozoa and nematodes), mesofauna (e.g. acari and springtails) and macrofauna (e.g. earthworms and termites). Plant roots can also be considered as soil organisms for their symbiotic relationships and interactions with other soil components. These diverse organisms interact with plants and animals in the ecosystem forming a complex network of biological activities. The component of soil organisms is involved in a wide range of essential services to the sustainable function of all ecosystems as well as to the sustainable management of agricultural systems. They act as the primary driving agents of nutrient cycling, regulating the dynamics of soil organic matter, soil carbon sequestration and greenhouse gas emission, modifying soil physical structure and water regimes, enhancing the amount and efficiency of nutrient acquisition by the vegetation and enhancing plant health (http://www.fao.org/soils-portal/soil-biodiversity).

1.3.1 Soil Microorganisms

Soil is a dynamic living system in which microbial community plays an important role in its quality assessment. The soil microbiota is responsible for the transformation of organic matter through biochemical reactions, as well as acting in nitrogen fixation, nutrients cycling, detoxification of pollutants and air and water purification, which are important processes for maintenance of soil

quality. Ideally, this microbiota allows nutrients to be gradually released, which contributes to the growth and nutrition of plants. The diversity of the microbial community, as well the functions within communities, influences the stability and resilience of the soil system and thus the health of the soil. The diversity and the functions of microorganisms can be more significant than the number of species in regulating ecosystem processes (Grime, 1997; Wardle et al., 1997; Bądgett and Shine, 1999). In particular, the variation in species richness, not discernible in many environments, can be indicative in stressed and altered systems (Yachi and Loreau, 1999). Although higher microbial diversity has an important role in soil quality, it may not lead to improved ecosystem health, thus diversity indices along with additional soil information are necessary to determine the status of a soil (Kennedy and Stubbs, 2006).

Under stress conditions caused by anthropogenic pollution, climate change, land use conversion or management regimes, soil microorganisms undergo several alterations. In particular, the great influence of the human action has been mainly ascribed to the influence of prensence of diverse xenobiotics in soil (Somerville and Greaves, 1987; Tòtola and Chaer, 2002; Moreira and Siqueira, 2006). Therefore, microbiologically-related parameters should be involved in soil quality assessment. Likewise, the use of microorganisms and their functioning for examination of environmental stress and declining biological diversity can be used at level of ecosystems (OTA, 1987). It has been estimated that only 0.1-1% of the total soil bacterial population can be cultured by applying standard cultivation techniques and many of soil enzymatic functions could not be identified. Thus, to circumvent some of the difficulties in enriching and isolating microorganisms, cultivation-independent molecular approaches have been developed (Rolf, 2004).

Catabolic genes encode for specific enzymes in metabolic pathways, such as key enzymes in xenobiotic degradation pathways. By assessing the abundance or the expression of key genes in environmental samples, one can get a potential measure of the degradation activity. The abundance and expression of specific catabolic genes can be analyzed from metagenomic DNA and RNA by using single gene PCR assays or by using sequences of known DNA as probes e.g. plotted onto an array on a microchip for multiple gene assessment. Specific real-time PCR methods have been used to calculate the original copy number of target genes in bioremediation studies focusing on naphthalene, nitro aromatics and arsenic degradation (Jørgensen, 2008). Due to the increasing number of known functional genes and the need of rapid screening analysis, efforts have been made to construct comprehensive microarrays that contain large numbers of DNA probes for the simultaneous probing of metagenomic DNA or cDNA. An oligoarray to detect hundreds of functions related to bacterial degradation of pollutants, including catabolic, regulatory, resistance and stress genes, has been reported (Rhee et al., 2004) and evolved as the so-called GeoChip (He et al., 2007).

As abovementioned, one of the main stress factor that influences soil microbial community is the presence of recalcitrant compounds. A representative recalcitrant compounds category is that of the hydrocarbons; their concentration is increased in the last decades principally because of processes involving petroleum extraction, transport or processing; these kinds of processes lead to extensive release in the environment of hydrocarbons mainly containing saturated hydrocarbons, polycyclic aromatic hydrocarbons and nitrogen-sulfur-oxygen compounds. A typical soil, sand or ocean sediment contains significant amounts of hydrocarbon-degrading microorganisms, and their numbers increase considerably in oil-polluted sites.

The microorganisms involved in the degradation of these recalcitrant compounds are often bacteria belonging to following genera: *Pseudomonas*, *Alcaligenes*, *Flavobacterium*, *Arthrobacter*, *Nocardia*, *Rhodococcus* and *Corynebacterium* (Wagner-Döbler et al., 1998) (**Table 1.1**).

Table 1.1. Aerobic microorganisms able to degrade different kind of hydrocarbons

Bacteria	*Achromobacter* - *Acinetobacter* - *Alcanivorax* - *Alcaligenes* - *Bacillus* - *Brevibacterium* - *Burkholderia* - *Corynebacterium* - *Flavobacterium* - *Mycobacterium* - *Nocardia* - *Pseudomonas* - *Rhodococcus* - *Sphingomonas* - *Streptomyces*
Yeast	*Candida* - *Cryptococcus* - *Debaryomyces* - *Hansenula* *Pichia* - *Rhodotorula* - *Saccharomyces* - *Sporobolomyces* *Torulopsis* - *Trichosporon* - *Yarrowia*
Fungi	*Aspergillus* - *Cladosporium* - *Corollasporium* - *Cunninghamella* - *Dendryphiella* – *Fusarium* – *Gliocladium* - *Lulworthia* – *Penicillium* - *Varicospora* - *Verticillium*
Algae	*Prototheca*

Indeed, these genera of bacteria have a remarkable versatile metabolism; they are able to completely or partially oxidize different organic substances.

In particular, bacteria belonging to *Pseudomonas* genus seem to be widespread and versatile (Harayama et al., 1989); indeed, some of them are able to produce biosurfactants with proper nutrients available, which improve the degradation of crude oil by decreasing the surface tension of medium. Among the other, the *Rhodococcus* genus is regarded as one of the most promising groups of organisms able to degrade a wide range of recalcitrant compounds that cannot be easily transformed by other organisms.

1.4 *Rhodococcus* genus

Rhodococcus genus members are taxonomically placed in the mycolic-acid-forming sub-order *Corynebacterineae* in the family *Nocardiaceae* within in the phylum *Actinobacteria*, which contains the Gram-positive, high mol% G+C (63-73%). They are non-motile, aerobic chemoorganotrophic organisms with an oxidative metabolism, able generally to use a wide range of organic compounds as sole carbon and energy source. *Rhodococcus* strains can be subjected to morphological changes, the cocci can convert to rods and filaments in some strains. The colonies colours are not always red, they can vary from colourless to buff, cream, yellow, orange and red. *Rhodococcus* species are particularly interesting for their structural cell envelope genomes and plasmids, along with metabolic and catabolic abilities. Their cell envelopes are dominated by the presence of large branched chain fatty acids, the mycolic acids and they often present arabinogalactan cell wall polysaccharide that is covalently attached to the cell wall peptidoglycan and in turn provides a scaffold for the covalent anchoring of mycolic acids (Larkin et al., 2010). The *Rhodococcus* peptidoglycan contains glycolated muramic acid residues in the glycan chain and meso-2,6-diaminopimelic acid as the dibasic amino acid in the tetrapeptide. Arabinose and galactose are major wall sugars. Mycolic acids are 30-54 carbon atoms in length and contain as many as four double bonds. Major phospholipids are diphosphatidylglycerol, phosphoethanolamine and phosphatidyly inositol mannosides, and fatty acids are straight-chain saturated and monosaturated fatty acids and 10-methyloctadecanoic branched fatty acids. *Rhodococcus* members are often metabolically versatile actinobacteria frequently found in the environment, distributed in soil, water and marine sediments (Finnerty, 1992; Martínková et al., 2009). *Rhodococci* have been isolated from many different habitats, initially from faecally contaminated soil and aquatic sediments (Goodfellow et al., 1998; Jones and Goodfellow, 2010), where they often have enriched from persistent contamination with xenobiotic and other organic compounds that they can use as sole carbon and energy source. Some of them, like *R. equi* are animal and human pathogens, especially in immuno-compromised patients, while others are members of the plant rhizosphere community, gall-forming plant pathogens or symbionts in the gastrointestinal tract of blood-sucking arthropods (Gürtler and Seviour, 1998).

Rhodococcus genus is regarded as one of the most promising groups of organisms suitable for the biodegradation of compounds that cannot be easily transformed by other organisms (Warhurst and Fewson, 1994). Indeed, these heterotrophs commonly occur in soil where they degrade a wide range of organic compounds, some of which are likely to pose long-term environmental risks. Thanks to their broad catabolic diversity and their tolerance to various environmental stress, *Rhodococcus* spp. play an important role in nutrient cycling and have potential applications in bioremediation,

biotransformations, and biocatalysis (de Carvalho, 2012). Moreover, they are also known to produce metabolites of industrial potential including carotenoids, bio-surfactants and bioflocculation agents and especially acrylamide (Goodfellow et al., 1998; Jones and Goodfellow, 2010). Therefore, *Rhodococcus* characteristics and "lifestyle" reflect their biotechnological importance. These capabilities are often due to their mobile, large and usually linear plasmids carrying genes encoding for a variety of enzymatic activities (de Carvalho and da Fonseca, 2005). *Rhodococci* are able to metabolize a wide range of xenobiotic compounds, including aliphatic and aromatic hydrocarbons, oxygenated, halogenated compounds, polychlorinated biphenyls, nitroaromatics, heterocyclic compounds, nitriles, sulphuric compounds, steroids and various herbicides. Their abilities have been also attributed to their mycolic acids, proposed to facilitate the uptake of hydrophobic compounds (Gürtler et al., 2004), determining an immense potential in bioremediation (Van der Geize and Dijkhuizen, 2004). They also tolerate various concentrations of water miscible (ethanol, butanol, and dimethylformamide, up to 50% v/v) and water immiscible solvents (dodecane, bis(2-ethylhexyl) phthalate and toluene, up to 5% v/v) (de Carvalho et al., 2004; de Carvalho and da Fonseca, 2005). The *Rhodococcus* strains tolerance could be related to several sequences located in their genomes coding for efflux pumps consenting the selective active transport of different substrates (de Carvalho et al., 2014; Kim et al., 2002). Moreover, they possess the ability to continue biodegradation even under potentially adverse conditions such as low temperatures, which is important for effective bioremediation in very cold climates (Larkin et al., 2005; Whyte et al., 1998). The metabolism of *Rhodococcus* has evolved to adapt to a wide range of nutritional conditions. This adaptation often involves the flexibility of the central metabolism, which usually provides energy and precursors for the biosynthesis processes, either during growth or during non-replicative metabolically active periods (Alvarez, 2010). Under stress conditions, the physiology of *rhodococci* seems to depend on the metabolism of diverse storage compounds. The members of the *Rhodococcus* genus possess an extensive capacity to synthesize and metabolize diverse storage compounds, such as triacylglycerols, wax esters, polyhydroxyalkanoates, glycogen, and polyphosphate (Alvarez et al. 1997; Alvarez, 2003; Hernandez et al., 2008).

Among the remarkable variety of organic compounds utilized by rhodococci as growth substrates there are the aromatic compounds and steroids. This degradation helps maintain the global carbon cycle and has increasing applications ranging from the biodegradation of pollutants to the biocatalytic production of drugs and hormones. The catabolism of aromatic compounds and steroids converge to aromatic intermediates; from the peripheral pathways common intermediates such as biphenyl and phthalate are obtained, while central pathways transform these intermediates in central metabolites, such as catechol and phenylacetate. Steroid degradation appears to be a very common, potentially

ubiquitous characteristic of rhodococci. It has been established that the catabolism of aromatic compounds is organized in a large number of peripheral aromatic pathways that converge into a restricted number of central aromatic pathways. The pathways of central metabolism are close to identical across widely divergent organisms, which share essentially the same metabolic network. However, this network possesses species-specific components, which depend on the biology of *rhodococci*. Various central aromatic pathways have been identified in rhodococci to date: *β*-ketoadipate (*pca-cat* genes), 2-hydroxypentadienoate (*hpd* genes), gentisate (*gen* genes), homogenisate (*hmg*), hydroxyquinol (*dxn* genes), homoprotocatechuate (*hpc* genes) (Yam et al., 2010).

The characteristic of *Rhodococcus* species are reflected by their large and complex genomes rich in multiple and often redundant genes.

1.4.1 *Rhodococcus* genus genomes

Consistent with the immense catabolic diversity of rhodococci, they often possess large genomes, which contain a multiplicity of catabolic genes, a high genetic redundancy of biosynthetic pathways and a sophisticated regulatory network. Many *Rhodococcus* genomes possess genes encoding multiple catabolic enzymes and pathways, often carried on plasmid ranging from small circular plasmids to large linear ones. The presence of these multiple catabolic pathways and gene homologues seems to be the basis of their catabolic versatility. Many of the genes associated with these pathways are dispersed around the genome, and the co-regulation of gene expression is a feature of how the rhodococci adapt to utilize many substrates. However, the majority of the molecular genetic mechanisms underlying the flexibility of the *Rhodococcus* genome are still to be elucidated.

In many cases, plasmids contain genes involved in the catabolism of solvents and aromatic compounds, such as trichloroethene, naphthalene, toluene, alkylbenzene, biphenyl and chloroaromatic compounds; likewise, genes associated with pathogenesis have been found in plasmids of pathogen strains. *Rhodococcus* genome flexibility is also due to a system that promotes high-frequency illegitimate recombination and the presence of transposons (Larkin, Kulakov and Allen, 2010). Events of illegitimate recombination, combined with homologous recombination, have been recorded and they may promote the introgression of DNA in their genomes without the help of mobile genetic elements. However, most mobile genetic elements in *Rhodococcus* spp. genomes have been identified. Moreover, regarding the gene expression and regulation, some studies on the regulation of *Rhodococcus* gene clusters have revealed many examples of both positive gene regulators (Komeda et al. 1996a) and repressors (Barnes et al. 1997; Nga et al. 2004). However, there

is evidence that the genes are not organized in a single cluster, and different strains have several homologous transcriptional units separated by non-homologous sequences containing direct and inverted repeats. Until now, around 182 *Rhodococcus* strains genomes are completely or partially sequenced.

The first sequenced genome belongs to *Rhodococcus jostii* RHA1; the 9.7 Mb genome with a G+C content of 67%, is distributed between a linear chromosome and three mega linear plasmids (pRHL1 (1100 kb), pRHL2 (450 kb) and pRHL3 (330 kb)) that act as storage for multiple copies of many degradative genes (McLeod et al. 2006). It is a polychlorinated biphenyl (PCB) degrading bacterium originally isolated on lindane as a substrate (Seto et al. 1995).

R. opacus B4 was isolated from petroleum-contaminated soil and is especially resistant to solvents (Honda et al. 2008; Grund et al. 1992; Na et al. 2005). As for most environmental rhodococci isolated, it also utilizes a wide variety of aromatic and aliphatic hydrocarbons. Its catabolic abilities are divided in a large linear chromosome (7.9 Mb) (G+C 67.9%), two linear plasmids pROB01 (558.192 kb) and pROB02 (244.9 kb) and three circular plasmids pKNR (111.2 kb), pKNR01 (4.4 kb) and pKNR02 (2.8 kb).

Another strain well known for its ability to degrade a wide variety of aromatic compounds is *Rhodococcus* sp. DK17. It is able to utilize various derivatives of benzene and bicyclics containing both aromatic and alicyclic moieties as sole carbon and energy source. The analysis of its genome (9.1 Mb) revealed that the G+C content is 67.15% and that it possesses three linear megaplasmids (380 kb pDK1, 330 kb pDK2, and 750 kb pDK3), which contain the gene cluster for alkylbenzene and phthalate degradation.

Rhodococcus opacus PD630 is well known for its ability to bioconvert a diverse range of organic substrates through lipid biosynthesis. It has the peculiarity to accumulate close to 80% of its cellular dry weight in oil. It has a large genome (9.27 Mb - 67.2% G+C), shared in a chromosome and 9 plasmids (170 kb; 100 kb; 80 kb; 60 kb; 100 kb; 50 kb; 40 kb; 160 kb; 40 kb), that contain many homologous genes dedicated to lipid metabolism. Analysis of the genome sequencing and metabolic-reconstruction assigned 261 genes, which are implicated in the *R. opacus* PD630 TAGs cycle.

In contrast to the terrestrial origins of the abovementioned strains, *Rhodococcus erythropolis* PR4 was isolated from the Pacific Ocean at a depth of about 1 km near Japan. This bacterium was found to utilize a wide variety of hydrocarbons and was solvent-resistant (Sekine et al. 2006; Peng et al. 2006). However, the large genome (6.5 Mb) consists of a circular chromosome (G+C 62.31%), a single linear plasmid (pREL1, 271.6 kb) and two circular plasmids (pREC1, 104 kb; pREC2, 3.6 kb). The chromosome and linear plasmid encode most of the genes associated with the ability of this bacterium to degrade a wide variety of alkanes.

The human pathogen *Rhodococcus equi* that causes bronchopneumonia in horses has a genome consisting of 5 Mb (68.82% G+C) and a virulence-associated plasmid of 80.6 kb that carries genes associated with the pathogenicity of the bacterium.
In particular, this work focused the attention on three *Rhodococcus* strains: *R. opacus* R7 (Di Gennaro et al., 2001), *R. aetherivorans* BCP1 (Frascari et al., 2006) and *R. erythropolis* MI2 (Wübbeler et al., 2010).

1.4.1.1 *Rhodococcus opacus* R7

Rhodococcus opacus R7 was isolated from a soil contaminated with polycyclic aromatic hydrocarbons for its ability to grow on naphthalene. The strain is also able to degrade a wide range of substrates, such as aliphatic, polycyclic and monoaromatic hydrocarbons; a peculiarity of this strain is its capability to degrade *o*-xylene, the xylene isomer most recalcitrant to microbial degradation. R7 genome sequencing was performed using a 454-sequencing technology (Roche GS FLX Titanium). It resulted in one shotgun library made of 312384 sequence reads and one paired-end library of 380920 sequence reads. All the reads were assembled into 223 contigs by using Newbler 2.6, with an N50 length of 184729 bp and an average genome coverage of 17X. The preliminary annotation was performed by using the RAST (Rapid Annotation using Subsystem Technology) server NCBI pipeline was used to annotate genome sequences and to make manual curation. A more complete metabolic reconstruction has been performed with Genomes KEGG (Di Gennaro et al., 2014). The contigs were ordered into 6 scaffolds, giving a total genome size of 10.1 Mb with a G+C content of 66.0%. A total of 9,602 putative open reading frames (ORFs), 62 RNAs genes, 9 rRNAs and 53 tRNAs, were predicted. A total of 745 ORFs are involved in metabolism of fatty acids, lipids, and isoprenoids, while 267 ORFs participate in metabolism of aromatic compounds. A total of 129 oxygenases/hydroxylases among the 134 annotated are predicted to initiate oxidation of organic compounds with industrial and environmental relevance, such as linear alkanes, cyclic ketones, aromatic compounds (e.g., benzoate, catechol, gentisate, salycilate, and biphenyl), aminopolycarboxilic acids, nitroalkanes, and phenylalkanoic acids. Forty-five ORFs encode cytochrome P450 monooxygenases that catalyze regio- and stereospecific oxidation of a vast number of substrates. R7 genome is arranged in a chromosome of 8.5 Mb and in 5 plasmids (pDG1 of 656.4 kb, pDG2 of 426.4 kb, pDG3 of 352.3 kb, pDG4 of 191.4 kb and pDG5 of 251.8 kb), giving a total genome size of 10.1 Mb.

1.4.1.2 *Rhodococcus aetherivorans* BCP1

R. aetherivorans BCP1 was isolated from an aerobic butane-utilizing consortium able to co-metabolize chloroform, vinyl chloride, and trichloroethylene. BCP1 was also able to catabolize a wide range of aliphatic, alicyclic, and carboxylated alkanes. (Frascari et al., 2006; Cappelletti et al., 2011). BCP1 cells are cocci that give rise to multiple branched hyphae that fragment into rods and cocci; its colonies are rough, with pink–orange pigment. It is a nonmotile strict aerobic bacterium.
BCP1 genome sequencing was performed using 454 sequencing technology (Roche GS FLX Titanium). The total numbers of the sequence reads were 668686 from one shotgun library and 353,744 from one paired-end library (8-kb inserts). All the reads were assembled into 123 contigs using Newbler 2.6, with an N50 length of 237787 bp and an average genome coverage of 65X. Based on paired-end directional information, the contigs were further ordered into 3 scaffolds, giving a total genome size of 6.2Mb with a G+C content of 70.4%; it is shared in one chromosome and two plasmids (pBMC1 and pBMC2 of 120.4 kb and 103.3 kb, respectively). The annotation was performed by using the Rapid Annotations using Subsystems Technology (RAST) server. A total of putative 5,902 open reading frames (ORFs), 8 rRNA genes, and 50 tRNA genes were predicted. Four hundred fifteen ORFs are involved in the metabolism of fatty acids, lipids, and isoprenoids, while 135 ORFs participate in the metabolism of aromatic compounds. Seventy-three oxygenases/hydroxylases are predicted to initiate the oxidation of organic compounds with industrial and environmental relevancy, such as linear alkanes, cyclic ketones, aromatic compounds (e.g., benzoate, catechol, gentisate, salicylate), aminopolycarboxylic acids, nitroalkanes, and phenylalkanoic acids. Twenty-six ORFs encode cytochrome P450 monooxygenases that catalyze regio- and stereospecific oxidation of a large number of substrates (Cappelletti et al., 2013).

1.4.1.3 *Rhodococcus erythropolis* MI2

R. erythropolis MI2 was isolated from an oil-contaminated absorber in a garage and later applied in polythioester (PTE) research, because this strain exhibits the rare ability to use the synthetic disulfide 4,4-dithiodibutyric acid (DTDB) as sole carbon source and electron donor for aerobic growth (Khairy et al., 2015; 2016). The genome sequencing was performed using Illumina GAIIx technology; all the reads were assembled into 123 contigs, with an N50 length of 217225 bp and an average genome coverage of 206.0x. Its genome consists of approximately 7.2 Mb with an overall G+C content of 62.25%. It is composed of three replicons: one chromosome of 6.45 Mb and two megaplasmids with sizes of 400 kb and 350 Mb, respectively.

1.4.2 *R. opacus* R7 and *R. aetherivorans* BCP1 comparative analysis

In Orro et al. 2015 a comparative analysis of *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7 is reported. The phylogeny of the two *Rhodococcus* strains has been constructed by a multi-locus sequence analysis (MLSA) maximum likelihood (ML) tree based on four marker genes (*16SrRNA*, *secY*, *rpoC* and *rpsA*) of reference strains of *Rhodococcus* genus. The results showed that *R. opacus* R7 correlates to *Rhodococcus opacus* and *Rhodococcus wratislaviensis* species in a clade that also includes *R. jostii* RHA1; while, *R. aetherivorans* BCP1 clusters with *R. aetherivorans* strains in a clade including also *R. ruber* species (**Figure 1.3**).

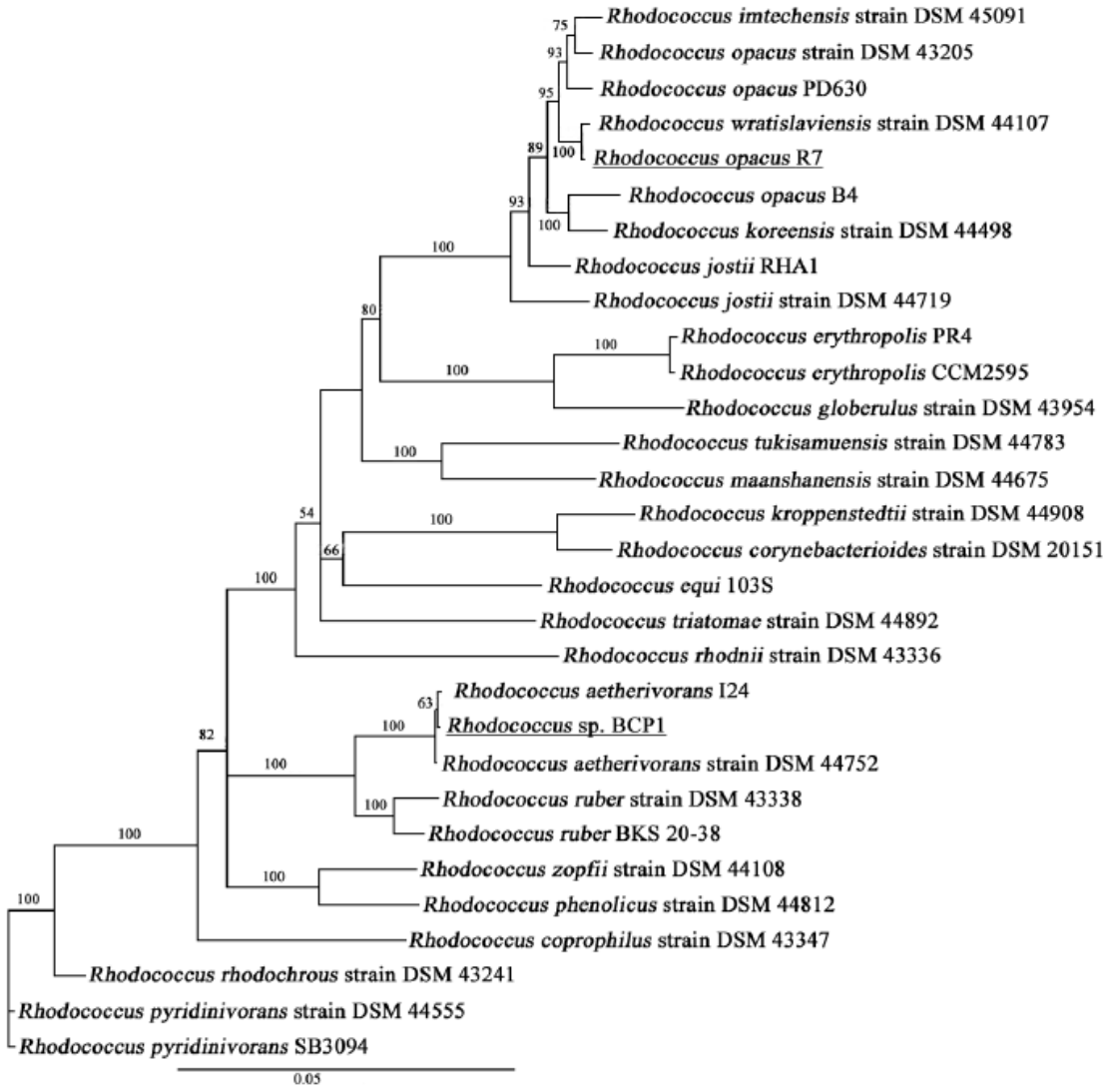


Figure 1.3. Phylogenetic Tree. Phylogenetic analysis of *R. opacus* R7 and *R.* sp. BCP1 (actually *R. aetherivorans* BCP1) based on sequence alignments with reference strains of *Rhodococcus* genus. The tree was constructed based on concatemer sequences of four marker genes of the 28 strains: *16S rRNA*, *secY*, *rpoC* and *rpsA* genes. (Orro et al., 2015)

In order to characterize their genomic diversity, the two strain whole genomes were compared to four other reference genomes (*R. jostii* RHA1, *R. opacus* PD630, *R. opacus* B4, *R. pyridinivorans* SB3094), using the Mauve program (**Figure 1.4**).

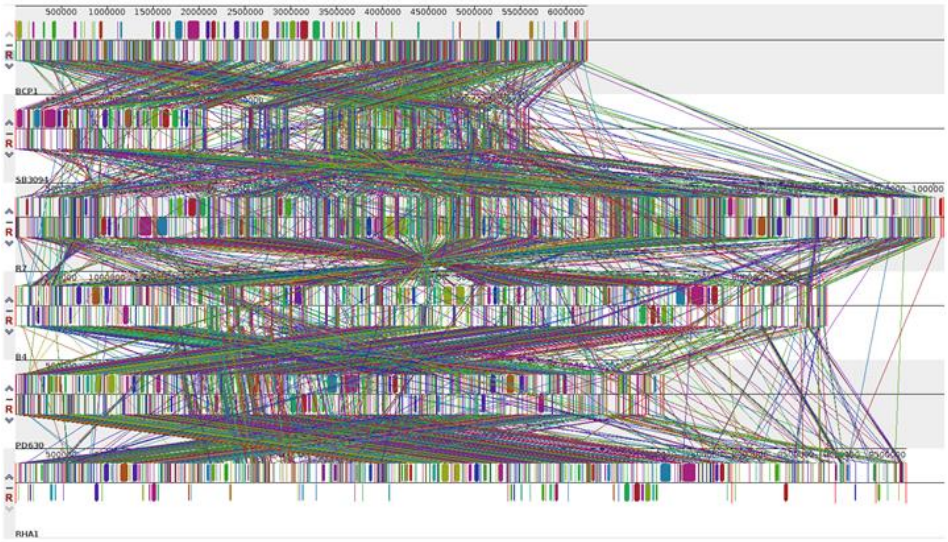


Figure 1.4. Whole genome sequence comparison of *R. opacus* R7 and *Rhodococcus* sp. BCP1 (actually *R. aetherivorans* BCP1) with a set of four other reference genomes: *R. jostii* RHA1, *R. opacus* PD630, *R. opacus* B4, *R. pyridinivorans* SB3094. For a global alignment of all six genomes the Mauve tool (2.3 Version) was used and the relative positions of the conserved regions found in more than one genome are presented in the same colored block. (Orro et al., 2015)

Results revealed that *Rhodococcus aetherivorans* BCP1 and *R. opacus* R7 share a total of 81% (BCP1 as reference) or 52% (R7 as reference) of similarity, calculated on all chromosomes and plasmids (**Table 1.2**). Moreover, the six strains share a range of similarity of 50-90%. Furthermore, every strain possesses a variety of unique regions containing characteristic genes encoding for enzymes involved in their peculiar metabolic capabilities.

Table 1.2. Similarity scores for *Rhodococcus* genomes under analysis. Values represent the percent of bases shared**.**

	R. sp. BCP1	*R. opacus* R7	*R. jostii* RHA1	*R. opacus* PD630	*R. opacus* B4	*R. pyridinivorans* SB3094
***R.* sp. BCP1**	-	81.27%	81.09%	65.43%	73.66%	65.07%
***R. opacus* R7**	51.67%	-	83.74%	71.39%	74.69%	44.67%
***R. jostii* RHA1**	50.90%	82.79%	-	68.61%	73.80%	44.48%
***R. opacus* PD630**	55.27%	91.54%	89.30%	-	80.25%	48.78%
***R. opacus* B4**	56.26%	89.91%	88.17%	73.15%	-	48.93%
***R. pyridinivorans* SB3094**	71.52%	74.36%	74.98%	60.36%	70.00%	-

These values indicate a high level of genetic variability amongst *Rhodococcus* genus strains and their genomic diversity in this group of bacteria was represented in the diagram of Venn shown in **Figure 1.5**.

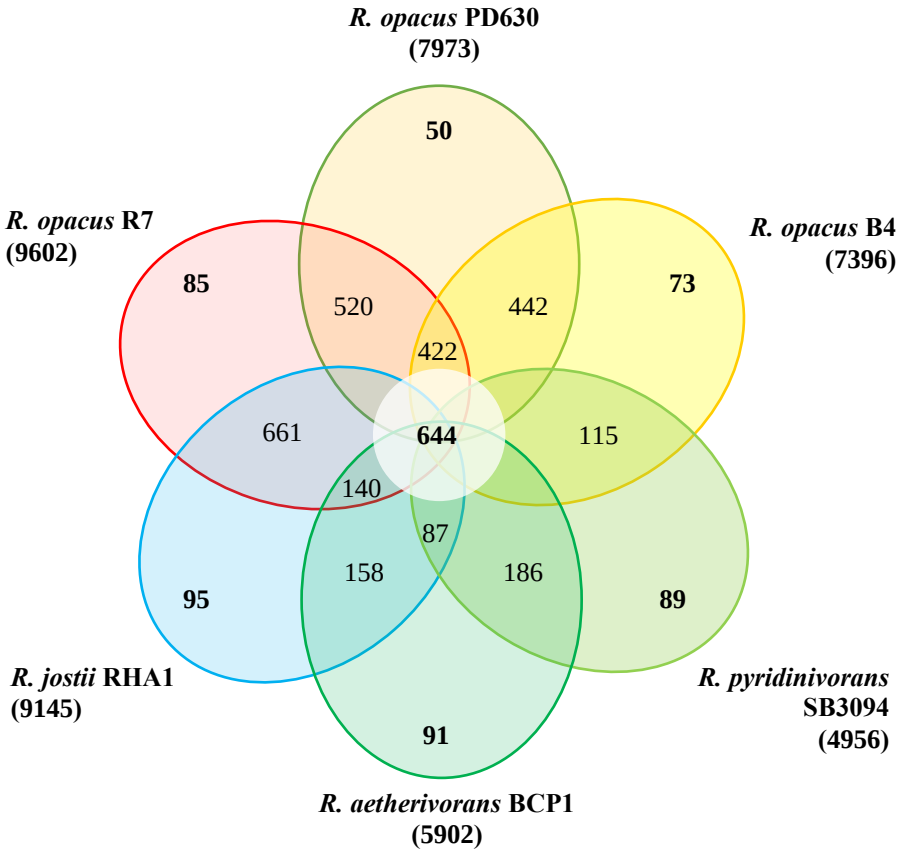


Figure 1.5. Genomic comparison of *R. opacus* R7 and *Rhodococcus* sp. BCP1 (actually *R. aetherivorans* BCP1) with other *Rhodococcus* strains, including *R. jostii* RHA1, *R. opacus* PD630, *R. opacus* B4, *R. pyridinivorans* SB4094. Each strain is represented by a colored oval. Number of predicted protein coding genes (CDSs) shared by all strains (i.e., the core genome) is in the centre. Overlapping regions show the number of CDSs conserved only within the specified genomes. Numbers in non-overlapping portions of each oval show the number of CDSs unique to each strain. The total number of protein coding genes within each genome is listed below the strain name. (Orro et al., 2015)

The core genome identified for the *Rhodococcus* strains under analysis contains 644 predicted protein-coding genes that represent around the 50% of the predicted proteome of each strain. The most shared regions identified in *R. aetherivorans* BCP1 and *R. opacus* R7 are located on the chromosomes; this result suggests that these regions could be characteristic of the *Rhodococcus* genus.

1.5 Scope of the thesis work

The aim of this project is the genomic characterization of bacteria belonging to *Rhodococcus* genus, in order to obtain molecular tools to use as "marker" sequences for the environmental quality evaluation of soils. The project was carried out by the identification of catabolic functions, gene clusters and metabolic pathways involved in the degradation of the most diffused environmental contaminants. In order to provide an example of the potentialities of this work, based on the identified catabolic functions, the assessment of contaminated soil in microcosms experiments was performed.

1.6 ***References***

Alexander M (1981). Biodegradation of chemicals of environmental concern. Science, 211:132–138.

Alvarez HM, Kalscheuer R and Steinbüchel A (1997). Accumulation of storage lipids in species of *Rhodococcus* and *Nocardia* and effect of inhibitors and polyethylene glycol. Fett/Lipid, 99:239–246.

Alvarez HM (2003). Relationship between b-oxidation pathway and the hydrocarbon-degrading profile in actinomycetes bacteria. Inter Biodeter Biodeg, 52:35–42.

Andrews SS, Karlen DL, Mitchell JP (2002). A comparison of soil quality indexing Agriculture, methods for vegetable production systems in Northern California. Agriculture, Ecosystems & Environment, 90:25–45.

Barnes MR, Duetz WA, Williams PA (1997). A 3-(3-hydroxyphenyl)propionic acid catabolic pathway in *Rhodococcus globerulus* PWD1: cloning and characterization of the *hpp* operon. J Bacteriol, 179:6145–6153.

Beare MH, Pohlad BR, Wright DH and Coleman DC (1993). Residue placement and fungicide effects on fungal communities in conventional and no-tillage soils. Soil Science Society of America Journal, 57:392–399.

Biology of *Rhodococcus* (2010) Volume 16, Editor: Héctor M. Alvarez

Bouchez T., Blieux AL, Dequiedt S, Domaizon I, Dufresne A, Ferreira S, Godon JJ, Hellal J, Joulian C, Quaiser A, Martin-Laurent F, Mauffret A, Monier JM, Peyret P, Schmitt-Koplin P (2016). Molecular microbiology methods for environmental diagnosis. Environmental Chemistry Letters, 14:423–441.

Bünemann EK, Bongiorno G, Bai Z, Creamer RE, De Deyn G, de Goede R, Fleskensd L, Geissend V, Kuyperb TW, Mädera P, Pullemanb M, Sukkelf W, van Groenigenb JW, Brussaard L (2018). Soil quality - A critical review. Soil Biology and Biochemistry, 120(January):105–125.

Burgess L (2013). Organic pollutants in soil. In Soils and Human Health (pp. 83–106). CRC Press.

Cappelletti M, Di Gennaro P, D'Ursi P, Orro A, Mezzelani A, Landini M, Fedi S, Frascari D, Presentato A, Zannoni D, Milanesi L, Di Gennaro P (2013). Genome Sequence of *Rhodococcus* sp. strain BCP1, a Biodegrader of Alkanes and Chlorinated Compounds. Genome Announcements, 1(6):e00657-13.

Cappelletti M, Fedi S, Frascari D, Ohtake H, Turner RJ and Zannoni D. (2011). Analyses of both the *alkB* gene transcriptional start site and *alkB* promoter-inducing properties of *Rhodococcus* sp. strain BCP1 grown on *n*-alkanes. Applied and Environmental Microbiology, 77:1619–1627.

Carter MR, Gregorich EG, Anderson DW, Doran JW, Janzen HH, Pierce FJ (1997). Concepts of soil quality and their significance. In Developments in Soil Science (pp. 1–19). Elsevier.

Daniel R (2004). The soil metagenome - a rich resource for the discovery of novel natural products. Current Opinion in Biotechnology, 15:199–204.

Davidson D (2000). Soil quality assessment: recent advances and controversies. Progress in Environmental Science, *2*, 342–350.

de Carvalho CCCR, da Cruz AA, Pons MN, Pinheiro HM, Cabral JM, da Fonseca MMR, Ferreira BS, Fernandes P (2004). *Mycobacterium* sp., *Rhodococcus erythropolis*, and *Pseudomonas* putida behaviour in the presence of organic solvents. Microsc Res Tech, 64:215–222.

de Carvalho CCCR and da Fonseca MMR (2005). Degradation of hydrocarbons and alcohols at different temperatures and salinities by *Rhodococcus erythropolis* DCL14. FEMS Microbiology Ecology, 51(3):389–399.

de Carvalho CCCR and da Fonseca MMR (2005). The remarkable *Rhodococcus erythropolis*. Appl Microbiol Biotechnol 67:715-726.

de Carvalho CCCR (2012) Adaptation of *Rhodococcus erythropolis* cells for growth and bioremediation under extreme conditions. Res Microbiol,163:125e36.

de Carvalho CCCR, Costa SS, Fernandes P, Couto I and Viveiros M (2014) Membrane transport systems and the biodegradation potential and pathogenicity of genus *Rhodococcus*. Front. Physiol., 5:133.

Di Gennaro P, Zampolli J, Presti I, Cappelletti M, D'Ursi P, Orro A, Mezzelani A, Milanesi L (2014) Genome sequence of *Rhodococcus opacus* strain R7, a biodegrader of mono- and polycyclic aromatic hydrocarbons. Genome Announc 2:e00827-14.

Di Gennaro P, Rescalli E, Galli E, Sello G and Bestetti G (2001). Characterization of *Rhodococcus opacus* R7, a strain able to degrade naphthalene and o-xylene isolated from a polycyclic aromatic hydrocarbon-contaminated soil. Research in Microbiology, 152(7):641–651.

Di Gennaro P, Terreni P, Masi G, Botti S, De Ferra F and Bestetti G (2010). Identification and characterization of genes involved in naphthalene degradation in *Rhodococcus opacus* R7. Applied Microbiology and Biotechnology, 87(1):297–308.

Doran JW and Parkin TB (1994). Defining and Assessing Soil Quality In: Doran JW, Coleman DC, Bezdicek DF, Stewart BA (Eds.). Soil Science Society of America Journal, 3–21.

Doran JW and Parkin TB (1996). Quantitative Indicators of Soil Quality: A Minimum Data Set, In: J. W. Doran and A. J. Jones, Eds., Methods for Assessing Soil Quality, Special Publication No. 49, Soil Science Society of America, Madison. 25–37.

FAO. http://www.fao.org/soils-portal/soil-biodiversity.

FAO. http://www.fao.org/home/en/.

Finnerty WR (1992) The biology and genetics of the genus *Rhodococcus*. Annu Rev Microbiol 46:193–218.

Frascari D, Pinelli D, Nocentini M, Fedi S, Pii Y and Zannoni, D. (2006). Chloroform degradation by butane-grown cells of *Rhodococcus aetherovorans* BCP1. Applied Microbiology and Biotechnology, 73(3):421–428.

Goodfellow M, Alderson G and Chun J (1998). Rhodococcal systematics: problems and developments. Antonie Van Leeuwenhoek. 74(1-3):3-20.

Grund E, Denecke B, Eichenlaub R (1992). Naphthalene degradation via salicylate and gentisate by *Rhodococcus* sp. strain B4. Appl. Environ. Microbiol., 58:1874–1877.

Hara H, Stewart GR and Mohn WW (2010). Involvement of a novel ABC transporter and monoalkyl phthalate ester hydrolase in phthalate ester catabolism by *Rhodococcus jostii* RHA1. Appl. Environ. Microbiol. 76:1516–1523.

Harayama S, Rekik M, Wubbolts M, Rose K, Leppik RA and Timmis KN (1989). Identification of five genes and their products in the upper pathway operon of TOL plasmid pWWO from *Pseudomonas putida*. J. Bacteriol., 171(9):5048–5055.

He Z, Gentry TJ, Schadt CW, Wu L, Liebich J, Chong SC, Huang Z, Wu W, Gu B, Jardine P, Criddle C and Zhou J (2007). GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J, 1:67–77.

Henriksen TM and Breland TA (1999). Nitrogen availability effects on carbon mineralization, fungal and bacterial growth, and enzyme activities during decomposition of wheat straw in soil. Soil Biology and Biochemistry, 31:1121–1134.

Honda K, Yamashita S, Nakagawa H, Sameshima Y, Omasa T, Kato J, Ohtake H (2008). Stabilization of water-in-oil emulsion by *Rhodococcus opacus* B-4 and its application to biotransformation. Appl Microbiol Biotechnol, 78:767–773.

Ivshina IB, Peshkur TA and Korobov VP (2002). Efficient uptake of cesium ions by *Rhodococcus* cells. Microbiology, 71:357–361.

Jones A and Goodfellow M (2010) Genus II. *Rhodococcus* (Zopf 1891) emend Goodfellow et al. 1998. In: Bergey's Manual of Systematic Bacteriology, vol 4, 2nd edn. Springer, Berlin. 1–65.

Kennedy AC and Stubbs TL (2006). Soil Microbial Communities as Indicators of Soil Health. Ann.Als of Arid Zone, 45(4):287–308.

Khairy H, Wubbeler JH, Steinbüchel A (2015). Biodegradation of the organic disulfide 4,4'-dithiodibutyric acid by *Rhodococcus* spp. Appl. Environ. Microbiol., 81:8294-8306.

Khairy H, Wübbeler JH and Steinbüchel A (2016). The NADH:flavin oxidoreductase Nox from *Rhodococcus erythropolis* MI2 is the key enzyme of 4,4′-dithiodibutyric acid degradation. Letters in Applied Microbiology, 63(6):434–441.

Kim IS, Foght JM and Gray RM (2002). Selective transport and accumulation of alkanes by *Rhodococcus erythropolis* S+14He. Biotechnol Bioeng, 80:650-659.

Komeda H, Hori Y, Kobayashi M, Shimizu S (1996). Transcriptional regulation of the *Rhodococcus rhodochrous* J1 *nitA* gene encoding a nitrilase. Proc Natl Acad Sci U S A, 93:10572-10577.

Jørgensen KS (2008). Advances in monitoring of catabolic genes during bioremediation. *Indian J Microbiol*, 48:152–155.

Larigauderie, Prieur-Richard, Mace GM, Lonsdale M, Mooney HA, Brussaard L, Cooper D, Cramer W, Daszak P, Díaz S, Duraiappah A, Elmqvist T, Faith DP, Jackson LE, Krug C, Leadley PW, Le Prestre P, Matsuda H, Palmer M, Perrings C, Pulleman M, Reyers B, Rosa EA, Scholes RJ, Spehn E, Turner II BL and Yahara T (2012). Biodiversity and ecosystem services science for a sustainable planet: the DIVERSITAS vision for 2012–20. Curr Opin Environ Sustain, 4(1):101–105.

Larkin MJ, Kulakov LA, Allen CCR (2005). Biodegradation and *Rhodococcus* - masters of catabolic versatility. Curr Opin Biotechnol. 16:282-290.

Larkin MJ, Kulakov LA, Allen CCR (2010). Genomes and plasmids in *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus.* In Alvarez HM, editor. *Biology of Rhodococcus*. Springer-Verlag Berlin Heidelberg. 73–90.

Lehman RM, Cambardella CA, Stott DE, Acosta-Martinez V, Manter DK, Buyer JS, Maul JE, Smith JL, Collins HP, Halvorson JJ, Kremer RJ, Lundgren JG, Ducey TF, Jin VL and Karlen DL (2015). Understanding and enhancing soil biological health: The solution for reversing soil degradation. Sustainability (Switzerland). 7:988-1027.

Lynch JM and Bragg E (1985). Microorganisms and soil aggregate stability. Advances in Soil Science, 2:133-171.

Martínková L, Uhnáková B, Pátek M, Nešvera J and Křen V (2009). Biodegradation potential of the genus *Rhodococcus*. Environment International, 35(1):162–177.

McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJ, Holt R, Brinkman FS, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD (2006). The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad SciUSA103:15582–15587

Moreira FMS, Siqueira JO (2006). Microbiologia e Bioquimica de Solo. *2a ed*. (U. E. Ufla, Ed.).

Na KS, Kuroda A, Takiguchi N, Ikeda T, Ohtake H, Kato J (2005). Isolation and characterization of benzene-tolerant *Rhodococcus opacus* strains. J Biosci Bioeng, 99:378–382.

Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G (2003). Microbial diversity and soil functions. Eur J Soil Sci, 54:655–670.

Nga DP, Altenbuchner J and Heiss GS (2004). NpdR, a repressor involved in 2,4,6-trinitrophenol degradation in *Rhodococcus opacus* HL PM-1. J Bacteriol, *186*, 98–103.

Orro A, Cappelletti M, D'Ursi P, Milanesi L, Di Canito A, Zampolli J, Collina E, Decorosi F, Viti C, Fedi S, Presentato A, Zannoni D, Di Gennaro P (2015) Genome and phenotype microarray analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: genetic determinants and metabolic abilities with environmental relevance. PLoS ONE. 10(10):e0139467.

OTA (Office of Technology Assessment Archive) (1987). https://ota.fas.org/otareports/year/1987a/

Peng X, Taki H, Komukai S, Sekine M, Kanoh K, Kasai H, Choi SK, Omata S, Tanikawa S and Harayama S, Misawa N (2006). Characterization of four *Rhodococcus* alcohol dehydrogenase genes responsible for the oxidation of aromatic alcohols. Appl Microbiol Biotechnol, 71:824–832.

Philippot L, Spor A, Hénault C, Bru D, Bizouard F, Jones CM, Sarr A, Maron PA (2013). Loss in microbial diversity affects nitrogen cycling in soil. ISME Journal, 7(8):1609–1619.

Pulleman M, Creamer R, Hamer U, Helder J, Pelosi C, Pérès G and Rutgers M (2012). Soil biodiversity, biological indicators and soil ecosystem services-an overview of European approaches. Current Opinion in Environmental Sustainability. 4(5):529–538.

Rhee S-K, Liu X, Wu L, Chong SC, Wan X, and Zhou J (2004). Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays. Appl. Environ. Microbiol., 70:4303–4317.

Rolf D (2004). The soil metagenome – a rich resource for the discovery of novel natural products. Current Opinion in Biotechnology,15:199–204.

Sekine M, Tanikawa S, Omata S, Saito M, Fujisawa T, Tsukatani N, Tajima T, Sekigawa T, Kosugi H, Matsuo Y, Nishiko R, Imamura K, Ito M, Narita H, Tago S, Fujita N, Harayama S (2006). Sequence analysis of three plasmids harboured in *Rhodococcus erythropolis* strain PR4. Environ Microbiol, 8:334–346.

Seto M, Masai E, Ida M, Hatta T, Kimbara K, Fukuda M, et al. (1995). Multiple polychlorinated biphenyl transformation systems in the Gram-positive bacterium *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol, 61:4510–4513.

Somerville L, Greaves PM. (1987) Efeitos de Pesticidas Sobre a Microflora do Solo. New York, NY: Taylor e Francis.

Tótola MR and Chaer GM (2002) Microrganismos e processos microbiológicos como indicadores da qualidade do solo. In: Alvarez VH, Schaefer CEGR, Barros NF et al. (eds). Tópicos em Ciência do Solo.Viçosa, MG: Sociedade Brasileira de Ciência do Solo. 195-276.

van der Geize R, Dijkhuizen L (2004). Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. Curr Opin Microbiol, 7:255-261.

Vidali M. (2001). Bioremediation - An overview. Pure and Applied Chemistry, 73(7):1163-1172.

Wagner-Döbler I, Bennasar A, Vancanneyt M, Störmpl C, Brümmer I, Eichner C, Grammel I, and Moore ERB (1998). Microcosm enrichment of biphenyl-degrading microbial communities from soils and sediments. Appl Environ Microbiol, 64:3014–3022.

Wang C, Liu D and Bai E (2018). Decreasing soil microbial diversity is associated with decreasing microbial biomass under nitrogen addition. Soil Biology and Biochemistry, 120(March):126-133.

Wardle DA, Zackrisson O, Hörnberg G, Gallet C (1997). The influence of island area on ecosystem properties. *Science*, 277:1296–1299.

Warhurst AW and Fewson CA (1994). Biotransformations catalyzed by the genus *Rhodococcus*. Critical Reviews in Biochemistry, 14:29-73.

Whyte LG, Hawari J, Zhou E, Bourbonnière L, Inniss WE, Greer CW (1998). Biodegradation of variable-chain-length alkanes at low temperatures by a psychrotrophic *Rhodococcus* sp. Appl Environ Microbiol, 64:2578-2584.

Wübbeler JH, Bruland N, Wozniczka M and Steinbüchel A (2010). Biodegradation of the xenobiotic organic disulphide 4,4′-dithiodibutyric acid by *Rhodococcus erythropolis* strain Ml2 and comparison with the microbial utilization of 3,3′-dithiodipropionic acid and 3,3′-thiodipropionic acid. Microbiology, 156(4):1221-1223.

Yachi S and Loreau M (1999). Biodiversity and ecosystem productivity in a fluctuating environment. The insurance hypothesis. Proceedings of the National Academy of Science of the United States of America, 96:1463-1468.

Yam KC, van der Geize R and Eltis LD (2010). Catabolism of Aromatic Compounds and Steroids by *Rhodococcus*. In Alvarez HM, editor. Biology of *Rhodococcus*. Springer-Verlag Berlin Heidelberg. 133-169.

Zampolli J, Collina E, Lasagni M and Di Gennaro P (2014). Biodegradation of variable-chain-length *n*-alkanes in *Rhodococcus opacus* R7 and the involvement of an alkane hydroxylase system in the metabolism. AMB Express, 4(1):1-9.

Zilli JE, Rumjanek NG, Xavier GR, Coutinho HL da C and Neves MCP (2003). Diversidade microbiana como indicador de qualidade do solo. Cadernos de Ciencia & Tecnología, 20(3):391-411.

Chapter 2.

Biodegradative potential of *R. opacus* R7 and *R. aetherivorans* BCP1: genome-based analysis for the identification of pathways and gene clusters

2.1 Introduction

Members of *Rhodococcus* genus show a remarkable ability to degrade and transform via several catabolic pathways a wide variety of natural organic and xenobiotic compounds, including aliphatic, aromatic, heterocyclic, and polycyclic aromatic compounds, alicyclic hydrocarbons, nitriles, cholesterol and lignins (Kim et al., 2018).

These metabolic capabilities, their tolerance to the presence of toxic substrates and solvents, in addition to a frequent mechanism of lack of catabolite repression, the production of biosurfactants and environmental persistence make them excellent candidates for bioremediation and bioconversion. The biodegradative pathways in rhodococci generally consist of many peripheral (upper) pathways and a few central (lower) pathways.

Rhodococci often harbor large catabolic linear plasmids, many of which increase their catabolic versatility and efficiency contributing to multiple pathways and gene homologous. Plasmids encode multiple catabolic enzyme systems; several degradative pathways are involved in the degradation of various organic compounds. Comparative analyses of rhodococcal genomes have suggested that the linear plasmids are possible determinants of the propagation of diverse degradative genes (Warren et al., 2004).

In addition to enzyme and pathway multiplicities, different induction mechanisms, substrate specificities, and activities of the catalytic enzymes are thought to contribute to their extraordinary capabilities to metabolize a wide range of compounds (Kim et al., 2018).

A previous work (Orro et al., 2015), showed the metabolic potential of *R. opacus* R7 and *R. aetherivorans* BCP1 to utilize several organic/xenobiotic compounds using a Phenotype Microarray (PM) approach. 41 organic/xenobiotic compounds (**Table 2.1**) belonging to four categories (aliphatic hydrocarbons and cycloalkanes, BTEX and aromatic compounds, polycyclic aromatic compounds (PAHs), naphthenic acids and other carboxylic acids), were tested as the only carbon and energy source.

In this chapter, the genome-based analysis performed to identify the gene clusters putatively involved in the degradation of the tested organic/xenobiotic compounds is reported.

Table 2.1. 41 organic/xenobiotic compounds tested on *R. opacus* R7 and *R. aetherivorans* BCP1

Category	Compounds
Aliphatic hydrocarbons and cycloalkanes	*Hexane; Heptane; Nonane; Decane; Dodecane; Tridecane; Tetradecane; Hexadecane; Heptadecane; Eicosane; Hentriacontane; Tetracosane; Octacosane; Cyclohexane; Cyclohexanone*
BTEX and aromatic compounds	*Benzene; Toluene; Ethylbenzene; m-Xylene; p-xylene; o-xylene; Dibenzothiophene; Fuel Oil*
Polycyclic aromatic compounds (PAHs)	*2-hydroxybiphenyl; Anthracene; Naphthalene; Phenanthrene*
Naphthenic acids and other carboxylic acids	*1,4-cyclohexane dicarboxylic acid; 1-adamantanecarboxylic acid; Cyclohexane butyric acid; Cyclohexaneacetic acid; Cyclopentanecarboxyl acid; Decanoic acid; Hexanoic acid; trans-1,2-cyclohexane dicarboxylic acid; Cyclohexane-1-carboxylate; 4-Phenylbutyric acid; 5,6,7,8-tetrahydro-2-naphthoic acid; Gentisic acid; Salicylic acid*

2.2 Material and methods

2.2.1 Identification of genetic aspects for xenobiotic degradation pathways

The whole-genome shotgun sequencing projects are deposited at DDBJ/EMBL/GenBank under accession numbers CM002177, CM002178, CM002179 for *R. aetherivorans* BCP1 and CP008947, CP008948, CP008949, CP008950, CP008951, CP008952 for *R. opacus* R7. The gene clusters involved in the degradation of some xenobiotic compounds were predicted in *R. opacus* R7 and *R. aetherivorans* BCP1 strains by manual alignment analyses of gene clusters previously identified or extracted from sequences of reference strains reported in databases. The identified gene clusters were described and compared with those from other *Rhodococcus* strains reported in literature and compared with the reference strain *R. jostii* RHA1. The NCBI pipeline (Altschul et al., 1990) and RAST (Aziz et al., 2008) server were used for genome predictions and comparisons. Moreover, Clustal Omega (Thompson et al. 1994), KEGG, (Kanehisa et al., 2017), UniProt (The UniProt Consortium, 2017) and Pfam (Bateman et al., 2004) were used for manual curation.

2.3 Results

2.3.1 Genetic aspects related to the great potential for organic/xenobiotic degradation in *R. opacus* R7 and *R. aetherivorans* BCP1

In order to characterize the great potential for organic/xenobiotic degradation of *R. opacus* R7 and *R. aetherivorans* BCP1 a genome-based approach was used to investigate the genetic aspects involved in this kind of metabolism. Four chemical categories of the 41 previously tested xenobiotic compounds were considered: 1) aliphatic hydrocarbons and cycloalkanes; 2) BTEX and aromatic compounds; 3) polycyclic aromatic compounds (PAHs); 4) naphthenic acids and other carboxylic acids. The gene-associated functions of the catabolic clusters were predicted on the basis of the NCBI genomic database, the Rapid Annotation using Subsystem Technology (RAST server), the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database and manual curation.

2.3.2 Aliphatic hydrocarbons and cycloalkanes degradation

The presence of catabolic genes involved in the growth on short-, medium- and long-chain *n*-alkanes was previously investigated in both the strains (Zampolli et al., 2014; Cappelletti et al., 2011; Cappelletti et al., 2015). Moreover, comparing a genome analysis of the strains with these previous studies, the presence of only one copy of the *alkB* gene was confirmed in *R. opacus* R7, while in *R. aetherivorans* BCP1, two copies of this gene (*alkB1* and *alkB2*) were found in the chromosome (**Figure 2.1**). One *alkB* gene in each strain was organized in a cluster associated in an operon with four consecutive coding sequences (CDSs): *alkB* coding for an alkane monooxygenase, *rubA* and *rubB* coding for two rubredoxins, and *rubred* coding for a rubredoxin reductase and the *tetR* gene coding for a regulator. Instead, the additional 1161-bp *alkB* gene (*alkB2*) in BCP1 was not organized in operon with *rubA/B* and *rubred* genes but was associated with genes coding for a putative esterase and a long-chain-fatty-acid-CoA ligase, involved in the fatty acid *β*-oxidation. The *alkB2* gene product showed a limited similarity (57% amino acid identity) with *alkB1* gene product of BCP1. The gene products of *alkB1* cluster of BCP1 showed high similarity with those coded by *alkB* cluster of R7 and with the ones of the most known bacteria belonging to the *Rhodococcus* genus, *R. jostii* RHA1.

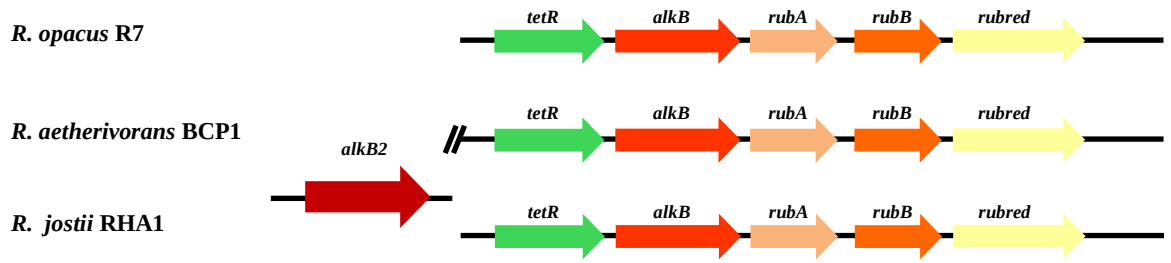


Figure 2.1. Comparative organization of *alk* gene cluster in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

Moreover, two *prm* gene clusters were identified in BCP1. The first, constituted by the *prmACBD* genes, was found in the chromosome and it has been proven to be specially involved in the degradation of propane and butane; the second cluster, also named *smo* gene cluster (soluble di-iron monooxygenase) (*smoABDC* genes), was found in the pBMC2 plasmid and it was described to be involved in the growth of BCP1 on a wide range of short-chain *n*-alkanes (Cappelletti et al., 2015). The comparison of the BCP1 *prm* gene cluster with R7 genome sequences allowed the identification of a *prm*-like gene cluster in R7 chromosome with a percentage of amino acid identities with the corresponding genic products of 90% (**Figure 2.2**). R7 and BCP1 strains showed also a considerable number of genes coding for P450 monooxygenases, 23 and 18 predicted coding sequences, respectively, that could be involved in the aliphatic hydrocarbon degradation.

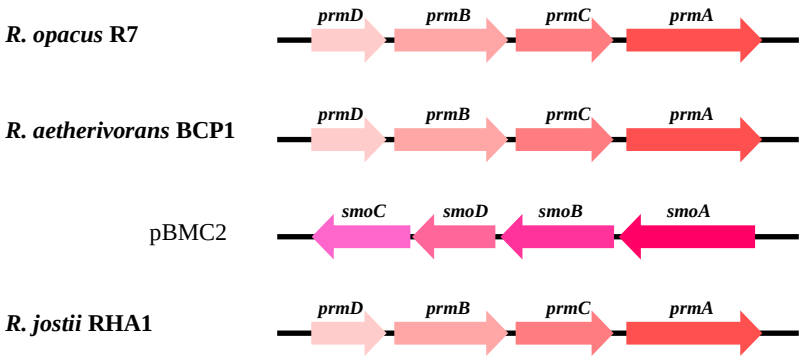


Figure 2.2. Comparative organization of *prm* and *smo* gene clusters in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

2.3.3 Aromatic hydrocarbons degradation

R7 and BCP1 strains are able to grow on different aromatic hydrocarbons, including BTEX compounds. The majority of the *Rhodococcus* genomes sequenced were characterized for the ability to grow on BTEX and in some cases also for their genetic organization (McLeod et al., 2006). In *R. jostii* RHA1, the *akb* gene cluster was identified as the cluster involved in the metabolism of alkylbenzenes. It contains two genes (*akbA1a* and *akbA1b*) encoding for an oxygenase component

large subunit, each one followed by genes (*akbA2a* and *akbA2b*) encoding for a small subunit oxygenase component. Moreover, the *akb* cluster consists of a reductase (AkbA4), a ferredoxin (AkbA3) component, an *akbB* gene and *akbCDEF* genes coding for the lower pathway enzymes of the ring cleavage. The same gene cluster is also found in R7 genome (exhibiting high amino acid identities) and it is allocated on two different plasmids: *akbA1a*, *akbA2a*, *akbA3*, *akbA4*, *akbB* genes were found on the pPDG5 plasmid, while *akbCDEF* genes, putatively coding for a complete meta-cleavage pathway, were identified on the pPDG2. These second part of the cluster is constituted by a meta-cleavage dioxygenase (AkbC), a metacleavage hydrolase product (AkbD), a hydratase (AkbE) and an aldolase (AkbF). A homologous sequence of *akb* cluster was also found on the chromosome and on the pPDG4 plasmid, but they showed a protein identity around 35% (**Figure 2.3**). Genome analysis of BCP1 revealed the presence of only one homologous cluster to RHA1 *akb* genes, that is putatively involved in the degradation of aromatic hydrocarbons; these sequences were found in pBMC2 plasmid with an amino acid identity of 36% for AkbA1 and AkbA2, while for AkbB of 47% (**Figure 2.3**). These data showed in R7 strain a redundancy of genes involved in the metabolism of the aromatic compounds probably due to its isolation from a contaminated polycyclic aromatic hydrocarbons soil. The presence of multiple copies in several plasmids can derive from transposition events and duplication of the same genes. These data are confirmed by the presence on the pPDG5 plasmid of five mobile elements and transposases upstream the *akb* genes.

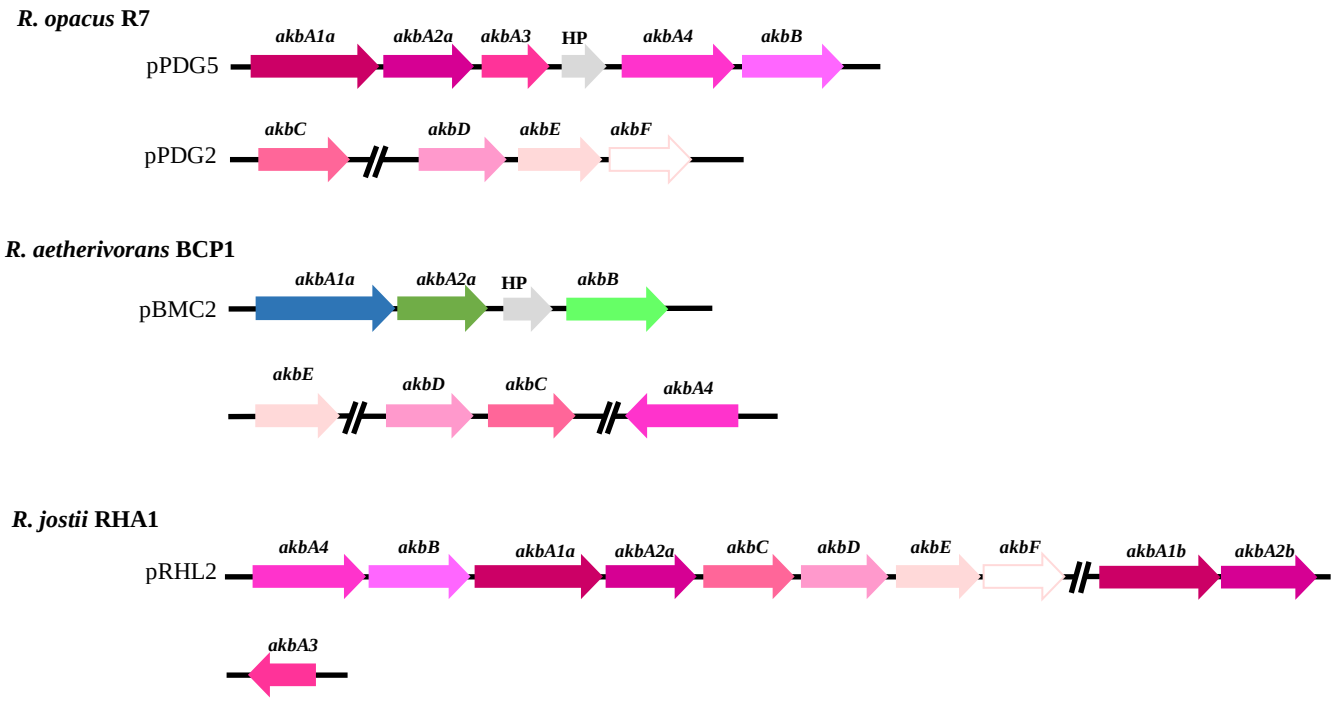


Figure 2.3. Comparative organization of *akb* gene cluster in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

In presence of dibenzothiophene (DBT) R7 strain showed high level of growth and the presence of putative genes involved in this metabolism was investigated. Some other *Rhodococcus* strains, including RHA1, are able to utilize dibenzothiophene as a sole source of sulphur due to the expression of *dsz* operon, which encodes three proteins, DszA, B and C. DszC catalyzes the stepwise S-oxidation of DBT, first to dibenzothiophene 5-oxide (DBTO) and then to dibenzothiophene 5,5-dioxide (DBTO2); DszA catalyzes the conversion of DBTO to 2-(2'-hydroxyphenyl) benzene sulphinate (HBPSi-) and DszB catalyzes the desulphation of HBPSi- to give HBP and sulphonate (Oldfield et al., 1997). A *dsz* gene cluster was found in R7 and BCP1 genomes. Comparing the protein sequence of DszA, two different oxygenases (*dszA1* and *dszA2*) were identified in R7 chromosome with amino acid identity of 98% and 97%. Sequence analysis revealed also two *dszC* genes (*dszC1* and *dszC2*) predicted to code for dibenzothiophene desulphurization enzymes, and not so far, only one copy of the *dszB* gene encoding an ABC sulfonate transporter protein was found. In BCP1 chromosome four coding sequences were identified as homologous to the *dsz* sequences of R7 genome. As in the R7 strain, two *dszC* genes (*dszC1* and *dszC2*) were found located close to each other; while only one *dszA* gene and one *dszB* gene copy were found clustered together on the chromosome. These preliminary indications suggest that R7 and BCP1 strains could have genes involved in the DBT degradation (**Figure 2.4**).

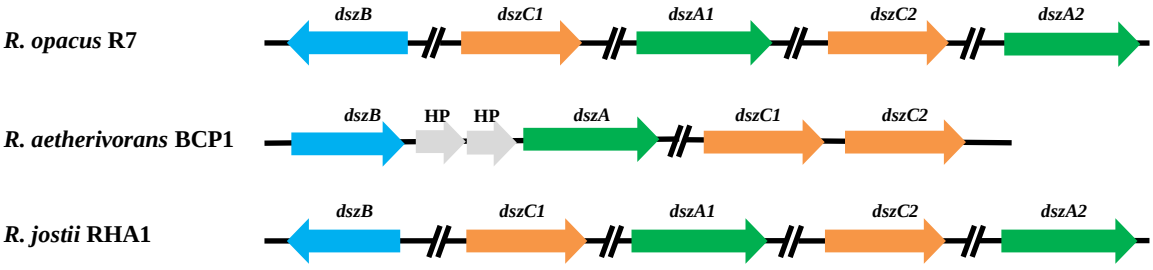


Figure 2.4. Comparative organization of *dsz* gene cluster in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

2.3.4 Polycyclic aromatic hydrocarbons degradation

In a previous work, genes involved in naphthalene (*nar* gene cluster) and salicylate (*gen* gene cluster) degradation were found in *R. opacus* R7 (Di Gennaro et al., 2010). The R7 whole genome sequence analysis exhibits that *nar* gene cluster is located in pPDG4 plasmid and that a similar cluster was found in pBMC2 in *R. aetherivorans* BCP1; this genomic region identified in pBMC2 is the same identified as *akb* gene cluster in BCP1. A homologous *nar* gene cluster was not found in the reference strain *R. jostii* RHA1. Moreover, two copies of the gene clusters involved in naphthalene lower pathway were found in R7: one in the pPDG4 plasmid distant 12.4 kb from *nar* gene cluster; the other

in the pPDG1 plasmid, but lacking *genL* gene. Comparison of this cluster with genome sequence of BCP1 revealed that some genes involved in gentisate oxidation were found: *genH* and *genI* genes were found in the chromosome (encoded proteins have an identity of around 80%), *genL* gene of which encoded protein has an identity of 48% (**Figure 2.5**). Instead, genes involved in salicylate degradation, *genABC* were found in different regions of the BCP1 chromosome showing a lower amino acid identity; these results confirm the BCP1 low growth level on salicylic acid.

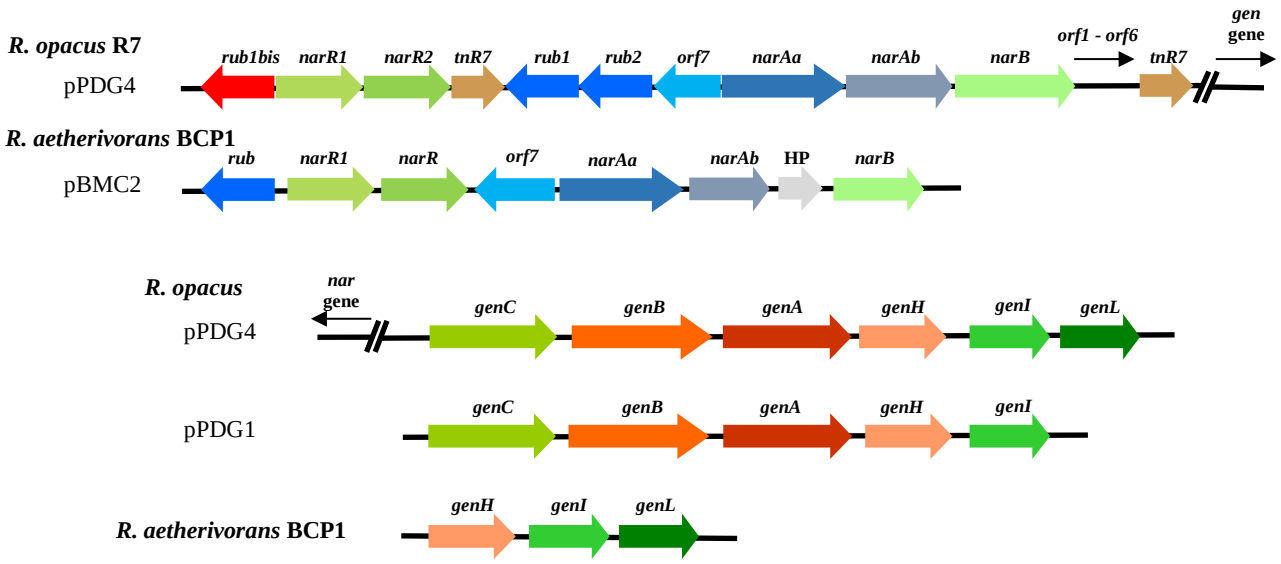


Figure 2.5. Comparative organization of *nar* and *gen* gene clusters in *R. opacus* R7 and *R. aetherivorans* BCP1.

A well-known microorganism characterized for its ability to degrade biphenyls is *R. jostii* RHA1. Genes involved in this pathway were identified in two different plasmids: pRHL1 and pRHL2 (Masai et al, 1997; Patrauchan et al., 2008). In the biphenyl metabolic pathway, it is transformed to 2,3-dihydroxy-1-phenylcyclohexa-4,6-diene (dihydrodiol) by a multicomponent biphenyl dioxygenase (BphA). Dihydrodiol is converted to 2,3-dihydroxybiphenyl (2,3DHBP) by dihydrodiol dehydrogenase (BphB). 2,3DHBP is cleaved at the 1,2 position (meta-ring cleavage) by 2,3DHBP dioxygenase (BphC). The ring cleavage product (2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate [HPDA]) is hydrolyzed to benzoate and 2-hydroxypenta-2,4-dienoate by HPDA hydrolase (BphD), and the resulting 2-hydroxypenta-2,4-dienoate is further converted to tricarboxylic acid cycle intermediates by 2-hydroxypenta-2,4-dienoate hydratase, 4-hydroxy-2-oxovalerate aldolase, and acetaldehyde dehydrogenase (BphE, BphF, and BphG, respectively). Therefore, the products of a set of catabolic genes, *bphAa,Ab,Ac,Ad,B,C,D,E,F,G* are responsible for the aerobic metabolism of biphenyl. The *bph* gene cluster of RHA1 was compared with the genome sequences of R7 and BCP1. Different genes of this cluster in R7 chromosome and in pPDG2 and pPDG5 plasmids were found. In particular, R7 showed genes encoding biphenyl large and small subunits and dihydrobiphenyldiol

dehydrogenase, involved in the first two steps of dioxygenation of biphenyl, on the pPDG5 plasmid; whereas genes encoding the ring-cleavage were found on the pPDG2 plasmid. Some homologous sequences of *bph* cluster were also found in R7 chromosome (**Figure 2.6**). Moreover, in BCP1 genome were identified CDSs homologous to *bph* genes in pBMC2. A large and a small subunit of a dioxygenase (*bphAa* and *bphAb* genes) and a dihydrobiphenyldiol dehydrogenase (*bphB* gene), coinciding with the region identified as *akb* gene cluster and *nar* gene cluster were found. The amino acid identity of the corresponding proteins is around 40%. Also, in BCP1, the genes encoding the ring-cleavage enzymes were not found near the cluster but in the chromosome with low protein identity (around 30%) (**Figure 2.6**). Results allowed to hypothesize that the same BCP1 cluster was involved in the first oxidation steps of several aromatic and polycyclic aromatic hydrocarbons, such as BTEX compounds, naphthalene and biphenyls.

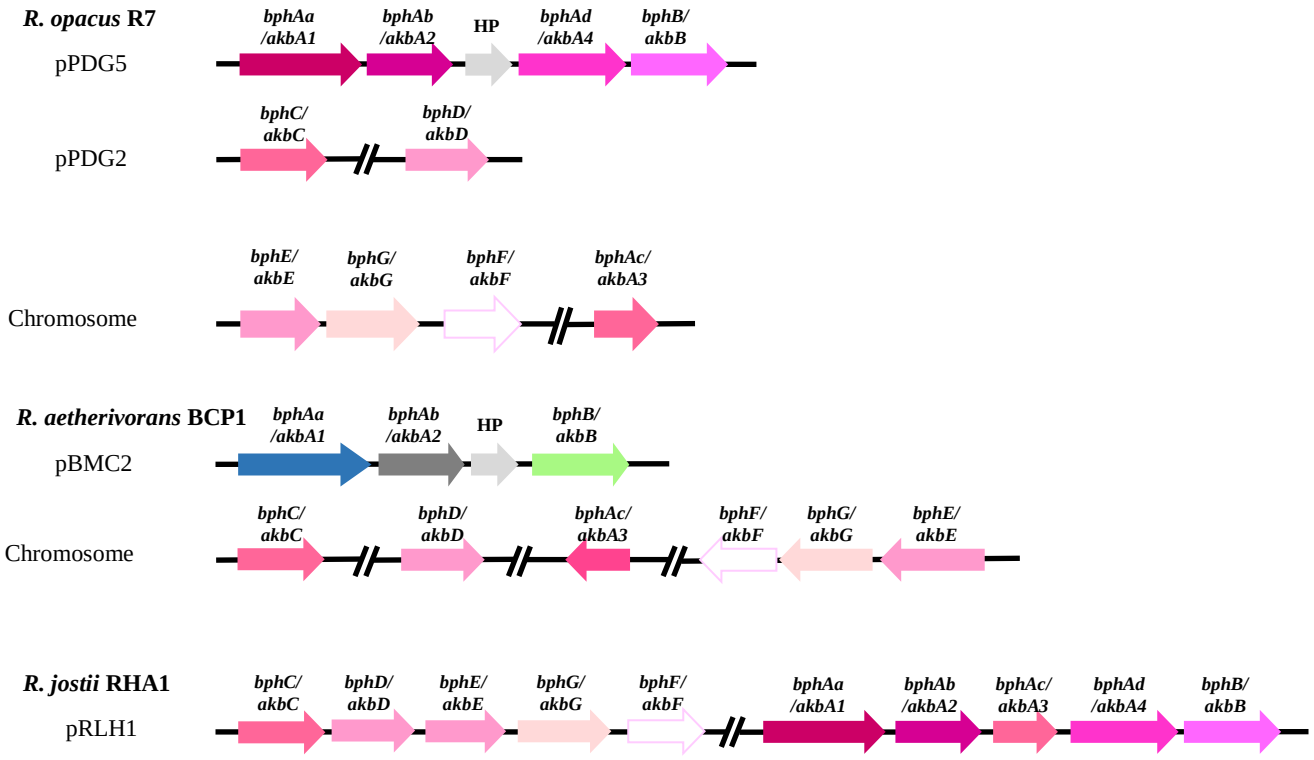


Figure 2.6. Comparative organization of *bph* gene cluster in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

2.3.5 Naphthenic acids degradation

Considering the ability of the R7 and BCP1 to grow on these contaminants, the putative gene clusters involved in this degradation were investigated. In literature, few metabolic studies are available on the biodegradation of naphthenic acids (NAs). Although no genetic information was provided in these

studies, the metabolism of CHCA was described to follow two main routes: (i) aromatization of the cycloalkane ring to produce hydroxybenzoate before ring opening (Iwaki et al., 2005), or (ii) activation of cycloalkane ring as CoA thioester-derivative that is further degraded through β-oxidation steps (Blakley et al., 1982). Iwaki and co-workers identified the *pobA* gene to have an essential role for the growth on CHCA by *Corynebacterium cyclohexanicum*. This gene codes for 4-hydroxybenzoate (4-HBA) 3-hydroxylase generally described to be responsible for the conversion of 4-hydroxybenzoate to protocatechuate, a common intermediate in the degradation of various aromatic compounds. One gene homologous to *pobA* was identified in both BCP1 and R7 genomes and its product was annotated as *p*-hydroxybenzoate hydroxylase in RAST server. Similarly, to what found in *Corynebacterium* strain, *pobA* gene was flanked by a gene coding for an IclR-type transcriptional regulator. The genomic region including *pobA* in R7 contained within also the *alkB* gene cluster, TetR-like regulator and a BenK transporter and the same organization was maintained between RHA1 and R7. On the contrary, the genomic organization of BCP1 region with *pobA* includes a permease coding gene and several oxidoreductases. The activation of the cycloalkane ring as Co-A thioester-derivative was firstly described during the anaerobic degradation of benzoic acid by *Rhodopseudomonas palustris*. CHCA is produced as metabolic intermediate during this pathway and it has been reported to be metabolized through *β*-oxidation. The enzymes responsible for CHCA degradation in *R. palustris* are encoded by the *bad* genes (Egland et al., 1997). These enzymes degrade CHCA by catalyzing the following reactions: (i) activation of cyclohexanecarboxylate as cyclohexanecarboxylate-CoA by a CoAligase (AliA); (ii) cyclohexanecarboxylate-CoA is dehydrogenated to the corresponding aldehyde cyclohex-1-ene-1-carboxylate by the dehydrogenase AliB; (iii) the hydratase BadK converts the aldehyde in the secondary alcohol 2-hydroxycyclohexane-1-carboxyl-CoA; (iv) the dehydrogenase BadH is responsible for the formation of 2-ketocyclohexane-1-carboxyl-CoA from the secondary alcohol; (v) the hydratase BadI catalyzes the cyclohexane ring opening with the formation of pimelyl-CoA. Based on the amino acid identity percentage, two genes (*badH1* and *badH2*) were found in BCP1 genome encoding for two enzymes annotated as 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase. Only one gene homologous to *bad*H was found in R7 and, compared to BCP1, it possesses conserved flanking regions including: a long-chain-fatty-acid-CoA-ligase, two dehydrogenases and a naphthoate synthase. This region is also maintained in RHA1 (**Figure 2.7**).

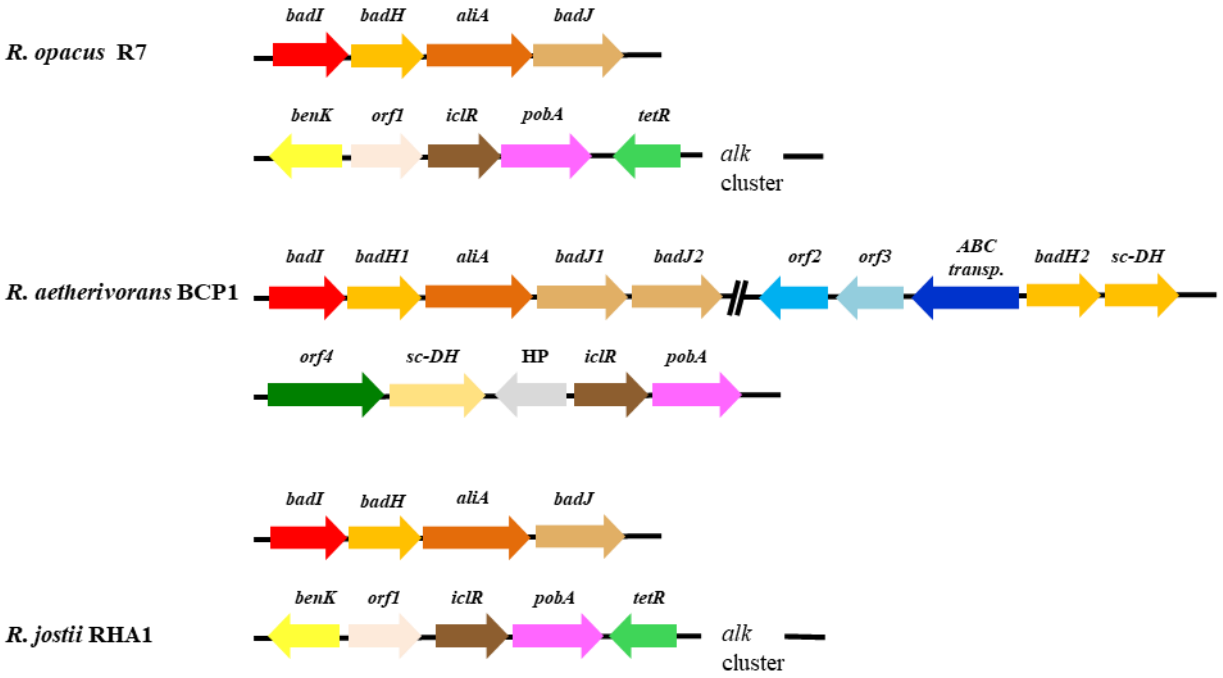


Figure 2.7. Comparative organization of *ali*, *bad* and *pob* gene clusters in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

2.3.6 Genetic aspects related to aromatic peripheral pathways in *R. opacus* R7 and *R. aetherivorans* BCP1

Considering the aromatic compounds that R7 and BCP1 can metabolize, four different peripheral pathways for the catabolism of several xenobiotics can be predicted, which include catechol (*cat* genes), protocatechuate (*pca* genes), phenylacetate (*paa* genes) and homogentisate (*hmg* genes) pathways. R7 genome contains several genes potentially involved in catechol catabolism. It shows six catechol 1,2-dioxygenases (five on the chromosome and one on pPDG2 plasmid), and three catechol 2,3-dioxygenases (one on the chromosome, one on pPDG2 and one on pPDG5 plasmid). BCP1 genome presents only two catechol 1,2-dioxygenases and one catechol 2,3-dioxygenase on the chromosome. Two catechol dioxygenase genes, amongst those identified on R7 chromosome, were organized in cluster. The first *cat* gene cluster presented *catA* (*catA1*), coding for a catechol 1,2-dioxygenase, *catB* (*catB1*) coding for a muconate cycloisomerase and *catC* coding for a muconolactone isomerase. The same gene cluster was identified in RHA1 and, compared to R7, it showed high protein identity (96–99%). The second *cat* gene cluster identified in R7 lacked *catC* gene; moreover, CatA2 (*catA2*) and CatB2 (*catB2*) were not found homologous to RHA1 genes. *R. aetherivorans* BCP1 presented only one copy of *cat* gene cluster with the same organization of RHA1 and R7, with which showed high similarity (70–90%) (**Figure 2.8**).

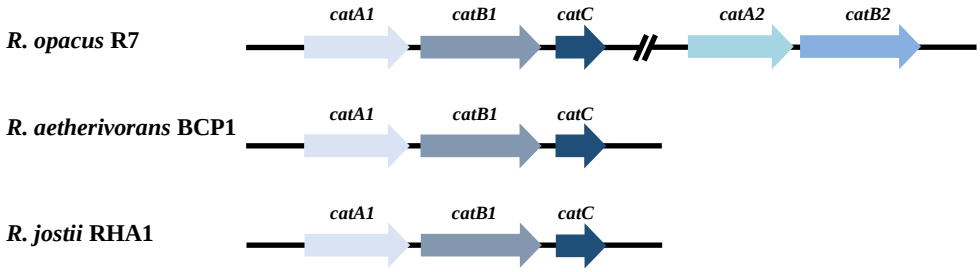


Figure 2.8. Comparative organization of *cat* gene cluster in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

Both R7 and BCP1 genomes contain several genes potentially involved in protocatechuate catabolism. The putative R7 and BCP1 *pca* clusters include genes predicted to encode the enzymes (PcaIJHGBLF and the regulator) required to convert protocatechuate to the TCA cycle intermediates. The predicted products of these R7 genes share high amino acid similarity (97–99%) with their homologous from RHA1. The predicted products of BCP1 *pca* genes share 57%-82% of similarity with the homologous genes from RHA1. The organization of the *pca* genes of both strains is similar to their organization in RHA1. Indeed, they are organized in two putative divergently transcribed operons, *pca*JI and *pcaHGBLRF* (**Figure 2.9**).

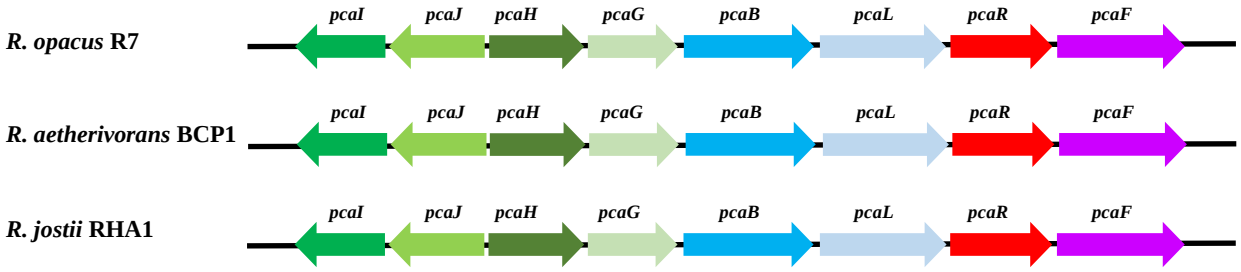


Figure 2.9. Comparative organization of *pca* gene cluster in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

Phenylacetate pathway for aerobic degradation of phenylacetic acid (PAA) can proceeds through the formation of phenylacetyl-coenzyme A (Co-A) and the subsequent hydrolytic ring fission. These metabolic steps in *R. jostii* RHA1 are catalyzed by enzymes coded by the *paa* gene cluster (Navarro-Llorens et al., 2005). The organization of the *paa* gene cluster differs amongst different bacteria, but some features are conserved. Genes encoding two core functional units of the pathway that are consistently clustered, include *paaGHIJK*, which encodes a ring-hydroxylating system, and *paaABCE*, which encodes a *β*-oxidation system. The *paa* gene cluster organization (composed by 15 genes) was conserved in *R. jostii* RHA1 and *R. opacus* R7 genomes. They also showed high percentage of similarity. On the contrary, the *paa* gene cluster was not conserved in *R. aetherivorans* BCP1 and few genes homologous to those of RHA1, are present without a co-localization (**Figure 2.10**).

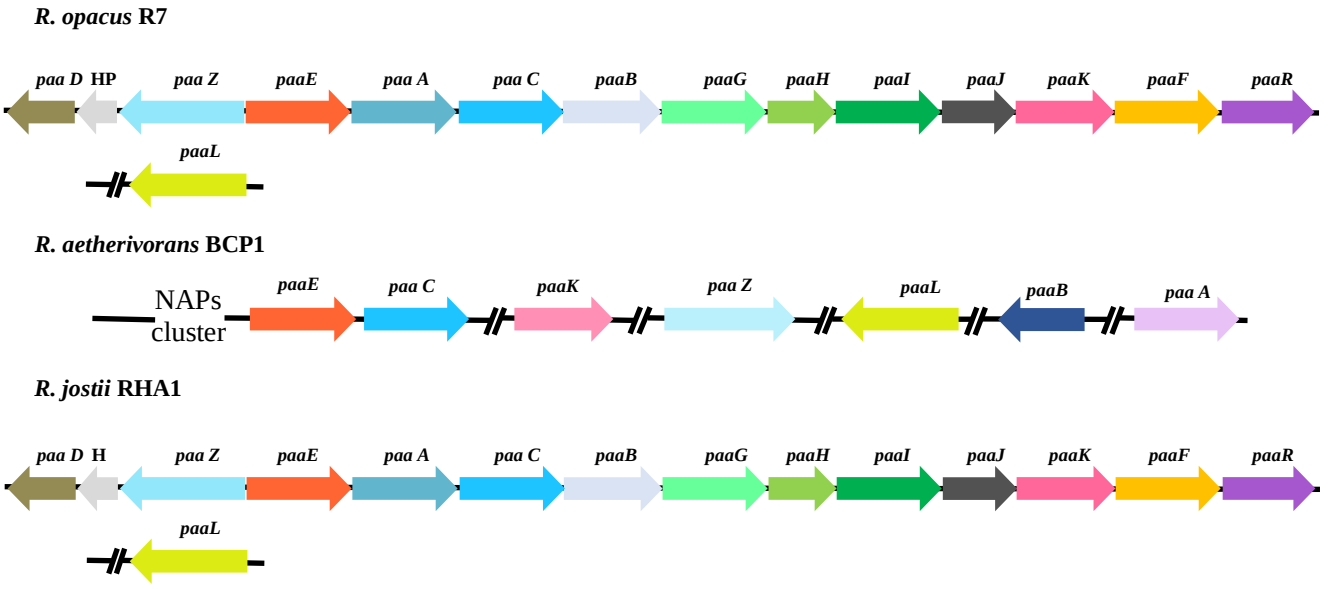


Figure 2.10. Comparative organization of *paa* gene cluster in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

R7 and BCP1 genomes analysis reported one homogentisate 1,2-dioxygenase (*hmgA*) in both strains (97% and 85% identity with RHA1 *hmgA*, respectively). The genetic organization of the *hmgA* flanking regions is quite similar between R7, BCP1 and RHA1, showing genes encoding an enoyl-CoA hydratase, fumarylacetoacetase and a glutaryl-CoA dehydrogenase. Similarly, to RHA1, R7 showed upstream and downstream of these genes, two CDSs coding for long-chain fatty- acid-CoA ligases (EC 6.2.1.3). In the same region, BCP1 genome reported only one gene coding for the same enzyme (**Figure 2.11**).

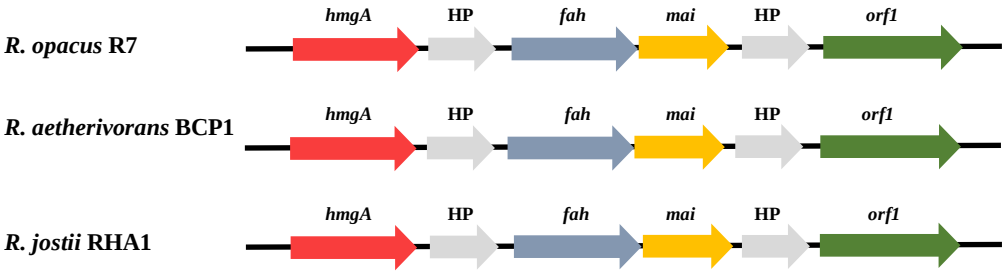


Figure 2.11. Comparative organization of *hmg* gene cluster in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

Moreover, the gene cluster characterized for its involvement in the lower pathway of phthalic acid is decribed in *R. jostii* RHA1. R7 and BCP1 genomes analysis revealed that this cluster is present only in R7 strain on pPDG2, indeed no homologous genes were identified in BCP1 genome. R7 genome presents the same gene cluster organization of RHA1, containing genes encoding for: a transcriptional regulator belonging to the IclR family protein (*padR*), a phthalate 3,4-dioxygenase alpha subunit (*padAa*), a phthalate 3,4-dioxygenase beta subunit (*padAb*), a phthalate 3,4-dihydrodiol

dehydrogenase (*padB*), a phthalate 3,4-dioxygenase ferredoxin (*padAc*), a phthalate 3,4-dioxygenase ferredoxin reductase (*padAd*) and a 3,4-dihydroxyphthalate decarboxylase (*padC*) (**Figure 2.12**). The genic products of these genes showed the following protein identity respect to RHA1: 78%, 84%, 77%, 80%, 76%, 61% and 74%, respectively.

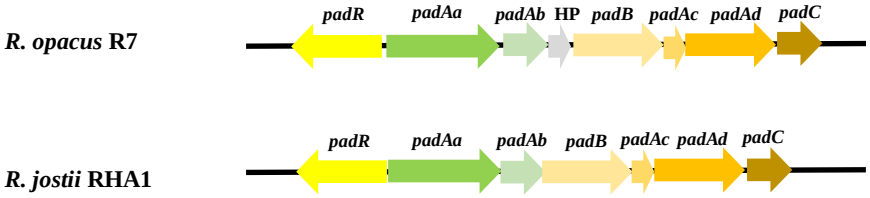


Figure 2.12. Comparative organization of *pad* gene cluster in *R. opacus* R7 and *R. aetherivorans* BCP1 with *R. jostii* RHA1 as reference strain.

2.4 Discussion

In a previous work (Orro et al., 2015) a Phenotype Microarray approach was used to test the catabolic potential of BCP1 and R7 strains towards different categories of xenobiotic compounds belonging to: aliphatic hydrocarbons, polycyclic aromatic hydrocarbons (PAHs), BTEX and other aromatic compounds, and carboxylated compounds. Both strains showed the ability to utilize activity on a wide range of these investigated compounds; consequently, a genome-based approach was used in order to examine the genetic potential of the two strains for the degradation of these compounds. This chapter reports the results of the genome-based analysis on *R. opacus* R7 and *R. aetherivorans* BCP1, respect to reference strains as *R. jostii* RHA1 and literature and database informations. *R. opacus* R7 contains multiple genes for the degradation of a large set of aromatic and polyaromatic hydrocarbons, while BCP1 shows a lower variability in terms of genes predicted to be involved. These genetic features can be related to the strong genetic pressure exerted by the specific carbon sources available in the different niches from which each of the two strains was isolated. Moreover, both strains possess a large number of alternative routes and gene clusters associated to the peripheral pathways of aromatic compound degradation. The relative dimension of the two genomes under analysis can also explain the lower degree of genetic redundancy present in BCP1 compared to R7; moreover, the induction of multiple redundant systems could be a strategy in these bacteria to degrade a wide range of recalcitrant compounds and to persist in the environments. The genome analysis of *R. aetherivorans* BCP1 and *R. opacus* R7 highlighted a great number of interesting features, underlying the peculiar capacities of these two rhodococci for biodegradation and biotransformation applications supported by both their extraordinary genetic repertoire and environmental persistence.

2.5 *References*

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. J Mol Biol, 215(3):403-10.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. (2008). The RAST server: rapid annotations using subsystems technology. BMC Genomics, 9:75.

Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marchall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR (2004). The Pfam protein families database. *Nucleic acids research*. *32* (Database issue), D138-D141.

Blakley ER, Papish B (1982). The metabolism of cyclohexanecarboxylic acid and 3-cyclohexenecarboxylic acid by *Pseudomonas putida*. Can J Microbiol, 28(12):1324–9. PMID: 7168830

Cappelletti M, Fedi S, Frascari D, Ohtake H, Turner RJ and Zannoni D (2011). Analyses of both the *alkB* gene transcriptional start site and *alkB* promoter-inducing properties of *Rhodococcus* sp. strain BCP1 grown on *n*-alkanes. Applied and Environmental Microbiology, 77:1619–1627.

Cappelletti M, Di Gennaro P, D'Ursi P, Orro A, Mezzelani A, Landini M, Fedi S, Frascari D, Presentato A, Zannoni D, Milanesi L, Di Gennaro P (2013). Genome Sequence of *Rhodococcus* sp. strain BCP1, a Biodegrader of Alkanes and Chlorinated Compounds. Genome Announcements, 1(6):e00657-13.

Cappelletti M, Presentato A, Milazzo G, Turner R J, Fedi S, Frascari D, et al. (2015). Growth of *Rhodococcus* sp. strain BCP1 on gaseous *n*-alkanes: new metabolic in sights and transcriptional analysis of two soluble di-iron monooxygenasegenes. Front Microbiol. 6:393. PMID:26029173

Di Gennaro P, Zampolli J, Presti I, Cappelletti M, D'Ursi P, Orro A, Mezzelani A, Milanesi L (2014) Genome sequence of *Rhodococcus opacus* strain R7, a biodegrader of mono- and polycyclic aromatic hydrocarbons. *Genome Announc* 2:e00827-14.

Di Gennaro P, Rescalli E, Galli E, Sello G and Bestetti G (2001). Characterization of *Rhodococcus opacus* R7, a strain able to degrade naphthalene and o-xylene isolated from a polycyclic aromatic hydrocarbon-contaminated soil. Research in Microbiology, 152(7):641–651.

Di Gennaro P, Terreni P, Masi G, Botti S, De Ferra F and Bestetti G (2010). Identification and characterization of genes involved in naphthalene degradation in *Rhodococcus opacus* R7. Applied Microbiology and Biotechnology, 87(1):297–308.

Egland PG, Pelletier DA, Dispensa M, Gibson J, Harwood CS (1997). A cluster of bacterial genes for anaerobic benzene ring biodegradation. Proc Natl Acad Sci U S A. 94(12): 6484–9. PMID: 9177244

Iwaki H, Saji H, Abe K, Hasegawa Y (2005). Cloning and sequence analysis of the 4-hydroxybenzoate 3-hydroxylase gene from a cyclohexanecarboxylate-degrading gram-positive bacterium, *Corynebacterium cyclohexanicum* strain ATCC 51369. Microb and Environ. 20(3):144-150.

Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 45(D1):D353-D361.

Kim D, Choi KY, Yoo M, Zylstra GJ, and Kim E (2018). Biotechnological Potential of *Rhodococcus* Biodegradative Pathways. J. Microbiol. Biotechnol, 28(7):1037-1051.

Masai E, Sugiyama, Iwashita N, Shimizu S, Hauschild JE, Hatta T, et al. (1997). The *bph*DEF meta-cleavage pathway genes involved in biphenyl/polychlorinated biphenyl degradation are located on a linear plasmid and separated from the initial *bph*ACB genes in *Rhodococcus* sp. strain RHA1. Gene, 187(1):141-149. PMID: 9073078

McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJ, Holt R, Brinkman FS, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD (2006). The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA, 103:15582-15587.

Navarro-Llorens JM, Patrauchan MA, Stewart GR, Davies JE, Eltis LD, Mohn WW (2005). Phenylacetate catabolism in *Rhodococcus* sp. strain RHA1: a central pathway for degradation of aromatic compounds. J Bacteriol, 187(13):4497-504. PMID: 15968060

Oldfield C, Pogrebinsky O, Simmonds J, Olson ES, Kulpa CF (1997). Elucidation of the metabolic pathway for dibenzothiophene desulphurization by *Rhodococcus* sp. strain IGTS8 (ATCC 53968). Microbiology, 143(9):2961-73.

Orro A, Cappelletti M, D'Ursi P, Milanesi L, Di Canito A, Zampolli J, Collina E, Decorosi F, Viti C, Fedi S, Presentato A, Zannoni D, Di Gennaro P (2015). Genome and phenotype microarray analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: genetic determinants and metabolic abilities with environmental relevance. PLoS ONE, 10(10):e0139467.

Patrauchan MA, Florizone C, Eapen S, Gómez-Gil L, Sethuraman B, Fukuda M, et al. (2008). Roles of ringhydroxylating dioxygenases in styrene and benzene catabolism in *Rhodococcus jostii* RHA1. J Bacteriol, 190(1):37-47. PMID: 17965160

The UniProt Consortium (2017). UniProt: the universal protein knowledgebase, Nucleic Acids Research, Volume 45, Issue D1, Pages D158–D169

Thompson, JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22(22):4673-80.

Warren R, Hsiao WW, Kudo H, Myhre M, Dosanjh M, Petrescu A, et al. (2004). Functional characterization of a catabolic plasmid from polychlorinated-biphenyl-degrading *Rhodococcus* sp. strain RHA1. J. Bacteriol, 186:7783-7795.

Zampolli J, Collina E, Lasagni M and Di Gennaro P (2014). Biodegradation of variable-chain-length *n*-alkanes in *Rhodococcus opacus* R7 and the involvement of an alkane hydroxylase system in the metabolism. AMB Express, 4(1):1–9.

Chapter 3.

Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7

In the recent years the ecological importance of the bacteria belonging to *Rhodococcus* genus has been assumed for their peculiar capability to resist under stress conditions, generating several stress responses. Indeed, they are able to oxidize natural and xenobiotic hydrocarbons thanks to their distinctive aptitude to cause significant structural and functional rearrangements in their cells, generating the development of resistance to hostile factors. Indeed, *Rhodococcus* strains maintain the characteristic of the membrane under adverse conditions, also regulating the net surface charge. Moreover, the strains can change the composition of fatty acids of the cellular membrane to support critical conditions.

In the present chapter, the response of *R. aetherivorans BCP1* and *R. opacus* R7 to various stress conditions and several antimicrobials was examined by a Phenotype Microarray (PM) approach in relation with genetic determinants, as revealed by annotation analysis of the two genomes. Comparison between metabolic activities and genetic features of BCP1 and R7 provided new insight into the environmental persistence of these two members of the genus *Rhodococcus*.

3.1 Introduction

Members of the genus *Rhodococcus*, which are non-sporulating aerobic bacteria with a high G-C content, degrade a variety of pollutants (Martínková et al., 2009). In this respect, the practical use of *Rhodococcus* spp. in bioremediation, biotransformation and biocatalysis is further enhanced by their persistence in the environment (LeBlanc et al., 2008; de Carvalho, 2012; de Carvalho et al., 2014). Indeed, *Rhodococcus* spp. are able to survive in presence of high doses of toxic compounds, as well as under desiccation conditions, carbon starvation, a wide range of temperatures (from 4°C to 45°C), UV irradiation and osmotic stress (NaCl up to 7.5%) (LeBlanc et al., 2008; de Carvalho et al., 2012; Bequer Urbano et al., 2013; Patrauchan et al., 2012; Alvarez et al., 2004). Survival mechanisms of *Rhodococcus* spp. in the presence of these environmental challenges rely on: i) the modification of the cell wall and the cell membrane fatty acid composition, ii) the accumulation of intracellular lipids as energy storage, iii) the synthesis of compatible solutes (ectoine, hydroxyectoine) for osmotic adjustment under saline stress, and iv) production of bio-molecules (e.g. biosurfactants and pigments) supposedly acting as chelants (Alvarez and Steinbüchel, 2010). Although physiological adaptation strategies regarding membrane composition and intracellular lipid accumulation were established, much less is known about the role of specific genetic traits in the *Rhodococcus* spp. stress response. Among the toxic compounds, numerous studies focused on the resistance of *Rhodococcus* spp. to diverse contaminants and xenobiotics, whereas their resistance to heavy metals and antibiotics remain poorly studied. In this respect, conjugative plasmids were shown to harbor genetic traits involved in resistance to cadmium, thallium, arsenate and arsenite in *Rhodococcus erythropolis* and *Rhodococcus fascians* (Dabbs, Sole, 1988; Desomer et al., 1988). The unique tolerance of *Rhodococcus* spp. to heavy metals as compared to other actinobacteria was also related to non-specific resistance mechanisms, including metal binding to biomolecules and cellular non-diffusing pigments (Ivshina et al., 2013). Further, antibiotic resistance was almost exclusively evaluated in *Rhodococcus equi*, a pathogen for animals and immunosuppressed humans (Cisek et al., 2014). A Phenotype Microarray (PM) approach has recently been used to assess the metabolic activity of two *Rhodococcus* strains, *R. aetherivorans* BCP1 and *R. opacus* R7, under different conditions of growth in plates PM1-PM20 and in plates supplied with xenobiotics (Orro et al., 2015). The PM high-throughput approach highlighted the wide metabolic abilities of BCP1 and R7 and pointed out the peculiar feature of the two strains for resisting hundreds of different stress conditions, although their genome analyses revealed a significant difference in genome size (Ivshina et al., 2013; Cisek et al., 2014). The R7 genome proved to be one of the largest bacterial genomes sequenced to date, with a total size of 10.1 Mb (1 chromosome and 5 plasmids), whereas, the BCP1 genome was 6.1 Mb in size (1 chromosome

and 2 plasmids). The present study sought to evaluate whether the difference in BCP1 and R7 genome size reflects significant differences in their stress tolerance profiles. For this purpose, this work compared the metabolic activities of BCP1 and R7 measured in the presence of various stressors, such as high osmolarity and pH stress, diverse toxic compounds and antibiotics, in plates PM9-PM20. Further, the genetic determinants involved in resistance mechanisms were analyzed in order to support metabolic similarities/differences and provide genetic insight into the peculiar environmental persistence shown by *R. aetherivorans* BCP1 and *R. opacus* R7.

3.2 Materials and methods

R. aetherivorans BCP1 (DSM44980) (also termed *Rhodococcus* sp. BCP1) and *R. opacus* R7 (CIP107348) were tested on inhibitor sensitivity arrays (PM9 to PM20) under more than 1100 different conditions (reagents used are listed at http://www.biolog.com). In particular, for each strain, part of the biomass grown at 30°C on BUG agar was suspended in 15 mL of salt solution. Cell density was adjusted to 85% transmittance before being diluted in IF-10 medium supplied with 1% of dye G (tetrazolium dye) for inoculation in PM plates (Orro et al., 2015; Viti et al., 2007). All PM plates were incubated at 30°C in an Omnilog reader (Biolog). Readings were recorded for 72 h and data were analyzed with Omilog-PM software (Biolog).

3.2.1 Statistical analysis of PM data

PM kinetic curves were analyzed as previously described (Orro et al., 2015). IC50 values were determined for the chemicals tested from PM11 to PM20 by the Omnilog-PM software (Viti et al., 2007). The comparison between metabolic activities of the two strains was visualized by plotting all their activity values (from PM9 to PM20) in a 2D graph as previously described (Viti et al., 2007).

3.2.2 Identification of genetic aspects

The whole-genome shotgun sequencing projects are deposited at DDBJ/EMBL/GenBank under accession numbers CM002177, CM002178, CM002179 for *R. aetherivorans* BCP1 and CP008947, CP008948, CP008949, CP008950, CP008951, CP008952 for *R. opacus* R7. By RAST (Rapid Annotations using Subsystems Technology; Aziz et al., 2008) server and BLAST analyses, we identified coding sequences (CDSs) associated with features predicted to be involved in the bacterial stress response (osmotic and pH stress) and in resistance to antibiotics, metals and other toxic compounds. Kyoto Encyclopedia of Genes and Genomes database (KEGG)

(http://www.genome.jp/kegg/; Kanehisa et al., 2017) and Transporter Classification Database (TCDB) (http://www.tcdb.org/; Saier et al., 2016) were used to assign Enzyme Commission numbers (EC) and TCDB numbers to the coding sequences (CDSs).

3.3 Results

3.3.1 Tolerance to osmotic stress

R. aetherivorans BCP1 and *R. opacus* R7 are metabolically active under a wide range of osmotic stress conditions and under different concentrations (Orro et al., 2015). *R. opacus* R7 cells showed higher osmotic resistance compared to *R. aetherivorans* BCP1 in terms of both number of osmolytes and osmolyte concentrations (circles numbered 1-17 in **Figure 3.1**) with the exception of urea (dots numbered from 18 to 20 in Figure 1, **Table 3.1**). The differences between the two strains are represented in the scatter plot in **Figure 3.1**, where comparison between metabolic activities in the presence of all stressors tested from PM9 to PM20 is visualized.

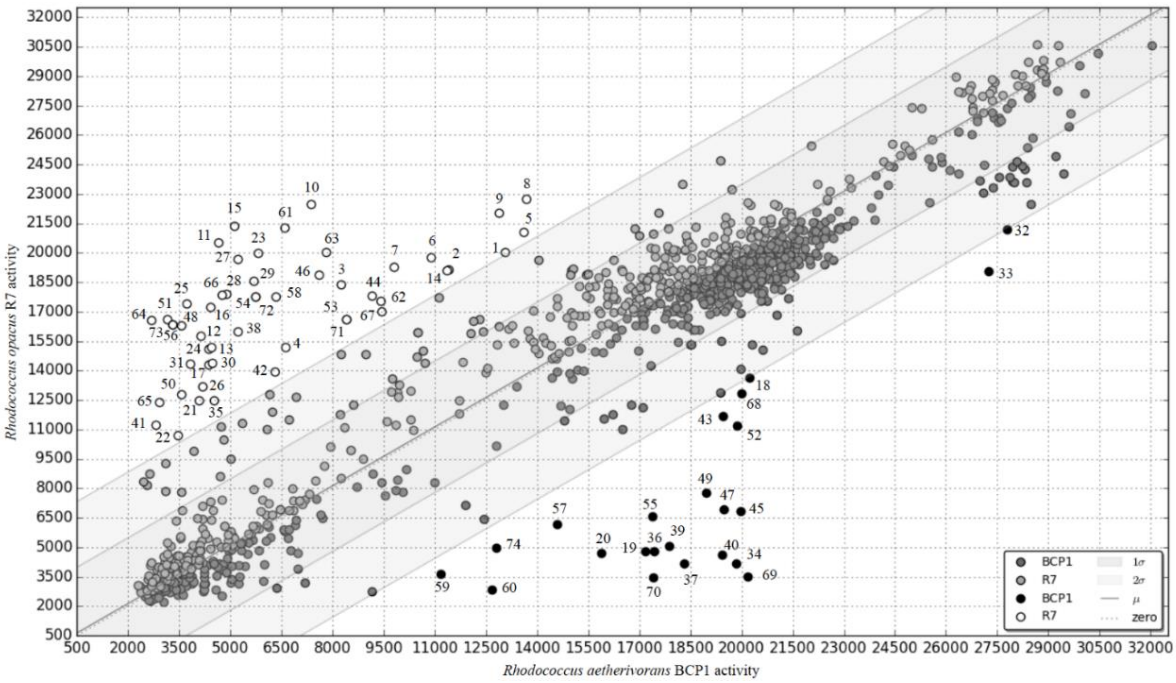


Figure 3.1. Scatter plot comparing the metabolic activities of *R. opacus* R7 and *R. aetherivorans* BCP1 from PM plates 9-20. The metabolic activities of BCP1 and R7 are represented as dots on a 2D scatterplot. Plot bands represent mean (μ_A), standard deviation (σ_A) and double standard deviation ($2\sigma_A$) from the mean of the activity difference (ΔA). The dashed line (named as "zero" in the legend) correlates the points corresponding to compounds inducing the same metabolic response in the two strains. The points located outside of the double standard deviation band represent compounds inducing significantly different metabolic response in BCP1 and R7. In particular, dots below the bands represent PM compounds inducing a higher metabolic activity in BCP1 compared to R7 (full black dots in the legend). Circles above the bands indicate compounds inducing a higher metabolic activity in R7 compared to BCP1 (white dots in the legend). Light grey and dark grey dots, reported in the legend, represent compounds inducing not significantly different metabolic responses in the two strains. Numbers in the plot represent the PM chemicals listed in the Table 1.

In particular, unlike BCP1, the activity of R7 was maintained at a high level up to NaCl 8% without the supply of any osmoprotectant; further, R7 could grow up to the highest dose of sodium formate, sodium lactate and sodium nitrite. Conversely, the BCP1 strain was more resistant to increasing concentrations of urea (up to 6%), while the highest dose of urea (7%) was toxic for both strains. Coding sequences (CDSs) involved in synthesis of compatible solutes such as operons for the synthesis of ectoine (*ectABC*), trehalose (*otsAB*), and glycogen (*glg* genes) were found in both BCP1 and R7 genomes (**Table 3.2**). In addition to one *ectABC* gene cluster, two ectoine hydroxylase coding genes (*ectD*) are present in R7 in two separate chromosomal regions and associated with CDSs coding for transporters, while one *ectD* gene is present in BCP1 and it is associated with trehalose biosynthesis genes. Ectoine hydroxylase (EctD) catalyzes the conversion of ectoine into hydroxyectoine. Differences were observed between BCP1 and R7 in CDSs predicted to be involved in biosynthesis, transport and catabolism of glycine betaine (betaine), which is considered one of the main compatible solutes (**Table 3.2**). Betaine is synthesized through the activity of a choline dehydrogenase (BetA) and a betaine aldehyde dehydrogenase (BetB) (Wargo et al., 2008). Among the 9 CDSs encoding BetA in R7, 2 *betA* genes are colocalized with *betB* (**Panel A in Figure 3.2**). In both cases, a CDS encoding a urea-carboxylase-related amino acid permease (UctT) is placed between the two genes. One *betA-uctT-betB* gene cluster is included in a chromosomal region putatively involved in choline uptake and catabolism through the pathway choline/glycine betaine/dimethylglycine/sarcosine/glycine. Indeed, in addition to *betA* and *betB* genes, this region includes CDSs encoding: i) choline uptake protein BetT, ii) a large subunit of a phenylpropionate dioxygenase that shows similarity (35% aa identity) to the GbcA protein of *Pseudomonas aeruginosa* converting glycine betaine to dimethylglycine (Meskys et al., 2001), iii) a dimethylglycine oxidase that oxidizes the dimethylglycine to sarcosine (DmgO), and iv) a heterotetrameric sarcosine oxidase (*soxBDAG* gene cluster) that catalyzes oxidative demethylation of sarcosine to glycine (**Panel A in Figure 3.2**). The salt tolerance of R7 can be further supported by the presence in R7 genome of CDSs encoding both a proline/ glycine betaine transporter ProP and a multi-component proline permease ProU (composed of the three components, ProV, ProX and Prox), the latter absent in BCP1. Interestingly, this chromosomal region also includes CDSs predicted to encode enzymes involved in tetrahydrofolate-dependent assimilation of methyl groups related to the ability to utilize betaine as carbon and energy source in *Arthrobacter* spp. (Cholo et al., 2015). The lack of a spatial association between *betA*, *betB* and *betT* genes in BCP1 and the absence of a *soxBDAG* gene cluster in BCP1 support the difference between the two strains in terms of osmolyte tolerance. The salt tolerance of R7 can be further supported by the presence in R7 genome of CDSs encoding both a proline/glycine betaine transporter ProP and a multi-component pro-line permease ProU (composed of the three

components, ProV, ProX and Prox), the latter absent in BCP1. Both R7 and BCP1 possess CDSs coding for components of Trk and Kdp K^+ transport systems that allow efflux of K^+ during the osmo-adaptation process (Cotter et al., 2003). In both strains, a *kdpABCDE* operon codes for the three components of a membrane-associated P-type ATPase (KdpA-C), the osmo-sensitive K^+ channel histidine kinase KdpD and the response regulator KdpE. In both BCP1 and R7, the Trk system is composed of two TrkA proteins highly similar to CeoB and CeoC (70 e 75% aa identity) of *Mycobacterium tuberculosis* (Cotter et al., 2003). In BCP1, the association of a *trkH* gene with an additional *trkA* gene suggests a functional association between integral membrane protein TrkH and regulatory protein TrkA (**Panel B in Figure 3.2**). Trk systems are considered secondary transporters, as the contribution of TrkH to K^+ transport is limited compared to other K^+ transport systems, e.g. Kpd (Cotter et al., 2003).

3.3.2 Tolerance to acid and alkaline stress

R. aetherivorans BCP1 and *R. opacus* R7 showed metabolic activities at pH values ranging from 5 to 10, suggesting the presence of mechanisms to persist in both acidic and alkaline environments. Figure 1 shows that, unlike BCP1, R7 was metabolically active at pH 4.5 in the presence of several amino acids such as L-methionine, L-phenilalanine, L-valine, L-homoserine, L-norvaline and a-amino-N-butyric acid (circles numbered from 21 to 31 in **Figure 3.1**). CDSs coding for transporters are involved in mechanisms of resistance to pH stress, i.e. multisubunit F1F0-ATPase and other monovalent cation/proton antiporters (**Table 3.3**). Other gene products supporting bacterial survival at acidic environment are the amino acid decarboxylase systems that control cellular pH by consuming H^+ (Kanamori et al., 2004). CDSs encoding arginine, lysine and glutamate decarboxylases were found in each genome, as well as one CDS encoding a glutamate/g-amino-butyrate antiporter (composing the GAD system with gluta-mate decarboxylase). Differences between BCP1 and R7 were observed in genes linked to urea degradation (**Table 3.3**). Only R7 genome possesses a complete urease gene cassette encoding three structural proteins (UreA, UreB and UreC), along with CDSs coding for three accessory proteins (UreF, UreG, UreD) (**Panel C in Figure 3.2**). Conversely, it shows both urea carboxylase (UcaA) and an allophanate hydrolase (UcaB) (composing the urea amidolyase activity, **Panel C Figure 3.2**) along with a specific urea channel UreI that is missing in R7 and is not clustered with urea-decomposition-related genes (**Panel C Figure 3.2**). CDSs encoding arginine deiminase and ornithine carbamoyltransferase are present, which are not clustered into the same chromosomal region (**Panel D in Figure 3.2, Table 3.3**). Interestingly, in both BCP1 and R7, the *arcB* gene is flanked by genes involved in arginine biosynthesis (*arg* genes) within a chromosomal

region well conserved among Actinomycetales. Conversely, the CDSs flanking the *arcA* gene are only partially conserved. In both BCP1 and R7 as well as in other *Rhodococcus* spp., the arginine deiminase coding gene (*arcA*) is clustered with CDSs encoding a protein containing the DUF1794 domain with unknown function, a choline uptake protein BetT, an arsenate reductase ArsC and a GntR transcriptional regulator. In BCP1, *arcA* is also associated with a CDS coding for an MscL mechano-sensitive channel and a two-component regulatory system. In R7, *arcA* colocalizes with a CDS encoding a SoxR transcriptional regulator (**Panel D Figure 3.2**).

3.3.3. Resistance to antibiotics, metals and other toxic compounds

BCP1 and R7 showed resistance to antibiotics of the following classes (they are represented by circles or dots from 34 to 74 in **Figure 3.1**): quinolones, fluoroquinolones, glycopeptides and sulfonamides. They were also resistant to all the tested doses of polymyxin B, phosphomycin, carbenicillin and nitrofurantoin. Conversely, BCP1 and R7 were sensitive to macrolides, rifampicin, rifamycin, vancomycin, puromycin, phenethicillin, oxacillin, fusidic acid, blasticidin S, chelerythrine and novobiocin. Several genetic determinants were found in both genomes that can contribute to antibiotic resistance mechanisms (**Table 3.4**). Fluoroquinolone resistance may be related to *gyrA* and gyrB genes, while resistance to *β*-lactams can be ascribed to CDSs encoding *β*-lactamases, penicillin-binding proteins and metal-dependent hydrolase of the *β*-lactamase superfamily III. CDSs encoding GCN5-related N-acetyltransferases may support aminoglycoside antibiotic resistance (Burian et al., 2013). Further, CDSs coding for proteins involved in antibiotic resistance in the two strains included a vancomycin resistance protein, quaternary ammonium compound resistance protein SugE and several transport systems (e.g. multidrug transporters and efflux pumps) (**Table 3.4**). Some transcriptional regulators are predicted to activate antibiotic resistance mechanisms. In particular, one CDS encoding a WhiB family transcriptional regulator is included in a chromosomal region conserved in BCP1, R7 and in *M. tuberculosis*. **Table 3.5** reports the antimicrobials of various kinds (from PM11 to PM20) able to discriminate between the two strains in PM experiments on the basis of the IC50 value. Among these, the higher resistance of R7 to penicillin G could be related to the presence of a CDS encoding the penicillin G acylase precursor that is absent in BCP1. Higher metabolic activities were shown in BCP1 as compared to R7 in the presence of ethionamide, which targets the enoyl-acyl reductase InhA involved in mycolic acid biosynthesis (Morlock et al., 2003). BCP1 and R7 possess CDSs encoding InhA-like enoyl-acyl reductases and EthA-like flavin-binding monooxygenases (**Table 3.4**). Both *R. aetherivorans* BCP1 and *R. opacus* R7 showed wide resistance to metalloids and some transition metals (Orro et al., 2015). In particular, the strains showed tolerance

to the tested con-centrations of alkali metals (Li^+, Ce^+), metalloids (TeO_3^{2-}, SeO_3^{2-}, SiO_3^{2-}, AsO_2^-, AsO_4^{3-}) and some transition metals (Cd^{2+},Co^{2+},Cu^{2+},Cr^{3+},$Cr_2O_7^{2+}$,Fe^{3+},Mn^{2+},WO_4^{2+}, Zn^+). The metal ions toxic for both R7 and BCP1 were Ni^+, VO_4^{3-} (*o*-vanadate), and VO_3^- (*m*-vanadate). A significant difference between the two strains was observed in the resistance to thallium acetate (Tl^+) (Table 1, dots from 69 to 70 in **Figure 3.1**). Tellurium (Te) resistance genes (*terA*, *terD*) were found in both genomes, along with arsenic-related genes involved in bacterial arsenic resistance and transformation. Multiple copies of *arsR*, *arsB* and *arsC* are present in both genomes. In BCP1, one gene cluster *arsCDA* has a chromosomal localization and it is associated with *arsR*, *acr3* and two additional *arsC* genes, while other clustered CDSs coding for uncharacterized metal-resistance proteins are localized on plasmid pBMC1 (**Figure 3.2, Panel E**). In R7, one gene cluster *arsCDAR* has a plasmid position (pPDG3) and it is associated with two additional *arsC*, one *acr3*, one *arsR* and a CDS encoding an aquaporin Z (**Figure 3.2, Panel E**). In particular, the *arsC* and *acr3* genes have been suggested to play a key role in resistance to arsenic (Li et al., 2014). Lastly, CDSs involved in oxidative stress response and other detoxification processes were analysed (**Table 3.6**). In particular, three and six genes were annotated as catalase KatE in BCP1 and R7, respectively, and one CDS in each strain was predicted to encode the heme-containing enzyme catalase-peroxidase KatG. CDSs coding for superoxide dismutases belonging to Cu/Zn and Mn families were found in both strains. Both genomes possess CDSs encoding transcriptional regulators involved in the stress response (SoxR, Fur, Zur and Rex), as well as CDSs encoding glutathione peroxidases, ferroxidases, and proteins involved in biosynthesis of mycothiol. This low-molecular-weight thiol is typically produced by Actinobacteria for protection against the hazards of aerobic metabolism (Garg et al., 2013). The CDSs encoding NADPH:quinone reductases in BCP1 and R7 show high similarity (>70% aa identity) with the NADPH:quinone reductase overexpressed in *R. jostii* RHA1 during carbon starvation (Patrauchan et al., 2012). Interestingly, in R7, BCP1 and RHA1 genomes, the position of a CDS coding for a redox-sensing transcriptional regulator QorR is conserved upstream of the NADPH:quinone reductase coding gene, as well as the proximity of CDSs encoding allophanate hydrolase and urea carboxylase. Additional stress-response-related CDSs in BCP1 and R7 encode polyols transporters, sulfate and thiosulfate import proteins CysA, Nudix proteins, DedA, D-Tyrosyl-tRNA deacylase (**Table 3.6**) and carbon starvation protein A. Similarly, to the high number of universal stress protein family genes (USPs) found in the RHA1 genome (Patrauchan et al., 2012), R7 and BCP1 possess 14 and 11 USP genes, respectively. Only in BCP1, a CDS encoding a phytochrome-like two-component sensor histidine kinase was found. Interestingly, this CDS was flanked by a gene predicted to code for serine phosphatase RsbU that activates the σ^B factor in response to various stress conditions (Shin et al., 2010).

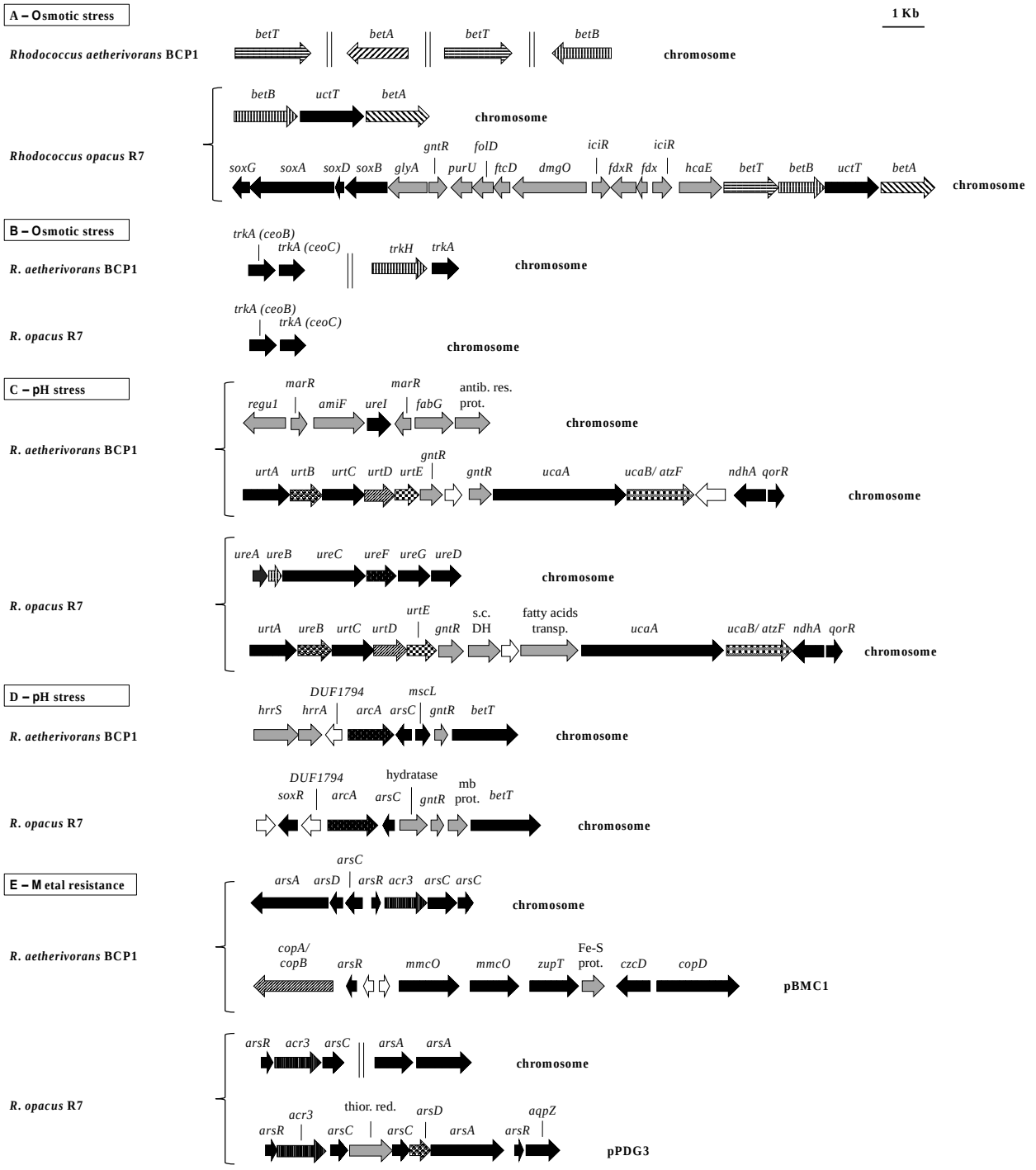


Figure 3.2. Representative gene clusters involved in *R. opacus* R7 and *R. aetherivorans* BCP1 stress response whose presence or organization differentiates the two strains. Genes are visualized as arrows. The predicted protein products of the genes are listed in Tables S2-S5 except for the genes (represented in light grey) following described: *hcaE*, phenylpropionate dioxygenase, large subunit; *iciR*, IclR family transcriptional regulator; *fdx*, ferredoxin; *fdxR*, ferredoxin reductase; *dmgO*, dimethylglycine oxidase; *ftcD*, formiminotetrahydrofolate cyclodeaminase; *folD*, methylenetetrahydrofolate dehydrogenase; *purU*, formyltetrahydrofolate deformylase; *gntR*, GntR family transcriptional regulator; *glyA*, serine hydroxymethyltransferase; *regu1*, putative nitrile hydratase regulator clustered with urea transport; *marR*, MarR family transcriptional regulator; *amiF*, formamidase; *fabG*, 3-oxoacyl-(acyl-carrier-protein) reductase; antib. res. prot., probable antibiotic resistance protein; *hrrA*, hemoglobin-dependent two component system response regulator HrrA; *hrrS*, sensory histidine kinase HrrS; DUF1794, protein including a DUF1794 domain with unknown function; s.c. DH, short chain dehydrogenase; mb prot., membrane protein; Fe-S prot., iron-sulphur protein; thior. red., thioredoxin reductase. Genes encoding hypothetical proteins are represented in white. Vertical lines are reported between gene clusters not close to each other.

3.4 Discussion

R. aetherivorans BCP1 and *R. opacus* R7 are metabolically active under a wide range of osmotic stress conditions (Orro et al., 2015). Concerning the tolerance to osmolitic stress the two strains revealed several differences; R7 cells showed higher osmotic resistance compared to BCP1 in terms of both number of osmolytes and osmolyte concentrations with the exception of urea. Conversely, the BCP1 strain was more resistant to increasing concentrations of urea up to the highest tested dose that was toxic for both strains. Coding sequences (CDSs) involved in synthesis of compatible solutes were found in both BCP1 and R7 genomes. However, the independent localization of *ectD* in BCP1 and R7 with respect to *ectABC* might be related to different mechanisms generating hydroxyectoine instead of ectoine. In line with this, in *R. opacus* PD630 the production of ectoine or hydroxyectoine under osmotic stress is affected by the carbon source (Lamark et al., 1991). Differences were also observed between BCP1 and R7 in CDSs predicted to be involved in biosynthesis, transport and catabolism of glycine betaine (betaine). The lack of a spatial association between *betA*, *betB* and *betT* genes in BCP1 and the absence of a *soxBDAG* gene cluster in BCP1 support the difference between the two strains in terms of osmolyte tolerance. In both strains K^+ transport systems and also the amino acid decarboxylase systems that support bacterial survival in acidic environment (Kanamori et al., 2004) are present in the genomes of the two strains. Urease and urea amidolyase, are known to degrade urea into carbonic acid and ammonia that raises the pH and counteracts acid stress (Van Vliet et al., 2004). R7 genome possesses a complete urease gene cassette, related to its capacity to resist to pH 4.5, unlike BCP1. BCP1 can tolerate increasing concentrations of urea maybe because of it possesses the *ureI* gene that codes for a protein that mediates uptake of urea and it colocalizes with a formamidase (AmiF), leading to production of ammonia from amides (Zúñiga et al., 2002). Ammonia is also produced from fermentative degradation of L-arginine by three enzymes belonging to the arginine deiminase (ADI) pathway: arginine deiminase (ADI or ArcA), ornithine carbamoyltransferase (OTC or ArcB), and carbamate kinase (CK or ArcC); these enzymes are present in both strains despite the ADI pathway is apparently confined to anaerobic or facultative anaerobic bacteria organisms. These genes are not clusterized into the same chromosomal and this is in contrast with the cluster organization typically found on the bacterial chromosome (Zúñiga et al., 2002). Additionally, in PM analysis *R. aetherivorans* BCP1 and *R. opacus* R7 showed the ability to counteract a wide range of toxic compounds and antibiotics belonging to different classes. In spite of the significant diversity in genome size, a similar resistance/sensitivity pattern was seen in the two strains. CDSs encoding GCN5-related N-acetyltransferases may support aminoglycoside antibiotic resistance (Burian et al., 2013). The difference in ethionamide sensitivity between R7 and BCP1

might be due to different EthA regulatory mechanisms occurring in the two strains and/or to different protein sequences. Both *R. aetherivorans* BCP1 and *R. opacus* R7 showed wide resistance to metalloids and some transition metals (Orro et al., 2015); a significant difference between the two strains was observed in the resistance to thallium, even if tellurium (Te) resistance genes were found in both genomes. CDSs predicted to be involved in mechanisms of metal resistance are copper chaperones, copper-resistance proteins, multicopper oxidases, heavy-metal-translocating P-type ATPases, Mg and Co efflux proteins, Co/Zn/Cd resistance protein CzcD and several transcriptional regulators. In both strains, some scattered and clustered genes related to heavy metal resistance are found on plasmids, representing part of the genetic pool involved in horizontal gene transfer (HGT) events. In line with this, native plasmids of several *Rhodococcus* strains were reported to mediate heavy metal resistance such as arsenate, arsenite, cadmium, and thallium (Dabbs and Sole,1988; Desomer et al., 1988). Lastly, CDSs involved in oxidative stress response and other detoxification processes can support the ability of the two strains to resist oxidizing agents, DNA synthesis inhibitors, folate antagonists, metal chelators and other toxicants. Superoxide dismutase and catalase are central antioxidant enzymatic scavengers. The limited knowledge of the genetic basis supporting oxidative stress response and of detoxification processes in *Rhodococcus* spp., hampers the interpretation of metabolic differences between the two strains. However, in addition to genetic traits predicted to play a role in antimicrobial resistance mechanisms, a fundamental contribution to the wide resistance ability of the two strains can be ascribed to the peculiar features of the mycolic-acid-containing cell wall typical of *Rhodococcus* spp., as well as to phenomena of intracellular sequestration, adsorption of metals/antimicrobials to cell wall and binding to carotenoid pigments (Ivshina et al., 2013; Kuyukina et al., 2000; Schmidt et al., 2005). In line with this, some differences in resistance/sensitivity to antimicrobials shown by the two strains might also be related to differences in cell membrane lipids, mycolic acid composition and cell pigment production.

In summary, the results show that *R. opacus* R7 has a higher resistance than *R. aetherivorans* BCP1 to acidic pHs and to a few osmolytes; conversely, in spite of the considerable difference in genome sizes, the two strains have similar behaviours in counteracting the toxicity of metals, antibiotics, oxidative agents and several other toxic compounds. From the point of view of environmental remediation technologies and transformation of toxic compounds, the two strains are expected to have similar behaviour in response to stressors that can occur in contaminated soils.

Table 3.1. List of compounds associated to the numbers reported in Figure 1

Index	Plate	Compound
1	PM09 A07	6% NaCl
2	PM09 A08	6.5% NaCl
3	PM09 A09	7% NaCl
4	PM09 A10	8% NaCl
5	PM09 B01	6% NaCl
6	PM09 B12	6% NaCl + L-Carnitine
7	PM09 C12	6% NaCl + Trigonelline
8	PM09 F06	6% Sodium Lactate
9	PM09 F07	7% Sodium Lactate
10	PM09 F09	9% Sodium Lactate
11	PM09 F10	10% Sodium Lactate
12	PM09 F11	11% Sodium Lactate
13	PM09 F12	12% Sodium Lactate
14	PM09 H12	100mM Sodium Nitrite
15	PM09 E03	3% Sodium Formate
16	PM09 E04	4% Sodium Formate
17	PM09 E05	5% Sodium Formate
18	PM09 E09	4% Urea
19	PM09 E10	5% Urea
20	PM09 E11	6% Urea
21	PM10 A03	pH 4.5
22	PM10 B07	pH 4.5+L-Glutamine
23	PM10 C01	pH 4.5+L-Methionine
24	PM10 C02	pH 4.5+L-Phenylalanine
25	PM10 C08	pH 4.5+L-Valine
26	PM10 C09	pH 4.5+Hydroxy-L-Proline
27	PM10 C12	pH 4.5+L-Homoserine
28	PM10 D03	pH 4.5+L-Norvaline
29	PM10 D04	pH 4.5+α-Amino-N-Butyric acid
30	PM10 D08	pH 4.5+5-Hydroxy-L-Lysine
31	PM10 D10	pH 4.5+D,L-Diamino-α,ε-Pimelic acid
32	PM10 F03	pH 9.5+L-Proline
33	PM10 G06	pH 9.5+Putrescine
34	PM12 C01	Paromomycin (1st conc.)
35	PM11 A09	Lincomycin (1st conc.)
36	PM11 F11	Neomycin (3rd conc.)
37	PM11 A04	Amikacin (4th conc.)

38	PM11 G08	Gentamicin (4th conc.)
39	PM17 B08	Hygromycin B (4th conc.)
40	PM20 A08	Apramycin (4th conc.)
41	PM11 B04	Amoxicillin (4th conc.)
42	PM13 B04	Azlocillin (4th conc.)
43	PM16 D11	Cinoxacin (3rd conc.)
44	PM11 E07	Enoxacin (3rd conc.)
45	PM17 B09	Ethionamide (1st conc.)
46	PM12 A02	Penicillin G (2nd conc.)
47	PM17 B10	Ethionamide (2nd conc.)
48	PM12 A03	Penicillin G (3rd conc.)
49	PM17 B11	Ethionamide (3rd conc.)
50	PM12 A04	Penicillin G (4th conc.)
51	PM14 G08	Carbenicillin (4th conc.)
52	PM17 B12	Ethionamide (4th conc.)
53	PM15 C10	1,10-phenanthroline (2nd conc.)
54	PM15 B11	5,7-Dichloro-8-hydroxyquinaldine (3rd conc.)
55	PM14 B08	Fusaric acid (4th conc.)
56	PM15 C04	5,7-Dichloro-8-hydroxyquinoline (4th conc.)
57	PM20 E10	Hexachlorophene (2nd conc.)
58	PM18 H07	2-Phenylphenol (3rd conc.)
59	PM20 E11	Hexachlorophene (3rd conc.)
60	PM16 H03	Glycine hydroxamate (3rd conc.)
61	PM18 H10	Plumbagin (2nd conc.)
62	PM18 D11	Lidocaine (3rd conc.)
63	PM18 H11	Plumbagin (3rd conc.)
64	PM14 C08	1-Hydroxypyridine-2-thione (4th conc.)
65	PM20 E06	Dodine (2nd conc.)
66	PM18 C10	Pentachlorophenol (2nd conc.)
67	PM15 G04	Carbonyl cyanide m-chlorophenyl hydrazone (CCCP) (4th conc.)
68	PM19 D08	Iodonitro Tetrazolium Violet (4th conc.)
69	PM13 F09	Thallium (I) acetate (1st conc.)
70	PM13 F10	Thallium (I) acetate (2nd conc.)
71	PM14 C10	Sodium cyanite (2nd conc.)
72	PM14 B11	Sodium arsenate (3rd conc.)
73	PM14 C04	Boric acid (4th conc.)
74	PM19 H04	Hexammine cobalt (III) chloride (4th conc.)

Table 3.2 A selection of protein functions annotated in *R. aetherivorans* BCP1 and *R. opacus* R7 genomes associated to osmotic stress response (RAST annotation analysis).

Subsystem[a]	**Function**	**EC #[b]**	**Gene**	**BCP1[a]**	**R7[a]**
Choline and Betaine Uptake and Betaine Biosynthesis	Betaine aldehyde dehydrogenase	1.2.1.8	*betB*	1	5
Choline and Betaine Uptake and Betaine Biosynthesis	Choline dehydrogenase	1.1.99.1	*betA*	1	9
Choline and Betaine Uptake and Betaine Biosynthesis	Glycine betaine ABC transport system permease protein	3.A.1.12.1	*opuCD/ opuCB*	-	2
Choline and Betaine Uptake and Betaine Biosynthesis	High-affinity choline uptake protein	2.A.15.1.4	*betT*	2	2
Choline and Betaine Uptake and Betaine Biosynthesis	L-Proline/Glycine betaine transporter	2.A.1.6.4	*proP*	3	19
Choline and Betaine Uptake and Betaine Biosynthesis	L-Proline/Glycine betaine ABC transport system permease	3.A.1.12.1	*proV*	1	3
Choline and Betaine Uptake and Betaine Biosynthesis	L-Proline/Glycine betaine binding ABC transporter	3.A.1.12.1	*proX*	-	3
Choline and Betaine Uptake and Betaine Biosynthesis	L-Proline/Glycine betaine ABC transport system permease	3.A.1.12.1	*proW*	-	2
Choline and Betaine Uptake and Betaine Biosynthesis	Monomeric sarcosine oxidase	1.5.3.1	*msox*	1	1
Choline and Betaine Uptake and Betaine Biosynthesis	Sarcosine oxidase α-subunit	1.5.3.1	*soxA*	-	2
Choline and Betaine Uptake and Betaine Biosynthesis	Sarcosine oxidase α-subunit	1.5.3.1	*soxB*	1	3
Choline and Betaine Uptake and Betaine Biosynthesis	Sarcosine oxidase α-subunit	1.5.3.1	*soxD*	-	1
Choline and Betaine Uptake and Betaine Biosynthesis	Sarcosine oxidase α-subunit	1.5.3.1	*soxG*	-	1
Ectoine biosynthesis and regulation	Diaminobutyrate-pyruvate aminotransferase	2.6.1.46	*ectB*	1	1
Ectoine biosynthesis and regulation	Ectoine hydroxylase	1.14.11.-	*ectD*	1	2
Ectoine biosynthesis and regulation	L-2,4-diaminobutyric acid acetyltransferase	2.3.1.178	*ectA*	1	1
Ectoine biosynthesis and regulation	L-ectoine synthase	4.2.1.108	*ectC*	1	1
Glycogen metabolism	Glucose-1-phosphate adenylyltransferase	2.7.7.27	*glgC*	1	1
Glycogen metabolism	Predicted glycogen synthase, ADP-glucose transglucosylase, Actinobacterial type	2.4.1.21	*glgA*	1	1
Glycogen metabolism, Maltose and Maltodextrin Utilization	4-α-glucanotransferase (amylomaltase)	2.4.1.25	*malQ*	1	1
Glycogen metabolism, Maltose and Maltodextrin Utilization	Glycogen phosphorylase	2.4.1.1	*glgP*	1	1
Osmoregulation	Aquaporin Z	1.A.8.3.1	*aqpZ*	1	2

Potassium homeostasis	DNA-binding response regulator	-	*kdpE*	1	1
Potassium homeostasis	Large-conductance mechanosensitive channel	1.A.22.1.-	*mscL*	2	1
Potassium homeostasis	Mini-conductance mechanosensitive channel	1.A.23.4.-	*mscM*	1	1
Potassium homeostasis	Osmosensitive K^+ channel histidine kinase	2.7.3.-	*kdpD*	3	5
Potassium homeostasis	Potassium efflux system KefA protein/Small-conductance mechanosensitive channel	1.A.23.3.-	*kefA/ mscS*	1	1
Potassium homeostasis	Potassium-transporting ATPase A chain	3.6.3.12/ 3.A.3.7.1	*kdpA*	1	1
Potassium homeostasis	Potassium-transporting ATPase B chain	3.6.3.12/ 3.A.3.7.1	*kdpB*	1	1
Potassium homeostasis	Potassium-transporting ATPase C chain	3.6.3.12/ 3.A.3.7.1	*kdpC*	1	1
Potassium homeostasis	Potassium uptake protein TrkH	2.A.38.4.-	*trkH*	1	-
Potassium homeostasis	Trk system potassium uptake protein TrkA	3.A.3.1.8/ 2.A.38.4.-	*trkA*	3	2
Trehalose Biosynthesis	1,4-α-glucan (glycogen) branching enzyme, GH-13-type	2.4.1.18	*glgB*	1	1
Trehalose Biosynthesis	α,α-trehalose-phosphate synthase	2.4.1.15	*otsA*	2	2
Trehalose Biosynthesis	Glucoamylase	3.2.1.3	*gaa*	5	2
Trehalose Biosynthesis	Glycogen debranching enzyme	3.2.1.-	*glgX*	2	2
Trehalose Biosynthesis	Malto-oligosyltrehalose synthase	5.4.99.15	*glgY/ treY*	1	1
Trehalose Biosynthesis	Malto-oligosyltrehalose trehalohydrolase	3.2.1.141	*glgZ/ treZ*	1	1
Trehalose Biosynthesis	Putative glucanase	3.2.1.-	*glgE*	1	1
Trehalose Biosynthesis	Trehalose-6-phosphate phosphatase	2.4.1.15	*tppb*	5	2
Trehalose Biosynthesis	Trehalose synthase	2.4.1.245	*treS*	1	1
Trehalose Biosynthesis, Maltose and Maltodextrin Utilization	Trehalose phosphorylase/ Maltose phosphorylase	2.4.1.8	*treP*	-	1

[a]The number of CDSs predicted to encode each protein function in BCP1 and R7 genomes is reported along with the functional categories defined by RAST.
[b] Either EC or TCDB number is reported (see Material and Methods for details)

Table 3.3. A selection of protein functions annotated in *R. aetherivorans* BCP1 and *R. opacus* R7 genomes associated to pH stress response (RAST annotation analysis)

Subsystem[a]	**Function**	**EC #[b]**	**Gene**	**BCP1[a]**	**R7[a]**
Arginine Deiminase Pathway	Arginine deiminase	3.5.3.6	*arcA*	1	1
Arginine Deiminase Pathway	Arginine pathway regulatory protein ArgR	-	*argR*	1	1
Arginine Deiminase Pathway	Ornithine carbamoyltransferase	2.1.3.3	*arcF*	1	1
Arginine, Ornithine and Lysine Degradation	Arginine decarboxylase/ Ornithine decarboxylase/ Lysine decarboxylase	4.1.1.18	*idcC*	1	1
G3E family of P-loop GTPases	Urease accessory protein UreD	-	*ureD*	-	1
G3E family of P-loop GTPases	Urease accessory protein UreF	-	*ureF*	-	1
G3E family of P-loop GTPases	Urease accessory protein UreG	-	*ureG*	-	1
G3E family of P-loop GTPases	Urease α-subunit	3.5.1.5	*ureC*	-	1
G3E family of P-loop GTPases	Urease β-subunit	3.5.1.5	*ureB*	-	1
G3E family of P-loop GTPases	Urease γ-subunit	3.5.1.5	*ureA*	-	1
GAD system	Glutamate decarboxylase	4.1.1.15	*gadB*	1	1
GAD system	Glutamate/γ-aminobutyrate antiporter	2.A.3.7.-	*gadC*	1	1
Multi-subunit cation antiporter	Calcium/proton antiporter	2.A.19.1.1	*chaA*	-	1
Multi-subunit cation antiporter	Na^+/H^+ antiporter, CPA1 family	2.A.36.-.-	*nhaP*	2	3
Multi-subunit cation antiporter	Na^+/H^+ antiporter NhaA type	2.A.33.1.-	*nhaA*	2	2
Multi-subunit cation antiporter	$Na^{(+)}$ $H^{(+)}$ antiporter subunit A/ $Na^{(+)}$ $H^{(+)}$ antiporter subunit B	2.A.63.1.4	*mnhA B/ mrpAB*	1	1
Multi-subunit cation antiporter	$Na^{(+)}$ $H^{(+)}$ antiporter subunit C	2.A.63.1.4	*mrpC/ mnhC*	1	1
Multi-subunit cation antiporter	$Na^{(+)}$ $H^{(+)}$ antiporter subunit D	2.A.63.1.4	*mrpD/ mnhD*	1	1
Multi-subunit cation antiporter	$Na^{(+)}$ $H^{(+)}$ antiporter subunit E	2.A.63.1.4	*mrpE/ mnhE*	1	1
Multi-subunit cation antiporter	$Na^{(+)}$ $H^{(+)}$ antiporter subunit F	2.A.63.1.4	*mrpF/ mnhF*	1	1
Multi-subunit cation antiporter	$Na^{(+)}$ $H^{(+)}$ antiporter subunit G	2.A.63.1.4	*mrpG/ mnhG*	1	1
Urea carboxylase and Allophanate hydrolase cluster	Allophanate hydrolase	3.5.1.54	*ucaB/ atzF*	1	4
Urea carboxylase and Allophanate hydrolase cluster	Biotin carboxyl carrier protein	6.4.1.3	*bccA*	1	1

Urea carboxylase and Allophanate hydrolase cluster	Urea carboxylase	6.3.4.6	*ucaA*	2	2
Urea decomposition	Urea ABC transporter, substrate binding protein UrtA	3.A.1.4.-	*urtA*	1	1
Urea decomposition	Urea ABC transporter, permease protein UrtB	3.A.1.4.-	*urtB*	1	1
Urea decomposition	Urea ABC transporter, permease protein UrtC	3.A.1.4.-	*urtC*	1	1
Urea decomposition	Urea ABC transporter, ATPase protein UrtD	3.A.1.4.-	*urtD*	1	1
Urea decomposition	Urea ABC transporter, ATPase protein UrtE	3.A.1.4.-	*urtE*	1	1
Urea decomposition	Urea carboxylase-related amino acid permease	2.A.3.-.-	*uctT*	1	3
Urea decomposition	Urea carboxylase-related aminomethyltransferase	2.1.2.10	*amt*	2	2
Urea decomposition	Urea channel UreI	1.A.29.1.-	*ureI*	1	-

[a] The number of CDSs predicted to encode each protein function in BCP1 and R7 genomes is reported along with the functional categories defined by RAST.
[b] Either EC or TCDB number is reported (see Material and Methods for details)

Table 3.4. A selection of protein functions annotated in *R. aetherivorans* BCP1 and *R. opacus* R7 genomes associated to antibiotic and metals transport/resistance (RAST annotation analysis)

Subsystem [a]	Role	EC #	Gene	BCP1 [a]	R7 [a]
Arsenic resistance	Arsenic efflux pump protein	3.A.4.1.1, 2.A.45.-	*arsB*	1	1
Arsenic resistance	Arsenical pump-driving ATPase	3.6.3.16	*arsA*	4	4
Arsenic resistance	Arsenate reductase	1.20.4.1	*arsC*	4	4
Arsenic resistance	Arsenical resistance operon trans-acting repressor ArsD	-	*arsD*	1	1
Arsenic resistance	Arsenical resistance operon repressor	-	*arsR*	2	3
Arsenic resistance	Arsenical-resistance protein ACR3	-	*acr3*	2	2
Beta-lactamase	Beta-lactamase	3.5.2.6	*blaP*	5	4
Beta-lactamase	Beta-lactamase class C and other penicillin binding proteins	-	*ampC*	1	2
Beta-lactamase	Metal-dependent hydrolase of the beta-lactamase superfamily III	-	*atsH/ elaC*	1	1
Cobalt-zinc-cadmium resistance	Cobalt-zinc-cadmium resistance protein CzcD	-	*czcD*	2	3
Cobalt-zinc-cadmium resistance	Transcriptional regulator, MerR family	-	*merR*	5	13
Copper homeostasis	Copper chaperone	3.A.3.5.-	*copZ*	1	5
Copper homeostasis	Copper resistance protein CopC	9.B.62.-	*copC*	2	2
Copper homeostasis	Copper resistance protein CopD	9.B.62.-	*copD*	4	4
Copper homeostasis,	Lead, cadmium, zinc and mercury transporting ATPase/ Copper-translocating P-type ATPase	3.6.3.3 3.6.3.5 3.6.3.4	*copA/ copB*	8	10
Copper homeostasis	Multicopper oxidase	-	*mmcO*	5	7
Copper homeostasis	Multidrug resistance transporter, Bcr/CflA family	-	*ydhC*	1	1
Ferrous iron transporter EfeUOB	Ferrous iron transport periplasmic protein EfeO	2.A.108.2.3	*efeO*	1	1
Ferrous iron transporter EfeUOB	Ferrous iron transport permease EfeU	2.A.108.2.5	*efeU*	1	1
Ferrous iron transporter EfeUOB	Ferrous iron transport peroxidase EfeB	2.A.108.2.3	*efeB*	1	1
Magnesium transport	Magnesium and cobalt transport protein CorA	1.A.35.3.-	*corA*	3	6
Magnesium transport	Magnesium and cobalt efflux protein CorC	9.A.40.-	*corC*	4	3
Magnesium transport	Mg/Co/Ni transporter MgtE	-	*mtgE*	2	1
Mercuric reductase	PF00070 family, FAD-dependent NAD(P)-disulphide oxidoreductase	-	-	1	1
Queuosine-Archaeosine Biosynthesis	Permease of the drug/metabolite transporter (DMT) superfamily		*dmt*	4	5
Resistance to fluoroquinolones	DNA gyrase subunit A	5.99.1.3	*gyrA*	1	1
Resistance to fluoroquinolones	DNA gyrase subunit B	5.99.1.3	*gyrB*	1	1
Resistance to Vancomycin	Vancomycin B-type resistance protein VanW	-	*vanW*	1	1
Transport of Nickel and Cobalt	ATPase component NikO of energizing module of nickel ECF transporter	3.A.1.23.7	*nikO*	1	1

Transport of Nickel and Cobalt	Substrate-specific component NikM of nickel ECF transporter	3.A.1.23.7	*nikM*	1	1
Transport of Nickel and Cobalt	Transmembrane component NikQ of energizing module of nickel ECF transporter	3.A.1.23.7	*nikQ*	1	1
WhiB and WhiB-type regulatory proteins	WhiB-family transcriptional regulator	-	*whiB7*	4	5
no subsystem	Enoyl-[acyl-carrier-protein] reductase	1.3.1.9	*inhA*	1	1
no subsystem	GCN5-related N-acetyltransferase (GNAT)	2.3.1.-	*aac*	6	5
no subsystem	Macrolide-efflux protein	2.A.1.21.-	-	1	1
no subsystem	Manganese transport protein MntH	2.A.55.3.-	*mntH*	3	2
no subsystem	Monooxygenase, flavin-binding family	1.14.13.-	*ethA*	2	2
no subsystem	Multidrug resistance protein B	2.A.1.3.-	-	5	4
no subsystem	Penicillin-binding protein	-	*pbp*	1	1
no subsystem	Penicillin G acylase precursor	3.5.1.11	-		1
no subsystem	Quaternary ammonium compound-resistance protein SugE	2.A.7.1.4	*sugE*	1	1
no subsystem	Tellurium resistance protein TerA	-	*terA*	1	3
no subsystem	Tellurium resistance protein TerD	-	*terD*	4	5
no subsystem	Zinc ABC transporter, periplasmic-binding protein ZnuA	3.A.1.15.-	*znuA*	1	2
no subsystem	Zinc ABC transporter, inner membrane permease protein ZnuB	3.A.1.15.-	*znuB*	1	2
no subsystem	Zinc ABC transporter, ATP-binding protein ZnuC	3.A.1.15.-	*znuC*	1	2
no subsystem	Zinc transporter, ZIP family		*zupT*	1	1

[a] The number of CDSs predicted to encode each protein function in BCP1 and R7 genomes is reported along with the functional categories defined by RAST.
[b] Either EC or TCDB number is reported (see Material and Methods for details)

Table 3.5. List and IC_{50} values of the PM11-20 inhibitors/toxicants giving different metabolic response in *R. aetherivorans* BCP1 and *R. opacus* R7[a,b]

Substrate/Chemical	Panels (wells)	Mode of action	IC50	
			BCP1	R7
Antibiotics				
Paromomycin	PM12 (C1-C4)	protein synthesis, 30S ribosomal subunit, aminoglycoside	1.70	0.64
Apramycin	PM20 (A5-A8)	protein synthesis, aminoglycoside	>4.40	3.63
Amikacin	PM11 (A1-A4)	wall, lactam	>4.40	3.56
Ethionamide	PM17 (B9-B12)	inhibits mycolic acid synthesis	>4.40	4.12
Hygromycin B	PM17 (B5-B8)	protein synthesis, aminoglycoside	>4.40	3.59
Neomycin	PM11 (F9-F12)	protein synthesis, 30S ribosomal subunit, aminoglycoside	3.51	2.59
Carbenicillin	PM14 (G5-G8)	wall, lactam	3.41	4.40
Penicillin G	PM12 (A1-A4)	wall, lactam	1.76	4.24
Gentamicin	PM11 (G5-G8)	protein synthesis, 30S ribosomal subunit, aminoglycoside	3.63	4.36
Membrane permeability and transport				
Hexachlorophene	PM20 (E9-E12)	membrane, electron transport	3.14	1.36
2-Polyphenol	PM18 (H5-H8)	membrane	4.13	4.38
Niaproof	PM17 (E1-E4)	membrane, detergent, anionic	1.30	0.61
Toxic cations				
Thallium(I) acetate	PM13 (F9-D12)	toxic cation	2.49	0.60
Respiration/Ionophores/Uncouplers				
Pentachlorophenol	PM18 (C9-C12)	respiration, ionophore, H+	1.58	2.53
Oxidizing agents				
Plumbagin	PM18 (H9-H12)	oxidizing agent	4.10	4.15
Chelator, lipophilic				
Fusaric acid	PM14 (B5-B8)	chelator, lipophilic	>4.40	3.63
1-Hydroxy-Pyridine-2-thione	PM14 (C5-C8)	chelator, lipophilic	3.55	>4.40
5,7-Dichloro-8-hydroxyquinoline	PM15 (C1-C4)	chelator, lipophilic	3.46	4.36
5,7-Dichloro-8-hydroxyquinaldine	PM15 (B9-B12)	chelator, lipophilic	2.52	3.67

[a] IC_{50} values are expressed in well units and are defined as the well or fraction of a well at which a particular per-well parameter (*i.e.* the area of the curve) is at half of its maximal value over a concentration series.
[b] IC_{50} is reported only for compounds with which the difference between the areas of the kinetic curves of BCP1 and R7 strains was over 10,000 units in at least one of the concentrations tested, or over 7,000 units in at least two of the concentrations tested (only for hexachlorophene).

Table 3.6. A selection of protein functions annotated in *R. aetherivorans* BCP1 and *R. opacus* R7 genomes associated to oxidative stress response, detoxification and other stresses (RAST annotation analysis).

Subsystem [a]	Function	EC	Gene	BCP1 [a]	R7 [a]
Oxidation stress response					
Glutaredoxins	Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin) (Nitric oxide dioxygenase)	1.14.12.17	*hmp*	6	2
Glutathione analogs: mycothiol	Glycosyltransferase MshA involved in mycothiol biosynthesis	2.4.1.-	*mshA*	1	1
Glutathione analogs: mycothiol	L-cysteine:1D-myo-inosityl 2-amino-2-deoxy-α-D-glucopyranoside ligase MshC	6.3.1.13	*mshC*	1	1
Glutathione analogs: mycothiol	Maleylpyruvate isomerase, mycothiol-dependent	5.2.1.4	-	1	2
Glutathione analogs: mycothiol	Mycothiol S-conjugate amidase Mca	3.5.1.115	*mca*	1	1
Glutathione analogs: mycothiol	Mycothiol synthase MshD	2.3.1.189	*mshD*	2	1
Glutathione analogs: mycothiol	N-acetyl-1-D-myo-inosityl-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase MshB	3.5.1.103	*mshB*	1	1
Glutathione analogs: mycothiol	NAD/ mycothiol-dependent / S-nitrosomycothiol reductase MscR	1.1.1.306	*mscR*	1	1
Glutathione analogs: mycothiol	NADPH-dependent mycothiol reductase Mtr	1.8.1.15	*mtr*	1	1
Glutathione: Non-redox reactions	Hydroxyacylglutathione hydrolase	3.1.2.6	*gloB*	3	1
Glutathione: Redox cycle	Lactoylglutathione lyase	4.4.1.5	*gloA*	4	1
Glutathione: Redox cycle	Glutaredoxin-like protein NrdH, required for reduction of Ribonucleotide reductase class Ib	1.17.4.1	*nrdH*	1	2
Glutathione: Redox cycle	Glutathione peroxidase	1.11.1.9	*btuE*	1	1
NADPH:quinone oxidoreductase 2	NADPH:quinone oxidoreductase 2	1.6.5.3	*ndhA*	1	1
NADPH:quinone oxidoreductase 2	Redox-sensing transcriptional regulator QorR	-	*qorR*	1	1
Oxidative stress	Hydrogen peroxide-inducible genes activator	-	*oxyR*	1	1
Oxidative stress	Nitrite-sensitive transcriptional repressor NsrR	-	*nsrR*	1	1
Oxidative stress	Non-specific DNA-binding protein Dps	1.16.3.1	*dps*	2	1
Oxidative stress	Organic hydroperoxide resistance protein	1.11.1.15	*ohrB*	2	2
Oxidative stress	Organic hydroperoxide resistance transcriptional regulator	-	*ohrR*	2	2
Oxidative stress	Phytochrome, two-component sensor histidine kinase cyanobacterial phytochrome 1	2.7.3.-	*bvgs*	1	-
Oxidative stress	Redox-sensitive transcriptional activator SoxR	-	*soxR*	2	2
Oxidative stress	Redox-sensitive transcriptional regulator	-	*soxS*	1	1

Oxidative stress	Transcriptional regulator, FUR family	-	*fur*	2	2
Oxidative stress	Zinc uptake regulation protein ZUR	2.7.7.6	*zur*	1	1
Protection from Reactive Oxygen Species	Catalase	1.11.1.6	*katE*	3	6
Protection from Reactive Oxygen Species	Catalase/Peroxidase	1.11.1.21	*katG*	1	1
Protection from Reactive Oxygen Species	Superoxide dismutase [Cu-Zn]	1.15.1.1	*sodC*	1	1
Protection from Reactive Oxygen Species	Superoxide dismutase [Mn]	1.15.1.1	*sodA*	1	1
Detoxification					
D-tyrosyl-tRNA(Tyr) deacylase	D-tyrosyl-tRNA(Tyr) deacylase	3.1.-	*dtd*	1	1
Housecleaning nucleoside triphosphate pyrophosphatases	5-nucleotidase SurE	3.1.3.5	*surE*	1	1
Housecleaning nucleoside triphosphate pyrophosphatases	Nucleoside 5-triphosphatase RdgB	3.6.1.15	*rdgB*	1	1
Housecleaning nucleoside triphosphate pyrophosphatases/ Nudix proteins	Deoxyuridine 5'-triphosphate nucleotidohydrolase	3.6.1.23	*dut/ yncF*	1	1
Nudix proteins (nucleoside triphosphate hydrolases)	ADP-ribose pyrophosphatase	3.6.1.13	*nudF*	1	1
Nudix proteins (nucleoside triphosphate hydrolases)	NADH pyrophosphatase	3.6.1.22	*nudC*	1	2
Uptake of selenate and selenite	DedA protein	-	*dedA*	2	2
Uptake of selenate and selenite	Sulfate and thiosulfate import ATP-binding protein CysA	3.6.3.25	*cysA*	1	1
Uptake of selenate and selenite	Various polyols ABC transporter components	3.A.1.1.-	-	3	3
Other stresses					
Carbon Starvation	Carbon starvation protein A	-	*cstA*	1	1
No subcategory	Universal stress protein family	-	*uspA*	11	14

[a] The number of CDSs predicted to encode each protein function in BCP1 and R7 genomes is reported along with the functional categories defined by RAST.
[b] Either EC or TCDB number is reported (see Material and Methods for details)

3.5 *References*

Alvarez HM, Silva RA, Cesari AC, Zamit AL, Peressutti SR, Reichelt R, et al. (2004). Physiological and morphological responses of the soil bacterium *Rhodococcus opacus* strain PD630 to water stress. FEMS Microbiol Ecol., 50:75e86.

Alvarez HM, Steinbuchel A (2010). Physiology, biochemistry and molecular biology of triacylglycerol accumulation by *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus*. Microbiology monographs series. Heidelberg: Springer Verlag, p. 263e90.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. (2008). The RAST server: rapid annotations using subsystems technology. BMC Genomics, 9:75.

Bequer Urbano S, Albarracín VH, Ordoñez OF, Farías ME, Alvarez HM (2013). Lipid storage in high-altitude Andean Lakes extremophiles and its mobilization under stress conditions in *Rhodococcus* sp. A5, a UV-resistant actinobacterium. Extremophiles, 17:217e27.

Burian J, Yim G, Hsing M, Axerio-Cilies P, Cherkasov A, Spiegelman GB (2013). The mycobacterial antibiotic resistance determinant WhiB7 acts as a transcriptional activator by binding the primary sigma factor SigA (RpoV). Nucleic Acids Res, 41:10062e76.

Cai L, Liu G, Rensing C, Wang G (2009). Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. BMC Microbiol, 9:4.

Cholo MC, van Rensburg EJ, Osman AG, Anderson R (2015). Expression of the genes encoding the Trk and Kdp potassium transport systems of *Mycobacterium tuberculosis* during growth in vitro. BioMed Res Int, Article ID 608682.

Cisek AA, Rzewuska M, Witkowski L, Binek M (2014). Antimicrobial resistance in *Rhodococcus equi*. Acta Biochim Pol, 61:633e8.

Cotter PD, Hill C (2003). Surviving the acid test: responses of gram-positive bacteria to low pH. Microbiol Mol Biol Rev, 67:429e53.

Dabbs ER, Sole GJ (1988). Plasmid-borne resistance to arsenate, arsenite, cadmium, and chloramphenicol in a *Rhodococcus* species. Mol Gen Genet, 211:148e54.

de Carvalho CCCR (2012). Adaptation of *Rhodococcus erythropolis* cells for growth and bioremediation under extreme conditions. Res Microbiol, 163:125e36.

de Carvalho CCCR, Costa SS, Fernandes P, Couto I and Viveiros M (2014). Membrane transport systems and the biodegradation potential and pathogenicity of genus *Rhodococcus*. Front. Physiol., 5:133.

De la Fuente JL, Martin JF, Liras P (1996). New type of hexameric ornithine carbamoyltransferase with arginase activity in the cephamycin producers Streptomyces clavuligerus and *Nocardia lactamdurans*. Biochem J, 320:173e9.

Desomer J, Dhaese P, van Montagu M (1988). Conjugative transfer of cadmium resistance plasmids in *Rhodococcus fascians* strains. J Bacteriol, 170:2401e5.

Garg A, Soni B, Singh Paliya B, Verma S, Jadaun V (2013). Low-molecular-weight thiols: glutathione (GSH), mycothiol (MSH) potential antioxidant compound from actinobacteria. J App Pharm Sci, 3:117e20.

Ivshina IB, Kuyukina MS, Kostina LV (2013). Adaptive mechanisms of nonspecific resistance to heavy metal ions in alkanotrophic actinobacteria. Russ J Ecol, 44:123e30.

Kalkus J, Dorrie C, Fischer D, Reh M, Schlegel HG (1993). The giant linear plasmid pHG207 from *Rhodococcus* sp. encoding hydrogen autotrophy: characterization of the plasmid and its termini. J Gen Microbiol, 139:2055e65.

Kanamori T, Kanou N, Atomi H, Imanaka T (2004). Enzymatic characterization of a prokaryotic urea carboxylase. J Bacteriol, 186:2532e9.

Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res, 45(D1):D353-D361.

Kim IS, Foght JM and Gray RM (2002). Selective transport and accumulation of alkanes by *Rhodococcus erythropolis* S+14He. Biotechnol Bioeng, 80:650-659.

Kuyukina MS, Ivshina IB, Rychkova MI, Chumakov OB (2000). Effect of cell lipid composition on the formation of nonspecific antibiotic resistance in alkanotrophic rhodococci. Microbiology, 69:51e7.

Lamark T, Kaasen I, Eshoo MW, Falkenberg P, McDougall J, Strøm AR (1991). DNA sequence and analysis of the bet genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. Mol Microbiol, 5:1049e64.

LeBlanc JC, Gonçalves ER, Mohn WW (2008). Global response to desiccation stress in the soil actinomycete *Rhodococcus jostii* RHA1. Appl. Environ. Microbiol., 74:2627e36.

Li X, Zhang L, Wang G (2014). Genomic evidence reveals the extreme diversity and wide distribution of the arsenic-related genes in *Burkholderiales*. PLoS One, 9:e92236.

Martínková L, Uhnáková B, Pátek M, Nešvera J and Křen V (2009). Biodegradation potential of the genus *Rhodococcus*. Environment International, 35(1):162-177.

McGarvey KM, Queitsch K, Fields S (2012). Wide variation in antibiotic resistance proteins identified by functional metagenomic screening of a soil DNA library. Appl. Environ. Microbiol., 78:1708e14.

Meskys R, Harris RJ, Casaite V, Basran J, Scrutton NS (2001). Organization of the genes involved in dimethylglycine and sarcosine degradation in *Arthrobacter* spp.: implications for glycine betaine catabolism. Eur J Biochem, 268:3390e8.

Morlock GP, Metchock B, Sikes D, Crawford JT, Cooksey RC (2003). *ethA*, *inhA*, and *katG* loci of ethionamide-resistant clinical *Mycobacterium tuberculosis* isolates. Antimicrob Agents Chemother, 47:3799e805.

Orro A, Cappelletti M, D'Ursi P, Milanesi L, Di Canito A, Zampolli J, Collina E, Decorosi F, Viti C, Fedi S, Presentato A, Zannoni D, Di Gennaro P (2015). Genome and Phenotype Microarray analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: genetic determinants and metabolic abilities with environmental relevance. PLoS One, 10:e0139467.

Patrauchan MA, Miyazawa D, LeBlanc JC, Aiga C, Florizone C, Dosanjh M, et al. (2012). Proteomic analysis of survival of *Rhodococcus jostii* RHA1 during carbon starvation. Appl. Environ. Microbiol., 78: 6714e25.

Saier MH, Reddy VS, Tsu BV, Ahmed MS, Li C, Moreno-Hagelsieb G (2016). The Transporter Classification Database (TCDB): recent advances. Nucleic Acids Res, 44:D372–9.

Schmidt A, Haferburg G, Sineriz M, Merten D, Büchel G, Kothe E (2005). Heavy metal resistance mechanisms in actinobacteria for survival in AMD contaminated soils. Chem Erde Geochem, 65:131e44.

Shin J-H, Brody MS, Price CW (2010). Physical and antibiotic stresses require activation of the RsbU phosphatase to induce the general stress response in *Listeria monocytogenes*. Microbiology, 156:2660e9.

Van Vliet AHM, Kuipers EJ, Stoof J, Poppelaars SW, Kusters JG (2004). Acid-responsive gene induction of ammonia-producing enzymes in *Helicobacter pylori* is mediated via a metal-responsive repressor cascade. Infect Immun, 72:766e73.

Villegas-Torres MF, Bedoya-Reina OC (2011). Horizontal *arsC* gene transfer among microorganisms isolated from arsenic polluted soil. Int Biodeterior Biodegrad, 65:147e52.

Viti C, Decorosi F, Tatti E, Giovannetti L (2007). Characterization of chromate-resistant and -reducing bacteria by traditional means and by a high-throughput phenomic technique for bioremediation purposes. Biotechnol Prog, 23:553e9.

Wargo MJ, Szwergold BS, Hogan DA (2008). Identification of two gene clusters and a transcriptional regulator required for *Pseudomonas aeruginosa* glycine betaine catabolism. J Bacteriol, 190:2690e9.

Zúñiga M, Pérez G, González-Candelas F (2002). Evolution of arginine deiminase (ADI) pathway genes. Mol Phylogenet Evol, 25:429-44.

Chapter 4.

Genome-based analysis for the identification of genes involved in *o*-xylene degradation in *Rhodococcus opacus* R7

Bacteria belonging to the *Rhodococcus* genus play an important role in the degradation of many contaminants, including methylbenzenes. The present chapter aimed to identify genes involved in the *o*-xylene degradation in *R. opacus* strain R7 through a genome-based approach.

Among all the studied and aforementioned metabolisms in R7 strain (previous chapters), the *o*-xylene catabolism is one of the most interesting. The reason of the interest is related to R7 rare ability to grow in presence of this compound, that is the most recalcitrant xylene isomer. *o*-Xylene is an oily flammable liquid aromatic hydrocarbon, which is contained in the petroleum mixtures and it is used in several industrial productions (i.e. anti-freeze goods, biocides, lubricants, adhesives, plasters, fuels). This compound is considered dangerous for human and environment health, because of its toxicity and great persistence.

Bacteria able to degrade *m*-xylene or *p*-xylene under aerobic conditions have been well known; at the opposite, bacteria able to degrade *o*-xylene are not so common and their metabolic pathways have been poorly examined. In literature are described in *Pseudomonas* and *Rhodococcus* strains three possible routes of the *o*-xylene degradation. These pathways involve oxygenases, therefore in this chapter is described the genome-based approach used to examine all the R7 oxygenases, in order to investigate their putative involvement in this pathway. Moreover, functional analysis such as, mutant analysis, cloning experiments, identification of metabolites, besides gene expression analyses allowed to validate the hypothesis deriving from the bioinformatic investigations.

4.1 Introduction

Methylbenzenes are pollutants of great relevance for their toxic properties and their wide spread in environment, commonly present in crude petroleum and in various industrial processes (Barbieri et al., 1993). Different methylbenzenes, including the three xylene isomers, can be degraded by several bacterial strains, with a degradation pathway that depends on the position of the methyl groups on the aromatic ring (Kim et al., 2002). These bacteria can be divided into two groups: i) microorganisms that can degrade both *m*- and *p*-xylene; and ii) microorganisms that can only degrade the *ortho* isomer. The two degradation pathways are rarely found simultaneously in the same microorganism; *Rhodococcus* sp. YU6 is an exception because it is able to degrade both the *ortho* and *para* isomers of xylene (Jang et al., 2005). Although bacteria capable of degrading *m*-xylene or *p*-xylene under aerobic conditions and their corresponding catabolic pathways have been well documented (Galli et al., 1992; Harayama et al., 1992; Williams et al., 1994), bacteria capable of degrading *o*-xylene are not so common and their metabolic pathways have been poorly investigated (Jindrovà et al., 2002; Schraa et al., 1987; Bickerdike et al., 1997; Vardar et al., 2004; Kim et al., 2004). Three *o*-xylene degradation pathways (leading to their corresponding catechols) have been reported (Barbieri et al., 1993; Bickerdike et al., 1997; Kim et al., 2010; Kukor et al., 1990) (**Figure 4.1**).

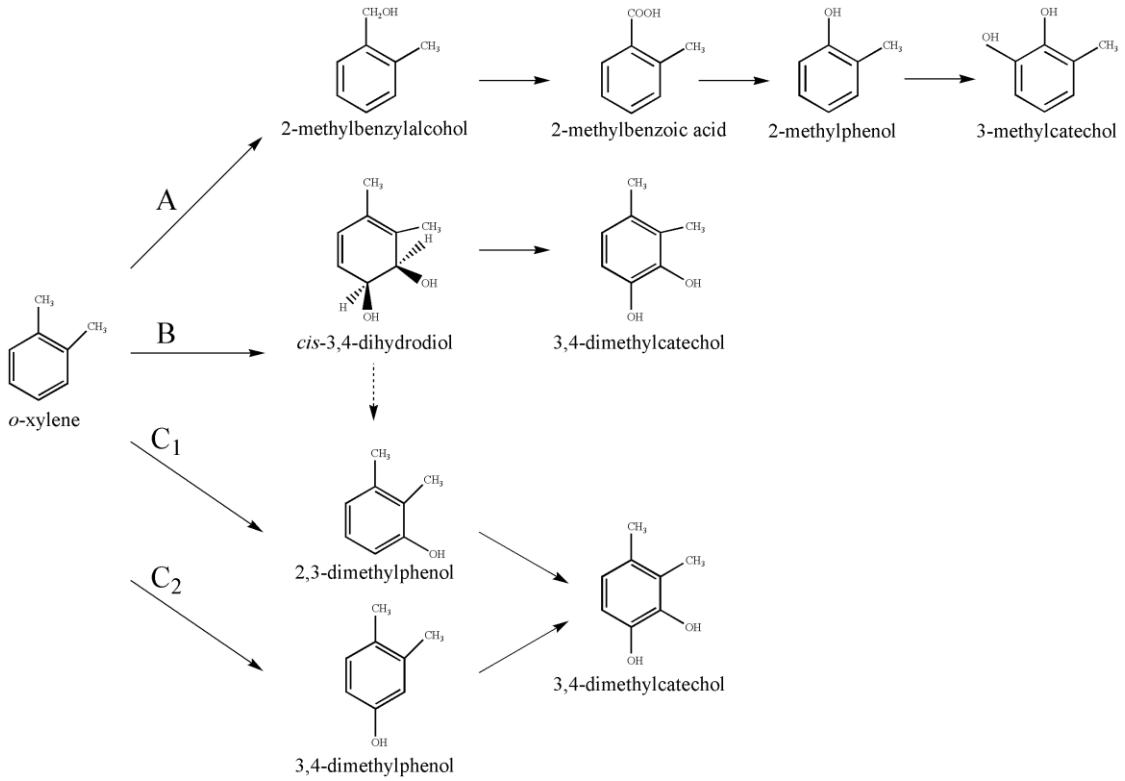


Figure 4.1. Overview of different *o*-xylene degradation pathways in bacteria. Monooxygenation of the methyl group and subsequent oxidations in *Rhodococcus* sp. strain B3 (**route A**); dioxygenation and subsequent dehydrogenation of the aromatic ring in *Rhodococcus* sp. strain DK17 (**route B**); two successive monooxygenation reactions to the aromatic ring in *P. stutzeri* strain OX1(**routes C1 and C2**). Dashed arrow indicates a spontaneous dehydration.

The first pathway is initiated by oxidation of the methyl group of *o*-xylene to form 2-methylbenzylalcohol, subsequently metabolized to form 3-methylcatechol as reported in the case of *Rhodococcus* sp. strain B3 (Bickerdike et al., 1997) (**Figure 4.1, Route A**). The second degradation pathway of *o*-xylene is initiated by a ring-hydroxylating 2,3-dioxygenase leading to the 3,4-dimethylcatechol in *Rhodococcus* sp. strain DK17 (**Figure 4.1, Route B**) (Kim et al., 2004). While the third pathway is initiated by a ring-hydroxylating monooxygenase at the different position 3, or 4, as performed by the Toluene *o*-xylene Monooxygenase (ToMo) in *P. stutzeri* strain OX1 (**Figure 4.1, Routes C1 and C2**), leading to the formation of 3,4-dimethylcatechol (Barbieri et al., 1993). One of the main roles in the degradation of many contaminants, including methylbenzenes, is played by bacteria belonging to the *Rhodococcus* genus. These bacteria, widely distributed in the environment, are characterized as a powerhouse of numerous degradation functions, since they are able to metabolize a wide range of organic molecules including aliphatic, aromatic and polycyclic aromatic compounds (PAHs), phenols, and nitriles (Martinkovà et al., 2009). In accordance with their immense catabolic diversity, *Rhodococcus* spp. are characterized to possess large and complex genomes, which contain a multiplicity of catabolic genes, a high gene redundancy of biosynthetic pathways and a sophisticated regulatory network (Alvarez et al., 2010). Many of them also possess a variety of large linear plasmids and smaller circular plasmids that contribute to and also explain the immense repertoire of catabolic abilities (Larkin et al., 2010). The most known example is represented by the genome of *R. jostii* strain RHA1 (McLeod et al., 2006), isolated for its ability to aerobically degrade polychlorinated biphenyls (PCBs) (Seto et al., 1995), and also able to utilize a wide range of compounds as sole carbon and energy source. Analyses of the 9.7 Mb large genome of RHA1 provided the evidence of catabolic pathway redundancy and horizontal gene transfer events (McLeod et al., 2006). To date, several gene clusters involved in the degradation of multiple aromatic compounds have been identified from genome analysis of several *Rhodococcus* spp. strains, including genes for biphenyl (Sakai et al., 2003), isopropylbenzene and ethylbenzene (Kesseler et al., 1996) and methylbenzenes (Kim et al., 2002). However, few not in depth genetic studies have been reported regarding the abilities of *Rhodococcus* strains to degrade *o*-xylene. The only data regarding the genes involved in *o*-xylene degradation in *Rhodococcus* strains derives from *R.* sp. strain DK17 (Kim et al., 2004) and *R. opacus* TKN14 (Maruyama et al., 2005). In *R. opacus* TKN14 a gene cluster homologous to the naphthalene-inducible dioxygenase system has been identified; while in *R.* sp. strain DK17 a multicomponent *o*-xylene dioxygenase (*akb* genes) has been identified and characterized (Kim et al., 2004; Kim et al., 2003). This strain is able to grow on *o*-xylene (and toluene, ethylbenzene, isopropylbenzene) trough the *akb* genes; the DK17 *o*-xylene dioxygenase is described to perform a ring-oxidizing pathway leading to the 3,4-dimethylcathechol formation either by a

deoxygenation or two monooxygenations, which can introduce two oxygen atoms successively. Thus, a deeper analysis concerning *o*-xylene degradation in *Rhodococcus* is necessary to understand which genes and enzymes could be involved in this metabolism. In this context, the metabolically versatile *R. opacus* R7 (Di Gennaro et al., 2001), known for its ability to grow on naphthalene, several long- and medium-chain n-alkanes, and aromatic hydrocarbons belonging to BTEX group (benzene, toluene, ethylbenzene and xylenes) (Di Gennaro et al., 2010; Zampolli et al., 2014), and also able to grow on *o*-xylene, can be used to add information of the metabolism of this compound. The whole genome of R7 strain was completely sequenced and it revealed to possess multiple genes for the degradation of a large set of aliphatic, aromatic and PAHs compounds (Orro et al., 2015). Moreover, the genome analysis revealed the presence, beside the chromosome, of five plasmids (pPDG1, pPDG2, pPDG3, pPDG4, pPDG5) that provided the evidence of high catabolic pathway redundancy. Through a genome-based analysis, the present work aimed to identify genes and molecular mechanisms involved in *o*-xylene degradation in *R. opacus* R7. Based on the previous identification of 2,3-dimethylphenol (2,3-DMP) and 3,4-dimethylphenol (3,4-DMP) in the R7 culture medium and the fact that these intermediates were metabolized by R7 strain, when supplied as the sole carbon and energy source, these compounds were identified as intermediates of R7 *o*-xylene degradation pathway (Di Gennaro et al., 2001; Di Gennaro et al., 2010). However, literature data suggested that the formation of dimethylphenol could be attributed to the dehydration of dihydrodiol deriving from the deoxygenation activity when *o*-xylene is supplied to *R.* sp. strain DK17 (Kim et al., 2002). For this reason, we searched and identified in the R7 genome all the sequences annotated as dioxygenases or monooxygenases/hydroxylases and clustered them into two different trees in order to select the oxygenases putatively involved in the *o*-xylene oxidation. Moreover, we demonstrated that the selected genes were involved in *o*-xylene degradation in *R. opacus* R7 by RT-PCR/qPCR, mutant analysis, cloning and expression experiments, thus revealing the complexity of this metabolic network.

4.2 Materials and methods

4.2.1 Bacterial strain and growth conditions

The bacterial strain used in this study is *R. opacus* R7, isolated for its ability to grow on naphthalene and *o*-xylene as previously described (Di Gennaro et al., 2001) (deposited to the Institute Pasteur Collection, CIP identification number 107348). The strain was grown at 30°C in M9 mineral medium (Maniatis et al. 1982), supplemented with the following carbon sources as only carbon and energy

source: *o*-xylene, toluene (final concentration of 1 g/l) or 2,3-dimethylphenol, 3,4-dimethylphenol (final concentration of 5-10 mM), or 2-dimethylbenzylalcohol, 3-dimethylbenzylalcohol or malate (final concentration of 10 mM). The *R. opacus* R7 growth on *o*-xylene, toluene, took place on M9 mineral medium in an atmosphere saturated with these aromatic compounds in a sealed system. The mutant R7-50 strain used in this paper was grown in the same conditions utilized for the wild type R7 strain. *Rhodococcus erythropolis* AP, isolated in our laboratory (CIP 110799) for its ability to grow on diesel fuel, was maintained on M9 mineral medium in a saturated atmosphere of diesel fuel at 30°C.

4.2.2 Bioinformatic analysis: nucleotide sequence determination and protein sequence analysis

The preliminary annotation of *R. opacus* R7 genome sequences was performed using the RAST (Rapid Annotation using Subsystem Technology) service (Aziz et al., 2008). BLASTn tool (Altschul et al., 1990) of NCBI pipeline was used to determine nucleotide sequence homology and to make manual curation. *R. opacus* R7 putative gene clusters for *o*-xylene catabolism were identified on chromosome and megaplasmids using BLAST tool and Clustal Omega (Thompson et al., 1994).
R. opacus R7 protein sequences were preliminary annotated using the RAST that allowed to identify potential monooxygenases/hydroxylases and dioxygenases using text string searching. These sequences annotated as monooxygenases/hydroxylases and dioxygenases were aligned separately against PDB (RCSB Protein Data Bank) database to identify reference sequences. Reference proteins were selected on the basis of the highest similarity or literature data. If no match was identified against PDB database, the same procedure was applied using BLASTp of NCBI pipeline. Afterwards, the identified reference sequences were aligned against R7 genome using the NCBI pipeline in order to verify to have considered all R7 putative monooxygenases/hydroxylases and dioxygenases. The retrieved sequences were aligned using the multiple sequence alignment (MSA) tool of Clustal Omega program using the default parameters (neighbour joining method, the Gonnet transition matrix, gap opening penalty of 6 bits, maintain gaps with an extension of 1 bit, used bed-like clustering during subsequent iterations, and zero number of combined iterations). For each group of oxygenases, the MSA was used for the cluster analysis inferred using the maximum likelihood (ML) method selected from the package MEGA version 6 (Tamura et al., 2013). The following parameters were used: JTT matrix, used all sites and gamma distribution of mutation rates with gamma optimized to 2. As a test of inferred phylogeny, 100 bootstrap replicates were used. The resulting groups allowed to define two different trees, one for all the dioxygenases and one for all the monooxygenases/hydroxylases of R7 showing clades with putative functions identified by InterPro/UniProt databases.

4.2.3 Preparation, analysis, and DNA manipulation

Total DNA from *R. opacus* R7 was extracted according to method reported by Di Gennaro et al. (Di Gennaro et al., 2010). The extract was precipitated by 0.1 volume of 3M sodium acetate and after centrifugation, the DNA was isolated and purified. Standard methods of DNA manipulation were used in this work (Maniatis et al., 1982). For the recovery and purification of DNA fragments from agarose, Extraction Kits by Machery-Nagel (Fisher Scientific, Germany) were used. Amplification of fragment containing genes target was achieved by PCR performed using primers designed *ad hoc* (**Table 4.1 and 4.2**) to amplify the sequences of interest.

Table 4.1. List of oligonucleotides used for RT-PCR and qRT-PCR

Oligonucleotide name	**Sequence (5' – 3')**	**Melting Temperature (T_m)**
27f	AGAGTTTGATCCTGGCTCAG	55°C
1495r	CTACGGCTACCTTGTTACGA	55°C
akbA1-GTG-for	GTGAATCCGCAGGACGGGTGG	51°C
akbA2-TGA-rev	TCAGAGGAAGAGGTTGAGATT	51°C
akbB-ATGfor	ATGGGTTGGCTGGAAGACAATG	54°C
akbB-TAGrev	CTAAAGGTCATTTCCGCCAGC	54°C
akbC-ATGfor	ATGACAAAAGTGACCGAGCTC	53°C
akbC-TAGrev	CTACTTGAGGGGGATATCCAAG	53°C
prmAf-TTG	TTGAGTAGGCAAAGCCTGACA	57°C
prmAr-TGA	TCAGGCCGGAACTGTGCCGCC	57°C
PheA1-ATG-for	ATGACCACCACCGAATCCGCC	60°C
PheA1-CTA-rev	CTAGCTGCGGCCGAAGTAGGA	60°C
Oligonucleotide name	**Sequence (5' – 3')**	**Melting Temperature (T_m)**
RT-16S-R7f	TCGTGAGATGTTGGGTTAAG	55°C
RT-16S-R7r	CCTCTGTACCGGCCATTGTAG	62°C
RT-AkbA1-f	ATATGATCTTGGACAATGAGG	54°C
RT-AkbA1-r	ATTCTCCATATCAATCTCGGG	56°C
RT-PrmA-f	AACATCTACCTGACCGTGGT	57°C
RT-PrmA-r	GCCATCAGCAGGATCGAATA	54°C
RT-PheA1-f	TGAAACTCGACTTCATCGCC	57°C
RT-PheA1-r	TATTCGAGCTTCGGGATGAC	57°C

Table 4.2. List of oligonucleotides used for Two-Step gene walking PCR amplification

Oligonucleotide name	Sequence (5' – 3')	Melting Temperature (T_m)
Walking_thio_1	GGAAAAGGACTGCTGTCGCTGCC	63°C
i-pTipThio-70-rev	CAAGGGGAAGTCGTCGCTCTCTGG	66°C
pTNR884-for	TTGGTAGCTCTTGATCCGGCAAAC	72°C
Walking_3-for	AACAACTGGCCGCCACC	68°C

4.2.4 RNA extraction and RT-PCR, quantitative real-time RT-PCR (qPCR)

Total RNA was extracted from bacterial cultures of *R. opacus* R7 (100 ml) grown at 30°C on M9 mineral medium supplemented with different substrates supplied (as described above) as the only carbon and energy source: *o*-xylene, toluene at the concentration of 1 g/l, 2,3-dimethylphenol and 3,4-dimethylphenol at the concentration of 0.5 g/l and 1 g/l malate used as reference.

RNA extraction protocol was performed using the RNA-Total RNA Mobio Isolation Kit (Qiagen Italia, Italy) according to the manufacturer's instructions and at the end the DNase treatment was performed. Reverse transcription was performed with iScript cDNA Synthesis kit (BIO-RAD, Italy) to obtain the corresponding cDNAs. For the cDNA synthesis 200 ng of total RNA was reverse-transcribed as follows: after denaturation for 5 min at 25°C, reverse transcription was performed for 1 h at 42°C and then 5 min of elongation at 85°C.

RT-PCR experiments were performed by amplification of the cDNA samples, each in 25-µl PCR volume containing 2 µl of the reverse-transcribed RNA samples.

Amplifications of the *akbA1A2*, *akbB*, *akbC*, *prmA*(P59), *pheA1a* (P115), *pheA3a* (P143) genes and 16SrDNA were performed using 2.5 U/µl of Long Range DNA Rabbit Polymerase (Eppendorf, Germany).

Thermo cycling conditions were as follows: 3 min at 95°C, 95°C for 30 sec, specific T_m for 45 sec, 72°C for 4 min, for 35 cycles; and 72°C for 3 min. Amplification of 16SrDNAwas performed using the universal bacterial primers 27f and 1495r as described in Di Gennaro et al. (Di Gennaro et al., 2001). The internal housekeeping gene (16S rDNA) was used as reference to evaluate relative differences in the integrity of individual RNA samples.

Quantitative real-time Reverse Transcriptase-PCR (qPCR) analyses were performed on the same samples used for RT-PCR. The reverse-transcribed samples were amplified using the StepOnePlus Real-Time PCR System (Applied Biosystem, Italy). Each 10-µl qPCR volume contained 4.4 µl of the reverse-transcribed RNA samples, 5 µl of PowerUp SYBR Green Master Mix (Applied

Biosystem, Thermo Scientific, Italy), and 300 nM of each primer, listed in **Table 4.1**. Thermocycling conditions were as follows: 30 sec at 95°C, followed by 40 cycles of 5 sec at 95°C, 10 sec at 60°C and 45 sec at 72°C and one cycle 15 sec at 95°C, 1 min at 60°C and 15 sec at 60°C. Expression of the housekeeping gene, 16SrDNA, was used as reference gene to normalize tested genes in *R. opacus* R7. The ΔΔCt method with 16S rDNA as reference gene was used to determine relative abundance of target transcripts in respect to malate as control. Data are expressed as mean ± standard deviation derived from at least three independent experiments. In order to exclude DNA contamination, negative controls were performed by omitting the reverse transcriptase in RT-PCR experiments, which were conducted with the same temperature program and the same primer sets for 35 cycles of amplification. The primers used in the RT-PCR analysis and qPCR are described in **Table 4.1**.

4.2.5 Mutant preparation

Transposon-induced mutagenesis in R. opacus R7 using IS1415 (pTNR-TA vector)

Plasmid pTNR-TA (Sallam et al., 2006) was transferred into *R. opacus* strain R7 by electroporation as described by Treadway et al. (Treadway et al., 1999), using a Gene Pulser II (BIO-RAD, Italy) set at 2.50 kV, 600 Ω, 25 μF in presence of maximum 1 μg DNA in a 2 mm-gap electro-cuvette (BIO-RAD, Italy). Afterwards, the electroporation mixture was suspended in 2.5 ml LB and it was incubated for 4 h at 30°C under shaking. Cells were plated on Luria-Bertani (LB) supplemented with 12.5 μg/ml thiostrepton and they were grown at 30°C for 5 days to select thiostrepton-resistant cells. Transposon-induced mutants were transferred to M9 mineral medium agar plates with 12.5 μg/ml thiostrepton and 10 Mm malate. The transposon-induced mutants obtained were tested on M9 mineral medium agar plates with the following carbon sources as the only carbon and energy source at the final concentration of 1 g/l: *o*-xylene; toluene; 2,3-dimethylphenol and 3,4-dimethylphenol (5-10 mM); 2-dimethylbenzylalcohol.

Analysis of pTNR-TA insertion sites

The genomic DNA of each transposon-induced mutant was extracted and the Two-Step gene walking PCR method was applied (Pilhofer et al., 2007). Insertions of IS1415 into the genomes of these mutants were confirmed by PCR using primers reported in **Table 4.2**. Genomic DNA of the wild type strain was used as a negative control. Homology searches of the interrupted DNA sequences from mutants were conducted by BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul et al., 1990).

4.2.6 Construction of the recombinant strain *R. erythropolis* AP (pTipQC2-*prmACBD*-R7)

The *prmACBD* gene cluster was ligated as *Nde*I/*Hind*III fragment into a shuttle-vector *E. coli-Rhodococcus*, pTipQC2 (Nakashima et al., 2004). The ligation mixture was used to transform *E. coli* DH5α by electroporation with standard procedures (Sambrook et al., 1989) and the recombinant clones were selected on LB agar supplemented with ampicillin (100 µg/ml) at 37°C. Ampicillin-resistant clones were selected and the recombinant plasmid (pTipQC2-*prmACBD*-R7) was isolated. The same recombinant plasmid was used to transform *R. erythropolis* strain AP by electroporation according to Zampolli et al. (Zampolli et al., 2014). Immediately after electroporation, 2.5 ml recovery broth (LB medium with 1.8% sucrose) were added and cells were incubated at 30°C for 4 h. Cells were plated on LB supplemented with chloramphenicol 50 µg/ml and grown at 30°C for 3-4 days. Recombinant strain *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) was used for bioconversion experiments in presence of *o*-xylene to evaluate the activity of the *prmACBD* system in comparison to the activity of the wild type strain.

4.2.7 Bioconversion experiments of the recombinant strain *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) in presence of *o*-xylene

Cells of *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) were grown in 500-mL Erlenmeyer flasks containing 100 mL of LB at 30°C. When the culture reached the $O.D._{600}$ of 0.6, they were induced with thiostrepton. After overnight growth, the cells were collected by centrifugation for 10 min at 20,000 g, washed and re-suspended in the mineral medium M9 to have final OD_{600} of 1.
Bioconversion experiments were performed in a biphasic system with *o*-xylene (1g/l) dissolved in isoctan at 20% of the final volume of the culture. At fixed times, 2 h, 4 h, 6 h, and 24 h, a flask was sacrificed to determine the metabolites produced during growth on *o*-xylene. It occurred by HPLC analysis after drawning the aqueous phase from the biphasic system. The aqueous phase was acidificatied with H_2SO_4 and subsequently extracted with ethyl acetate. After settling of the suspension, the aqueous phase was drawn from the organic phase (ethyl acetate). The organic phase was stripped under a gentle stream of nitrogen and resuspended in acetonitrile:water (50:50) in order to concentrate each culture 1:20 and filtered with 0.45 µm filters for HPLC analysis.

4.2.8 Analytical methods

HPLC analyses were performed with a Waters 600E delivery system (Waters Corporation, Milford, MA, USA) equipped with a Waters 486 UV-Vis detector and a Waters reverse-phase µBondapak 3.9

× 300 mm C18 column. The mobile phase was acetonitrile:water (50:50) at a flow rate of 1 mL/min in isocratic conditions. The detection was carried out at 254 nm and the retention times of the identified peaks were compared with those of authentic compounds. Co-elution experiments were also performed in which the culture broth samples were supplemented with authentic compounds. In these conditions, the retention times were 4.33 min for 2-methylbenzylalcohol, 6.39 min for 2,3-dimethylphenol, and 5.79 min for 3,4-dimethylphenol.

4.3 Results

4.3.1 R7 genome sequence analysis

The formation of 2,3-dimethylphenol and 3,4-dimethylphenol (and the oxidation to 2-methylbenzylalcohol) in the R7 *o*-xylene degradation pathway (Di Gennaro et al., 2001; Di Gennaro et al., 2010) suggests the involvement of monooxygenases able to oxidize *o*-xylene. However, the only *o*-xylene degradation pathway described in literature for bacteria of *Rhodococcus* genus is through the dioxygenase system of *R.* sp. DK17 leading to the corresponding dihydrodiol that could be dehydrated to DMPs (Kim et al., 2002). Therefore, we hypothesize that in the case of *R. opacus* strain R7 the formation of the identified intermediates could be explained with the involvement of different oxygenase systems for *o*-xylene degradation. For this reason, first we identified all the sequences annotated as dioxygenases or monooxygenases/hydroxylases in the R7 genome and clustered them into two different trees in order to select all the oxygenases putatively involved in *o*-xylene oxidation.

Analysis and clusterization of dioxygenases

A preliminary genome RAST annotation of R7 allowed the identification of 83 potential dioxygenases. Among those, only 57were selected as catalytic subunits of R7 dioxygenases. In order to cluster the R7 catalytic subunits, 22 reference sequences putatively involved in the aromatic compound degradation were considered. The generated tree revealed that these amino acid sequences are divided into eight clades (**Figure 4.2**).

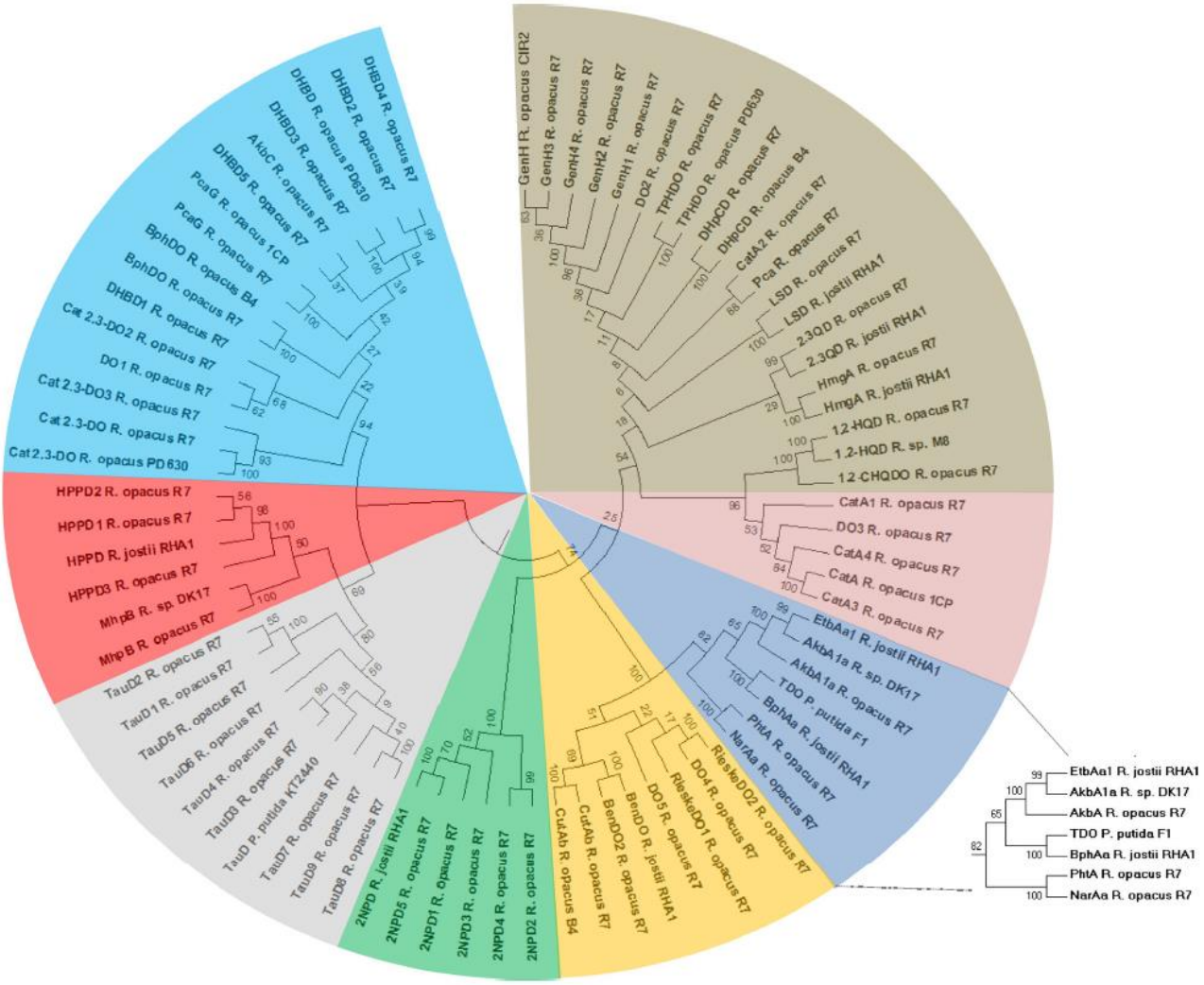


Figure 4.2. Phylogenetic tree of *R. opacus* R7 proteins containing the catalytic subunit of dioxygenases. The tree was constructed aligning selected reference protein sequences with R7 sequences. Numbers represent the bootstrap values on the branches calculated for maximum likelihood (ML) method selected from the package MEGA version 6 with 100 bootstraps. Color scheme for dioxygenases: brown, clade 1; pink, clade 2; blue, clade 3; yellow, clade 4; green, clade 5; grey, clade 6; red, clade 7; light blue, clade 8. The extended clade includes the AkbA1a of R. opacus R7 homologous to the AkbA1a of *R.* sp. DK17 involved in the oxidation of *o*-xylene. Abbreviation of dioxygenase names are reported in Table S1.

All the sequences are listed in **Table 4.3**, including the reference sequences with the relative source strain and all the sequences belonging to R7 strain.

Clade number 3 included dioxygenases putatively involved in the upper pathways of BTEX compounds and polycyclic aromatic hydrocarbons (**Figure 4.2, blue box**), while the other seven clades included all the dioxygenases putatively involved in the peripheral pathways of different aromatic compounds (**Figure 4.2**). Among the sequences in clade number 3 of the dioxygenase-tree (**Figure 4.2, extended clade**), the catalytic subunit of ethylbenzene dioxygenase (EtbAa1) of *R. jostii* RHA1 (Gonçalves et al., 2006) and the *o*-xylene dioxygenase (AkbA1a) of *R.* sp. DK17 (Kim et al., 2004), were shown to cluster near the only homologous dioxygenase sequence (AkbA1a) of R7.Multiple alignments of AkbA1a of R7 with proteins belonging to Bacterial Rieske non-heme iron

oxygenases revealed the coordination of the center iron-sulfur (Fe-S) (CxH - CxxH) with the amino acids that coordinate the iron atom of the active site (H - H - D). Moreover, the *akbA1a* gene encoding for the AkbA1a dioxygenase showed a nucleotide identity around 90% with the *etbAa1* of *R. jostii* RHA1 and *akbA1a* of *R.* sp. DK17. For these features, the AkbA1a was taken into consideration for its involvement in *o*-xylene catabolism. The remaining clades of the tree analysis included several dioxygenases involved in the mechanism of aromatic ring cleavage (e. g. gentisate 1,2-dioxygenases, catechol 1,2-dioxygenases, catechol 2,3-dioxygenases, homogentisate dioxygenase and protocatechuate dioxygenase), that are not taken into consideration.

Analysis and clusterization of monooxygenases/hydroxylases

From all the sequences derived from the whole genome of R7, the attention was also focused on putative sequences annotated as monooxygenases/hydroxylases. Then, a multiple amino acidic sequence alignment and a clusterization analysis were performed to predict protein functions using characterized monooxygenases/hydroxylases from different bacteria as reference (**Figure 4.3**). All the sequences are listed in **Table 4.4**, including the reference sequences with the relative source strain and all the sequences belonging to the R7 strain. The sequences obtained by this analysis are clustered into 10 clades.

Clade number 1included the phenol hydroxylasesP164, P165, P166, P167, P149 and P150 used as reference sequences. Amongst these sequences, solely the P149 of *R. opacus* 1CP (Gröning et al., 2014) was found to be similar to P115 (PheA1a) (98%), P122 (PheA2a) (92%), and P143 (PheA3a) (92%) of *R. opacus* R7. Accordingly, from this clade we selected the P115, P122 and P143sequences (encoded by the *pheA1a* gene, *pheA2a* gene, and *pheA3a* gene, respectively) for further molecular analysis. In clade number 4, we found only the P59 (PrmA) (encoded by the *prmA* gene) of R7 annotated as *alpha* chain methane monooxygenase component, which clustered with all the phenol hydroxylases/monooxygenases used as reference sequences. Among the reference sequences, the TouAcomponent (P151) of Toluene *o*-xylene Monooxygenase (ToMo) and the PhN component (P152) of Phenol Hydroxylase (PH) from *P. stutzeri* OX1, were mainly considered because they are the most phenol hydroxylases/monooxygenases described in literature for *o*-xylene oxidation (Bertoni et al., 1998; Cafaro et al., 2005; Notomista et al., 2003; Leahy et al., 2003). In fact, comparing the amino acid sequences of the TouA component (P151) of ToMo, the PhN component (P152) of PH, and of PrmA (P59), the most residues of the catalytic site of the three proteins were found to be conserved. Thus, P59 was selected to investigate its involvement in *o*-xylene degradation.

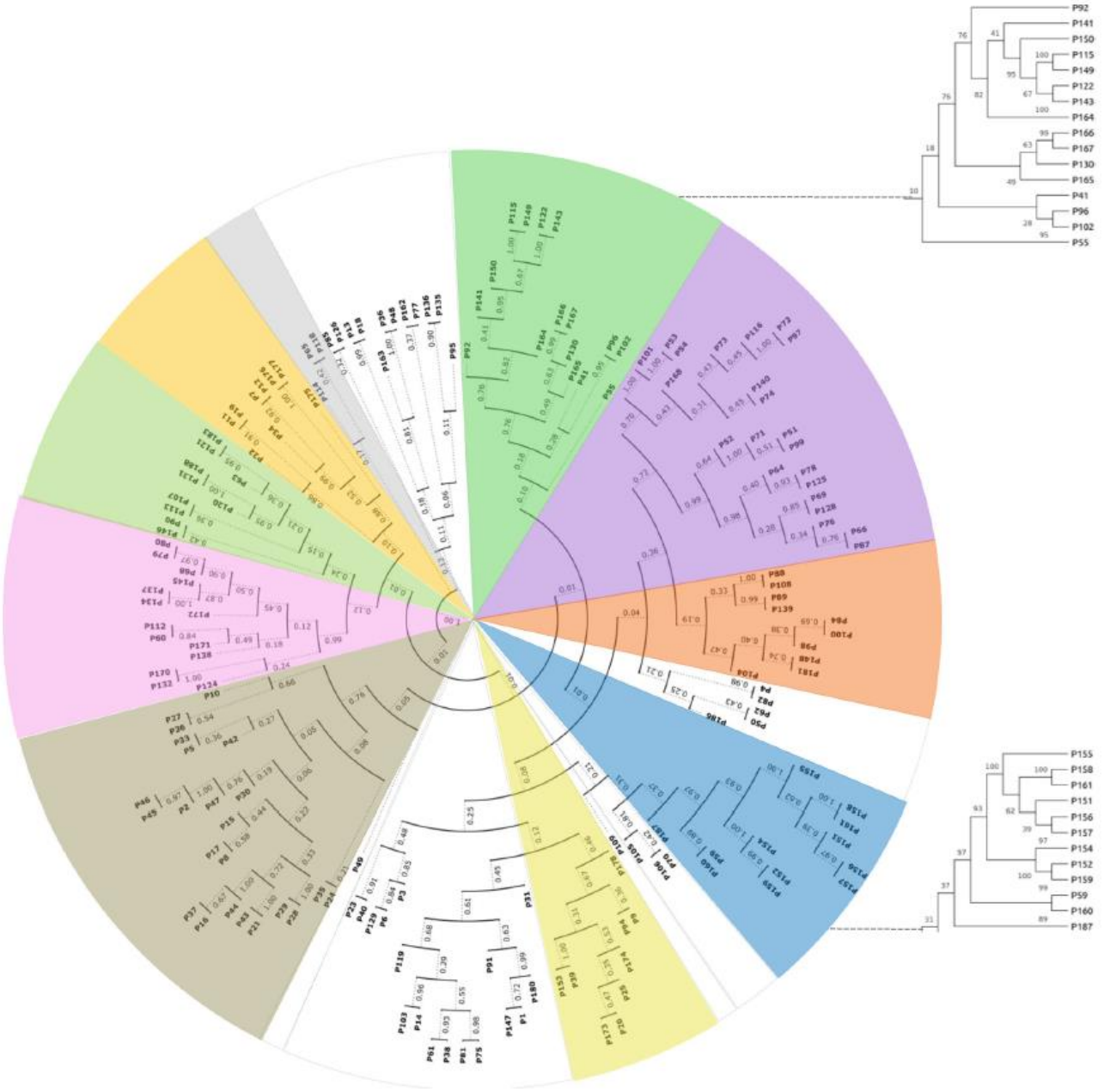


Figure 4.3. Phylogenetic tree of *R. opacus* R7 proteins containing the catalytic subunit of hydroxylases and monooxygenases. The tree was constructed aligning selected reference protein sequences with R7 sequences. Numbers represent the bootstrap values on the branches calculated for maximum likelihood (ML) method selected from the package MEGA version 6 with 100 bootstraps. Distinct clades are labeled with different colors: green, clade 1; purple, clade 2; red, clade 3; blue, clade 4; yellow, clade 5; brown, clade 6; pink, clade 7; light green, clade 8; orange, clade 9; grey, clade 10. White clades surrounded by grey lines correspond to not classified proteins: clade of not classified protein 1-6. The extended clade 1 includes the R7 proteins PheA1a, PheA2a, PheA3a; the extended clade 4 includes the R7 PrmA protein. The trees images were built with the ETE Toolkit using the circular plot function, with equal branch lengths and labeling each node with bootstrap support. Abbreviation of hydroxylase and monooxygenase names are reported in Table S2.

R7 proteins of clades number 2, 3 and 5, 6 including cyclohexanone monooxygenases (Yachnin et al., 2014), FAD-dependent monooxygenases (Jensen et al., 2014), salicylate hydroxylases (Montersino et al., 2012), and putative P450 hydroxylases respectively (Bernhardt et al., 2014), were

excluded from further molecular analysis in this paper, as little is known regarding their putative involvement in *o*-xylene oxidation.

Regarding the other clades number 7, 8, 9, 10 including, respectively, putative nitrilotriacetate monooxygenases (Uetz et al., 1992), alkanesulfonate monooxygenases (Eichhorn et al., 2002), other putative monooxygenases/hydroxylases, and ubiquinone monooxygenases (Gin et al., 2003), we excluded their involvement in the hydroxylation/monooxygenation of *o*-xylene, because their function seems to be very distant from our search.

Moreover, the proteins of the non-classified clades 1-6 were excluded because they were lacking in reference sequences.

4.3.2 Involvement of the *akb* genes in *o*-xylene degradation by RT-PCR experiments

Analyses of the R7 genome sequences evidenced the presence of the *akbA1a* gene in the *akb* gene cluster allocated on the megaplasmid pPDG5 (**Table 4.5**).This gene, coding for a large subunit dioxygenase component (AkbA1a), clustered with the following: the *akbA2a*gene coding for a small subunit dioxygenase component (AkbA2a), the *akbA3*gene for a ferredoxin component (AkbA3), the *HP* sequence for an hypothetical protein (HP) of unknown function, the *akbA4*gene for a reductase component (AkbA4), and the *akbB* gene coding for a dihydrodiol dehydrogenase (AkbB) (**Figure 4.4, Panel A**). Downstream (in the opposite direction) of these sequences, we found two sequences homologous (near the 80%) to the *akbS* and *akbT* sequences encoding for the sensor and regulator elements of DK17 strain, potentially involved in the regulatory mechanism. Moreover, we found a second group of genes (*akbCDEF* genes) coding for complete *meta*-cleavage enzymes of the lower pathway allocated on the pPDG2 plasmid, including: a *meta*-cleavage dioxygenase (AkbC), a *meta*-cleavage hydrolase (AkbD), an hydratase component (AkbE), and an aldolase (AkbF), respectively.

The involvement of the *akb* genes in the *o*-xylene degradation of R7 strain was analyzed by RT-PCR experiments. For this, RT-PCR were performed with RNA derived from *R. opacus* R7 cells grown in presence of *o*-xylene, or 2,3-DMP, or toluene, or malate as control.

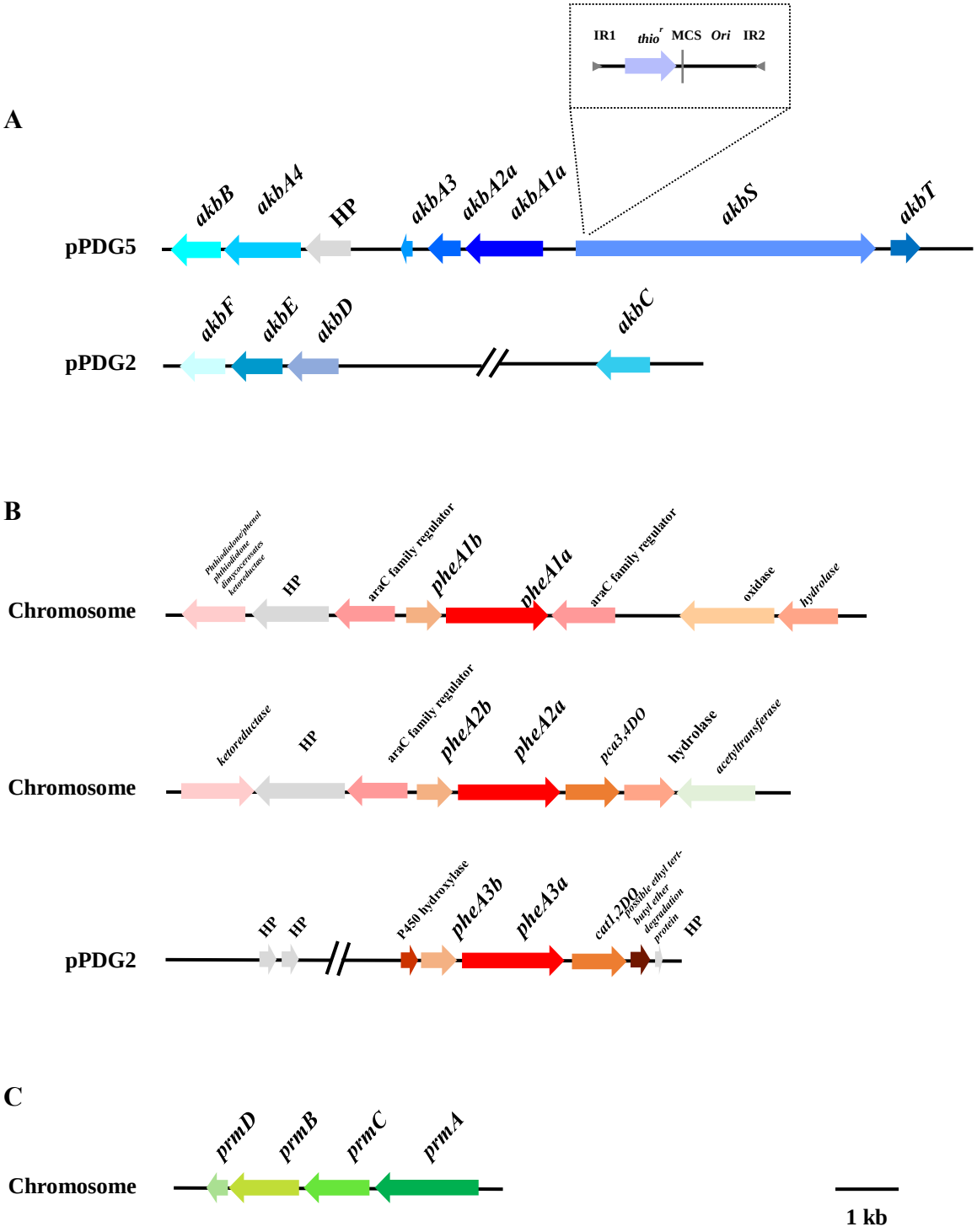


Figure 4.4. Genetic organization of the *akb*, *phe*, and *prm* genesin *R. opacus* R7. (**A**) On pPDG2 plasmid: *akbA1a*, large subunit of *o*-xylene dioxygenase; *akbA2a* small subunit of *o*-xylene dioxygenase; *akbA3*, ferredoxin component; *akbA4*, reductase component; *akbB*, dihydrodiol dehydrogenase; *akbS*, sensor kinase; and *akbT*, response regulator. On pPDG5 plasmid: *akbC*, meta-cleavage dioxygenase, *akbD*, meta-cleavage hydrolase product, *akbE*, hydratase and *akbF*, aldolase. Dashed box reported the IS1415 insertion deriving from the transposon mutagenesis and dashed lines localized the insertion element within the *akbS* gene of R7. (**B**) Three *phe* gene clusters: *pheA1a* (P115), phenol hydroxylase, and pheA1b, phenol hydroxylase-reductase component located on the chromosome; *pheA2a* (P122), phenol hydroxylase, and *pheA2b*, phenol hydroxylase-reductase component located on the chromosome; *pheA3a* (P143), phenol hydroxylase, and *pheA3b*, phenol hydroxylase-reductase component located on the pPDG2 plasmid. (**C**) *prm* gene cluster located on the chromosome: *prmA* (P59), large hydroxylase subunit of a monooxygenase, *prmC*, small hydroxylase subunit of a monooxygenase, *prmB*, reductase component, and *prmD*, regulatory coupling protein. Genes with unknown or hypothetical functions were reported as HP. Identified genes (listed in **Table 1**) and their orientation are shown by arrow.

Separate cDNA synthesis reactions were performed, and cDNA was then amplified with primer pairs used to amplify the target genes. The target genes were *akbA1A2a*coding for the small and the large components of the dioxygenase, or *akbB* coding for the dihydrodiol dehydrogenase, or *akbC* coding for the *meta*-cleavage dioxygenase. RT-PCR analysis showed that the *akbA1A2a*, *akbB*, and *akbC* genes were transcribed in R7 cells after growth on *o*-xylene (or toluene) (**Figure 4.5, Panel A**). These results indicate that *o*-xylene induced the transcription of the *akb* genes, suggesting the involvement of a deoxygenation route for R7 *o*-xylene degradation. But, as described above, the analyses of intermediates revealed that 2,3-dimethylphenol and 3,4-dimethylphenol were non-inducers of the pathway. Indeed, RT-PCR experiments with the same *akbA1A2a*, *akbB*, and *akbC* genes, after growth in presence of 2,3-dimethylphenol and 3,4-dimethylphenol, did not show any amplification. This propelled us in the direction to search in the R7 genome for other sequences encoding for monooxygenases/hydroxylases and to demonstrate their subsequent involvement in alternative pathways for *o*-xylene degradation leading to dimethylphenols.

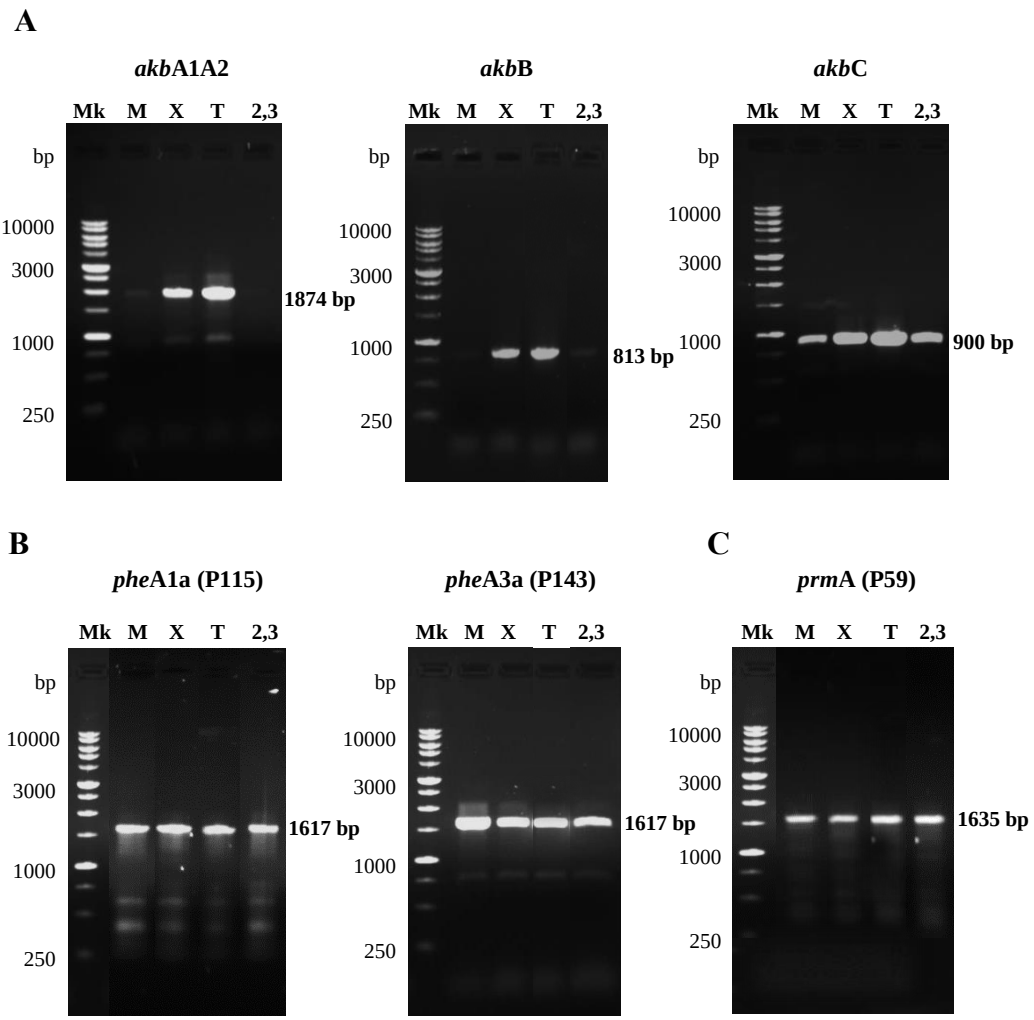


Figure 4.5. RT-PCR experiments from RNA extracted from *R. opacus* R7. RT-PCR experiments with *akbA1A2*, *akbB* and *akbC* genes (**A**); RT-PCR experiments with *pheA1a*, *pheA3a* (**B**); and with *prmA* genes (C**).** M 100 to 10,000-bp markers, M growth on mineral medium M9 and malate used as control, X growth on *o*-xylene, T growth on toluene, 2,3 growth on 2,3-dimethylphenol.

4.3.3 Involvement of the *akb* genes in *o*-xylene degradation by the identification of *R. opacus* R7 mutants in this cluster

Random mutagenesis performed after electroporation of R7 cells with the pTNR vector generated mutants of R7 unable to growth on *o*-xylene. We investigated the growth phenotypes and substrate transforming capabilities of R7 mutants by the transposon insertion detection. Among the single transposed mutant, the clearest phenotype was observed in the R7-50 mutant strain, in which the mutation was constituted by the insertion of the transposon in the *akbS* gene (**Figure 4.4, dashed box**). This strain was considered a leaky mutant for the growth on *o*-xylene, as it is reported in **Figure 4.6** (**Panel A** growth on malate, **Panel B** growth on *o*-xylene) in comparison to the R7 wild type strain. In fact, **Figure 4.6** displays a lower rate of growth for the mutant in respect to the wild type strain when grown on *o*-xylene, while there is a similar trend when both strains are grown on malate.

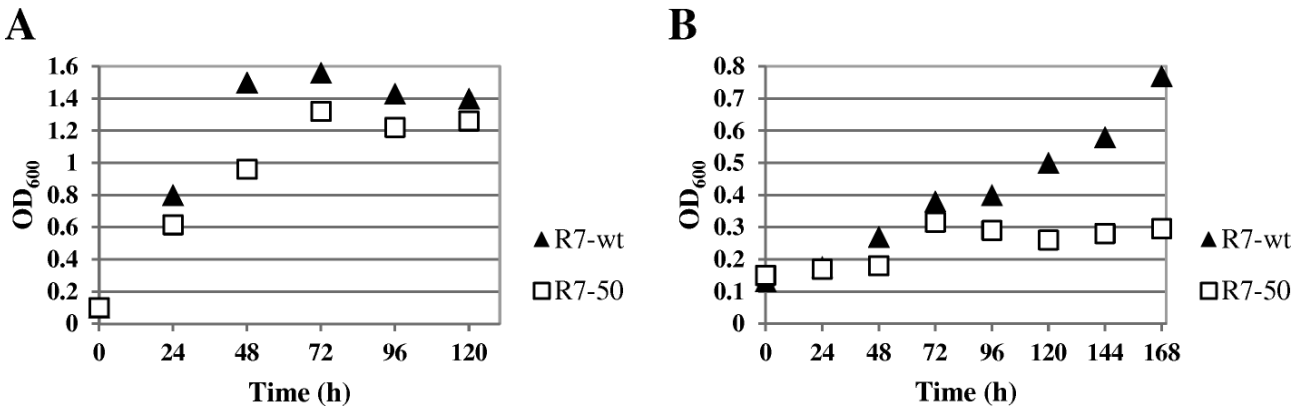


Figure 4.6. Growth curves on *o*-xylene in wild type *R. opacus* R7 and R7-50 mutant strain. Cells were grown in presence of M9 mineral medium and malate (**A**) or M9 mineral medium supplied in an atmosphere saturated with *o*-xylene (**B**).

These data are in accordance with what was observed by Kim et al. (Kim et al., 2005) when the ATP-binding motif of the sensor *akbS* gene was mutated in the DK17 strain. The mutation in the *akbS* gene allowed the incapacity of DK17 strain to grow well on *o*-xylene. So, our results indicate that *akbS* is necessary for the growth on *o*-xylene also in the R7 strain. Moreover, these data confirm the results of RT-PCR and suggest the kind of regulation involved in the *o*-xylene degradation pathway in the R7 strain. Indeed, this degradation process is likely mediated by the system sensor-regulator AkbS-AkbT through the binding of *o*-xylene.

4.3.4 Identification of the involvement of the *phe* genes and the *prm* genes in *o*-xylene degradation by RT-PCR experiments

Based on the identified sequences from genome analysis and the previous metabolic intermediates of R7 *o*-xylene degradation pathway, we analyzed the involvement of some putative sequences encoding for monooxygenases/phenol hydroxylases in this pathway. In particular, we selected the sequences deriving from clade number 1 (called *phe* sequences) and the sequences from clade number 4 (called *prm* sequences) (**Table 4.5**). We identified a first *phe* sequence (*pheA1a*) encoding for the monooxygenase PheA1a (P115) that showed a nucleotide identity of 98% with the sequences of the *pheA1(3)* gene of *R. opacus* 1CP, involved in the phenol hydroxylation (Groöning et al., 2014). In the R7genome, this gene (*pheA1a*) clustered with another gene (*pheA1b*) encoding for a phenol hydroxylase-reductase component and other open reading frames (ORFs) encoding for unknown functions (**Figure 4.4, Panel B**). From the same group of sequences, we also selected other two sequences, *pheA2a* (PheA2a) (P122) homologous to the *pheA1(2)* gene (99%) of *R. opacus* 1CP and *pheA3a* (PheA3a) (P143) homologous to the *pheA1(2)* gene (97%) of *R. opacus* 1CP. The *pheA2a* gene clustered with the *pheA2b* gene coding for a phenol hydroxylase-reductase component and a *pca* 3,4-dioxygenase; while the *pheA3a* gene clustered with the *pheA3b* gene coding for a phenol hydroxylase-reductase component and a catechol 1,2 dioxygenase. The last *phe* gene cluster was found allocated on the pPDG2 plasmid. We performed RT-PCR experiments on these identified sequences after growth of R7 cells in presence of *o*-xylene, or toluene, or 2,3-DMP, or malate as control. Separate cDNA synthesis reactions were performed, and cDNA was then amplified with primer pairs used to amplify the target genes. RT-PCR analysis showed that in R7 cells grown on *o*-xylene (or on toluene), both the *pheA1a* gene (P115) and the *pheA3a* gene (P143) were amplified as well as on malate (**Figure 4.5, Panel B**). Among the identified *phe* sequences, the *pheA2a*sequence (P122) was not tested because we decided to test only the main representatives *pheA1a* (P115) and *pheA3a* (P143) as they are allocated on the chromosome and on the pPDG2 plasmid, respectively. Moreover, as R7 strain was also able to oxidize *o*-xylene leading to the corresponding 2-methylbenzylalchol, and data on RHA1 strain indicated the presence of genes up-regulated on propane coding for components of ethylbenzene dioxygenase (Sharp et al., 2007), we decided to include the *prm* genes in the analysis (**Figure 4.4, Panel C**). The *prm* genes were found in a cluster constituted by the *prmA,C,B,D* genes, allocated in the chromosome with a percentage of amino acid identity near the 90% with the corresponding gene products of RHA1 strain. The *prmA* gene and *prmC* gene coded the large hydroxylase and the small hydroxylase subunits of a monooxygenase (annotated as propane monooxygenase), as *prmB* for the reductase component and *prmD* for the regulatory coupling protein, respectively. Also concerning these genes, we tested their involvement

in *o*-xylene degradation by expression of the *prmA* gene (P59) in presence of the same substrates reported above for *phe* genes (**Figure 4.5, Panel C**). This gene was expressed in the presence of *o*-xylene as well as on malate, toluene and DMPs. These results indicated that the *prm* genes were amplified similarly to the *phe* genes, suggesting that they can occur even when the strain is in absence of the hydrocarbon or phenols. Moreover, the gene redundancy of several monooxygenases/phenol hydroxylases supported the hypothesis of alternative pathways for *o*-xylene degradation in R7 strain. At the same time, the amplification of the PrmA (P59) could explain the formation of the corresponding 2-methylbenzyalchol.

Table 4.5. List of the identified genes.

Gene name	Protein name	Homologous protein	Bacterium	aa Identity	Function	Accession number	Reference
akbB **(813 bp)**	AkbB	AkbB	*R.* sp. DK17	85%	Dihydrodiol dehydrogenase	AII11489.1	Kim et al., 2004
akbA4 **(1263 bp)**	AkbA4	AkbA4	*R.* sp. DK17	81%	Ferredoxin reductase	AII11490.1	Kim et al., 2004
akbA3 **(150 bp)**	AkbA3	AkbA3	*R.* sp. DK17	69%	Ethylbenzene dioxygenase ferredoxin	CP008952.1	Kim et al., 2004
akbA2a **(549 bp)**	AkbA2a	AkbA2a	*R.* sp. DK17	84%	Ethylbenzene dioxygenase small subunit	AII11492.1	Kim et al., 2004
akbA1a **(1323 bp)**	AkbA1a	AkbA1a	*R.* sp. DK17	92%	Ethylbenzene dioxygenase large subunit	AII11493.1	Kim et al., 2004
akbS **(4812 bp)**	AkbS	AkbS	*R.* sp. DK17	76%	Sensor kinase	AII11494.1	Kim et al., 2004
akbT **(534 bp)**	AkbT	AkbT	*R.* sp. DK17	86%	Responseregulator	AII11495.1	Kim et al., 2004
akbF **(762 bp)**	AkbF	AkbF	*R.* sp. DK17	63%	4-Hydroxy-2-oxovalerate aldolase	AII11049.1	Kim et al., 2004
akbE **(900 bp)**	AkbE	AkbE	*R.* sp. DK17	64%	2-Hydroxypenta-2,4-dienoate hydratase	AII11050.1	Kim et al., 2004
akbD **(858 bp)**	AkbD	AkbD	*R.* sp. DK17	67%	2-Hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	AII11051.1	Kim et al., 2004
akbC **(900 bp)**	AkbC	AkbC	*R.* sp. DK17	87%	2,3-Dihydroxybiphenyl 1,2-dioxygenase	AII11058.1	Kim et al., 2004
pheA1a **(1617 bp)**	PheA1a	PheA1 (1) PheA1 (2) PheA1 (3)	*R. opacus* 1CP	83% 93% 98%	Phenol hydroxylase	AII08653.1	Gröning et al., 2014
pheA1b **(573 bp)**	PheA1b	PheA2 (1) PheA2 (2) PheA2 (3)	*R. opacus* 1CP	73% 100% 74%	Phenol hydroxylase reductase component	AII08654.1	Gröning et al., 2014
pheA2a **(1617 bp)**	PheA2a	PheA1 (1) PheA1 (2) PheA1 (3)	*R. opacus* 1CP	83% 100% 92%	Phenol hydroxylase	AII08806.1	Gröning et al., 2014
pheA2b **(561 bp)**	PheA2b	PheA2 (1) PheA2 (2) PheA2 (3)	*R. opacus* 1CP	69% 75% 94%	Phenol hydroxylase reductase component	AII08807.1	Gröninget al., 2014
pheA3a **(1617 bp)**	PheA3a	PheA1 (1) PheA1 (2) PheA1 (3)	*R. opacus* 1CP	83% 97% 91%	Phenol hydroxylase	AII10865.1	Gröning et al., 2014
pheA3b **(654 bp)**	PheA3b	PheA2 (1) PheA2 (2) PheA2 (3)	*R. opacus* 1CP	61% 64% 58%	Phenol hydroxylase reductase component	CP008949.1	Gröning et al., 2014

***prmA* (1635 bp)**	PrmA	PrmA	*R. jostii*RHA1	97%	Methane monooxygenase component A *alpha* chain	AII03499.1	Sharp et al., 2007
***prmC* (1044 bp)**	PrmC	PrmC	*R. jostii* RHA1	94%	Methane monooxygenase component C	AII03498.1	Sharp et al., 2007
***prmB* (1107 bp)**	PrmB	PrmB	*R. jostii* RHA1	97%	Methane monooxygenase component A *beta* chain	AII03497.1	Sharp et al., 2007
***prmD* (342 bp)**	PrmD	PrmD	*R. jostii* RHA1	98%	Methane monooxygenase regulatory protein	AII03496.1	Sharp et al., 2007

4.3.5 Quantitative real-time RT-PCR (qPCR) analysis

Quantitative real-time reverse transcription-PCR (qPCR) experiments were performed to quantify the levels of transcription of *akbA1a* (AkbA1a), *prmA* (P59) and *pheA1a* (P115)genes of R7 strain, representative of the selected catalytic subunit of different oxygenase systems putatively involved in *o*-xylene oxidation.qPCR experiments were performed after growth of R7 cells in presence of *o*-xylene, toluene and 2,3-DMPs or malate as control. The values of transcription after R7 malate-grown cells were used as a basal level for comparison with the quantities determined with the substrates of interest. The level of *akbA1a*gene was approximately 19 ± 7.5-fold higher after growth on *o*-xylene (with a similar trend on toluene) than on malate. On the other hand, this analysis confirms that *prmA* and *pheA1a*gene transcription levels increased much less, which probably reflects their constitutive expression. In fact, the transcription levels of *akbA1a* gene in respect to *prmA* and *pheA1a* genes after growth on *o*-xylene (and on toluene), were found to be significantly different, with respective values of 0.23 ± 0.04 and0.44 ± 0.11 (**Table 4.6**). A different trend was observed for the expression of the *prmA* and the *pheA1a* genes (*akbA1a* gene was not tested as it was not amplified in RT-PCR) after growth on 2,3-DMP. In this case, results showed an increase of the *pheA1a* transcription levels 5.18± 0.91-fold higher after growth on the corresponding dimethylphenol, whereas *prmA* was not induced. These results demonstrated that *o*-xylene was able to activate mainly the transcription of the*akbA1a* gene whilst a very low level of the other two genes during the aerobic growth of R7 cells on *o*-xylene. Meanwhile, in presence of 2,3-DMP a higher level of expression of phenol hydroxylase was observed.

Table 4.6. qPCR analysis. Relative gene expression of *R. opacus* R7 grown on *o*-xylene, toluene and 2,3-dimethylphenol (2,3-DMP). Values are means of three replicates ± standard deviation.

Substrate	**Normalized genes amount relative to malate condition ($2^{-\Delta\Delta Ct}$)**					
	akbA1a		***prmA***		***pheA1a***	
***o*-xylene**	19.14	±7.55	0.23	±0.04	0.44	±0.11
toluene	15.14	±6.56	1.53	±0.39	2.14	±0.65
2,3-DMP	-	-	0.04	±0.01	5.18	±0.91

4.3.6 Involvement of the *prm* genes by cloning and expression of the activity in *R. erythropolis* AP

In order to evaluate the role of the *prmACBD* gene cluster in the *o*-xylene metabolism, the region of 4.3 kb was isolated from R7 genomic DNA as *Nde*I/*Hind*III fragment. The PCR product was cloned into the shuttle-vector *E. coli-Rhodococcus* pTipQC2 to obtain pTipQC2-*prmACBD*-R7.

The recombinant plasmid pTipQC2-*prmACBD*-R7 was isolated from *E. coli* DH5a and transferred by electroporation into *R. erythropolis* AP, which was not able to use the *o*-xylene as only carbon and energy source. The *prmACBD* gene cluster was expressed under the inducible thiostrepton promoter (P*tipA*) through experiments with resting cells of *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) exposed to *o*-xylene to identify the metabolites. The activity of the recombinant strain was compared to the activity of wild type AP strain treated in the same conditions as control. *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) cells, which were pre-grown on LB and washed in mineral medium M9, were then exposed to *o*-xylene dissolved in isoctan in a biphasic system. The water phase was analyzed at different incubation times by reverse-HPLC analysis; 3,4-dimethylphenol and 2-methylbenzylalcohol were identified by comparison with reference compounds (standard mixture) (**Figure 4.7**). These compounds were observed in the first 2h of exposure, then they were progressively metabolized and disappeared from the organic phase after 6 h. It was not possible to confirm the formation of the 2,3-dimethylphenol. None of these metabolites was identified in the wild type host strain. These results suggested that the *prmACBD* gene cluster could have a role within the *o*-xylene metabolism, in particular in the first step of oxidation.

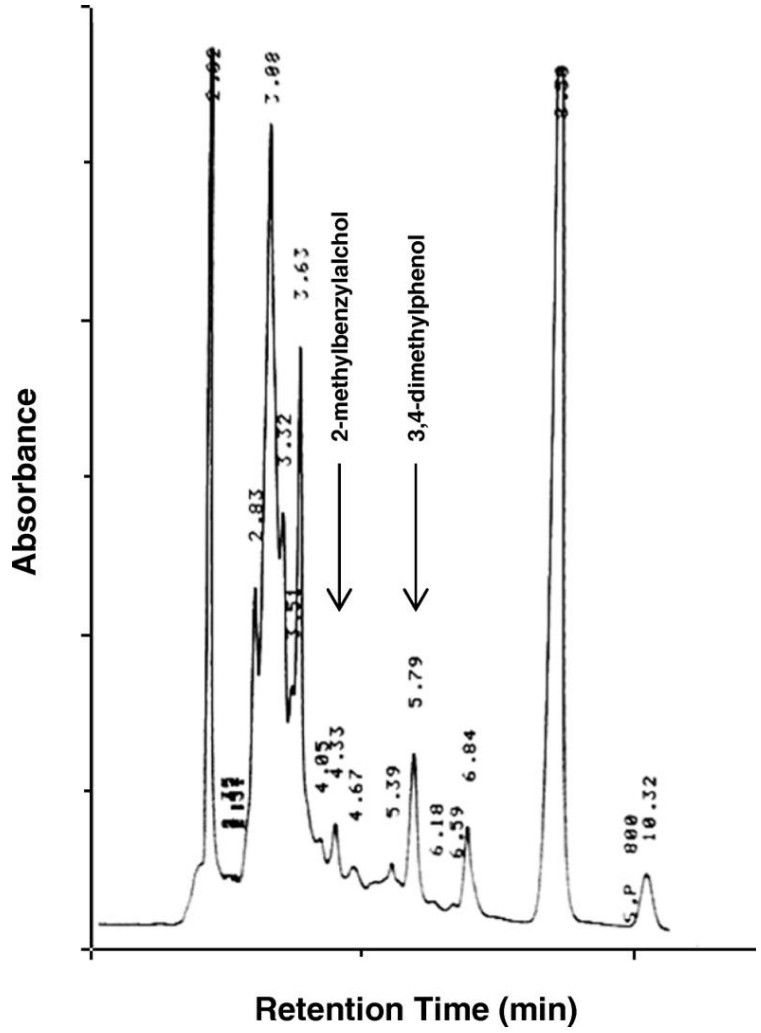


Figure 4.7. Reverse-Phase HPLC analysis of culture broth of *R. erythropolis* AP (pTipQC2-*prmACBD*-R7). The analysis was performed after exposition to *o*-xylene of the strain on the water phase of the biphasic system with isoctane. The retention times were 4.33 min for 2-methylbenzylalcohol, 5.79 min for 3,4-dimethylphenol.

4.4 Discussion

The genome-based analysis of *R. opacus* strain R7 revealed a considerable multiplicity of genes potentially involved in *o*-xylene catabolism. Although much is known about the ability of *Rhodococcus* strains to grow on toluene and ethylbenzene (Sharp et al., 2007; Patrauchan et al., 2007; Pathak et al., 2016), little is known about the catabolism of *o*-xylene in bacteria belonging to the *Rhodococcus* genus (Kim et al., 2010). *R. opacus* R7 is a strain isolated for its ability to grow on *o*-xylene as the only carbon and energy source. The strain is able to grow on *o*-xylene but not on *m*- and *p*-xylene. The inability of R7 strain to grow on the latter two compounds reinforced the hypothesis that the xylenes are metabolized at least by two different pathways (Kim et al., 2003). Moreover, we have previously identified (Di Gennaro et al., 2010) the 2,3- and 3,4-dimethylphenols, as the main

intermediates in the culture medium of R7 exposed to *o*-xylene, which are used by the strain as the only carbon and energy source, and not the corresponding dihydrodiol. Otherwise, in literature is reported by Kim et al. (Kim et al., 2004) that *o*-xylene is oxidized to the corresponding dihydrodiol. Moreover, Kim and co-authors also reported the direct formation of dimethylphenols in presence of *m*- and *p*-xylenes by *R.* sp. strain DK17. This suggests that alternative oxidation mechanisms of xylenes are possible (Kim et al., 2010). Whether this is through the action of a dioxygenase, forming a dihydrodiol, which dehydrates to a phenolic intermediate, or through the action of a monooxygenase which can directly hydroxylate the aromatic ring (or a combination of the two steps), it remains to be investigated.

In this context, a genome-based approach was used to better understand the R7 peculiar *o*-xylene pathway. Consequently, we decided to investigate the role of some selected genes and to demonstrate their involvement in this catabolism (**Figure 4.8**). As a first step we analyzed and clustered all the R7 genome oxygenase sequences generating two phylogenetic trees (**Figure 4.2 and 4.3**).

From the dioxygenase tree analysis, we selected the AkbA1a dioxygenase coded by the *akbA1a* gene (included in the *akb* gene cluster), whose sequences are 90% homologous to the sequence of the DK17 strain. In this paper, we demonstrated that the *akb* genes are induced by the presence of *o*-xylene supplied as the only carbon and energy source, both by RT-PCR/qPCR experiments and by selection of the R7-50 leaky mutant on *o*-xylene. RT-PCR analysis showed that *o*-xylene activated the transcription of the *akbA1A2* genes coding for a *o*-xylene dioxygenase and the *akbB* gene coding for a dihydrodiol dehydrogenase, suggesting the deoxygenation route for the *o*-xylene oxidation. These data confirmed what was reported for DK17 strain (Kim et al., 2004). However, these data are apparently in disagreement with what we observed in the R7 metabolic analysis (Di Gennaro et al., 2001), because we identified the 2,3- and 3,4-DMPs and no literature data supports enough that DMPs could derive from the dehydration of the corresponding dihydrodiol. Moreover, R7 strain was also able to grow on the corresponding 2,3- and 3,4-DMPs as the only carbon and energy source, suggesting an alternative pathway for the *o*-xylene oxidation through a monooxygenation. Since the focus of the present work was the identification of genes involved in the initial oxidation of *o*-xylene, we wanted to verify the formation mechanism of 2,3- and 3,4-DMPs from *o*-xylene. To support this hypothesis, we analyzed all the R7 monooxygenases/phenol hydroxylases sequences (**Figure 4.3**). As a result, on the basis of sequence identities with other *o*-xylene monooxygenases/phenol hydroxylases, genome sequences of other bacteria, and comparison of the protein catalytic site, we found putative sequences that could be involved in *o*-xylene degradation. In clade number 4, only one R7 protein sequence (PrmA) (P59) showed a significant aminoacid identity with respect to reference *o*-xylene monooxygenases (like Toluene *o*-xylene Monooxygenase, ToMo). As the ToMo is the best-

known monooxygenase able to oxidize *o*-xylene with the formation of the corresponding DMPs, we hypothesized that PrmA could also be involved in this monooxygenation/hydroxylation. Thus, we investigated the role of the *prmA* gene in the *o*-xylene (or toluene) degradation by RT-PCR and quantitative real time RT-PCR experiments on R7 cells grown in presence of *o*-xylene. Results showed that although the levels of the *prmA* gene increased little with respect to the growth on malate, it could be a first indication of the involvement of PrmA in this pathway. Then, the activity of the PrmACBD multicomponent monooxygenase was examined after the cloning of the corresponding genes in another *Rhodococcus* strain unable to use *o*-xylene like *R. erythropolis* AP. Results evidenced that *o*-xylene was slowly transformed in 3,4-DMP and 2-methylbenzylalcohol, that could be then metabolized by other monooxygenases/phenol hydroxylases. Indeed, R7 strain is able to grow also in presence of the DMPs when supplied as sole carbon and energy source. To support these data, we have also looked for R7 monooxygenases/phenol hydroxylases. From R7 genome analysis, we identified three sets of two component phenol hydroxylases, constituted of an oxygenase component and a reductase component; two sets were allocated on the chromosome on two different regions, and the third one on the pPDG2 plasmid of R7, respectively. In this case, *phe* genes were selected for RT-PCR experiments to evaluate their involvement in *o*-xylene oxidation. In all the growth conditions utilized we observed an amplification of the corresponding genes. The *pheA1a* encoding for the PheA1a (P115) was also tested in the quantitative real time RT-PCR on R7 cells grown in presence of *o*-xylene, toluene, 2,3-DMP and malate. Results showed a significant increase of *pheA1a* gene transcriptional levels during the growth on 2,3-DMP in respect to the growth on malate. This suggested that the *phe* genes could be involved in the second step of *o*-xylene degradation.

R7 strain showed a substrate versatility in respect to different substituted phenols, including 2,3-DMP and 3,4-DMP. This substrate versatility could likely be the result of gene redundancy and the presence of several phenol hydroxylase (iso)enzymes. These data are in accordance with literature, where it is reported that: three phenol hydroxylases in *R. opacus* B4, four in *R. opacus* M213, five in *R. opacus* PD630 and four in the reference strain *R. jostii* RHA1 have evident activities and expression profiles for this class of enzymes in these bacteria (Gröning et al., 2014).

Thus, considering such metabolic diversity of R7 strain, we would deduce that, although the *akb* genes are the specific activated genes for *o*-xylene degradation, other genes such as *prm* genes can induce an increase of levels of phenols that can converge towards the phenol oxidation route. The co-activation of multiple oxygenases could contribute to such strategy in these kinds of bacteria particularly resistant to environmental stress. Indeed, it has been demonstrated (Iino et al., 2012) that large genome with multiple broad-specificity catabolic enzymes such as those reported in RHA1 strain could have a competitive advantage in environmental changing soil conditions.

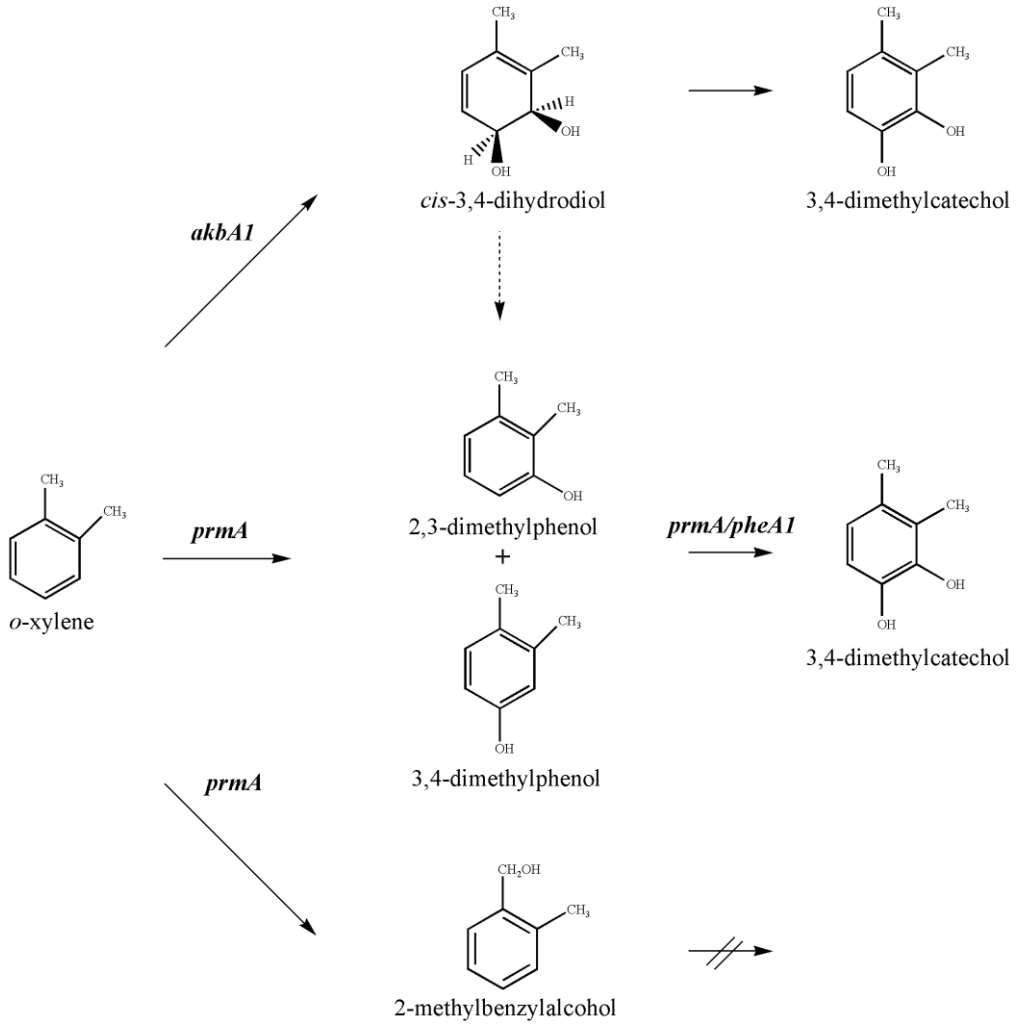


Figure 4.8. Proposed metabolic pathways involved in *o*-xylene degradation in *R. opacus* R7. *akbA1*, *o*-xylene dioxygenase route; *prmA*, monooxygenase route, *pheA1*, phenol hydroxylase route. Predicted genes are listed in Table 4.5. Dashed arrow indicates a spontaneous dehydration.

In conclusion, in this paper we demonstrate that *R. opacus* R7 is able to degrade *o*-xylene by the activation of the *akb* genes leading to the production of the corresponding dihydrodiol. Likewise, the redundancy of sequences encoding for several monooxygenases/phenol hydroxylases, supports the involvement of other genes that can induce an increase of levels of phenols that can converge towards the phenol oxidation route. The activation of multiple converging oxygenase systems represents a strategy in bacteria of *Rhodococcus* genus to degrade a wide range of recalcitrant compounds and to persist in severely contaminated environments.

Table 4.3. Sequences used to generate dioxygenase tree

Protein		Location on Genome	Function	Accession Number
C1	GenH		Gentisate 1,2-dioxygenase - *R. opacus* CIR2	BAD35142.1
	GenH3	pPDG1	Gentisate 1,2-dioxygenase (EC 1.13.11.4)	AII10780.1
	GenH4	pPDG4	Gentisate 1,2-dioxygenase (EC 1.13.11.4)	AII11451.1
	GenH2	chromosome	Gentisate 1,2-dioxygenase (EC 1.13.11.4)	AII09311.1
	GenH1	chromosome	Gentisate 1,2-dioxygenase (EC 1.13.11.4)	AII05708.1
	DO2	chromosome	Putative dioxygenase	AII08015.1
	TPHDO	chromosome	Tryptophan 2,3-dioxygenase (EC 1.13.11.11)	AII09387.1
	TPHDO		Tryptophan 2,3-dioxygenase (EC 1.13.11.11) - *R. opacus* PD630	EHI40897.1
	DHpCD	chromosome	2,3-Dihydroxy-*p*-cumate dioxygenase	AII08480.1
	DHpCD		2,3-Dihydroxy *p*-cumate dioxygenase - *R. opacus* B4	BAH52889.1
	CatA2	chromosome	Probable catechol 1,2-dioxygenase	AII07174.1
	Pca	chromosome	Protocatechaute dioxygenase	AII05947.1
	LSD	chromosome	Carotenoid oxygenase	AII05595.1
	LSD		Carotenoid oxygenase - *R. jostii* RHA1	WP_054247633.1
	2.3QD	chromosome	Quercetin 2,3-dioxygenase (EC 1.13.11.24)	AII07182.1
	2.3QD		Quercetin 2,3-dioxygenase (EC 1.13.11.24) - *R. jostii* RHA1	ABG95706.1
	HmgA	chromosome	Homogentisate 1,2-dioxygenase (EC 1.13.11.5)	AII08874.1
	HmgA		Homogentisate 1,2-dioxygenase (EC 1.13.11.5) - *R. jostii* RHA1	WP_073365188.1
C2	1.2-HQD	chromosome	Hydroxyquinol 1,2-dioxygenase	AII05734.1
	1.2-HQD		Hydroxyquinol 1,2-dioxygenase - *R.* sp. M8	WP_072634577.1
	1.2-CHQDO	chromosome	6-Chlorohydroxyquinol-1,2-dioxygenase	AII09320.1
	CatA1	chromosome	Catechol 1,2-dioxygenase (EC 1.13.11.1)	CP008947.1
	DO3	chromosome	Intradiol ring-cleavagedioxygenase	AII08811.1
	CatA4	pPDG2	Catechol 1,2-dioxygenase (EC 1.13.11.1)	AII10864.1
	CatA		Catechol 1,2-Dioxygenase - *R. opacus* 1CP	CAA67941.1
	CatA3	chromosome	Catechol 1,2-dioxygenase (EC 1.13.11.1)	AII08813.1
C3	EtbAa1		Ethylbenzene dioxygenase *alpha* subunit - *R. jostii* RHA1	BAC92712.1
	AkbA1a		Alkylbenzene dioxygenase - *R.* sp. DK17	AAR90131.2
	AkbA1a	pPDG5	Ethylbenzene dioxygenase large subunit	AII11493.1
	TDO		Toluene dioxygenase *alpha* subunit - *P. putida* F1	ABQ79012.1
	BphAa		Biphenyl 2,3-dioxygenase - *R. jostii* RHA1	ABG99107.1
	PhtA	pPDG2	Phthalate 3,4-dioxygenase *alpha* subunit	AII10987.1
	NarAa	pPDG4	Biphenyl dioxygenase *alpha* subunit (EC 1.14.12.18)	AII11432.1
C4	Rieske DO2	chromosome	Phenylpropionate dioxygenase and related ring-hydroxylating dioxygenases, large terminal subunit	AII09375.1
	DO4	pPDG1	Putative dioxygenase *alpha* subunit YeaW	AII10625.1
	Rieske DO1	chromosome	Ring hydroxylating dioxygenase, *alpha* subunit/Rieske (2Fe-2S) protein (EC 1.14.12.18)	AII08355.1
	DO5	pPDG2	Putative dioxygenase hydroxylase component	AII10950.1
	BenDO		Benzoate 1,2-dioxygenase *alpha* subunit - *R. jostii* RHA1	ABG94189.1
	BenDO2	chromosome	Benzoate 1,2-dioxygenase *alpha* subunit (EC 1.14.12.10)	AII08802.1

	CutAb	chromosome	*p*-Cumatedioxygenase	AII08474.1
	CutAb		*p*-Cumate dioxygenase - *R. opacus* B4	BAH52881.1
C5	2NPD2	chromosome	2-Nitropropane dioxygenase	AII07138.1
	2NPD4	chromosome	2-Nitropropane dioxygenase	AII08898.1
	2NPD3	chromosome	2-Nitropropane dioxygenase NPD	AII08140.1
	2NPD1	chromosome	2-Nitropropane dioxygenase (EC 1.13.11.32)	AII06420.1
	2NPD5	chromosome	2-Nitropropane dioxygenase (EC 1.13.11.32)	AII10318.1
	2NPD		2-Nitropropane dioxygenase - *R. jostii* RHA1	ABG92647.1
C6	TauD8	chromosome	Taurine dioxygenase	AII09447.1
	TauD9	chromosome	Taurine dioxygenase	CP008947.1
	TauD7	chromosome	Taurine dioxygenase	AII09095.1
	TauD		Taurine dioxygenase-dependent sulfonate dioxygenase - *P. putida* KT2440	NP_742398.1
	TauD3	chromosome	Taurine dioxygenase	AII05751.1
	TauD4	chromosome	Taurine dioxygenase	AII07329.1
	TauD6	chromosome	Taurine dioxygenase	AII08691.1
	TauD5	chromosome	Taurine dioxygenase	AII07522.1
	TauD1	chromosome	Taurine dioxygenase	AII03182.1
	TauD2	chromosome	Taurine dioxygenase	AII03183.1
C7	MhpB	pPDG2	3-Carboxyethyl catechol 2,3-dioxygenase (EC 1.13.11.16)	AII11046.1
	MhpB		3-Carboxyethyl catechol 2,3-dioxygenase - *R.* sp. DK17	WP_029538291.1
	HPPD3	chromosome	4-Hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	AII09141.1
	HPPD		4-Hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27) - *R. jostii* RHA1	WP_054245951.1
	HPPD1	chromosome	4-Hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	AII05993.1
	HPPD2	chromosome	4-Hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	AII07999.1
C8	Cat 2.3-DO		Catechol 2,3-dioxygenase (EC 1.13.11.2) - *R. opacus* PD630	EHI47566.1
	Cat 2.3-DO	chromosome	Catechol 2,3-dioxygenase (EC 1.13.11.2)	AII07256.1
	Cat 2.3-DO3	pPDG5	Catechol 2,3-dioxygenase (EC 1.13.11.2)	AII11500.1
	DO1	chromosome	2,3-Dihydroxybiphenyl 1,2-dioxygenase (EC 1.13.11.39)	AII06164.1
	Cat 2.3-DO2	pPDG2	Catechol 2,3-dioxygenase (EC 1.13.11.2)	AII11025.1
	DHBD1	chromosome	2,3-Dihydroxybiphenyl 1,2-dioxygenase	AII04187.1
	BphDO	chromosome	Biphenyl-2,3-diol 1,2-dioxygenase 2 (EC 1.13.11.39) (Biphenyl-2,3-diol 1,2-dioxygenase II) (23OHBP oxygenase II) (2,3-dihydroxybiphenyl dioxygenase II) (DHBD II)	AII03287.1
	BphDO		Biphenyl 2,3-dioxygenase - *R. opacus* B4	BAH52868.1
	PcaG	chromosome	Protocatechuate 3,4-dioxygenase *alpha* chain (EC 1.13.11.3)	AII09801.1
	PcaG		Protocatechuate 3,4-dioxygenase *alpha* chain (EC 1.13.11.3) - *R. opacus* 1CP	ANS29566.1
	DHBD5	chromosome	2,3-Dihydroxybiphenyl 1,2-dioxygenase	CP008947.1
	AkbC	pPDG2	1,2-Dihydroxynaphthalene dioxygenase	AII11058.1
	DHBD3	chromosome	2,3-Dihydroxybiphenyl 1,2-dioxygenase (EC 1.13.11.39)	AII06551.1
	DHBD		2,3-Dihydroxybiphenyl dioxygenase – *R. opacus* PD630	AHK28227.1
	DHBD2	chromosome	2,3-Dihydroxybiphenyl 1,2-dioxygenase	AII05250.1
	DHBD4	chromosome	2,3-Dihydroxybiphenyl 1,2-dioxygenase (EC 1.13.11.39)	AII08680.1

Table 4.4. Sequences used to generate monooxygenase tree

Protein			Location on Genome	Function	Accession Number
C1 – Phenol Hydroxylases and similar enzymes	P92	HpaB1	chromosome	4-Hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.2)	AII07219.1
	P141	HpaB2	pPDG2	4-Hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.2)	CP008949.1
	P150	PheA1		Phenol hydroxylase - *Rhodococcus erythropolis* UPV-1	ABS30825.1
	P115	PheA1a	chromosome	4-Hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.2)	AII08653.1
	P149	PheA1		(Chloro)Phenol hydroxylase - *Rhodococcus opacus* 1CP	CAT00453.1
	P122	PheA2a	chromosome	4-Hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.2)	AII08806.1
	P143	PheA3a	pPDG2	4-Hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.2)	AII10865.1
	P164	SgcC		4OO2 Chlorophenol-4-monooxygenase- *Streptomyces globisporus*	AAL06674.1
	P166	TftD		4G5E 2,4,6-Trichlorophenol 4-monooxygenase - *Cupriavidus pinatubonensis*	AAM55214.1
	P167	TftD		3HWC Chlorophenol 4-monooxygenase - *Burkholderia cepacia* AC1100	AAC23548.2
	P130	HpaB3	chromosome	4-Hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.2)	AII09405.1
	P165	HpaB		2YYG 4-Hydroxyphenylacetate 3-monooxygenase - *Thermus thermophilus*	YP_144226.1
	P41	FaH1P	chromosome	Fatty acid hydroxylase, putative	AII10087.1
	P96	DavB1	chromosome	Lysine 2-monooxygenase (EC 1.13.12.2)	AII07581.1
	P102	DavB2	chromosome	Lysine 2-monooxygenase (EC 1.13.12.2)	AII07957.1
	P55	HmoA	chromosome	Antibiotic biosynthesis monooxygenase	AII03486.1
C2 – Cyclohexanone monooxygenase s and similar enzymes	P101	CzcO1	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII07936.1
	P53	CzcO2	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII03309.1
	P54	CzcO3	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII03325.1
	P168	CzcO		4RG3 Cyclohexanone monooxygenase - *Rhodococcus* sp. HI-31	BAH56677.1
	P73	CzcO4	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII05723.1
	P116	CzcO5	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII08676.1
	P72	CzcO6	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII05694.1
	P97	CzcO7	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII07703.1
	P140	CzcO8	pPDG2	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII10811.1
	P74	CzcO9	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	CP008947.1
	P52	CzcO10	chromosome	4-Hydroxyacetophenone monooxygenase	AII03268.1
	P71	CzcO11	chromosome	4-Hydroxyacetophenone monooxygenase	AII05631.1
	P51	CzcO12	chromosome	4-Hydroxyacetophenone monooxygenase	AII03242.1

	P99	CzcO13	chromosome	4-Hydroxyacetophenone monooxygenase	AII07784.1
	P64	CzcO14	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII04180.1
	P78	CzcO15	chromosome	4-Hydroxyacetophenone monooxygenase	AII05843.1
	P125	CzcO16	chromosome	4-Hydroxyacetophenone monooxygenase	AII09069.1
	P69	CzcO17	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII05547.1
	P128	CzcO18	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII09304.1
	P76	CzcO19	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	CP008947.1
	P66	CzcO20	chromosome	4-Hydroxyacetophenone monooxygenase	AII05035.1
	P87	CzcO21	chromosome	Cyclohexanone monooxygenase (EC 1.14.13.22)	AII06778.1
C3 – FAD-dependent monooxygenases and similar enzymes	P88	EthA1	chromosome	FAD-containing monooxygenase EthA	AII06822.1
	P108	EthA2	chromosome	FAD-containing monooxygenase EthA	AII08233.1
	P89	Fmo1	chromosome	Flavin-containing monooxygenase	AII06850.1
	P139	Fmo2	chromosome	Flavin-containing monooxygenase	AII10303.1
	P84	Fmo3	chromosome	Monooxygenase, putative	AII06597.1
	P100	Fmo4	chromosome	FAD-dependent oxidoreductase	AII07843.1
	P98	Fmo5	chromosome	Monooxygenase	AII07779.1
	P148	Fmo6	pPDG3	FAD-dependent oxidoreductase	AII11264.1
	P181	Fmos		4USR Flavin-containing monooxygenase - *Pseudomonas stutzeri* NF13	EME00352.1
	P104	Dma	chromosome	FAD-dependent oxidoreductase	AII08101.1
NC 1	P4	DavB1	chromosome	Lysine N6-hydroxylase (EC:1.14.13.59)	AII04026.1
	P82	DavB2	chromosome	L-lysine 6-monooxygenase	AII06354.1
	P62	Fmo7	chromosome	Monooxygenase	AII04027.1
	P50	Fmo8	chromosome	Monooxygenase, FAD-binding	AII03165.1
	P186	HpaC		2ECR 4-Hydroxyphenylacetate 3-monooxygenase *Thermus thermophilus*	Q5SJP7.1
C4 – Methane monooxygenases and similar enzymes	P155	TmoA		Toluene monooxygenase hydroxylase - *Pseudomonas mendocina* KR1	AAS66660.1
	P158	TbuA1		Toluene-3-monooxygenase oxygenase - *Ralstonia pickettii* PKO1	AAB09618.1
	P161	PH		Phenol hydroxylase four component - *Ralstonia eutropha* JMP134	AAC77380.1
	P151	TouA		Toluene *o*-xylene monooxygenase - *Pseudomonas* sp. OX1	CAA06654.1
	P156	BmoA		Benzene monooxygenase - *Pseudomonas aeruginosa* JI104	BAA11761.1
	P157	TbhA		Toluene-3-monooxygenase oxygenase – *Burkholderia cepacia* AA1	AAB58740.1
	P154	TbmD		α subunit-terminal oxygenase component - *Pseudomonas* sp. JS150	AAA88459.1
	P152	PhN		Phenol hydroxylase - *Pseudomonas* sp. OX1	AAO47358.1
	P159	DmpN		Phenol Hydroxylase - *Pseudomonas putida* CF600	BAP28469.1
	P59	PrmA	chromosome	Methane monooxygenase component A *alpha* chain (EC 1.14.13.25)	AII03499.1
	P160	Mmox		Methane monooxygenase, A subunit - *Methylococcus capsulatus*	AAU92736.1

	P187	MmoH		1MHY Methane monooxygenase hydroxylase -*Methylosinus trichosporium*	ATQ70364.1
NC 2	P70	SsuD1	chromosome	Monooxygenase-luciferase	AII05616.1
	P106	LuxA1	chromosome	Alkanal monooxygenase *alpha* chain (EC 1.14.14.3)	AII08184.1
	P105	LuxA2	chromosome	Alkanal monooxygenase *alpha* chain (EC 1.14.14.3)	AII08179.1
NC 3	P109	-	chromosome	Putative monooxygenase	AII08275.1
C5 – Salicylate hydroxylases and similar enzymes	P178	Thi4		4Y4M Thiazole synthase *Methanocaldococcus jannaschii*	AAB98592.1
	P9	Sal1	chromosome	Salicylate hydroxylase (EC 1.14.13.1)	AII05373.1
	P94	Fmo9	chromosome	Aromatic ring hydroxylase	AII07390.1
	P174	Sal		5EVY Salicylate hydroxylase - *Pseudomonas putida*	BAA61829.1
	P25	Sal2	chromosome	Salicylate hydroxylase (EC 1.14.13.1)	AII07743.1
	P20	Sal3	chromosome	Salicylate hydroxylase (EC 1.14.13.1)	AII07112.1
	P173	NicC		5EOW 6-Hydroxynicotinic Acid 3-monooxygenase - *Pseudomonas putida* KT2440	AAN69538.1
	P39	Hba1	chromosome	Putative *n*-hydroxybenzoate hydroxylase	AII09308.1
	P153	Hba		Salicylate monooxygenase – *Rhodococcus jostii* RHA1	ABG93680.1
NC 4	P31	HpaB4	chromosome	4-Hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.2)	AII08627.1
	P180	HbpA		4CY6 2-Hydroxybiphenyl 3-monooxygenase - *Pseudomonas nitroreducens*	AAB57640.1
	P1	-	chromosome	Aromatic ring hydroxylase	AII03285.1
	P147	Fmo10	pPDG2	Monooxygenase, FAD-binding	CP008949.1
	P91	Fmo11	chromosome	2-polyprenyl-6-methoxyphenol hydroxylase	AII07178.1
	P75	Fmo12	chromosome	FAD-dependent monooxygenase	CP008947.1
	P81	Fmo13	chromosome	Monooxygenase	AII06159.1
	P38	MhaA1	chromosome	2-Polyprenyl-6-methoxyphenol hydroxylase and related FAD-dependent oxidoreductases	AII09237.1
	P61	Fmo14	chromosome	FAD-dependent oxidoreductases	AII03943.1
	P14	MhaA2	chromosome	Pentachlorophenol monooxygenase	AII06154.1
	P103	MhpA1	chromosome	3-(3-hydroxyphenyl)propinate hydroxylase	AII08000.1
	P119	Fmo15	chromosome	Aromatic ring hydroxylase	AII08743.1
	P3	MhpA2	chromosome	3-(3-Hydroxy-phenyl)propionate hydroxylase (EC 1.14.13.-)	CP008947.1
	P6	MhpA3	chromosome	3-(3-Hydroxy-phenyl)propionate hydroxylase (EC 1.14.13.-)	CP008947.1
	P129	MhpA4	chromosome	3-(3-hydroxyphenyl)propionate hydroxylase	AII09317.1
	P40	HpnW	chromosome	Probable aromatic ring hydroxylase	AII09332.1
	P23	MhaA3	chromosome	Pentachlorophenol monooxygenase	AII07569.1
NC 5	P49	-	chromosome	Putative dioxygenasehydroxylase component	AII10950.1

C6 – Putative cytochrome P450 hydroxylases	P24	CypX1	chromosome	Putative cytochrome P450 hydroxylase	AII07617.1
	P35	CypX2	chromosome	Putative cytochrome P450 hydroxylase	AII08744.1
	P28	CypX3	chromosome	Putative cytochrome P450 hydroxylase	AII08421.1
	P29	CypX4	chromosome	Putative cytochrome P450 hydroxylase	AII08505.1
	P21	CypX5	chromosome	Putative cytochrome P450 hydroxylase	AII07296.1
	P43	CypX6	chromosome	Putative cytochrome P450 hydroxylase	AII10559.1
	P44	CypX7	chromosome	Putative cytochrome P450 hydroxylase	AII10827.1
	P16	CypX8	chromosome	Putative cytochrome P450 hydroxylase	AII06401.1
	P37	CypX9	chromosome	Putative cytochrome P450 hydroxylase	AII08831.1
	P8	CypX10	chromosome	Putative cytochrome P450 hydroxylase	AII05330.1
	P17	CypX11	chromosome	Putative cytochrome P450 hydroxylase	AII06482.1
	P15	CypX12	chromosome	Putative cytochrome P450 hydroxylase	AII06399.1
	P30	CypX13	chromosome	Putative cytochrome P450 hydroxylase	AII08534.1
	P47	CypX14	chromosome	Putative cytochrome P450 hydroxylase	AII10868.1
	P2	CypX15	chromosome	Putative cytochrome P450 hydroxylase	AII03447.1
	P45	CypX16	chromosome	Putative cytochrome P450 hydroxylase	AII10849.1
	P46	CypX17	chromosome	Putative cytochrome P450 hydroxylase	AII10849.1
	P42	CypX18	chromosome	Putative cytochrome P450 hydroxylase	AII10517.1
	P5	CypX19	chromosome	Putative cytochrome P450 hydroxylase	AII04108.1
	P33	CypX20	chromosome	Putative cytochrome P450 hydroxylase	AII08657.1
	P26	CypX21	chromosome	Putative cytochrome P450 hydroxylase	AII07750.1
	P27	CypX22	chromosome	Putative cytochrome P450 hydroxylase	AII08206.1
	P10	CypX23	chromosome	Putative cytochrome P450 hydroxylase	CP008947.1
C7 – Nitrilotriacetate monooxygenases and similar enzymes	P124	NmoA1	chromosome	Nitrilotriacetate monooxygenase component A (EC 1.14.13.-)	AII08959.1
	P132	NmoA2	chromosome	Nitrilotriacetate monooxygenase component A (EC 1.14.13.-)	AII09422.1
	P170	Nmo		3SDO Nitrilotriacetatemonooxygenase - *Burkholderia pseudomallei*	ABA52898.1
	P138	NmoA3	chromosome	Nitrilotriacetatemonooxygenase component A (EC 1.14.13.-)	AII09502.1
	P171	LadA		3B9N Long-chain alkane monooxygenase - *Geobacillus thermodenitrificans*	ABO68832.1
	P60	NmoA4	chromosome	Nitrilotriacetate monooxygenase component A	AII03608.1
	P112	NmoA5	chromosome	Nitrilotriacetate monooxygenase component A	AII08556.1
	P172	EmoA		5DQP EDTA monooxygenase - *Chelativorans* sp. BNC1	AAG09252.1
	P134	NmoA6	chromosome	Putative nitrilotriacetate monooxygenase, component A	AII09433.1
	P137	NmoA7	chromosome	Probable nitrilotriacetate monooxygenase, component A	AII09445.1
	P145	NmoA8	pPDG2	Nitrilotriacetate monooxygenase component A (EC 1.14.13.-)	AII10963.1
	P68	NmoA9	chromosome	Nitrilotriacetate monooxygenase component A (EC 1.14.13.-)	AII05536.1
	P79	NmoA10	chromosome	Nitrilotriacetate monooxygenase component A (EC 1.14.13.-)	AII06110.1
	P80	-	chromosome	Putative monooxygenase	AII06121.1

C8 – Alkanesulfonate monooxygenases and similar enzymes	P146	SsuD2	pPDG2	Alkanal monooxygenase *alpha* chain (EC 1.14.14.3)	AII10970.1
	P90	Llm	chromosome	F420-dependent oxidoreductase	AII07075.1
	P113	AlkB	chromosome	Alkane-1-monooxygenase (EC 1.14.15.3)	AII08632.1
	P107	RutA	chromosome	Predicted monooxygenase RutA in novel pyrimidine catabolism pathway	AII08186.1
	P120	SfnG1	chromosome	Alkanesulfonate monooxygenase	AII08749.1
	P131	SfnG2	chromosome	Alkanesulfonate monooxygenase	AII09418.1
	P188	Lux		3RAO Luciferase-like Monooxygenase - *Bacillus cereus* ATCC 10987	AAS39998.1
	P63	SsuD3	chromosome	Alkanesulfonate monooxygenase (EC 1.14.14.5)	CP008947.1
	P121	SsuD4	chromosome	Alkanesulfonate monooxygenase (EC 1.14.14.5)	AII08753.1
	P183	SsuD		1M41 Alkanesulfonate monooxygenase - *Escherichia coli*	NP_415455.1
C9 – Putative monooxygenases/hydroxylases and similar enzymes	P22	-	chromosome	Putative hydroxylase	AII07316.1
	P11	-	chromosome	Putative hydroxylase	AII05855.1
	P19	-	chromosome	Putative hydroxylase	AII06974.1
	P34	-	chromosome	Hydroxylase	AII08679.1
	P7	-	chromosome	Hydroxylase	AII05251.1
	P12	-	chromosome	Hydroxylase	AII05997.1
	P176	DszC		3X0X Dibenzothiophene monooxygenase - *Rhodococcus erythropolis* D-1	BAQ25859.1
	P177	DszC		4NXL Dibenzothiophene monooxygenase - *Rhodococcus erythropolis*	AAU14822.1
	P175	DnmZ		4ZXV Acyl-CoA dehydrogenase - *Streptomyces peucetius*	ATW50550.1
C10– Ubiquinone monooxygenases and similar enzymes	P114	HmoA1	chromosome	Antibiotic biosynthesis monooxygenase	AII08639.1
	P65	UbiB1	chromosome	Ubiquinone biosynthesis monooxygenase	AII04651.1
	P118	UbiB2	chromosome	Ubiquinone biosynthesis monooxygenase	AII08725.1
NC 6	P85	HmoA2	chromosome	Antibiotic biosynthesis monooxygenase	AII06649.1
	P126	Fmo16	chromosome	Possible flavin binding monooxygenase	AII09070.1
	P13	EctD1	chromosome	Hydroxylase	AII06134.1
	P18	EctD2	chromosome	Hydroxylase	AII06497.1
	P163	Mmo		Oxidoreductase FAD-binding domain protein - *Pseudomonas* sp. Chol1	EKM96312.1
	P36	DdhD1	chromosome	2-Polyprenylphenol hydroxylase and related flavodoxin oxidoreductases / CDP-6-deoxy-delta-3,4-glucoseen reductase-like	AII08803.1
	P48	DdhD2	chromosome	2-Polyprenylphenol hydroxylase and related flavodoxin oxidoreductases / CDP-6-deoxy-delta-3,4-glucoseen reductase-like	CP008949.1
	P162	TomA1		Toluene/phenol hydroxylase – *Burkholderia cepacia* G4	AAK07409.1
	P77	AmoA	chromosome	Putative ammonia monooxygenase	AII05831.1
	P136	NmoA11	chromosome	Nitrilotriacetatemonooxygenase	AII09443.1
	P135	NmoA12	chromosome	Probable nitrilotriacetate monooxygenase, component A	AII09437.1
	P95	-	chromosome	Putative monooxygenase	AII07476.1

4.6 *References*

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

Alvarez HM (2010). Central metabolism of the species of the genus *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus*. Berlin Heidelberg: Springer. p. 91-108.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. (2008). The RAST server: rapid annotations using subsystems technology. BMC Genomics, 9:75.

Barbieri P, Palladino L, Di Gennaro P, Galli E (1993). Alternative pathways for *o*-xylene or *m*-xylene and *p*-xylene degradation in *Pseudomonas stutzeri* strain. Biodegradation, 4:71-80.

Bernhardt R and Urlacher VB (2014). Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. Appl Microbiol Biotechnol, 98:6185-6203.

Bertoni G, Martino M, Galli E, Barbieri P (1998). Analysis of the gene cluster encoding toluene/*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1. Appl Environ Microbiol. 64:3626-3632.

Bickerdike SR, Holt RA, Stephens GM (1997). Evidence for metabolism of *o*-xylene by simultaneous ring and methyl group oxidation in a new soil isolate. Microbiology. 143:2321-2329.

Cafaro V, Notomista E, Capasso P, Di Donato A (2005). Regiospecificity of two multicomponent monooxygenases from *Pseudomonas stutzeri* OX1: molecular basis for catabolic adaptation of this microorganism to methylated aromatic compounds. Appl Environ Microbiol. 71:4736-4743.

Di Gennaro P, Rescalli E, Galli E, Sello G, Bestetti G (2001). Characterization of *Rhodococcus opacus* R7, a strain able to degradate naphthalene and *o*-xylene isolated from polycyclic aromatic hydrocarbon-contaminated soil. Res Microbiol. 152:641-651.

Di Gennaro P, Terreni P, Masi G, Botti S, De Ferra F, Bestetti G (2010). Identification and characterization of genes involved in naphthalene degradation in *Rhodococcus opacus* R7. Appl. Microbiol. Biotechnol., 87:297-308.

Eichhorn E, Davey CA, Sargent DF, Leisinger T, Richmond TJ (2002). Crystal structure of *Escherichia coli* alkanesulfonate monooxygenase SsuD. J Mol Biol, 324:457-468.

Galli E, Silver S, Witholt E, editors. (199). *Pseudomonas*: molecular biology and biotechnology. Washington, DC: American Society for Microbiology. p. 268-276.

Gin P, Hsu AY, Rothman SC, Jonassen T, Lee PT, Tzagoloff A, Clarke CF (2003). The *Saccharomyces cerevisiae* COQ6 gene encodes a mitochondrial flavin-dependent monooxygenase required for coenzyme Q biosynthesis. J Biol Chem, 278:5308-25316.

Gonçalves ER, Hara H, Miyazawa D, Davies JE, Eltid LD, Mohn WW (2006). Transcriptomic assessment of isozymes in the biphenyl pathway of *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol., 72:6183-6193.

Gröning JAD, Eulberg D, Tischler D, Kaschabek SR, Schlömann M (2014). Gene redundancy of two-component (chloro)phenol hydroxylases in *Rhodococcus opacus* 1CP. FEMS Microbiol Lett, 361:68-75.

Harayama S, Kok M, Neidle EL (1992). Functional and evolutionary relationships among diverse oxygenases. Annu Rev Microbiol, 46:565-601.

Iino T, Wang Y, Miyauchi K, Kasai D, Masai E, Fujii T, et al. (2012). Specific gene responses of *Rhodococcus jostii* RHA1 during growth in soil. Appl. Environ. Microbiol., 78:6954-6962.

Jang JY1, Kim D, Bae HW, Choi KY, Chae JC, Zylstra GJ, Kim YM, Kim E (2005). Isolation and characterization of a rhodococcus species strain able to grow on *ortho*- and *para*-xylene. J Microbiol, 43(4):325-30.

Jensen CN, Ali ST, Allen MJ, Grogan G (2014). Exploring nicotinamide cofactor promiscuity in NAD(P)H-dependent flavin containing monooxygenases (FMOs) using natural variation within the phosphate binding loop. Structure and activity of FMOs from *Cellvibrio* sp. BR and *Pseudomonas stutzeri* NF13. J MolCatal B Enzym, 109:191-198.

Jindrovà E, Chocovà M, Demnerovà K, Brenner V (2002). Bacterial aerobic degradation of benzene, toluene, ethylbenzene and xylene. Folia Microbiol, 47:83-93.

Kesseler M, Dabbs ER, Averhoff B, Gottschalk G (1996). Studies on the isopropylbenzene 2,3-dioxygenase and the 3-isopropylcatechol 2,3-dioxygenase genes encoded by the linear plasmid of *Rhodococcus erythropolis* BD2. Microbiology, 142:3241-3251.

Kim D, Kim YS, Kim SK, Kim SW, Zylstra GJ, Kim YM, et al. (2002). Monocyclic aromatic hydrocarbon degradation by *Rhodococcus* sp. strain DK17. Appl. Environ. Microbiol., 6:3270-3278.

Kim D, Kim Y-S, Jung JW, Zylstra J, Kim YM, Kim S-K, et al. (2003). Regioselective oxidation of xylene isomers by *Rhodococcus* sp. strain DK17. FEMS Microbiol Lett, 223:211-214.

Kim D, Chae JC, Zylstra GJ, Kim YS, Kim SK, Nam MH, et al. (2004). Identification of a novel dioxygenase involved in metabolism of o-xylene, toluene, and ethylbenzene by *Rhodococcus* sp. strain DK17. Appl. Environ. Microbiol., 70:7086-7092.

Kim D, Chae JC, ZylstraGJ, Sohn HY, Kwon GS, Kim E (2005). Identification of two-component regulatory genes involved in *o*-xylene degradation by *Rhodococcus* sp. strain DK17. J Microbiol, 43:49-53.

Kim D, Choi KY, Yoo M, Choi JN, Lee CH, Zylstra GJ, et al. (2010). Benzylic and aryl hydroxylations of m-xylene by *o*-xylene dioxygenase from *Rhodococcus* sp. strain DK17. Appl Microbiol Biotechnol, 86:1841-1847.

Kukor JJ, Olsen RH (1990). Molecular cloning, characterization and regulation of a *Pseudomonas pikettii* PKO1 gene encoding phenol hydroxylase and expression of the gene in *Pseudomonas aeruginosa* PA01c. J Bacteriol, 172:4624-4630.

Larkin MJ, Kulakov LA, Allen CCR (2010). Genomes and Plasmids in *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus*. Berlin, Heidelberg: Springer. p. 73-90.

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

Maniatis T, Fritsch EF, Sambrook J (1982). Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory, Cold Spring Harbor.

Martìnkovà L, Uhnàkovà B, Pàtek M, Nesvera J, Kren V (2009). Biodegradation potential of the genus *Rhodococcus*. Environ Int, 35:162-77.

Maruyama T, Ishikura M, Taki H, Shindo K, Kasai H, Haga M, Inomata Y and Misawa N (2005). Isolation and characterization of *o*-xylene oxygenase genes from *Rhodococcus opacus* TKN14. Appl. Environ. Microbiol., 7705–7715.

McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, et al. (2006). The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA, 103:15582-15587.

Montersino S, van Berkel WJH (2012). Functional annotation and characterization of 3-hydroxybenzoate 6-hydroxylase from *Rhodococcus jostii* RHA1. Biochim Biophys Acta, 1824:433-442.

Nakashima N, Tamura T (2004). A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35°C. Biotechnol Bioeng, 86:136-148.

Notomista E, Lahm A, Di Donato A, Tramontano A (2003). Evolution of bacterial and archaeal multicomponent monooxygenases. J MolEvol, 56:435-445.

Orro A, Cappelletti M, D'Ursi P, Milanesi L, Di Canito A, Zampolli J, et al. (2015). Genome and phenotype microarray analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: genetic determinants and metabolic abilities with environmental relevance. PlosOne; doi:10.1371/journal.pone.0139467

Pathak A, Chauhan A, Blom J, Indest KJ, Jung CM, Stothard P, et al. (2016). Comparative genomics and metabolic analysis reveals peculiar characteristics of *Rhodococcus opacus* strain M213 particularly for naphthalene degradation. PlosOne; 17:1-32.

Patrauchan MA, Florizone C, Eapen S, Gòmez-Gil L, Sethuaman B, Fukuda M, et al. (2007). Roles of ring-hydroxylating dioxygenases in styrene and benzene catabolism in *Rhodococcus jostii* RHA1. J Bacteriol. 190:37-47.

Pilhofer M, Bauer AP, Schrallhammer M, Richter L, Ludwig W, Schleifer K-H, et al. (2007). Characterization of bacterial operons consisting of two tubulins and a kinesin-like gene by the novel Two-Step Gene Walking method. Nucleic Acids Research. doi:10.1093/nar/gkm836.

Sakai M, Miyauchi K, Kato N, Masai E, Fukuda M (2003). 2-Hydroxypenta-2,4-dienoate metabolic pathway genes in a strong polychlorinated biphenyl degrader, *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol., 69:427-433.

Sallam KI, Mitani Y, Tamura T (2006). Construction of random transposition mutagenesis system in *Rhodococcus erythropolis* using IS1415. J Biotechnol. 121:13-22.

Sambrook J, Russell DW (1989). Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory, Cold Spring Harbor.

Schraa G, Bethe BM, van Neerven ARW, Van Den Tweel WJJ, Van Der Wende E, Zehnder AJB (1987). Degradation of 1,2-dimethylbenzene by *Corynebacterium* strain C125. Antoine van Leewenhock, 53:159-170.

Seto M, Masai E, Ida M, Hatta T, Kimbara K, Fukuda M, et al. (1995). Multiple polychlorinated biphenyl transformation systems in the Gram-positive bacterium *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol., 61:4510-4513.

Sharp JO, Sales CM, LeBlanc JC, Liu J, Wood TK, Eltis LD, et al. (2007). An inducible propane monooxygenase is responsible for N-nitrosodimethylamine degradation by *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol., 73:6930-6938.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. MolBiolEvol, 30:2725-2729.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids Res, 22:4673-4680.

Treadway SL, Yanagimachi KS, Lankenau E, Lessard PA, Stephanopoulos G, Sinskey AJ (1999). Isolation and characterization of indene bioconversion genes from *Rhodococcus* strain I24. Appl. Microbiol. Biotechnol., 51:786-793.

Uetz T, Schneider R, Mario S, Egli T (1992). Purification and characterization of a two-component monooxygenase that hydroxylates nitrilotriacetate from "*Chelatobacter*" strain ATCC 29600. J Bacterio, 174:1179-1188.

Vardar G, Wood TK (2004). Protein engineering of toluene-*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1 for synthesizing 4-methylresorcinol, methylhydroquinone, and pyrogallol. Appl. Environ. Microbiol., 70:3253-3262.

Williams PA, Sayers JR (1994). The evolution of pathways for aromatic hydrocarbon oxidation in *Pseudomonas*. Biodegradation, 5:195-217.

Yachnin BJ, McEvoy MB, MacCuish RJD, Morley KL, Lau PCK, Berghuis AM (2014). Lactone-bound structures of cyclohexanone monooxygenase provide insight into the stereochemistry of catalysis. ACS Chem Biol, 9:2843-2851.

Zampolli J, Collina E, Lasagni M, Di Gennaro P (2014). Biodegradation of variable-chain-length *n*-alkanes in *Rhodococcus opacus* R7 and the involvement of an alkane hydroxylase system in the metabolism. AMB Express, 4:73.

Chapter 5.

Soil quality assessment by the identified sequences and degradative functions of *R. opacus* R7

This chapter focuses on the assessment of depuration potential of an hydrocarbon contaminated soil using a culture-independent approach in slurry-phase microcosms. Kinetic microcosm experiments were performed to optimize the condition to evaluate the degradation rate of the cyclohexane carboxylic acid (CHCA) in presence of autochthonous bacteria in comparison to bioaugmented microcosms, in presence of *R. opacus* strain R7. Microcosm experiments were established to monitor the degradation of a model naphthenic acid compound as cyclohexane carboxylic acid (CHCA) using marker sequences previously identified and selected for their involvement.

5.1 Introduction

Bioremediation processes are considered the best tool for decontamination of polluted natural ecosystems. In particular, *in situ* bioremediation is an approach particularly attractive for both environmental and economic reasons for the reclamation of polluted soils. Bioaugmentation has been proposed as the most promising option to reduce contamination in a relatively short time period under different circumstances, such as in acute events of pollution, when great masses of contaminants are spilled on pristine soils (Adams et al., 2015). As degrading strains can be poor survivors or lose catabolic activity when inoculated into mixed microbial ecosystems, bioaugmentation requires inocula that, besides degrading pollutants in pure culture, are also able to survive in non-sterile soil for long periods of time (Schwartz et al. 2000). In such situations, the absence of adaptation of the allochthonous microorganisms prevents effective biodegradation activity. Bioaugmentation also proved to be successful when there is the transfer of genetic information from the introduced donor strain to the competitive, indigenous bacterial population of soil; such transfer represents a valuable approach to broaden the soil biodegradation potential (Ruberto et al., 2005). In some situations, bioremediation can be limited to biostimulation techniques, when nutrients are added to balance the C:N:P ratio and they facilitate the growth of the hydrocarbon degrading members of the autochthonous microflora (Ruberto et al., 2005). Introduced and indigenous strains can be monitored using probes based on conserved housekeeping and catabolic genes. Genetic databases now include not only the well-known genes coding for degradative systems of Proteobacteria but also genes such as naphthalene dioxygenase and alkane hydroxylase of Gram-positive bacteria. These probes have widened the possibility of describing the biodiversity of degrading microbial communities in polluted environments (Cavalca et al., 2002).

Among emerging contaminants that are nowadays attracting more attention, naphthenic acids are recalcitrant compounds deriving from both natural and anthropogenic processes. They are naturally found in hydrocarbon deposits (petroleum, oil sands, bitumen, and crude oils) and they have also

widespread industrial uses. Naphthenic acids (NAs) are defined as a complex mixture of organic acids comprising both saturated aliphatic and alicyclic carboxylic acids (Whitby, 2010). They are the major contributors to the toxicity of oil sands-process affected waters (OSPW), produced during bitumen extraction and together with NAs release exert acute and chronic toxicity to a wide range of aquatic organisms, birds and mammals (Allen, 2008; Whitby, 2010). Their toxicity effects are a result of their surfactant properties because they can disrupt the lipid bilayer of membranes and change membrane properties (Franket al., 2008). The degree of toxicity is directly related to NA molecular structure; indeed, lower molecular weight acids (particularly acyclic compounds) often demonstrate higher toxicity because they can easily penetrate and disrupt cell membranes compared to higher molecular weight compounds (Jones et al., 2011). Nevertheless, oil sands residues ponds contain microbial communities able to biodegrade NAs (Lai et al., 1996). Several abiotic factors have proven to affect the NA biodegradation efficiency: temperature, oxygen, phosphate and nitrogen concentrations, contaminant structure along with the presence of additional nutrients (Lai et al., 1996; Clemente et al., 2004; Smith et al., 2008; Johnson et al., 2010; Demeter et al., 2014; 2015).

Members of the genera *Corynebacterium*, *Arthrobacter*, *Acinetobacter*, *Alcaligenes*, *Alkanivorax*, and *Pseudomonas* are able to utilize naphthenic acids (NAs) based on cyclohexane ring i.e. cyclohexanecarboxylic acid (CHCA) as carbon sources for growth (Blakley, 1974; 1978; Rho and Evans, 1975; Blakley and Papish, 1982; Whitby, 2010). In these strains, the specific metabolic intermediates were detected, suggestingtwo possible metabolic pathways for CHCA utilization: 1) aromatization of the cyclohexane ring and 2) activation as CoA thioester (Pelletier and Harwood, 2000; Wang et al., 2015; Iwaki et al., 2005). *Rhodopseudomonas palustris* is able to degrade CHCA through metabolic reactions involved in anaerobic metabolism of benzoate, implying the involvement of *bad/ali* gene products (*aliA*, *badH*, *badI*, *badJ* coding for a long-chain-fatty-acid-CoA ligase,2-hydroxycyclohexanecarboxyl-CoA dehydrogenase, naphthoate synthase and acyl-CoA dehydrogenase, respectively) (Pelletier and Harwood, 2000). Homologs to *bad/ali* genes were predicted to be involved in the peripheral ring-cleavage pathway for NA degradation also in *Cupriavidus gilardii* strain C3 (Wang et al., 2015). Instead, the *pob*A gene coding for a 4-hydroxybenzoate hydroxylase was described to be involved in CHCA metabolism and the flanking genes encoding for a regulator in *Arthrobacter* sp. strain ATCC51369 (Iwaki et al., 2005).

Recently, members of the genus *Rhodococcus* were isolated from OSPW microbial communities (Demeter et al., 2015; Golby et al., 2012; Koma et al., 2004). *Rhodococcus* spp. are known for their broad catabolic diversity and their extraordinary tolerance to various environmental stresses (Cappelletti et al., 2016). In Presentato et al. 2018, the capacity of *R. aetherivorans* BCP1 to grow on representative aliphatic and alicyclic NAs was evaluated. The identified *chcpa* cluster (**Figure 5.1**) is

transcriptionally induced in BCP1 cells utilizing cyclopentane carboxylic acid (CPCA) and cyclohexane carboxylic acid (CHCA) as the only carbon and energy source. On the contrary, the only *pob*A gene detected in BCP1 genome (Orro et al., 2015) was not transcriptionally induced during the growth on CHCA and CPCA, excluding the possible involvement of the aromatization pathway in their degradation by BCP1.

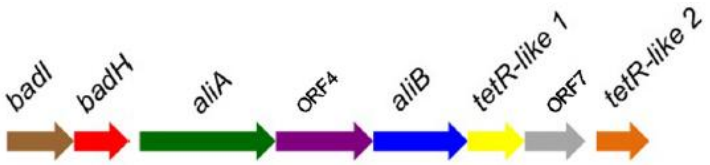


Figure 5.1 The organization of the *chcpca* gene cluster in *R. aetherivorans* BCP1 (Adapted from Presentato et al. 2018)

R. opacus R7 was tested for its ability to grow on CHCA as only carbon and energy source chosen as model compound among naphthenic acids. A genome-based approach was used to identify *R. opacus* R7 *chcpa* gene cluster putatively involved in the NAs degradation. Therefore, a gene expression analysis was used to evaluate the transcriptional induction of R7 genes after growth on CHCA, in order to obtain marker sequences to monitor bacteria degradation. Microcosms experiments were used to assess the depuration potential of an hydrocarbon contaminated soil using a culture-independent approach in slurry-phase system.

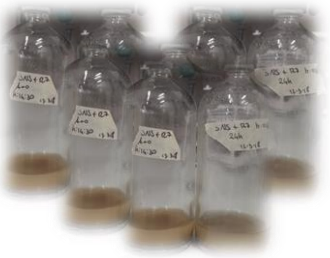
5.2 Materials and Methods

5.2.1 Characterization of sand sample

The sand was collected from a non-contaminated side of a river. The sand was subsequently dried at room temperature and passed through a sieve to obtain < 500 µm grain fraction after removal of visible plant debris and stones. It was homogenized and stored at room temperature. Sand humidity was measured evaluating the dry weight before and after heating sand samples at 105°C for 4h; whilst, the total organic carbon (TOC) at 550°C for 4h.

The total heterotrophic and CHCA degrading bacteria were evaluated using colony-forming unit counting method diluting the soil bacteria coltures in a solution of mineral medium M9. Each dilution was plated on Luria-Bertani (LB) agar medium or M9 mineral medium (Maniatis et al., 1982) supplemented with CHCA (0.5 g/l) as the only carbon and energy source. The plates were incubated at 30°C.

5.2.2 Preparation of slurry microcosms



Microcosms were established as slurry phase with 12.5 g of sand (particle size <500 µm) transferred to a 250-ml glass bottle with screw cap (previously sterilized for 4 hours at 165°C in a heater). For each glass bottle, 37.5 ml M9 mineral medium and 24 µl CHCA were added to reach 600 ppm final concentration of the contaminant. Bioaugmented microcosms were prepared adding 1 ml of a *R. opacus* R7 culture grown over-night until OD_{600}~ 1 - 1.5. The cells were washed once and resuspended in M9 mineral medium to obtain an initial optical density at 600 nm, OD_{600}, equal to 1. *R. opacus* R7 was maintained in mineral M9 (Maniatis et al., 1982) medium supplied with malate 1M (1 g/l) as the only carbon source at 30°C. Control microcosm experiments were established using sterilized sand (4 hours at 165°C in a heater) and transferred in the 250-ml glass bottle.

The microcosms bottles were closed and maintained under shaking (120 rpm) at 30°C and each bottle was entirely sacrificed for analysis at fixed times. For each condition tested (see section 5.2.3 Experimental plan and Table 5.2), fixed time microcosms were prepared and sampled (0, 16, 24, 30, 48 hours) in order to perform various determinations: 15 ml of slurry were collected, centrifuged and stored at -80°C to extract subsequently RNA, 1 ml to measure the bacterial count and the remaining amount of slurry phase was used to quantify the decrease of the contaminant using GC-MSD injections. Each replicate was independent, and each experiment was conducted in triplicate.

5.2.3 Extraction and analysis of CHCA

At fixed times, slurry microcosms were acidified with 5.2 M HCl (pH 2) and CHCA was extracted in 2 volumes of dichloromethane after 35 minutes of manual shacking. The organic phase was collected in a GC vial, properly diluted and it was supplied with N,O-Bis(trimethylsilyl)-trifluoroacetamide) (BSTFA) for the derivatization (20 µl each 25 mg L^{-1} CHCA) for 10 min at 60°C (adapted by Presentato et al., 2018). The extract was analysed by a 6890N Network gas chromatograph system (ZB-5MS column, 30 m x 0.25 mm, d_f 0.25 µm - Alltech) with He as gas carrier, coupled to an HP 5970 mass selective detector (230°C) with quadrupole mass analyser (150°C), 5973 Network Mass Selective Detector (280°C) in sim mode (Agilent Technologies). The GC injector temperature was 200°C. The temperature program was 5 min at 40°C, 8°C min^{-1} to 152°C, then 15°C min^{-1} to 250°C and hold for 10 min. The CHCA calibration curve was constructed by analysing standard CHCA ranging from 1.5 ppm to 18 ppm. Periodically CHCA calibration curve was re-established.

5.2.4 Experimental plan

The expression of catabolic functions was evaluated in various conditions: non-sterilized sand contaminated not inoculated (SNS-C), to determine the intrinsic degradative capacity of the indigenous sand bacterial community; non-sterilized sand not contaminated and not inoculated (SNS), to evaluate the basal activity of the sand; contaminated sand not sterilized and inoculated with *R. opacus* strain R7 (SNS-C-I), to evaluate the effect of bioaugmentation; non-sterilized sand not contaminated inoculated with *R. opacus* R7 (SNS-I), to evaluate the interactions between the inoculated strain and the autochthonous bacteria in absence of contamination; sterilized contaminated sand not inoculated (SS-C) to assess the abiotic loss; sterilized sand not contaminated and not inoculated (SS) to assess the sterilization process; sterilized and contaminated sand inoculated with R7 (SS-C-I), to evaluate the ability of the strain to degrade the contaminant in slurry sand microcosms (**Table 5.1**).

Table 5.1 - Experimental plan of sand microcosm experiments

Microcosm	Sand (S)	Inoculum (I)	Contaminant (C)
SNS	Not sterilized (NS)	-	-
SNS-C	Not sterilized (NS)	-	CHCA
SNS-I	Not sterilized (NS)	*R. opacus* R7	-
SNS-C-I	Not sterilized (NS)	*R. opacus* R7	CHCA
SS	Sterilized (SS)	-	-
SS-C	Sterilized (SS)	-	CHCA
SS-C-I	Sterilized (SS)	*R. opacus* R7	CHCA

5.2.5 *chpca* cluster identification in R7 genome

Sequence alignments and manual curation were performed in order to identify homologous sequences of *R. aetherivorans* BCP1 *chcpa* gene cluster (Presentato et al., 2018). The analyses were carried out using BLAST (Altschul et al., 1990) and Clustal Omega (Thompson et al., 1994) alignments.

5.2.6 RNA extraction, RT-PCR and RT-qPCR experiments

R. opacus R7 was maintained in mineral M9 (Maniatis et al., 1982) medium supplied with malate 1M (1 g/l) or CHCA (500 mg/l) as the only carbon source at 30°C. Total RNA was extracted both from *R. opacus* R7 pure culture and from 30 ml of the microcosm slurry phase (stored at -80°C) using RNA Power Soil Total RNA Isolation Kit (MoBio) according to the manufacturer's instructions with one modification: a second step of sample purification was performed adding 3.5 ml of chloroform),

mixing gently and centrifuging at 2500×g for 10 min. The upper phase was then transferred to a clean tube, then continuing to follow the instructions. The quality of RNA preparation was evaluated on agarose gel electrophoresis.

Reverse transcription was performed with iScript cDNA Synthesis kit (BIO-RAD) to obtain the corresponding cDNAs. The cDNA was synthesized from 0.5 µg of total RNA-DNase. The synthesis was carried out at 42°C for 60 min following the manufacturer instructions (Transcriptor High Fidelity cDNA Synthesis Kit, Roche Applied Science, Penzberg, Germany), then the cDNA solutions were stored at -20°C.RT-PCR semi-quantitative experiments were performed on the total synthesized cDNA, using a negative control on the negative reverse transcripted RNA. Two pair of primers for each gene (**Table 5.2**) were designed on the basis of the *ali*A1 and *pob*A1sequence genes of R7 strain, chosen as representative of the two different identified gene clusters. They were carried out using the following program: 95°C for 2 min; 95°C for 30 sec, 56°C for 45 sec, 72°C for 1 min 45 sec, for 35 cycles; and 72°C for 5 min (Mastercycler nexus GSX1, Eppendorf). PCR products were evaluated by gel-electrophoresis and 1Kb Direct Load (Sigma Aldrich) was used as marker of molecular weight.

Quantitative real-time Reverse Transcriptase-PCR (qPCR) analyses were performed on the same samples used for RT-PCR. The reverse-transcribed samples were amplified using the StepOnePlus Real-Time PCR System (Applied Biosystem, Italy). Each 10-µl qPCR volume contained 4.4 µl of the reverse-transcribed RNA samples, 5 µl of PowerUp SYBR Green Master Mix (Applied Biosystem, Thermo Scientific, Italy), and 300 nM of each primer. The specific primers were designed to amplify sequences of almost 400 bp (**Table 5.2**). Thermocycling conditions were as follows: 30 s at95 °C, followed by 45 cycles of 5 s at 95 °C, 10 s at 60 °C and 45 s at 72 °C and one cycle 15 s at 95 °C, 1 min at 60 °C and 15 s at 95°C. Expression of the housekeeping gene, 16S rDNA, was used as reference gene to normalize the tested gene. The ΔΔCt method with 16S rDNA as reference gene was used to determine relative abundance of target transcripts in respect to malate as control. Data are expressed as mean ± standard deviation derived from at least three independent experiments.

Table 5.2. List of oligonucleotides used for PCR amplifications and qRT-PCR analyses

Oligonucleotide name	sequence (5' – 3')	Tm (°C)
27f	AGAGTTTGATCCTGGCTCAG	55
1495r	CTACGGCTACCTTGTTACGA	55
aliA1-for	ATGATGTACACATACGATGCCCAC	58
aliA1-rev	CTTTACGAGCTGATGGTGCGG	59
pobA1-for	ATGAACACACAGGTCGGGATC	55
pobA1-rev	TCAGCCCAAAGGGGTGCCGAC	60
RT-16S-R7f	TCGTGAGATGTTGGGTTAAG	55
RT-16S-R7r	CCTCTGTACCGGCCATTGTAG	62
aliA1-40-for	TTCGAGCACGATTTCACCTAT	51
aliA1-500-rev	AGTTCGTCGAACCGGATCAC	54

5.3 Results

5.3.1 Growth kinetic experiments of *R. opacus* R7 on CHCA

R. opacus R7 capability to degrade naphthenic acids was evaluated testing its ability to utilize CHCA as sole carbon and energy sourcethat was chosen as model compound. Kinetic growth experiments were performed to measure the decrease of the contaminant in liquid pure culture analyzing the culture broth on GC-MSDat 0, 3, 6, 16, 24, 30, and 48 hours; pure CHCA (Sigma-Aldrich) was used as standard and to obtain a calibration curve. As shown in **Figure 5.2**, *R. opacus* R7 was able to degrade 90% of CHCA in 16 hours and bacterial increased from 6.7 x 10^6 CFU ml^{-1} to 2.5 x 10^7 CFU ml^{-1}. The degradation rate was evaluated until 48 hours and the contaminant resulted completely degraded and the R7 counts reached the value of 2.2 x 10^8 CFU ml^{-1}.

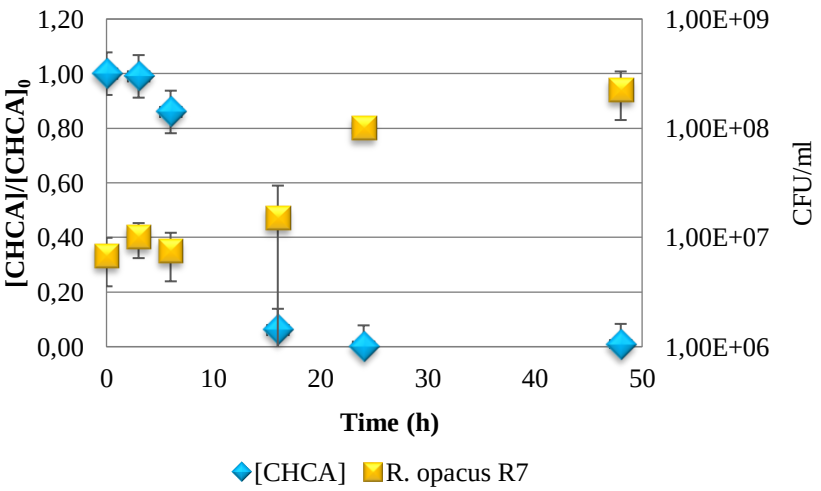


Figure 5.2 Kinetic growth curve of R7 strain pure culture in presence of CHCA. The cyan curve represents the CHCA degradation, while the yellow curve theR7 growth based on viable cell counts (logarithmic scale).

5.3.2 Identification of the *aliA1* and *pobA1* gene clusters and evaluation of their involvement in CHCA degradation in *R. opacus* R7

Genome-based analysis of *R. opacus* R7 allowed to identify the homologous gene clusters of *R. aetherivorans* BCP1, involved in the naphthenic acid degradation. The results revealed the presence in R7 genome of BCP1 homologous sequences allocated in two different regions of R7chromosome (**Figure 5.3**). The first region contains the *p*-hydroxybenzoate hydroxylasegene (*pobA*); PobA showed an aminoacidic identity of 67% respect to PobA gene of BCP1. Whereas, the second region contains homologous genes to the *chpca* genes of BCP1.

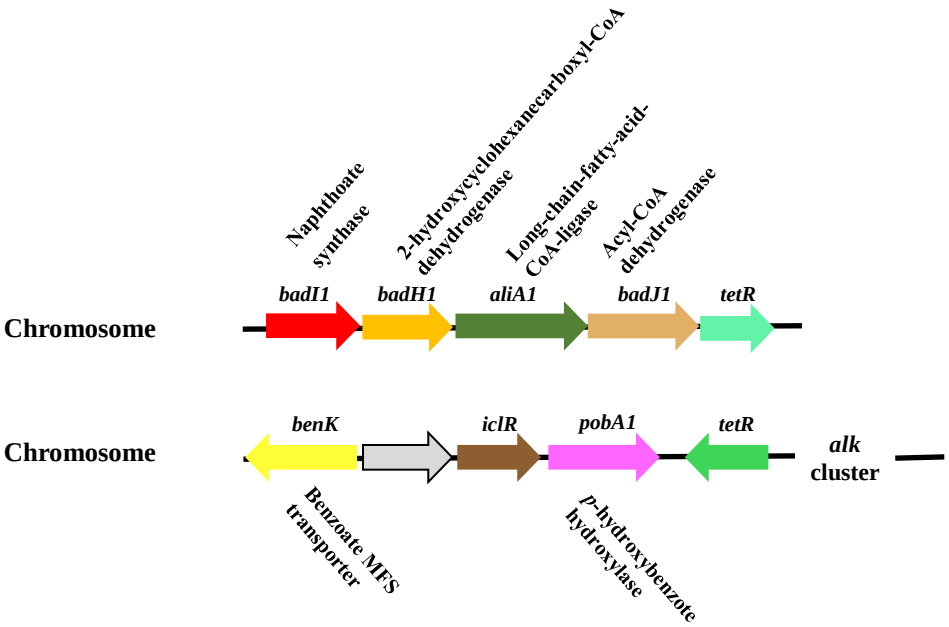


Figure 5.3 R7 chromosomal regions containing the homologous genes respect to BCP1 genes involved in NAs degradation.

In particular, *aliA1* gene encoding for a long-chain-fatty-acid-CoA ligase, showed an aminoacidic identity of 71% respect on the *aliA1* of BCP1, *badH*, *badI*, *badJ* encoding for 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase, naphthoate synthase and acyl-CoA dehydrogenase, showed a protein identity of 73%, 87% and 85% respect to BCP1, respectively.

The involvement of R7 genes in the CHCA degradation was evaluated by RT-PCR experiments after growth on CHCA or malate as control, selecting *aliA1* and *pobA1* genes as marker. These experiments were performed amplifying the target genes after RNA extraction and cDNA synthesis. The results showed that CHCA induced the transcription of the *aliA1* gene, while *pobA1* gene is transcript in a constitutive manner (**Figure 5.4**). Therefore, *aliA1* gene was chosen as marker sequence to represent R7 *chpca* gene cluster.

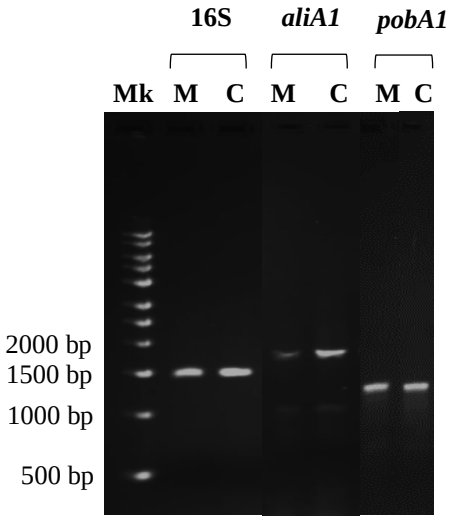


Figure 5.4 RT-PCR experiments. RT-PCR was performed after*R. opacus* R7 growth on CHCA (C) or malate (M) as control. Mk: 1Kb Direct Load (Sigma Aldrich). Amplified genes: 16SrDNA, *aliA1*, *pobA1*.

5.3.3 Characterization of sand samples

The microcosms were prepared using sand collected from the side of a river that was characterized considering agronomic, chemical and microbiological parameters. The humidity and TOC were around 0.21 % and 0.5%, respectively. The agronomic characterization showed that the sand had 0.5% of organic matter content, which, according to the USDA classification (United States Department of Agriculture, Natural Resources Conservation Service), is considered a low content; however, the concentration of organic carbon around 5000 ppm can sustain the growth of microbial community. Indeed, R7 in slurry microcosm with sterilized sand without adding other organic carbon sources showed a growth from 3 x 10^6 CFU ml^{-1} to 1 x 10^7 CFU ml^{-1}. Microbiological sand characterization showed that total heterotrophic bacteria counts were 7 x 10^4 - 10^5 CFU ml^{-1}, while CHCA degrading bacteria counts were equal to 3 x 10^3 CFU ml^{-1}. Soil chemical analysis showed the absence of naphthenic acid contamination. For this reason, we decided to contaminate it with a model compound belonging to naphthenic acids class, the cyclohexane carboxylic acid (CHCA).

5.3.4 Growth kinetic experiments in microcosms in presence of CHCA

Slurry-phase microcosms were established to evaluate the depuration potential of an hydrocarbon contaminated soil using a culture-independent approach. Microcosms experiments were performed to evaluate the degradation rate of NAs in presence of autochthonous bacteria in comparison to bioaugmented microcosms, in presence of *R. opacus* R7 strain. At fixed times (0, 16, 24, 30, 48 hours) different parameters were determined: the decrease of the contaminant (normalized on the initial CHCA amount), the bacterial count and the RNA extraction for *aliA1* gene expression in RT-PCR and RT-qPCR experiments. In order to evaluate the R7 CHCA degradative ability in the sand, a

microcosm with sterilized sand, contaminated and inoculated with R7, was used to evaluate the CHCA degradation rate and the R7 growth (**Figure 5.5**).

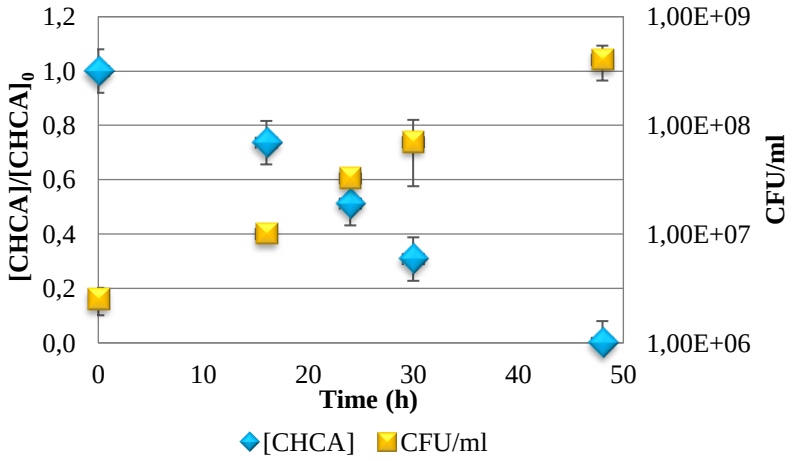


Figure 5.5. Kinetic growth curve of R7 strain in SS+CHCA microcosms. The cyan curve represents the CHCA degradation, while the yellow curve R7 growth based on viable cell counts (logarithmic scale).

The results revealed that the strain is able to degrade around 99% of the CHCA in 48 hours. This is in line with the growth curve of R7 strain in pure liquid culture; indeed, in the same period of time the values of CFU ml^{-1} vary from 2 x 10^6 to 4 x 10^8. Comparing R7 growth rate in pure culture and in slurry microcosms, it was able to degrade around 90% of CHCA after only 16 hours of incubation in a pure culture. The linear growth curve is justified by the presence in the microcosms of a large amount of carbon source due to the sum of CHCA contaminant and the intrinsic organic carbon of the sand; this curve corresponds to a linear biodegradation. The intrinsic NAs degradative capability of the indigenous soil bacteria was evaluated in the non-sterilized contaminated not inoculated (SNS-C) sand microcosm. As showed in **Figure 5.6**, autochthonous bacterial population was able to degrade over the 90% of CHCA in 48 hours, showing a slower degradation rate compared to R7 CHCA degradation. In the same incubation period the microbial community growth was measured revealing that it varies from 10^4 CFU ml^{-1} to 10^9 CFU ml^{-1}.

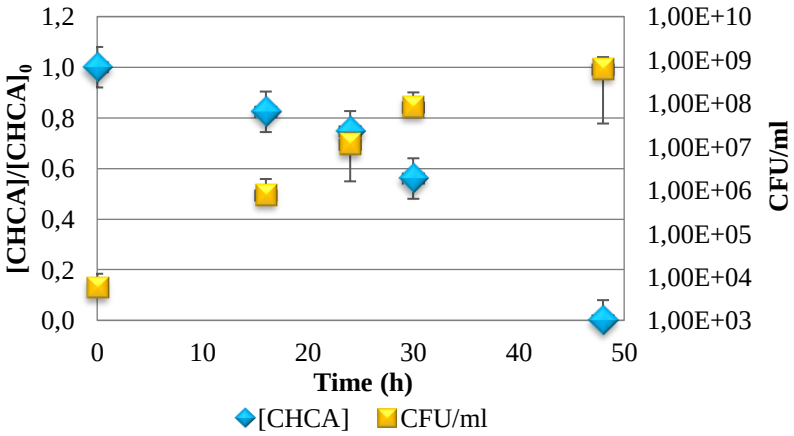


Figure 5.6. Kinetic growth curve of the autochthonous bacterial population in SNS+CHCA microcosms. The cyan curve represents the CHCA degradation, while the yellow curve the bacterial growth based on viable cell counts (logarithmic scale).

Moreover, bioaugmented microcosms containing not sterilized sand were established to evaluate the contribute of R7 strain in the CHCA degradation in presence of the microbial community and (**Figure 5.7**).

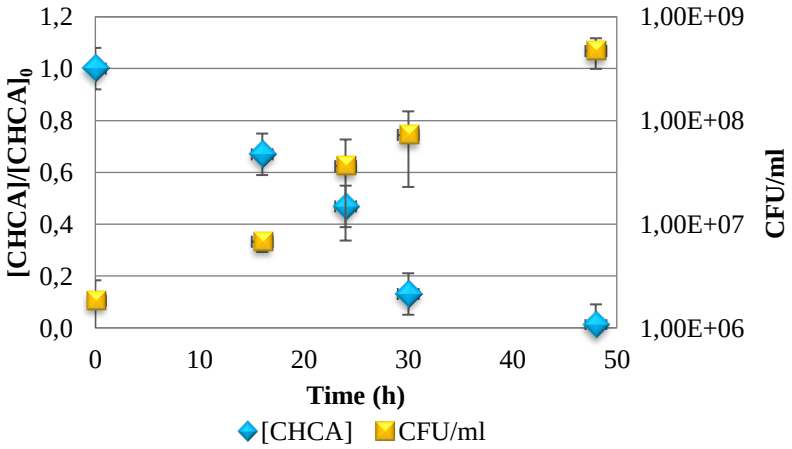


Figure 5.7 Kinetic growth curve of total bacteria in SNS+CHCA+R7 microcosms. The cyan curve represents the CHCA degradation, while the yellow curve the bacterial growth using viable cell counts (logarithmic scale).

The CHCA consumption in these conditions was around 99% in 48 hours, even if after 30 hours the percentage of the degradation was already around 70%. Additionally, in the same incubation period the microbial community varies from 1 x 10^6 CFU ml^{-1} to 5 x 10^8 CFU ml^{-1}.

As control, three different microcosm conditions were established and afterwards analyzed at the 0, 16, 24, 30, 48 hours. i) Microcosms containing not sterilized sand bioaugmented with R7 strain and in absence of the contaminant (SNS-I) were analyzed to investigate the capability of both autochthonous bacterial population and R7 strain to grow in presence of the naturally present organic carbon (0.5%). Indeed, this experimental condition without another organic carbon source, showed the growth of the microcosm microbial population from 6 x 10^5 CFU ml^{-1} to 2 x 10^7 CFU ml^{-1}. ii) Microcosms containing sand autochthonous bacterial population in absence of the contaminant (SNS) were tested to evaluate the microbial community potential in presence of the 0.5% of organic carbon. The autochthonous bacteria showed a growth from 6 x 10^3 CFU ml^{-1} to 2 x 10^7 CFU ml^{-1}. iii) Sterilized sand microcosms containing R7 strain without spiking with the CHCA (SS-I) showed that R7 grew slightly on the naturally present organic carbon (0.5%) from 3 x 10^6 CFU ml^{-1} to 1 x 10^7 CFU ml^{-1}.

In order to assess the abiotic loss, a microcosm with sterilized contaminated soil not inoculated with R7 strain was performed. The analysis of the CHCA degradation showed that the experimental conditions presented anabiotic loss of around 20%.

In conclusion, the microcosm experiments revealed that CHCA were completely degraded in 48 hours in all the established conditions. The CHCA concentration rapidly decrease in presence of R7 strain,

showing already after 16 hours the percentage of degradation of 47%; while, in absence of the strain the percentage of degradation was of 13% (**Figure 5.8**).

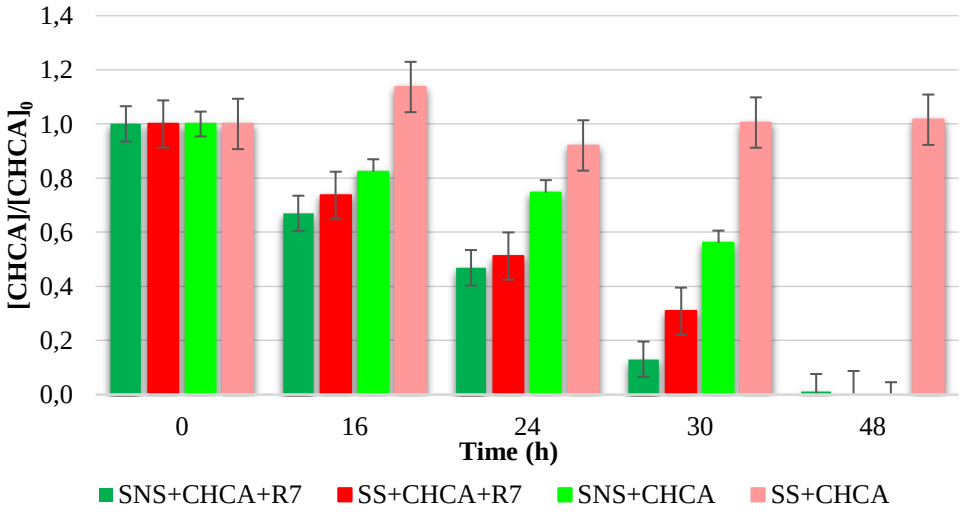


Figure 5.8 Comparison of CHCA degradation in SS+CHCA+R7, SNS+CHCA+R7, SNS+CHCA, and SS+CHCA microcosms.

The estimation of bacterial counts was performed using viable cell counts in LB agar plates (**Figure 5.9**), in order to evaluate the viability of CHCA degrading bacteria and to validate the degradation of the contaminant along the complete duration of microcosm experiments.

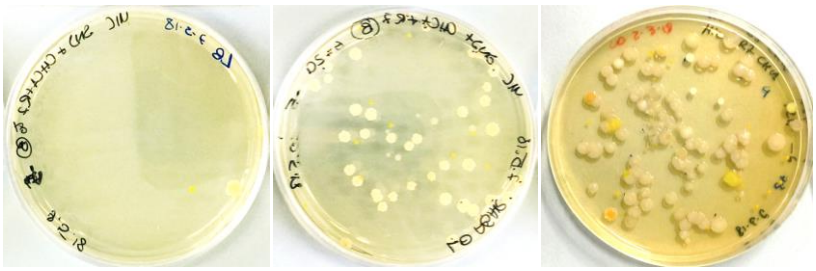


Figure 5.9. CHCA degrading bacteria from SNS+CHCA+R7 microcosm plated on LB Agar medium. Example of plates showing the CHCA degrading bacteria at 0, 24, 48 hours.

5.3.5 Gene expression of the marker sequence in microcosms by RT-PCR and RT-qPCR experiments

Among the identified and tested genes putatively involved in CHCA degradation, *aliA1* coding for a long-chain-fatty-acid-CoA ligase, was selected as a possible marker sequence to follow the degradation in microcosm experiments. The expression levels of *aliA1* gene were analysed by qualitative and quantitative RT-PCR experiments. Total RNAs were isolated as described in the paragraph 5.2.6 of Materials and Methods. The extractions were performed from SS+CHCA+R7, SNS+CHCA+R7 and SNS+CHCA microcosms after 0, 16, 24, 30 and 48 hours. RNA in all the initial

times (t_0) was absent, consequently the analysis of the *aliA1*gene expression was performed for the subsequent fixed times.

RT-PCR results in **Figure 5.10** reported both the 16S rDNA and *aliA1* gene expressions. The 16S expression was used as basal level of R7 transcripts after growth on CHCA. Results revealed that the expression of *aliA1* gene in the microcosms was strictly related to the presence of R7 strain, suggesting that it can be used to follow the CHCA degradation in R7 strain. Therefore, quantitative RT-PCR experiments were performed to verify and evaluate the transcription levels in different conditions. RT-qPCR experiments were performed after RNA extraction from SS+CHCA+R7, SNS+CHCA+R7 and SNS+CHCA microcosms.

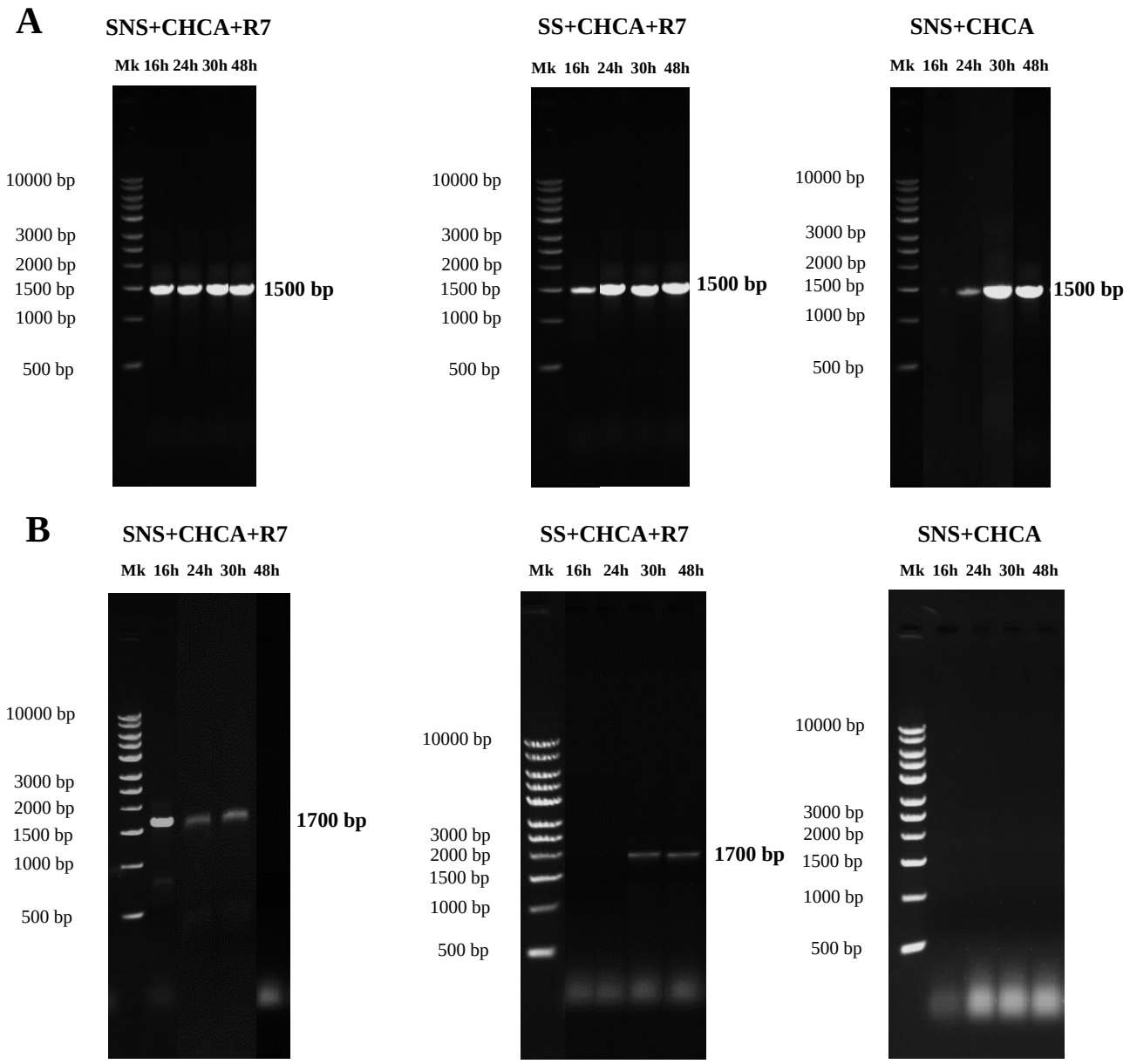


Figure 5.10. RT-PCR experiments from RNA extracted from SNS+CHCA+R7, SS+CHCA+R7, and SNS+CHCA microcosms. RT-PCR experiments with 16S rDNA (**Panel A**) and *aliA1* genes (**Panel B**) after 16, 24, 30 and 48 hours of incubation.

Quantitative reverse transcription-PCR (RT-qPCR) experiments were performed to define the differential levels of transcription of the *aliA1* gene in R7 strain during growth in presence of CHCA respect to malate condition (used as a control). The 16S rDNA was used as a housekeeping gene to normalize *aliA1* expression levels. The ΔΔCt method was applied to evaluate *aliA1* transcript expression levels analysing the $2^{-\Delta\Delta Ct}$ values for the different conditions. The Figure **5.10** reports the RT-qPCR results: *aliA1* transcription levels increased during the 48 hours in the microcosm containing sterilized sand, CHCA and R7 strain. While, the microcosm containing the not sterilized sand, the contaminant and R7 showed a different expression trend. In this condition the highest expression level was observed after 24 hours, successively the level decreased in the last fixed times. The analysis confirmed the results of the RT-PCR experiments, showing the *aliA1* strictly related to the presence of R7 strain.

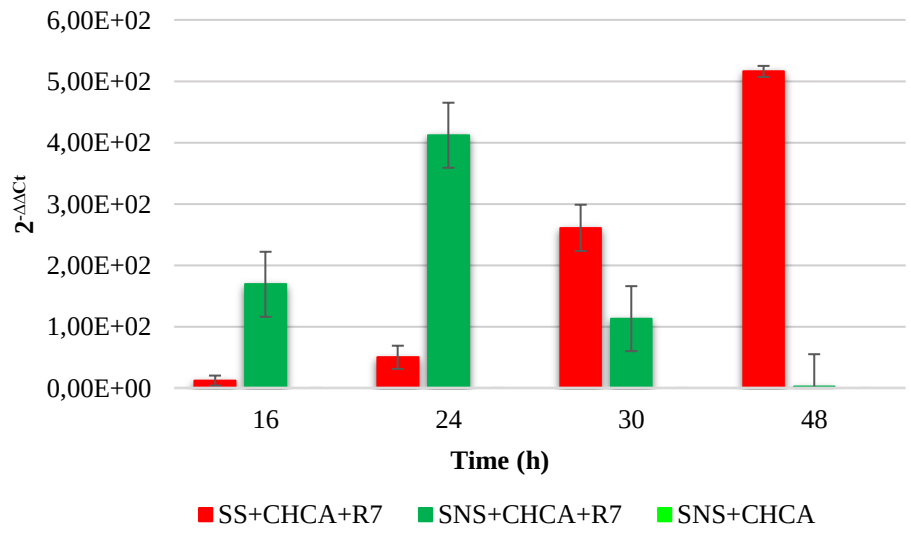


Figure 5.11. RT-qPCR analysis. Relative gene expression of *aliA1* in SS+CHCA+R7, SNS+CHCA+R7 and SNS+CHCA microcosms.

5.4 Discussion

The aim of the present chapter was to assess the depuration potential of naphthenic acid contaminated soil using a culture-independent approach in sand slurry-phase microcosm experiments. Preliminary analysis revealed that R7 strain was able to mineralize CHCA as only carbon and energy source in 24 hours in a pure culture. Genome-based analysis and RT-qPCR experiments indicated that the degradation of CHCA in R7 strain mainly occurs via the aromatization of the cyclohexane ring, involving the *chpca* gene cluster that includes *aliA1* gene (Pelletier and Harwood, 2000). Based on *aliA1* gene expression levels after R7 growth on CHCA, *aliA1* gene sequence was chosen as a marker sequence to monitor R7 activity in microcosm experiments. For this reason, the sand, collected from

a non-contaminated side of a river, was sampled and characterized considering agronomic, chemicals and microbiological parameters. The humidity and TOC were around 0.21 % and 0.5%, respectively. The agronomic characterization showed that the soil had 0.5% of organic matter content, that can sustain the growth of microbial community. Microbiological characterization showed the presence in microbial community of CHCA degrading bacteria. Sand chemical analysis showed the absence of naphthenic acid contamination. For this reason, microcosms were spiked with a model compound belonging to naphthenic acid class, the cyclohexane carboxylic acid (CHCA). In order to follow and monitor the CHCA degradation in microcosms, different conditions and controls were established resulting in all cases in a complete mineralization of CHCA within 48 hours. Results showed that CHCA mineralization was performed both by the autochthonous CHCA degrading bacteria and R7 strain, even if R7 strain can degrade the contaminant faster than the microbial community. Indeed, comparing the microbial community behaviour in presence or in absence of R7, we could observe a considerable difference up to 30 hours.

RT-PCR and RT-qPCR results revealed that the expression of *aliA1* gene in the microcosms was strictly related to the presence of R7 strain. These results highlighted the possibility to follow the R7 CHCA degradation using a specific marker sequence in a model microcosm system revealing the efficiency of the established method.

Moreover, RT-qPCR analysis indicated that the microbial community and R7 strain had a competitive interaction. It generated R7 growth inhibition likely occurring for substrate competition or for the production of toxic compounds during CHCA utilization by the autochthonous bacterial population. Considering this last aspect, the investigation of diverse metabolic pathways of naphthenic acid degradation should be deeply investigated in different *Rhodococcus* strains and in general, in different bacteria to gain more knowledge regarding new gene sequences or enzymes. Moreover, these studies can unveil the correspondent mineralization mechanisms and consequently mechanisms of bacteria interaction in presence of emerging compound as naphthenic acids.

5.5 *References*

Adams GO, Fufeyin PT, Okoro SE, Ehinomen I (2015). Bioremediation, Biostimulation and Bioaugmention: A Review. International Journal of Environmental Bioremediation & Biodegradation. 1:28-39

Allen EW (2008)**.** Process water treatment in Canada's oil sands industry: I. Target pollutants and treatment objectives. J. Environ. Eng. Sci., 7:123-138.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. J Mol Biol, 215:403–10.

Blakley ER (1974). The microbial degradation of cyclohexanecarboxylic acid: a pathway involving aromatization to form p-hydroxybenzoic acid. Can. J. Microbiol., 20:1297–1306.

Blakley ER (1978). The microbial degradation of cyclohexanecarboxylic acid by a beta-oxidation pathway with simultaneous induction to the utilization of benzoate. Can. J. Microbiol., 24:847–855.

Blakley ER and Papish B (1982). The metabolism of cyclohexanecarboxylic acid and 3-cyclohexenecarboxylic acid by *Pseudomonas putida*. Can. J. Microbiol., 28:1324–1329.

Cappelletti M, Fedi S, Zampolli J, Di Canito A, D'Ursi P, Orro A et al. (2016). Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7. Res. Microbiol., 167:1–9.

Cavalca L, Colombo M, Larcher S, Gigliotti C, Collina E, Andreoni V (2002) Survival and naphthalene-degrading activity of *Rhodococcus* sp. strain 1BN in soil microcosms. J Appl Microbiol, 92:1058-1065.

Clemente JS, Mackinnon MD and Fedorak PM (2004). Aerobic biodegradation of two commercial naphthenic acids preparations. Environ. Sci. Technol., 38:1009–1016.

Demeter MA, Lemire JA, Yue G, Ceri H and Turner RJ (2015). Culturing oil sands microbes as mixed species communities enhances ex situ model naphthenic acid degradation. Front. Microbiol., 6:936.

Demeter MA, Lemire J, George I, Yue G, Ceri H and Turner RJ (2014). Harnessing oil sands microbial communities for use in ex situ naphthenic acid bioremediation. Chemosphere, 97:78-85.

Frank RA, Kavanagh R, Kent Burnison B, Arsenault G, Headley JV, Peru KM et al. (2008). Toxicity assessment of collected fractions from an extracted naphthenic acid mixture. Chemosphere, 72:1309-1314.

Golby S, Ceri H, Gieg LM, Chatterjee I, Marques LLR and Turner RJ (2012). Evaluation of microbial biofilm communities from an Alberta oil sands tailings pond. FEMS Microbiol. Ecol., 79:240-250.

Iwaki H, Saji H, Abe K and Hasegawa Y (2005). Cloning and sequence analysis of the 4-hydroxybenzoate 3-hydroxylase gene from a cyclohexanecarboxylate degrading Gram-positive bacterium, "*Corynebacterium cyclohexanicum*" strain ATCC 51369. Microb. Environ. 20:144-150.

Johnson RJ, West CE, Swaih AM, Folwell BD, Smith BE, Rowland SJ et al. (2011). Aerobic biotransformation of alkyl branched aromatic alkanoic naphthenic acids via two different pathways by a new isolate of *Mycobacterium*. Environ. Microbiol., 14:872-882.

Jones D, Scarlett AG, West CE and Rowland SJ (2011). Toxicity of individual naphthenic acids to Vibrio fischeri. Environ. Sci. Technol., 45:9776-9782.

Koma D, Sakashita Y, Kubota K, Fujii Y, Hasumi F, Chung SY et al. (2004). Degradation pathways of cyclic alkanes in *Rhodococcus* sp. NDKK48. Appl. Microbiol. Biotechnol., 66:92-99.

Lai JWS, Pinto LJ, Bendell-Young LI, Moore MM and Kiehlmann E (1996). Factors that affect the degradation of naphthenic acids in oil sands wastewater by indigenous microbial communities. Environ. Toxicol. Chem., 15:1482-1491.

Maniatis T, Fritsch EF, Sambrook J (1982). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Orro A, Cappelletti M, D'Ursi P, Milanesi L, Di Canito A, Zampolli J et al. (2015). Genome and phenotype microarray analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: genetic determinants and metabolic abilities with environmental relevance. PLoS One 10:e0139467.

Pelletier DA and Harwood CS (2000). 2-Hydroxycyclohexanecarboxyl coenzyme A dehydrogenase, an enzyme characteristic of the anaerobic benzoate degradation pathway used by *Rhodopseudomonas palustris.* J. Bacteriol., 182:2753–2760.

Presentato A, Cappelletti M, Sansone A, Ferreri C, Piacenza E, Demeter MA, Crognale S, Petruccioli M, Milazzo G, Fedi S, Steinbüchel A, Turner RJ and Zannoni D (2018). Aerobic growth of *Rhodococcus aetherivorans* BCP1 using selected naphthenic acids as the sole carbon and energy sources. Front. Microbiol., 9:672.

Rho EM and Evans WC (1975). The aerobic metabolism of cyclohexanecarboxylic acid by *Acinetobacter anitratum*. Biochem. J., 148:11–15.

Ruberto LAM, Vazquez S, Lobalbo A, Maccormack WP (2005) Psychrotolerant hydrocarbon-degrading *Rhodococcus* strains isolated from polluted Antarctic soils. Antarctic Science, 17:47-56.

Schwartz E, Trinh SV and Scow KM (2000) Measuring growth of a phenantrene-degrading bacterial inoculum in soil with a quantitative competitive polymerase chain reaction method. FEMS Microbiology Ecology, 34:1-7.

Smith BE, Lewis CA, Belt ST, Whitby C and Rowland SJ (2008). Effects of alkyl chain branching on the biotransformation of naphthenic acids. Environ. Sci. Technol., 42: 9323–9328.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res, 22:4673–80.

Wang X, Chen M, Xiao J, Hao L, Crowley DE, Zhang Z et al. (2015). Genome sequence analysis of the naphthenic acid degrading and metal resistant bacterium *Cupriavidus gilardii* CR3. PLoS One 10:e0132881.

Whitby C (2010). Microbial naphthenic acid degradation. Adv. Appl. Microbiol., 70, 93-125.

Chapter 6.

Investigation of 4,4’-dithiodibutyric acid (DTDB) utilization of *Rhodococcus erythropolis* MI2 by generation of marker-free deletion mutants

(in collaboration with the Institut für Molekulare Mikrobiologie und Biotechnologie of the Westfälische Wilhelms-Universität of Münster, during the PhD period abroad)

6.1 Introduction

Rhodococcus erythropolis MI2 is a Gram-positive, aerobic and non-motile bacterium belonging to the family *Nocardiaceae*, which contains highly diverse genera and species. It was isolated from an oil-contaminated absorber in a garage and later applied in polythioester (PTE) research. Indeed, in Wübbeler et al. 2010, *R. erythropolis* MI2 ability to degrade a potential PTE precursor, 4,4'-dithiodibutyric acid (DTDB), was studied. The results revealed that MI2 is DTDB degrader; so, to elucidate DTDB catabolic pathway, the genome of MI2 was sequenced.

The genome consists of approximately 7.2 Mb with an overall G+C content of 62.25%. It is composed of three replicons: one chromosome of 6.45 Mb and two megaplasmids with sizes of 400 kb and 350 Mb, respectively. Analyzing the genome, 6911 predicted open reading frames (ORFs), 50 tRNA genes and two rRNA genes were identified. A putative function was assigned to 5236 (75.76%) of all predicted protein-coding sequences.

The aim of this work was the generation of *R. erythropolis* MI2 marker-free deletion mutants in order to investigate DTDB utilization, in particular about the degradation of 4-oxo-4-sulfanylbutyric acid. The first step of the work has been the deletion of the identified genes involved in the desulfhydration of the 4-oxo-4-sulfanylbutyric acid (**Table 6.1**), using the suicide vector pJQ200mp18Tc which contains the *oriT* of plasmid RP4 for conjugative mobilization and the *sacB* system as well as the tetracyclin resistance as primary selection marker.

Table 6.1. Striking proteins based on proteomic analysis (and *in silico* analysis)

Protein	Locus tag
Sulfide: quinone oxidoreductase (SQR)	RERY_02710
Putative Zn- metallo-*β*-lactamase/ rhodanese domain containing protein	RERY_02720
Putative rhodanese-related sulfurtransferase	RERY_02740
Putative transmembrane protein TauE	RERY_02750
Putative acyl-CoA dehydrogenase	RERY_66330
Desulfydrase	RERY_06500

6.1.1 PTEs production

PTEs are industrially and scientifically interesting materials, not only because they possess some unique characteristics regarding their biodegradability, microstructure, thermal behavior and diverse mechanical properties, but also because many of them can be produced from nonpetrochemical resources (Steinbüchel, 2001). However, many difficulties limit the large production scale of PTEs,

such as intensive costs, difficulties during preparations, requirement of toxic reagents and low polymer yields. PTEs are microbial synthesized, non-biodegradable polymers containing sulfur in the thioester linkages of the backbone (Kim et al., 2005; Lütke-Eversloh et al., 2002a). They are accumulated as hydrophobic inclusions in the cytoplasm of bacterial cells in a manner similar to polyhydroxyalkanoic acids (PHAs) (Lütke-Eversloh & Steinbüchel, 2004). For the generation of metabolically engineered PTE-production strains, it is necessary to investigate the microbial catabolism of potential precursor substrates. PTE precursor degradation is restricted to only few bacteria; likewise, PTE-producing strains were not able to utilize the precursors as the sole carbon and energy source for growth and for the improvement of polythioester (PTE) production. The main PTE precursor substrates are: 3,3'-thiodipropionic acid (TDP), 3,3'-dithiodipropionic acid (DTDP) and 4,4'-dithiodibutyric acid (DTDB) (**Figure 6.1**).

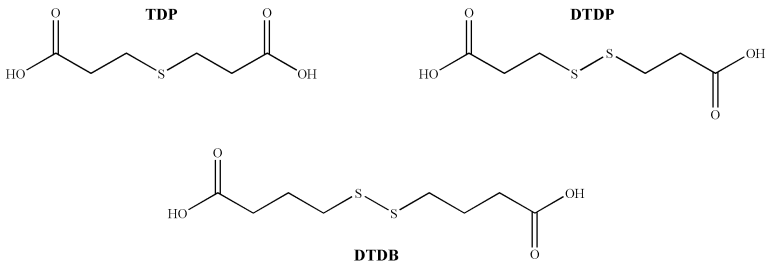


Figure 6.1. Chemical structures of TDP, DTDP, DTDB

The putative TDP degradation pathway has been proposed in 2009 (Bruland et al., 2009a). In the first step, the thioether is cleaved by an unidentified enzyme into 3-hydroxypropionate and 3-mercaptopropionic acid (3MP). 3MP is then sulfoxygenated by a cysteine dioxygenase, generating 3-sulfinopropionate (3SP). After linkage to coenzyme A (CoA) by a family III acyl-CoA transferase, the sulfur moiety is probably removed by a desulfinase, producing propionyl-CoA, which is then metabolized, most probable via the methylmalonyl-CoA pathway. Catabolism of DTDP begins with a symmetrical cleavage into two molecules of 3MP by a disulfide reductase. 3MP dioxygenase converts the sulfydryl group into a sulfinic group, thus yielding 3SP. A linkage between CoA and 3SP is performed by a thiokinase giving an activate product as 3SP-CoA. Afterward, the elimination of sulfite from 3SP-CoA probably requires three enzymes that catalyze the conversions to propionyl-CoA, which is then probably metabolized via the 2-methylcitric acid pathway (Wübbeler et al., 2008). In particular, in this chapter the attention is focused on DTDB degradation pathway in *R. erythropolis* MI2 because it possesses the rare ability to use it as sole carbon source and electron donor for aerobic growth (Khairy et al., 2016). DTDB is an organic sulfur compound (OSC) useful in diverse biochemical fields, as a cross-linker in the production of chemically epoxidized natural rubber (ENR) (Imbernon et al., 2015). It is also used as a monolayer for protein chips, based on a gold surface (Jang

& Keng, 2006), which were applied for the recognition of various sugars by surface-enhanced raman spectroscopy and cyclic voltammetry (Kanayama and Kitano, 2000). Furthermore, DTDB is considered as an alternative promising substrate for PTE synthesis; indeed, DTDB is the oxidized form of two 4-mercaptobutyric acid (4MB) molecules that are the first produced intermediates (**Figure 6.2**). 4MB is a building block for PTE synthesis; the use of DTDB as the precursor substrate could be useful to begin PTE production in bacterial strains showing a certain vulnerability to the toxic thiol 4MB.

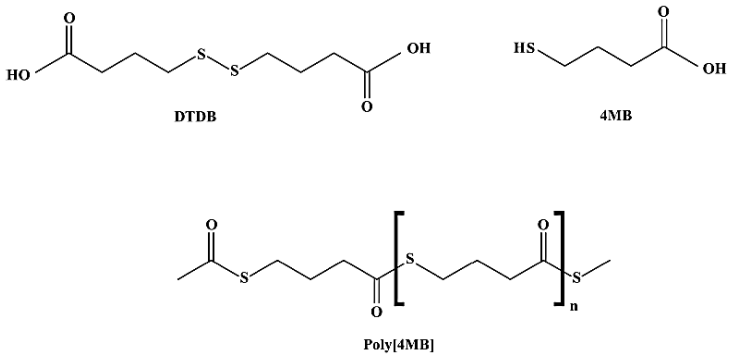


Figure 6.2. Chemical structures of DTDB, 4MB and Poly[4MB]

6.1.2 DTDB degradation in *R. erythropolis* MI2

In Wübbeler et al. 2010 *R. erythropolis* MI2 was subjected to random transposon mutagenesis to elucidate genes involved in DTDB catabolism; the genomic informations were compared to proteomic studies performed after cultivation with succinic acid and DTDB (Khairy et al., 2016a; 2016b). DTDB is transported by an unknown transport system into the cells; the transport could necessitate of an active transport system, due to its chemical and structural properties. Its metabolic intermediates were identified by GC/MS; the catabolism of DTDB started by the symmetrical cleavage of DTDB into two molecules of 4-mercaptobutyric acid (4MB) by an NADH:Flavin dependent oxidoreductase (Nox, Khairy et al., 2016). The second proposed step in the catabolism of DTDB is the oxidation of 4MB to 4-oxo-4-sulfanylbutyric acid by a luciferase-like monooxygenase (LLM_{MI2}). Finally, this compound is probably desulfurized to succinic acid by a putative desulfhydrase, releasing the sulfur moiety as a volatile hydrogen sulfide, which could be detected during growth of *R. erythropolis* MI2 (Wübbeler et al., 2010; Khairy et al., 2015). Two pathways were proposed regarding the degradation of 4-oxo-4-sulfanylbutyric acid. The molecule is also degraded to succinic acid and hydrogen sulfide by a putative desulfhydrase (DSH_{M12}, RERY_06500) or the reaction could be catalyzed by a sulfide:chinone oxidoreductase (SQR_{MI2}, RERY_02710). In that case, 4-sulfanyl-4-oxobutyric acid would be the reaction product. Two enzymes probably

degrade this compound: i. a Zn-metallo-*β*-lactamase and ii. a putative rhodanese-related sulfurtransferase (RSTMI2, Khairy et al., 2016). Therefore, more information is needed to verify the postulated enzymatic reactions. Based on the information acquired in these studies a putative catabolic pathway of DTDB was proposed (**Figure 6.3**).

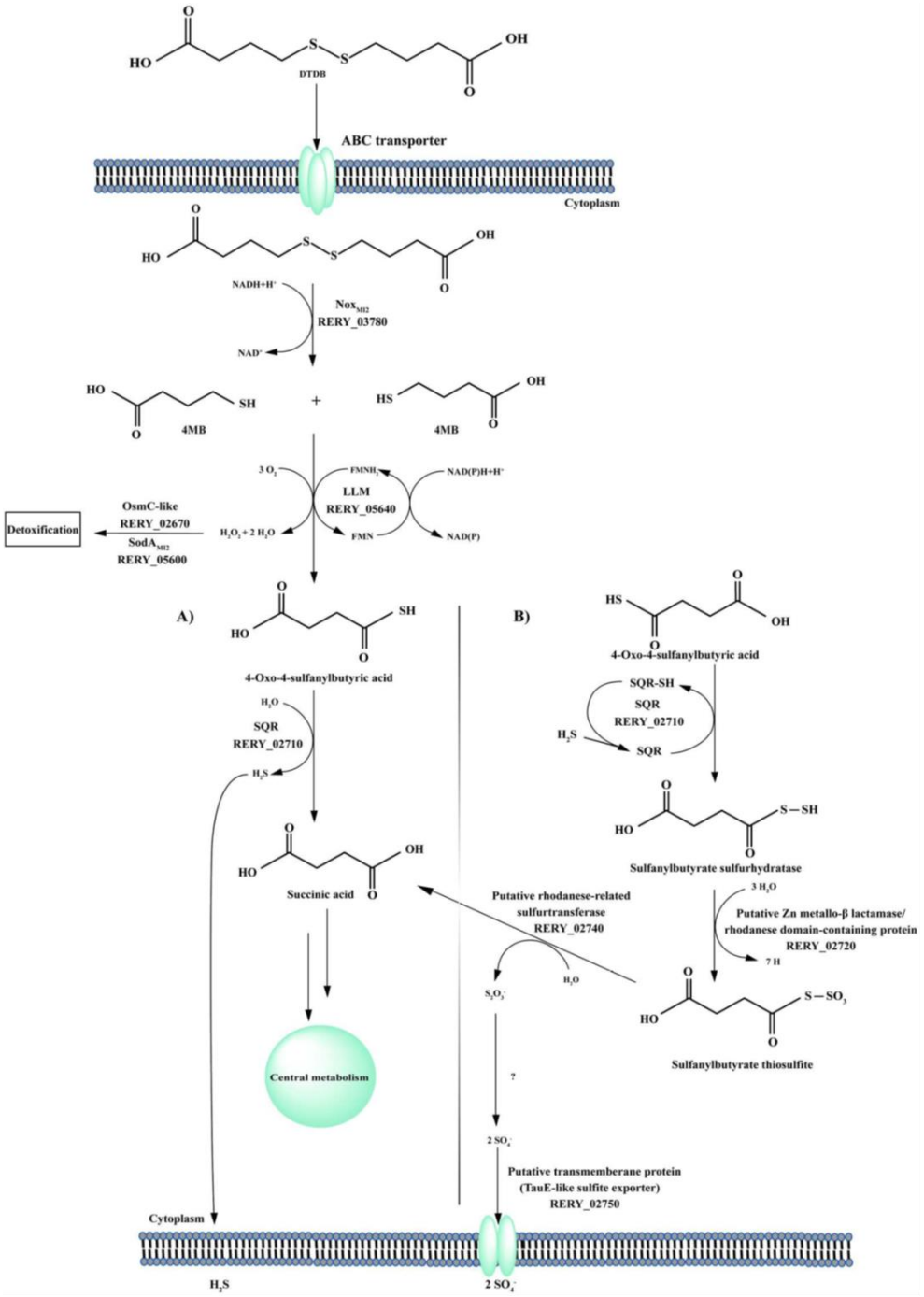


Figure 6.3. Updated degradation pathway of DTDB in *R. erythropolis* MI2. Nox, using NADH, cleaves DTDB into two molecules 4MB. Then, a putative oxygenase converts 4MB into 4-oxo-4-sulfanylbutyric acid. **A)** Desulfurization of 4-oxo-4-sulfanylbutyric acid by either the action of SQR or the putative desulfhydrase. **B)** Hypothesized alternative sulfur oxidation process. Abbreviations: Nox$_{MI2}$, NADH:flavin oxidoreductase, LLM, luciferase-like monooxygenase; SQR, sulfide:quinone oxidoreductase; OsmC-like, osmotically induced protein; SodA$_{MI2}$, superoxide dismutase. (Khairy et al., 2016)

6.1.3 Removal of the sulfur group

MI2 genome analysis revealed that only one gene is annotated as a desulfhydrase (RERY_06500); this enzyme is highly conserved in many *Rhodococcus* strains. However, in *R. jostii* RHA1 that is unable to grow in the presence of DTDB, an orthologue of RERY_06500 is completely missing (Khairy et al., 2015). RHA1 strain potentially can import and cleave DTDB yielding toxic 4MB and converting it to the presumably also highly toxic 4-oxo-4-sulfanylbutyrate. Indeed, its genome contains an orthologue of *nox*, involved in the cleavage of DTDB and the LLM F420-dependent oxidoreductase involved in the oxidation of 4MB; however, an orthologue to the putative desulfhydrase is absent. For this reason, the toxic 4-oxo-4-sulfanylbutyrate is accumulated in the cells and their growth ends. The presence of a desulfhydrase is necessary in the catabolism of DTDB. Moreover, an alternative pathway for the sulfur-removal was described; it needs three proteins: i. a sulfide:quinone oxidoreductase(SQRMI2) (RERY_02710) ii. a rhodanese-related sulfurtransferase (RERY_02740) iii. a protein consisting of two large domains: a putative zinc metallo-beta lactamase domain and a rhodanese-like domain (RERY_02720) iv. a putative sulfite exporter v. a putative acyl-CoA dehydrogenase (**Table 6.1**).

- The SQRMI2 was identified as one of the highly expressed proteins; this enzyme belongs to the flavin disulfide reductase family. It catalyzes sulfide oxidation and it donates electrons from sulfide to the electron transport chain at the level of quinone. The enzyme could likely accomplish the desulfhydration process of 4-oxo-4-sulfanylbutyric acid, or the oxidation of consequently produced hydrogen sulfide.
- Rhodanese-related sulfurtransferases catalyze the transfer of a sulfur atom from a donor to a nucleophilic acceptor. Additionally, reactions catalyzed by oxygenase in *Rhodococcus* strains facilitate the growth and degradation of a wide range of pollutants. Many *Rhodococcus* strains are known to perform such kind of reactions: sulfoxidation of sulfide to sulfoxide and later sulfoxidation of sulfoxide to sulfone.
- The Zn metallo-*beta* lactamase and the sulfur oxidizing (Sox) protein plays an important role in sulfate metabolism. This type of hybrid proteins could be involved in the transport of sulfate or in the metabolism of sulfur containing molecules. Therefore, this hybrid protein could have an important role in DTDB degradation.
- A putative sulfite exporter is a protein transmembrane exporter that could export the sulfite/sulfate formed as a final product outside of the cells.
- A putative acyl-CoA dehydrogenase which has the highest expression level in the exponential and the stationary phase of DTDB-grown cells. Eleven isoforms of this protein were detected;

moreover, in silico analysis revealed that this protein has no orthologues in genomes of other *R. erythropolis* strains.

6.2 Materials and Methods

6.2.1 Bacterial strains and culture conditions

The bacterial strains and plasmids used in this study are *Escherichia coli* TOP10 and *R. erythropolis* MI2. *E. coli* TOP10 was cultivated aerobically in LB (lysogeny broth) medium in a rotary shaker at 37°C x 120 rpm (Model G25, New Brunswick, Scientific Co. Inc). Antibiotics at the following concentrations were added to the required growth medium to maintain plasmids: ampicillin, 75 mg/ml; tetracycline 25 mg/ml. If required, X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranosid) and IPTG (Isopropyl-D-1-thiogalactopyranosid) were added to LB medium, in the concentration of 1 ml/l and 0.1 ml/l, respectively. *R. erythropolis* MI2 were cultivated aerobically on a rotary shaker at 30°C x 120 rpm (Pilot-Shake TM), in LB medium or in mineral salts medium (MSM) containing the required carbon source and adding the required antibiotics. All the nutrient media were sterilized by autoclaving for 20 min at 121°C and 2 x 10^5 Pa prior to use (Varioklav – Dampfsterilisator, HP Medizintechnik GmbH). All carbon sources were prepared as filter-sterilized 20% (wt/vol) stock solutions and added to the media at 2%. Solid medium contained 1.8% (wt/vol) purified agar. Cell growth was monitored by photometrical measuring the turbidity of liquid cultures (Genesys 20, Thermo Scientific); the optical density was measured at 600 nm (OD_{600}) against sterile medium. Conservation for long-term of the recombinant strains was carried out in glycerol stocks, stored at -80°C. For that purpose, cells were cultivated for overnight in LB medium with the appropriate antibiotics. An aliquot of the overnight culture (800 µl) was transferred to a sterile screw cab tube and 200 µl of sterile glycerol was added. In order to reactivate a strain an aliquot was streaked out on LB agar with the required antibiotic and incubated overnight.

6.2.2 Deletion constructs preparation

The generation of the *R. erythropolis* MI2 deletion mutants was accomplished following the cloning strategy described in **Figure 6.4**, generating a deletion construct for each target gene. The used strains and vectors generated during this work are listed in the **Table 6.2**.

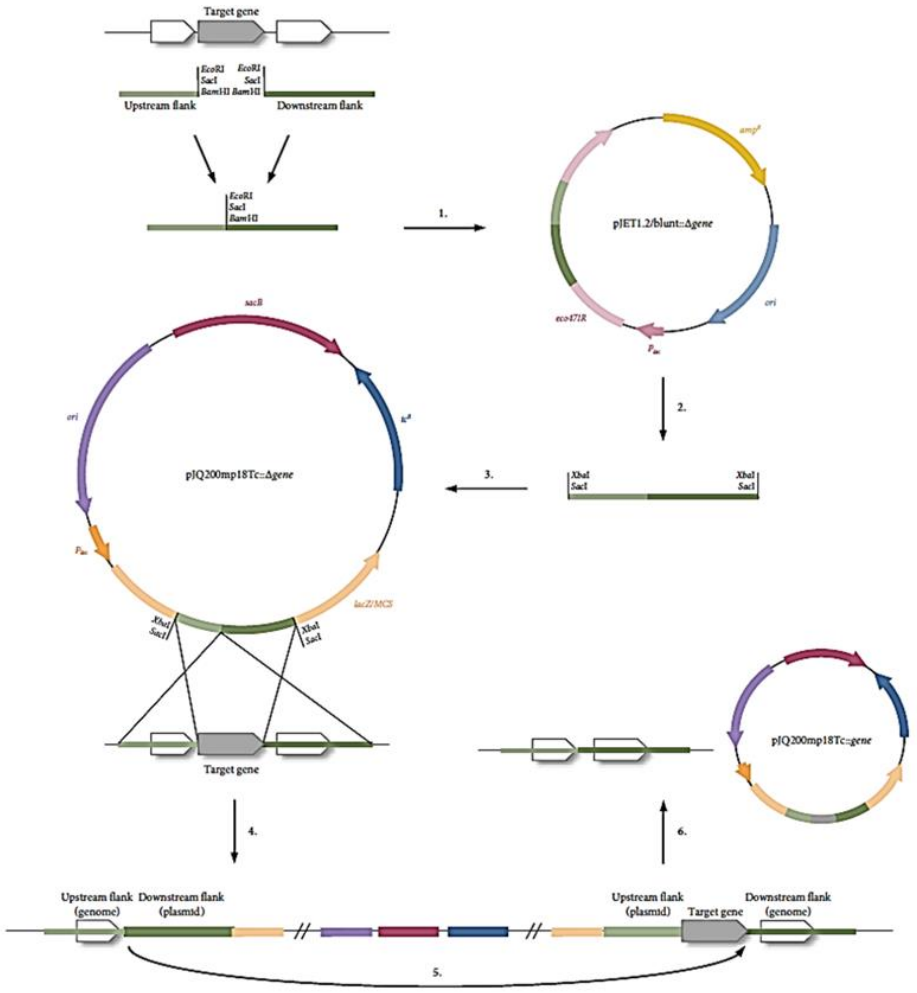


Figure 6.4. Deletion mutants scheme generation (Figure by Viktoria Heine Master Thesis)

Table 6.2. Strains and plasmids used in this work

Strain or plasmid	**Description**	**Reference or Source**
Strain		
Escherichia coli TOP10	Φ80*lac*ZΔM15, Δ*lac*X74, *deo*R, *rec*A1, *ara*D139, Δ(ara-leu)7697, *aal*U, *gal*K, *end*A1	Thermo Fisher Scientific
E. coli TOP10 (pJET1.2/blunt::02710)	*E. coli* TOP10 containing pJET1.2/blunt vector, 02710 gene flanks	This study
E. coli TOP10 (pJET1.2/blunt::02720)	*E. coli* TOP10 containing pJET1.2/blunt vector, 02720 gene flanks	This study
E. coli TOP10 (pJET1.2/blunt::02740)	*E. coli* TOP10 containing pJET1.2/blunt vector, 02740 gene flanks	This study
E. coli TOP10 (pJET1.2/blunt::02750)	*E. coli* TOP10 containing pJET1.2/blunt vector, 02750 gene flanks	This study
E. coli TOP10 (pJET1.2/blunt::66330)	*E. coli* TOP10 containing pJET1.2/blunt vector, 66330 gene flanks	This study
E. coli TOP10 (pJET1.2/blunt::06500)	*E. coli* TOP10 containing pJET1.2/blunt vector, 06500 gene flanks	This study
E. coli TOP10 (pJQ200mp18Tc::02710)	*E. coli* TOP10 containing pJQ200mp18Tc vector, 02710 gene flanks	This study
E. coli TOP10 (pJQ200mp18Tc::02720)	*E. coli* TOP10 containing pJQ200mp18Tc vector, 02720 gene flanks	This study
E. coli TOP10 (pJQ200mp18Tc::02740)	*E. coli* TOP10 containing pJQ200mp18Tc vector, 02740 gene flanks	This study
E. coli TOP10 (pJQ200mp18Tc::02750)	*E. coli* TOP10 containing pJQ200mp18Tc vector, 02750 gene flanks	This study

E. coli TOP10 (pJQ200mp18Tc::66330)	*E. coli* TOP10 containing pJQ200mp18Tc vector, 66330 gene flanks	This study
E. coli TOP10 (pJQ200mp18Tc::06500)	*E. coli* TOP10 containing pJQ200mp18Tc vector, 06500 gene flanks	This study
Rhodococcus erythropolis MI2		Wübbeler et al., 2010
R. erythropolis MI2 Δ02710	Δ02710	This study
R. erythropolis MI2 Δ02720	Δ02720	This study
R. erythropolis MI2 Δ02740	Δ02740	This study
R. erythropolis MI2 Δ02750	Δ02750	This study
R. erythropolis MI2 Δ66330	Δ66330	This study
R. erythropolis MI2 Δ06500	Δ06500	This study
Plasmids		
pJET1.2/blunt	Amp^r *eco*471R	Thermo Fisher Scientific
pJET1.2/blunt::02710	pJET1.2/blunt containing flanks of 02710 gene	This study
pJET1.2/blunt::02720	pJET1.2/blunt containing flanks of 02720 gene	This study
pJET1.2/blunt::02740	pJET1.2/blunt containing flanks of 02740 gene	This study
pJET1.2/blunt::02750	pJET1.2/blunt containing flanks of 02750 gene	This study
pJET1.2/blunt::66330	pJET1.2/blunt containing flanks of 66330 gene	This study
pJET1.2/blunt::06500	pJET1.2/blunt containing flanks of 06500 gene	This study
pJQ200mp18Tc	Tc^r, *sac*B	Pötter et al., 2005
pJQ200mp18Tc::02710	pJQ200mp18Tc containing flanks of 02710 gene	This study
pJQ200mp18Tc::02720	pJQ200mp18Tc containing flanks of 02720 gene	This study
pJQ200mp18Tc::02740	pJQ200mp18Tc containing flanks of 02740 gene	This study
pJQ200mp18Tc::02750	pJQ200mp18Tc containing flanks of 02750 gene	This study
pJQ200mp18Tc::66330	pJQ200mp18Tc containing flanks of 66330 gene	This study
pJQ200mp18Tc::06500	pJQ200mp18Tc containing flanks of 06500 gene	This study

6.2.3 Isolation of the target gene flanking regions of *R. erythropolis* MI2

The first procedure provided the construction of pJET1.2/blunt vector (**Figure 6.4**) containing the joined flanking regions of the target genes. PCR for the amplification of the six upstream and downstream target gene-flanking regions were carried out with Phusion High-Fidelity DNA polymerase (Thermo Scientific) using total genomic DNA of *R. erythropolis* strain MI2 and the oligonucleotides forward and reverse, described in Table 3. The PCR program was the following: 98°C x 1 min – (98°C x 30 sec – 67°C x 30 sec – 72°C x 30 sec) x 35 cycles – 72°C x 2 min (PeqStar 2x gradient, Peqlab). The PCR product was isolated from an agarose gel using the Monarch Gel Extraction kit (Peqlab) according to the manufacturer's instructions. Consequently, the upstream and downstream flanking regions were digested by the inner site enzymes (Fast Digest; Thermo Scientific): BamHI for 02710, 02720 and 66330, HindIII for 02740 and 02750, EcoRI for 06500. The reactions were incubated at 37°C x 4 hours. Ligation between upstream and downstream flanking regions of all the target genes were achieved with T4 DNA ligase (Thermo Scientific). The reactions were incubated in a beaker full of water applying a temperature shift starting from 25°C and then reducing the temperature to 4°C overnight. Subsequently, the reactions were inactivated at 65°C x 10 min.

6.2.4 Recombinant pJET1.2/blunt vector constructions

The entire flanking region fragments were amplified by PCR using the same program described above and the oligonucleotides described in **Table 6.3**. The fragment ligations in pJET1.2/blunt vector were performed by CloneJET PCR Cloning Kit (Thermo Scientific), according to the manufacturer's instructions. The pJET1.2/blunt vectors containing the obtained fragments were used to transform of *E. coli* TOP10 competent cells. Competent cells preparation was performed by $CaCl_2$ method and the transformation was performed by heat-shock method. Transformants were selected on LB medium containing ampicillin ($LB_{Ap,}$ 75 mg/ml), and hybrid plasmids were isolated by GeneJET Plasmid Miniprep kit (Thermo Scientific), according to the manufacturer's instructions. Then the plasmids were digested by XbaI enzyme, in order to prepare them for ligation into the pJQ200mp18Tc suicide vector; then they were isolated from an agarose gel by the Monarch Gel Extraction kit (Peqlab), according to the manufacturer's instructions.

6.2.5 *R. erythropolis* MI2 deletion mutant generations

The pJQ200mp18Tc suicide vector was used for the directed deletion of target genes without leaving any markers in the genome of *R. erythropolis* MI2. The deletion constructs were transferred in *R. erythropolis* MI2 via electroporation.

6.2.6 Recombinant pJQ200mp18Tc suicide vector constructions

The obtained flanking region fragments were ligated into the pJQ200mp18Tc suicide vector. The generated vectors were used to transform *E. coli* TOP10 competent cells, as described above. Transformants were selected by blue/white screening method on LB medium containing IPTG, X-Gal and Tetracycline (12.5 µg/ml), and hybrid plasmids were isolated by GeneJET Plasmid Miniprep kit (Thermo Fisher Scientific), according to the manufacturer's instructions. The isolated plasmids were sequenced using the M13 primers, in order to verify the correct sequences of the obtained constructs. Site direct mutagenesis by PCR was performed in order to change the single mutation present in the sequence of three plasmids: pJQ200mp18Tc::02710; pJQ200mp18Tc::66330; pJQ200mp18Tc::06500. The used primers are listed in the **Table 6.3**.

Table 6.3. Oligonucleotides used in this work

Name	Sequence (3´-5´)	Restriction site enzyme	Type
02710_upFl_fw	AAATCTAGAACCGATGAAGTATCCGGCTTGG	XbaI	PCR
02710_upFl_rv	AAAGGATCCATGGTCCTCCTGTTCGAGACACG	BamHI	PCR
02710_dwFl_fw	AAAGGATCCGCTGTCGACAGCCGAGGAGTCAC	BamHI	PCR
02710_dwFl_rv	AAATCTAGAGGTGACATCGAAAGCATCGCG	XbaI	PCR
02720_upFl_fw	AAATCTAGAGTGTGCGCGGACATATGGTCC	XbaI	PCR
02720_upFl_rv	AAAGGATCCCGAAATCTCCTTGTCGATCGTGG	BamHI	PCR
02720_dwFl_fw	AAAGGATCCCCGGGCACGCCCGAATC	BamHI	PCR
02720_dwFl_rv	AAATCTAGAGCAGCGAAGAGACTGTAGAGCACACC	XbaI	PCR
02740_upFl_fw	AAATCTAGAGCGTTGACGACGAGCTATGGC	XbaI	PCR
02740_upFl_rv	AAAAAGCTTGAAGAACGCTCCTTGGATCAGTGG	HindIII	PCR
02740_dwFl_fw	AAAAAGCTTCATGGTCGCCACCCTCGCGA	HindIII	PCR
02740_dwFl_rv	AAATCTAGACATGGAGGCCTGGACGGTGACG	XbaI	PCR
02750_upFl_fw	AAATCTAGACAATACCGTTGGCGCTCGACTATG	XbaI	PCR
02750_upFl_rv	AAAAAGCTTGTCAGATCACTGCGCCGGGG	HindIII	PCR
02750_dwFl_fw	AAAAAGCTTCGCATCGTTGCCACTCTAGGGT	HindIII	PCR
02750_dwFl_rv	AAATCTAGACCCGCCGTGATGACCGATC	XbaI	PCR
66330_upFl_fw	AAATCTAGACTGCGGTCGGGAACCTGC	XbaI	PCR
66330_upFl_rv	AAAGGATCCACGCCGTGCCGAGAACAATG	BamHI	PCR
66330_dwFl_fw	AAAGGATCCCCTTCAAGAATTCATCGCAAACCAG	BamHI	PCR
66330_dwFl_rv	AAATCTAGACGCACAGCGTCGTGGTCAC	XbaI	PCR
06500_upFl_fw	AAATCTAGAGGCGTAGTGGCTGGTCGAAGTG	XbaI	PCR
06500_upFl_rv	AAAGAATTCTCGACGAGTCTAGTTCGACCGG	EcoRI	PCR
06500_dwFl_fw	AAAGAATTCGCGAGCAGATCGACTACCCGC	EcoRI	PCR
06500_dwFl_rv	AAATCTAGAGGCGACCAACTGATCTGCGATC	XbaI	PCR
M13_fw	CGCCAGGGTTTTCCCAGTCACGAC		PCR
M13_rv	CAGGAAACAGCTATGAC		PCR
ssMut_02710_Fw_2	GGAGTTCGCGTGTCCGTcCATCGGAATGTA		5‘-PHO
ssMut_02710_Rv_2	AACTCCGCATCGGGCGTCAGAGATTGCCAG		5‘-PHO
ssMut_66330_Fw_2	CGGATGCCTACgCGTGGTTGCTCGAGCAGG		5‘-PHO
ssMut_66330_Rv_2	TGCAATCATTCACCGGCACCGGATGGGGGT		5‘-PHO
ssMut_06500_Fw	CTCcTCGACGAGCCGAGGCGTTTGCCATCG		5‘-PHO
ssMut_06500_Rv	CCGCGTCGGTCAAGTCGTCTCGGGTGAGGT		5‘-PHO
inSQR_02710_Fw	ATCGCGTACCTCGCAGCCGAC		PCR
inSQR_02710_Rv	ATCGTCTTGCGAGAGGTGGTC		PCR
inZnlact/rhod_02720_Fw	AACAGATCTTCCGCAGTCTGAA		PCR
inZnlact/rhod_02720_Rv	GATGCCGACCCGCAGCAGGTG		PCR
inSulfTr_02740_Fw	ATCCGGCCGCCGCACACGCAC		PCR
inSulfTr_02740_Rv	GTGCCGTCATCACGGACCAC		PCR
inTauE_02750_Fw	CGTCGGAATCACCTCGATCATC		PCR
inTauE_02750_Rv	CGACGATCTCCCACGGGATG		PCR
inAcCoADH_66330_Fw	AAGCTGATCGATTTCGGGGTGGG		PCR
inAcCoADH_66330_Rv	TCGGTGTACGAGAAGCCCATC		PCR
inDesulf_06500_Fw	CGCTCTCCGGCGGGATACTCG		PCR
inDesulf_06500_Rv	CGCATACAGTCGAGACATGGG		PCR
out02710_Fw	CGCCAGGTCCAGATACACACC		PCR
out02710_Rv	GCCACCATTATCACCTTTACC		PCR
out02720_Fw	CTCGCCGCAGACGTGCTGCGA		PCR
out02720_Rv	GCGACCGCAGGGTCGGTCCG		PCR
out02740_Fw	AACTGCATGCCGATCTGGTCC		PCR
out02740_Rv	TCCGGCGGGCAGCGGAGGGAG		PCR
out02750_Fw	GGTCGCTGCCACCCTGGAAG		PCR
out02750_Rv	CCTGCGGCCGTGGACCAGCG		PCR
out66330_Fw	ATCCTCCCTGCAATCATCAT		PCR
out66330_Rv	GTGATCATCATGTCGGTCCA		PCR
out06500_Fw	CTCGCATTGGCAATCAAAGGG		PCR
out06500_Rv	CGTGTGAGTGTCCTCGTACTG		PCR

6.2.7 *R. erythropolis* MI2 competent cells preparation

R. erythropolis MI2 competent cells were prepared starting by an inoculum in LB medium containing glycine (1,5%) and sucrose (1,8%). The inoculum was incubated at 30°C under shaking until the OD_{600} reached 0,5-0,7 values. After the supplement of lysozyme (40 mg/ml), the inoculum was incubated 1.5 h at 30°C under shaking. The cells are collected by centrifugation at 4°C (6000 rpm x 10 min); then, the pellet was resuspended in 1/2 of the initial volume of cold EPB1 (HEPES 40 mM pH 7.2 (50%), glycerol 10%). After centrifugation at 4°C (6000 rpm x 10 min) the pellet was resuspended in 1/4 of the initial volume of cold EPB1; then, the pellet deriving from the centrifugation at 4°C (6000 rpm x 10 min) was resuspended in 1/10 of the initial volume of cold EPB2 (HEPES 40 mM pH 7.2 (12.5%), glycerol 30%). Finally, after centrifugation at 4°C (6000 rpm x 10 min) the pellet was resuspended in 500 µl of EPB2 and then divided in aliquots of 100 µl; then they were stored at -80°C.

6.2.8 *R. erythropolis* MI2 transformation by the deletion constructs

The isolated recombinant plasmid (pJQ200mp18Tc), after dialysis by Millipore filters (0.025 µm), was used to transform *R. erythropolis* MI2 by electroporation. Plasmids were introduced into the strain using an Electroporator 2510 (Eppendorf) set at 2.50 kV in presence of 1 µg DNA. Immediately after electroporation, 2.5 ml recovery broth (LB medium with 1.8% sucrose) were added and cells were incubated at 30°C for 4 h. Cells were plated on LB supplemented with tetracycline 25 µg/ml and incubated at 30°C for 3–4 days. The transformants, also referred as “heterogenotes”, were checked for the presence of the plasmid in the chromosome by PCR on gDNA. The generation of the corresponding deletion mutants, so called “homogenotes”, implies that homologous recombination had occurred within the heterogenotes; in this case, target genes are carried over to the suicide vector before the plasmid is cured from the cells. Therefore, to evaluate if the homologous recombination happened, a double selection was performed: a positive selection on LB medium containing sucrose (15%) and a negative selection on MSM containing glucose (2%) and tetracycline (25 µg/ml). Colonies that grew on LB + sucrose while showing no growth on MSM + glucose + Tc were screened for the homogenates identification. To verify potential deletion mutants and distinguish those from re-established wild type cells, PCR from gDNA with two sets of primers were performed. On the one hand a fragment within the target gene was amplified which should not occur in the corresponding deletion mutant. Furthermore, oligonucleotides were used that bind to the gDNA some base pairs outside the flanking regions and the PCR with these primers should result in a shortened fragment for the deletion mutants compared to the wild type because the target gene is absent.

6.3 Results

The aim of this work was the generation of *R. erythropolis* MI2 marker-free deletion mutants in order to investigate DTDB utilization of the Gram-positive bacterium; the first step of the work has been the deletion of the genes involved in the desulfhydration of the 4-oxo-4-sulfanylbutyric acid (**Table 6.1**), using the suicide vector pJQ200mp18Tc.

6.3.1 *Rhodococcus erythropolis* MI2 characterization

In order to characterize the strain, kinetic growth curves on different basal nutrient medium were carried out in triplicate. The first kinetic growth curve of *R. erythropolis* MI2 was carried out on LB medium (**Figure 6.5**) and the growth was measured by spectrophotometer (experiments in triplicate):

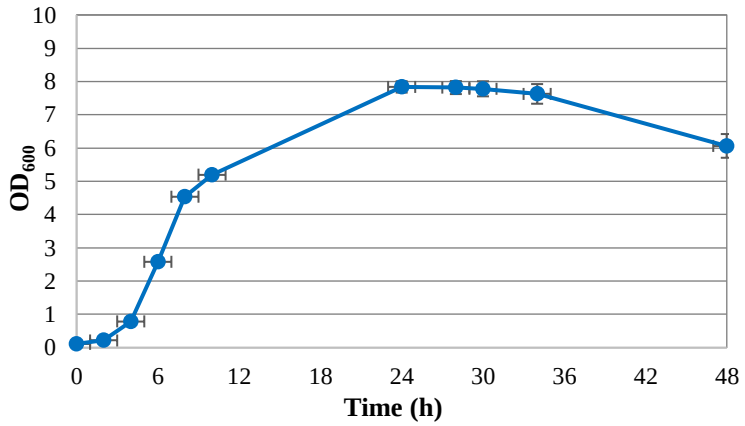


Figure 6.5. Kinetic growth curve of *R. erythropolis* MI2 on LB medium

The second kinetic growth curve of *R. erythropolis* MI2 was carried out on MSM medium (**Figure 6.6**) added with glucose 2%. In this case, the growth was slow and so it was measured every 24 hours by spectrophotometer (experiments in triplicate):

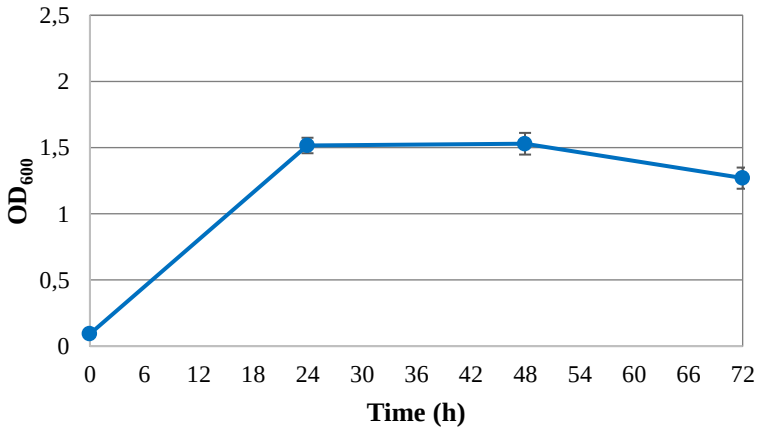


Figure 6.6. Kinetic growth curve of *R. erythropolis* MI2 on MSM + glucose 2%

6.3.2 Deletion mutants

Six genes putatively involved in the removal of the sulfur group in the DTDB pathway were the study objects of this work. Previously proteomic analysis revealed that they are highly up regulated during growth on DTDB. Therefore, a deletion mutant generation approach is necessary to verify their involvement in the pathway. The genomic regions containing the identified genes are described in **Figure 6.7 A-B**.

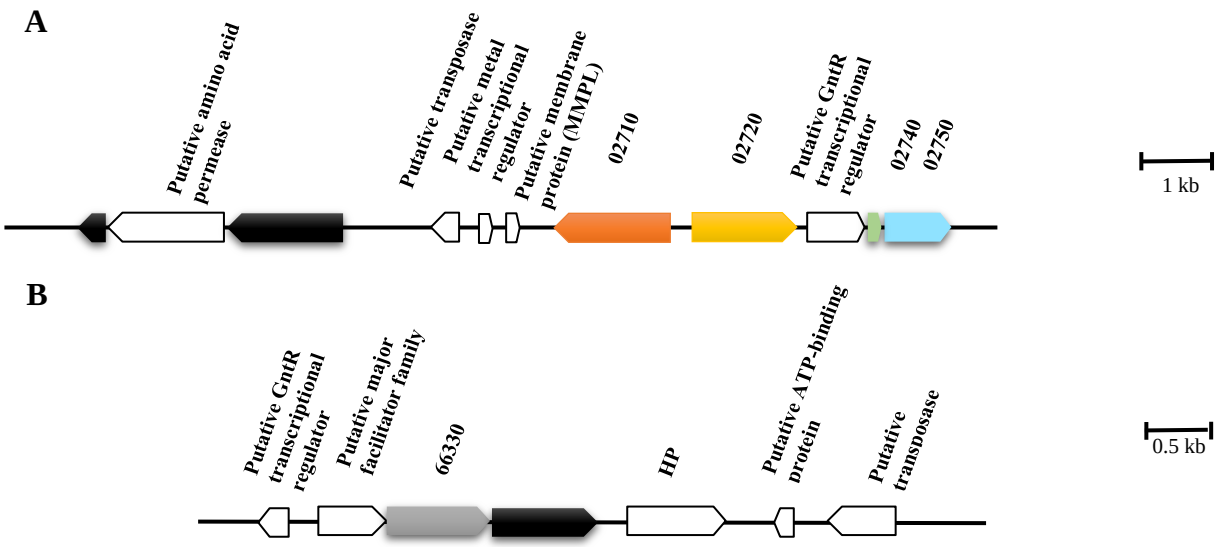


Figure 6.7. A-B Genomic regions containing the target genes of *R. erythropolis* MI2

6.3.3 Deletion of genes proposed to be involved in metabolism of DTDB

Supportive of generating marker-free deletion mutants, the upstream and downstream regions of the respective target gene were separately amplified and then joined in order to obtain the entire fragment to insert in the pJET1.2/blunt vector (02710, 1409 bp; 02720, 1330 bp; 02740, 1520 bp; 02750, 1284 bp; 66330, 1223 bp; 06500, 1242 bp) (**Figure 6.8 A-B**). This preliminary step is necessary because of the difficulty to manipulate the genome of *Rhodococcus* strains.

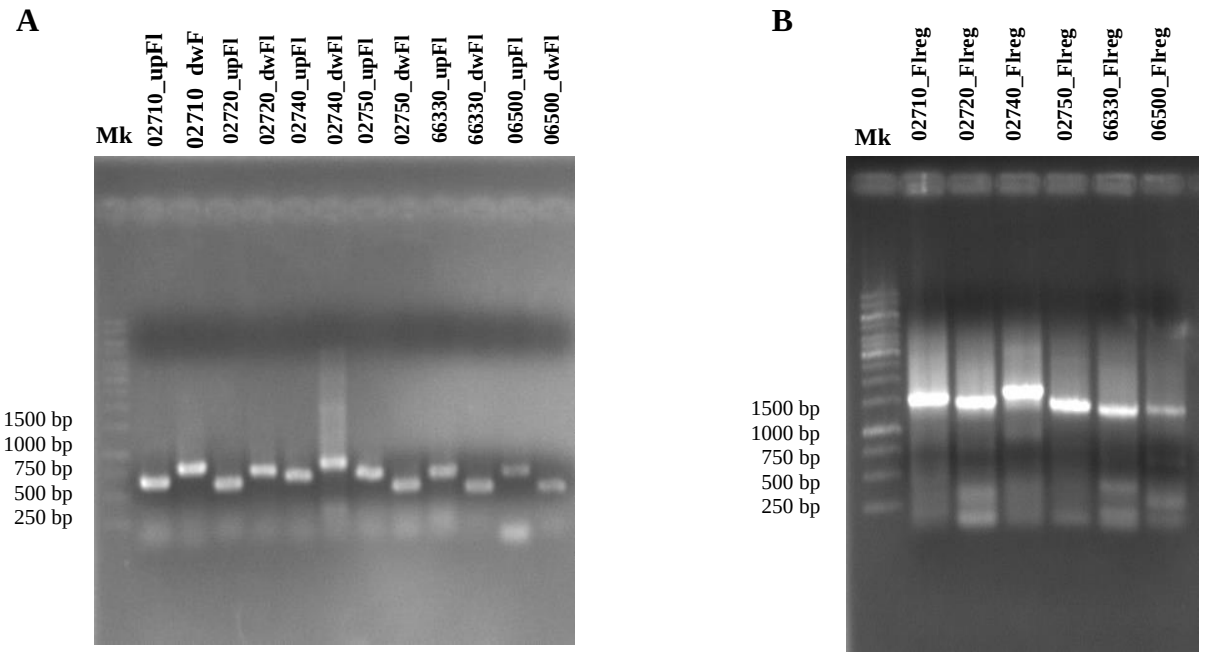


Figure 6.8. A) Amplified target gene upstream and downstream flanking regions **B**) Amplified whole fragments of interest Mk: 250 to 10000 bp marker

The vectors containing the obtained fragments were used to transform *E. coli* TOP10 in order to generate multiple copies of the plasmids. Therefore, after the *E. coli* TOP10 transformations and transformant screening, plasmid extractions were performed.

Consequently, in order to prepare them for ligation into the pJQ200mp18Tc suicide vector, a digestion by the outer site enzyme was performed (**Figure 6.9**).

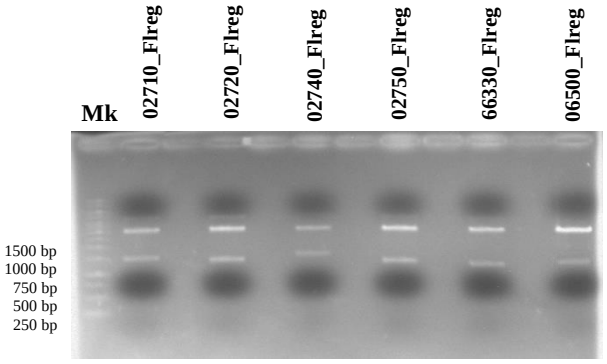


Figure 6.9. pJET1.2/blunt containing fragments of interest digestion (plasmids extracted from *E. coli* TOP10 transformants)

The digested fragments were amplified and cloned into the suicide vector pJQ200mp18Tc. The deletion constructs were subsequently transferred in *E.coli* TOP10. Then, after the *E. coli* TOP10 transformations and transformant screening, plasmid extractions were performed. The extracted plasmids were sequenced in order to verify the correct sequences of the inserted fragments. Three of the six constructs, particularly those containing the 02720, 02740 and 02750 flanking regions, possessed a completely correct sequence. The other three constructs presented a single mutation in

the sequence. In order to obtain the correct insert sequences, site-directed mutagenesis by PCR was performed. After the mutations, also the other three constructs were obtained.
Subsequently, the following deletion constructs were obtained: pJQ200mp18Tc::02710, pJQ200mp18Tc::02720, pJQ200mp18Tc::02740, pJQ200mp18Tc::02750, pJQ200mp18Tc::66330 and pJQ200mp18Tc::06500.

6.3.4 *Rhodococcus erythropolis* MI2 mutant generation

The generation of the deletion mutants was performed following the strategy described in **Figure 6.2**. For the identification of a deletion mutant and to verify that the target gene is absent from its genome, PCR was performed using the forward and the reverse primers of the upstream and the downstream flanking regions, respectively. In this way, the heterogenotes were identified (**Figure 6.10**).

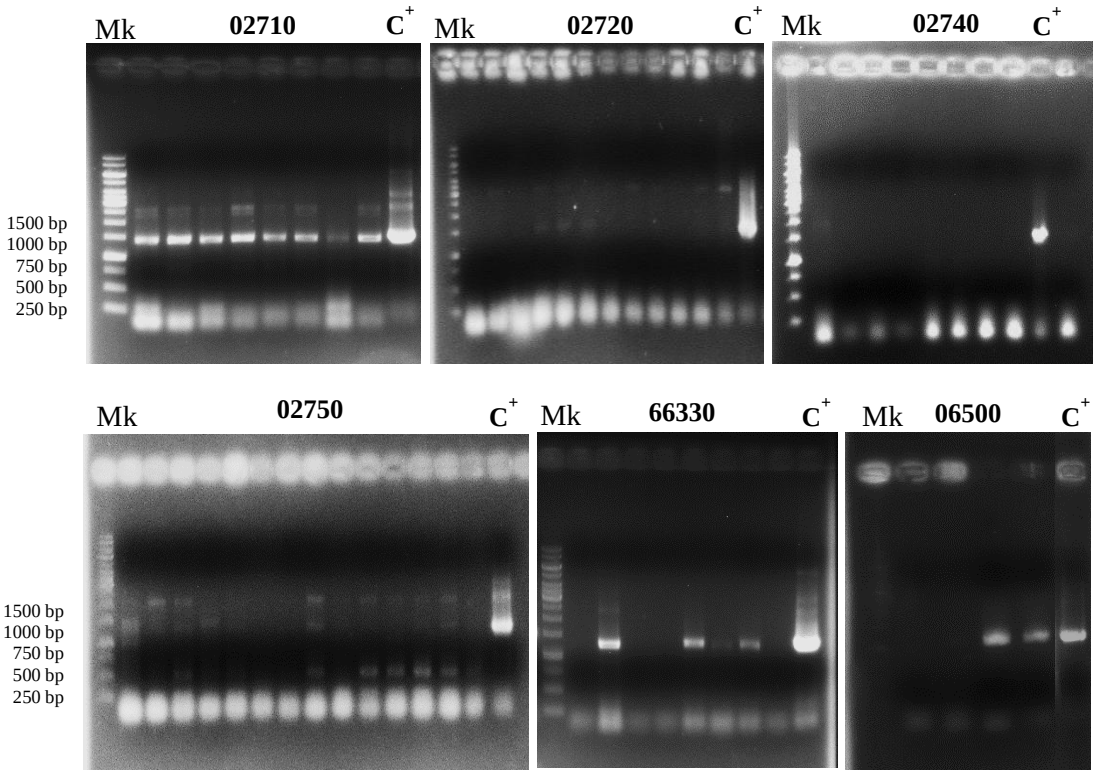


Figure 6.10. *R. erythropolis* MI2 heterogenotes identification for the six kinds of generated mutants. Mk: 250 to 10000 bp marker, C^+: positive control (amplification of the fragments from gDNA). Amplifications revealed the presence of the constructs in MI2 cells.

The identification of the homogenotes was carried out by a double selection on plates of LB medium containing sucrose (15%) and a negative selection on plates of MSM containing glucose (2%) and tetracycline (25 μg/ml) (**Figure 6.11**). No growth on the second type of plates is expected, because of homologous recombination would be happened.

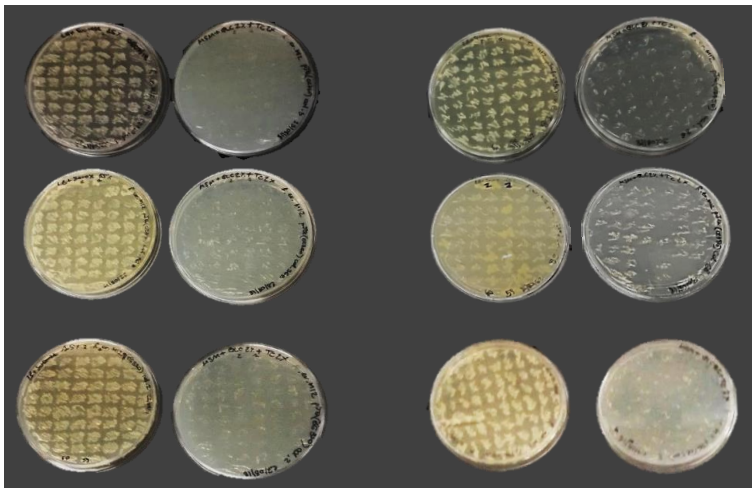


Figure 6.11. Probable homogenotes identification for the six kinds of generated mutants.

To verify potential deletion mutants and distinguish them from re-established wild type cells, PCR from gDNA with two sets of primers was performed. Internal primers binding within the target gene were used to assure its deletion from the genome of *R. erythropolis* MI2 (**Figure 6.12**).

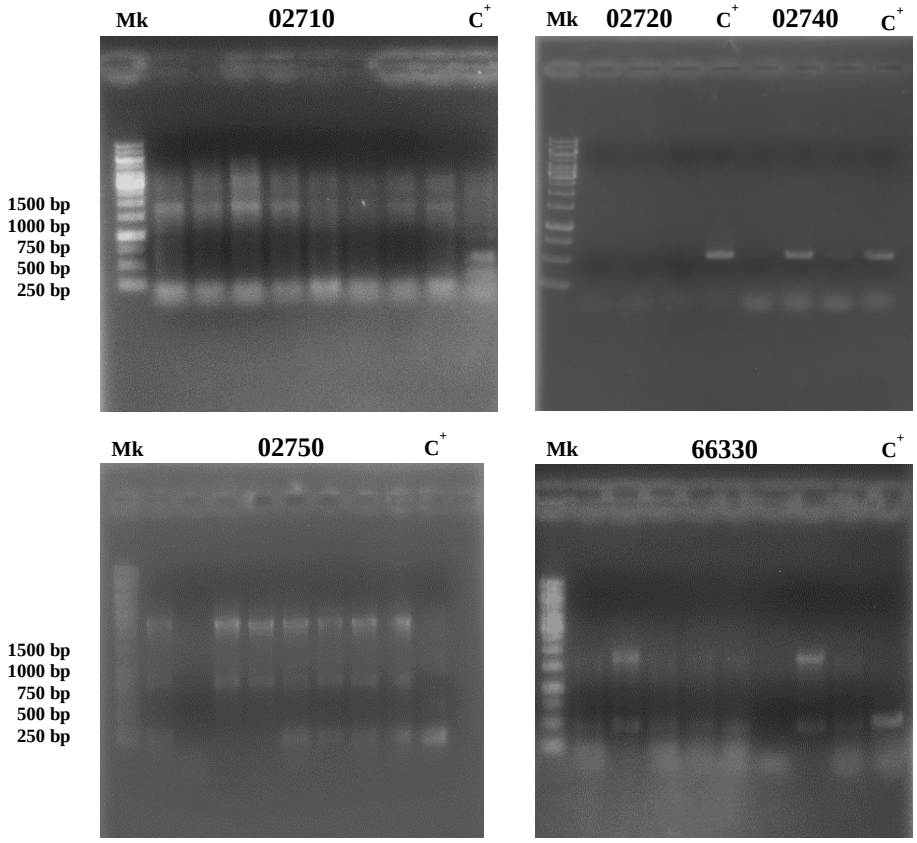


Figure 6.12. *R. erythropolis* MI2 homogenotes identification for five of the six kinds of generated mutants. Mk: 250 to 10000 bp marker, C^+: positive control (amplification of the fragments from gDNA). No amplification revealed the probable homogenotes.

External primers with a binding site outside of the flanking regions should yield an abbreviated PCR product compared to the wild type.

In summary, for a deletion mutant no signal was expected from internal PCR, but a fragment shortened by the size of the target gene should be amplified during external PCR. The genomic DNA of the wild type strain was always used as a control during PCR. The expected internal and external PCR product sizes for all target genes are summarized in **Table 6.4.**

Table 6.4. Sizes of expected internal and external PCR products for all target genes

Target gene	Internal fragments in wt (bp)	External fragments deletion mutant/wt (bp)
02710	504	1509/2686
02720	502	1433/2828
02740	254	1620/1959
02750	497	1385/2228
66330	504	1323/2937
06500	500	1344/2022

6.4 Discussion

R. erythropolis strains MI2 is able to catabolize DTDB and to use it as sole carbon and energy source. *Rhodococcus* species possess the enzymatic machinery for the utilization of DTDB. Many strains of *R. erythropolis* are metabolically flexible and able to perform to degrade xenobiotics and complex carbon sources, which are usually hard to be metabolized by bacteria. In addition, all members of the genus *Rhodococcus* have the capacity to cleave sulfur-sulfur bonds. Based on the information gained in previous studies a putative catabolic pathway of DTDB in *R. erythropolis* MI2 was postulated (Khairy et al., 2016). In this chapter, the study of the sulfur group removal step of the pathway was examined by generation of marker-free deletion mutants for six target genes previously identified by proteomic analysis. For all these six genes, a suicide vector containing the flanking regions of the respective genes were constructed. These vectors were used to transform *R. erythropolis* MI2 and six types of mutants were generated. The analyses revealed that the mutants could be defined as heterogenotes because the first homologous recombination had happened. The second recombination (homogenotes generation) was evaluated testing by a double selection in order to verify the ability of the mutants to grow in presence of tetracycline; moreover, testing the presence of the target genes in the mutant genomes. The expected result, that is the absence of the genes in their genome, was verified for five of the six types of the obtained mutants; consequently, it is predictable that these five mutants of the strain lost its capability to metabolize DTDB.

6.5 *References*

Imbernon L, Oikonomou EK, Norvez S and Leibler L (2015). Chemically crosslinked yet reprocessable epoxidized natural rubber via thermo-activated disulfide rearrangements. Polym Chem, 6:4271-4278.

Jang L and Keng H (2006). Development and characterization of as a monolayer for protein chips. Sens Mater, 18:367-380.

Kanayama N and Kitano H (2000). Interfacial recognition of sugars by boronic acid-carrying self-assembled monolayer. Langmuir, 16:577-583.

Khairy H, Wübbeler JH and Steinbüchel A (2015). Biodegradation of the organic disulfide 4,4'-dithiodibutyric acid by *Rhodococcus* spp. Appl. Environ. Microbiol., 81:8294-8306.

Khairy H, Wübbeler JH, Steinbüchel A (2016a). The NADH:flavin oxidoreductase Nox from *Rhodococcus erythropolis* MI2 is the key enzyme of 4,4'-dithiodibutyric acid degradation. Letters in Applied Microbiology. ISSN 0266-8254

Khairy H, Meinert C, Wübbeler JH, Poehlein A, Daniel R, Voigt B, et al. (2016b). Genome and Proteome Analysis of *Rhodococcus erythropolis* MI2: Elucidation of the 4,4'-Dithiodibutyric Acid Catabolism. PLoS ONE, 11(12):e0167539.

Kim DY, Elbanna K, Thakor N, Lütke-Eversloh T and Steinbüchel A (2005). Poly(3-mercaptopropionate): a non-biodegradable biopolymer? Biomacromolecules, 6:897-901.

Lütke-Eversloh T, Kawada J, Marchessault RH and Steinbüchel A (2002a). Characterization of biological polythioesters: physical properties of novel copolymers synthesized by *Ralstonia eutropha*. Biomacromolecules, 3:159-166.

Lütke-Eversloh T and Steinbüchel A (2004). Microbial polythioesters. Macromol Biosci, 4:166-174.

Steinbüchel A (2001). Perspectives for biotechnological production and utilization of biopolymers: metabolic engineering of polyhydroxyalkanoate biosynthesis pathways as a successful example. Macromol Biosci, 1:1-24.

Soil Survey Staff (2010). "Keys to Soil Taxonomy." Twelfth Edition, United States Department of Agriculture, Natural Resources Conservation Service.

Wübbeler JH, Bruland N, Wozniczka M and Steinbüuchel A (2010). Biodegradation of the xenobiotic organic disulfide 4,4'-dithiodibutyric acid by *Rhodococcus erythropolis* strain Ml2 and comparison with the microbial utilization of 3,3'-dithiodipropionic acid and 3,3'-thiodipropionic acid. Microbiology , 156:1221-1223.

Chapter 7.

Conclusions and Future Perspectives

The main purpose of this thesis was the genomic characterization of bacteria belonging to *Rhodococcus* genus, in order to obtain molecular tools to use as "marker" sequences for the soils environmental quality.

Rhodococcus genus members are particularly interesting for their metabolic versatility and their tolerance to various environmental stresses; for these reasons, they play an important role in nutrient cycling of soil and in bioremediation, biotransformation and biocatalysis potential applications. Consistent with the immense catabolic diversity of rhodococci, they often possess large genomes, which contain a multiplicity of catabolic genes, a high genetic redundancy of biosynthetic pathways often connected to sophisticated regulatory networks. This work focused the attention on three *Rhodococcus* strains: *R. opacus* R7, *R. aetherivorans* BCP1 and *R. erythropolis* MI2. Their genomes were characterized for the identification of several genes involved in the degradation of different compounds and in their persistence in stressed environments.

A genome-based analysis of *R. opacus* R7 and *R. aetherivorans* BCP1 genomes was performed to identify several gene clusters putatively involved in the degradation of organic/xenobiotic compounds, belonging to four categories: aliphatic hydrocarbons and cycloalkanes, BTEX and aromatic compounds, polycyclic aromatic compounds (PAHs), naphthenic acids and other carboxylic acids. The genome analysis of BCP1 and R7 highlighted a great number of interesting features, underlying the peculiar capacities of these two rhodococci for biodegradation and biotransformation applications supported by both their extraordinary genetic repertoire and environmental persistence. Results revealed that BCP1 shows a lower variability in terms of genes predicted to be involved in organic/xenobiotic compounds degradation compared to *R. opacus* R7. These genetic features can be related to the genetic pressure exerted in the different niches from which they were isolated and to the relative dimension of the two genomes, that can explain the different degree of genetic redundancy present in BCP1 compared to R7.

The use of *Rhodococcus* spp. in bioremediation, biotransformation and biocatalysis applications is further enhanced by their persistence in the environment (LeBlanc et al., 2008; de Carvalho, 2012). Indeed, the ecological importance of *R. aetherivorans* BCP1 and *R. opacus* R7 was investigated for their peculiar capability to resist under stress conditions. Their responses to various stress conditions and several antimicrobials was examined by a Phenotype Microarray (PM) approach that correlates different phenotypes to genetic determinants. Comparison between metabolic activities and genetic features of BCP1 and R7 provided new insight into the environmental persistence of these two members of the genus *Rhodococcus*. Results revealed that *R. opacus* R7 has a higher resistance than *R. aetherivorans* BCP1 to acidic pHs and to a few osmolytes; conversely, despite the different genome size, the two strains have similar behaviours in presence of metals, antibiotics, oxidative agents and

several other toxic compounds. Both strains are expected to have similar behaviour for environmental remediation technologies and transformation of toxic compounds, in response to stressors that can occur in contaminated soils.

Besides all the analysed metabolisms, an in-depth characterization of *o*-xylene and naphthenic acids degradation pathways was performed using a genome-based approach, supported by functional analyses.

o-Xylene metabolism is one of the most interesting because *o*-xylene is the most recalcitrant xylene isomer, not so commonly reported in bacteria and it has been poorly examined. Information deriving from literature combined to the computational analyses of all the R7 oxygenases allowed to identify the gene clusters putatively involved in the *o*-xylene pathway. Moreover, in order to investigate their putative involvement in *o*-xylene metabolism, functional analysis (mutant analysis, cloning experiments, identification of metabolites, gene expression experiments) allowed to validate the hypothesis deriving from the bioinformatic investigations. The genome-based analysis of *R. opacus* R7 revealed a considerable multiplicity of genes potentially involved in *o*-xylene catabolism. In particular, three systems were identified to be involved in this degradation route; we demonstrated that R7 is able to degrade *o*-xylene by the activation of the *akb* gene cluster (a dioxygenase system) leading to the production of the corresponding dihydrodiol. Also, the redundancy of sequences encoding for several monooxygenases/phenol hydroxylases, supports the involvement of the *prm* and *phe* gene clusters that induce the formation of phenols, successively converging to the phenol oxidation route. The activation of multiple converging oxygenase systems represents a strategy in bacteria of *Rhodococcus* genus to degrade a wide range of recalcitrant compounds and to persist in severely contaminated environments.

Naphthenic acids (NAs) are new emerging and recalcitrant contaminants deriving from both natural and anthropogenic processes. The analysis for the identification of gene clusters involved in their degradation was performed using cyclohexane carboxylic acid (CHCA) chosen as model compound. The results revealed the presence in R7 genome of *R. aetherivorans* BCP1 homologous sequences allocated in two different regions of R7 chromosome: the first region contains the *p*-hydroxybenzoate hydroxylase gene (*pobA*); the second region contains homologous genes to BCP1 (long-chain-fatty-acid-CoA ligase (*aliA1*), 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase (*badH*), naphthoate synthase (*badI*), acyl-CoA dehydrogenase (*badJ*)). The involvement of R7 genes in the CHCA degradation was evaluated by RT-PCR experiments, revealing that CHCA induced the transcription of the *aliA1* gene. Therefore, *aliA1* gene was chosen as marker sequence to represent R7 gene clusters in a culture-independent approach, to monitor the contaminant degradation and to evaluate the soil quality. Slurry-phase microcosms were established using sand collected from a non-contaminated

riverside spiked with CHCA. The microcosms were bioaugmented with *R. opacus* R7 in order to follow the CHCA degradation and to compare its activity to that of the autochthonous CHCA degrading bacteria. The biodegradation results indicated that R7 strain degrade the contaminant faster than the microbial community and that its contribute increased the CHCA degradation considerably up to 30 hours. Moreover, RT-PCR and qRT-PCR exeriments revealed that the expression of the *aliA1* gene in the microcosms was strictly related to the presence of R7 strain. These results demonstrated the possibility to follow the R7 CHCA degradation using *aliA1* gene as a marker sequence in a model microcosm system. In general, this work showed the helpfulness of the identification of marker sequences with a genome-based approach for the assessment of soil activity and quality, using microcosms experiments.

During my period abroad, in collaboration with the Institut für Molekulare Mikrobiologie und Biotechnologie of the Westfälische Wilhelms-Universität of Münster, a potential biotechnological application of *Rhodococcus* genomes analysis was examined. *R. erythropolis* MI2 genome was investigated for the degradation of a potential polythioester (PTE) precursor, 4,4'-dithiodibutyric acid (DTDB). Therefore, *R. erythropolis* MI2 marker-free deletion mutants were generated in order to investigate DTDB metabolic pathway, in particular studying the steps involved in the removal of sulfur group. The first step has been the deletion of the identified genes involved in the desulfhydration of the 4-oxo-4-sulfanylbutyric acid using the suicide vector pJQ200mp18Tc; it contains the *oriT* of plasmid RP4 for conjugative mobilization and the *sacB* system as well as the tetracyclin resistance as primary selection marker. The deletion mutant construction was performed preparing the deletion vectors for the six genes of interest. The deletion mutants (homogenotes) were tested by a double selection screening, testing the presence of the target genes in the mutant genomes. Upcoming analyses will be performed for the evaluation of the homogenotes to confirm the sequence insertion and successively the evaluation of mutant ability to grow on DTDB.

Overall, this thesis highlights the wide capabilities of *Rhodococcus* strains and the great potential of their genomes. Genome-based analyses in *Rhodococcus* strains led to characterize several interesting degradative functions and the related genetic determinants. Moreover, microcosm experiments allowed to demonstrate the potential application of the *Rhodococcus* "marker" sequences for soil quality assessment. However, since *Rhodococcus* strains could be defined as a powerhouse of numerous degradative and synthetic functions, in-depth analysis of their genomes is required to discover other molecular tools and degradative functions useful for both environmental and biotechnological applications.

Scientific Contributions

Research papers

- **Zampolli J, Zeaiter Z, Di Canito A and Di Gennaro P** (2018). Genome analysis and -omics approaches provide new insights into the biodegradation potential of *Rhodococcus*. Mini-Review. Applied Microbiology and Biotechnology.

- **Di Canito A, Zampolli J, Orro A, D'Ursi P, Milanesi L, Sello G, Steinbüchel A and Di Gennaro P** (2018). Genome-based analysis for the identification of genes involved in *o*-xylene degradation in *Rhodococcus opacus* R7. BMC Genomics.

- **Cappelletti M, Fedi S, Zampolli J, Di Canito A, D'Ursi P, Orro A, Viti C, Milanesi L, Zannoni D, Di Gennaro P** (2016). Phenotype Microarray analysis to unravel genetic determinants involved in stress response by *Rhodococcus aetherivorans* BCP1 and *R. opacus* R7. Research in Microbiology.

Conference contributions

- **Di Canito A, Zampolli J, D'Ursi P, Orro A, Milanesi L, Di Gennaro P**. Genome analysis of *Rhodococcus opacus* R7 strain unravel genetic determinants involved in its environmental persistence. Presented at SIMGBM Congress, Palermo, 2017.

- **Zampolli J, Di Canito A, Collina E, Di Gennaro P**. Gene redundancy for *o*-xylene degradation in *Rhodococcus opacus* R7. Presented at SIMGBM Congress, Palermo, 2017.

- **Di Canito A, Zampolli A, Cappelletti M, Orro A, D'Ursi P, Milanesi L, Di Gennaro P**. Development of new molecular tools from *Rhodococcus* strains to evaluate soil quality and to enhance microbial degradation in soil bioremediation. Presented at International Meeting on New Strategies in Bioremediation Processes, BioRemid, Granada, 2017.

- **Orro A, Cappelletti M, D'Ursi P, Milanesi L, Di Canito A, Zampolli A, Collina E, Decorosi F, Viti C, Fedi S, Presentato A, Zannoni D, Di Gennaro P**. Genome and Phenotype Microarray analyses of two *Rhodococcus* strains with environmental and industrial relevance. Conference Phenotype MicroArray Analysis of Cells, Firenze 2015.

Annex

- Cappelletti M et al., 2016
- Di Canito A et al., 2018
- Zampolli J et al., 2018

Institut Pasteur
Research in Microbiology 167 (2016) 766–773
www.elsevier.com/locate/resmic

Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7

Martina Cappelletti [a,*], Stefani Fedi [a], Jessica Zampolli [b], Alessandra Di Canito [b], Pasqualina D'Ursi [c], Alessandro Orro [c], Carlo Viti [d], Luciano Milanesi [c], Davide Zannoni [a], Patrizia Di Gennaro [b]

[a] *Department of Pharmacy and Biotechnology, University of Bologna, Via Irnerio 42, 40126 Bologna, Italy*
[b] *Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milano, Italy*
[c] *CNR, Institute for Biomedical Technologies, Via Fratelli Cervi 93, 20090 Segrate, Milan, Italy*
[d] *Department of Agrifood Production and Environmental Sciences, University of Firenze, Piazzale delle Cascine 24, 50144 Florence, Italy*

Received 24 March 2016; accepted 24 June 2016
Available online 7 July 2016

Abstract

In the present study, the response of *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7 to various stress conditions and several antimicrobials was examined by PM in relation with genetic determinants, as revealed by annotation analysis of the two genomes. Comparison between metabolic activities and genetic features of BCP1 and R7 provided new insight into the environmental persistence of these two members of the genus *Rhodococcus*.

Keywords: *Rhodococcus*; *Rhodococcus aetherivorans* BCP1; *Rhodococcus opacus* R7; Phenotype microarray; Toxic compounds; Stress response

1. Introduction

Members of the genus *Rhodococcus*, which are non-sporulating aerobic bacteria with a high G+C content, degrade a variety of pollutants [1]. In this respect, the practical use of *Rhodococcus* spp. in bioremediation, biotransformation and biocatalysis is further enhanced by their persistence in the environment [2,3]. Indeed, *Rhodococcus* spp. are able to survive in the presence of high doses of toxic compounds, as well as under desiccation conditions, carbon starvation, a wide range of temperatures (from 4 °C to 45 °C), UV irradiation and osmotic stress (NaCl up to 7.5%) [2–6]. Survival mechanisms of *Rhodococcus* spp. in the presence of these environmental challenges rely on: i) the modification of the cell wall and the cell membrane fatty acid composition, ii) the accumulation of intracellular lipids as energy storage, iii) the synthesis of compatible solutes (ectoine, hydroxyectoine) for osmotic adjustment under saline stress, and iv) production of biomolecules (e.g. biosurfactants and pigments) supposedly acting as chelants [7]. Although physiological adaptation strategies regarding membrane composition and intracellular lipid accumulation were established, much less is known about

* Corresponding author.
E-mail addresses: martina.cappelletti2@unibo.it (M. Cappelletti), stefano.fedi@unibo.it (S. Fedi), j.zampolli@campus.unimib.it (J. Zampolli), a.dicanito@campus.unimib.it (A. Di Canito), pasqualina.dursi@itb.cnr.it (P. D'Ursi), alessandro.orro@itb.cnr.it (A. Orro), carlo.viti@unifi.it (C. Viti), luciano.milanesi@gmail.com (L. Milanesi), davide.zannoni@unibo.it (D. Zannoni), patrizia.digennaro@unimib.it (P. Di Gennaro).

http://dx.doi.org/10.1016/j.resmic.2016.06.008
0923-2508/

the role of specific genetic traits in the *Rhodococcus* spp. stress response. Among the toxic compounds, numerous studies focused on the resistance of *Rhodococcus* spp. to diverse contaminants and xenobiotics, whereas their resistance to heavy metals and antibiotics remain poorly studied. In this respect, conjugative plasmids were shown to harbor genetic traits involved in resistance to cadmium, thallium, arsenate and arsenite in *Rhodococcus erythropolis* and *Rhodococcus fascians* [8,9]. The unique tolerance of *Rhodococcus* spp. to heavy metals as compared to other actinobacteria was also related to non-specific resistance mechanisms, including metal binding to biomolecules and cellular non-diffusing pigments [10]. Further, antibiotic resistance was almost exclusively evaluated in *Rhodococcus equi*, a pathogen for animals and immunosuppressed humans [11].

A Phenotype Microarray (PM) approach has recently been used to assess the metabolic activity of two *Rhodococcus* strains, *Rhodococcus aetherivorans* BCP1 (also known as *Rhodococcus* sp. BCP1) and *Rhodococcus opacus* R7, under different conditions of growth in plates PM1-PM20 and in plates supplied with xenobiotics [12]. The PM high-throughput approach highlighted the wide metabolic abilities of BCP1 and R7 and pointed out the peculiar feature of the two strains for resisting hundreds of different stress conditions, although their genome analyses revealed a significant difference in genome size [10,11]. The R7 genome proved to be one of the largest bacterial genomes sequenced to date, with a total size of 10.1 Mb (1 chromosome and 5 plasmids), whereas, the BCP1 genome was 6.1 Mb in size (1 chromosome and 2 plasmids). The present study sought to evaluate whether the difference in BCP1 and R7 genome size reflects significant differences in their stress tolerance profiles. For this purpose, this work compared the metabolic activities of BCP1 and R7 measured in the presence of various stressors, such as high osmolarity and pH stress, diverse toxic compounds and antibiotics, in plates PM9-PM20. Further, the genetic determinants involved in resistance mechanisms were analyzed in order to support metabolic similarities/differences and provide genetic insight into the peculiar environmental persistence shown by *R. aetherivorans* BCP1 and *R. opacus* R7.

2. Materials and methods

2.1. PM assays

R. aetherivorans BCP1 (DSM44980) (also termed *Rhodococcus* sp. BCP1) and *R. opacus* R7 (CIP107348) were tested on inhibitor sensitivity arrays (PM9 to PM20) under more than 1100 different conditions (reagents used are listed at http://www.biolog.com). In particular, for each strain, part of the biomass grown at 30 °C on BUG agar was suspended in 15 mL of salt solution. Cell density was adjusted to 85% transmittance before being diluted in IF-10 medium supplied with 1% of dye G (tetrazolium dye) for inoculation in PM plates [12,13]. All PM plates were incubated at 30 °C in an Omnilog reader (Biolog). Readings were recorded for 72 h and data were analyzed with Omilog-PM software (Biolog).

2.2. Statistical analysis of PM data

PM kinetic curves were analyzed as previously described [12]. IC_{50} values were determined for the chemicals tested from PM11 to PM20 by the Omnilog-PM software [13]. The comparison between metabolic activities of the two strains was visualized by plotting all their activity values (from PM9 to PM20) in a 2D graph as previously described [13].

2.3. Identification of genetic aspects

The whole-genome shotgun sequencing projects are deposited at DDBJ/EMBL/GenBank under accession numbers CM002177, CM002178, CM002179 for *R. aetherivorans* BCP1 and CP008947, CP008948, CP008949, CP008950, CP008951, CP008952 for *R. opacus* R7. By RAST (Rapid Annotations using Subsystems Technology) server and BLAST analyses, we identified coding sequences (CDSs) associated with features predicted to be involved in the bacterial stress response (osmotic and pH stress) and in resistance to antibiotics, metals and other toxic compounds. Kyoto Encyclopedia of Genes and Genomes database (KEGG) (http://www.genome.jp/kegg/) and Transporter Classification Database (TCDB) (http://www.tcdb.org/) were used to assign Enzyme Commission numbers (EC) and TCDB numbers to the coding sequences (CDSs).

3. Results and discussion

3.1. Tolerance to osmotic stress

R. aetherivorans BCP1 and *R. opacus* R7 are metabolically active under a wide range of osmotic stress conditions [12]. The differences between the two strains are represented in the scatter plot in Fig. 1, where comparison between metabolic activities in the presence of all stressors tested from PM9 to PM20 is visualized. R7 cells showed higher osmotic resistance compared to BCP1 in terms of both number of osmolytes and osmolyte concentrations (circles numbered 1–17 in Fig. 1) with the exception of urea (dots numbered from 18 to 20 in Fig. 1, Table S1). In particular, unlike BCP1, the activity of R7 was maintained at a high level up to NaCl 8% without the supply of any osmoprotectant; further, R7 could grow up to the highest dose of sodium formate, sodium lactate and sodium nitrite. Conversely, the BCP1 strain was more resistant to increasing concentrations of urea (up to 6%), while the highest dose of urea (7%) was toxic for both strains.

Coding sequences (CDSs) involved in synthesis of compatible solutes such as operons for the synthesis of ectoine (*ectABC*), trehalose (*otsAB*), and glycogen (*glg* genes) were found in both BCP1 and R7 genomes (Table S2). In addition to one *ectABC* gene cluster, two ectoine hydroxylase coding genes (*ectD*) are present in R7 in two separate chromosomal regions and associated with CDSs coding for transporters, while one *ectD* gene is present in BCP1 and it is associated with trehalose biosynthesis genes. Ectoine hydroxylase (EctD) catalyzes the conversion of ectoine into hydroxyectoine. The

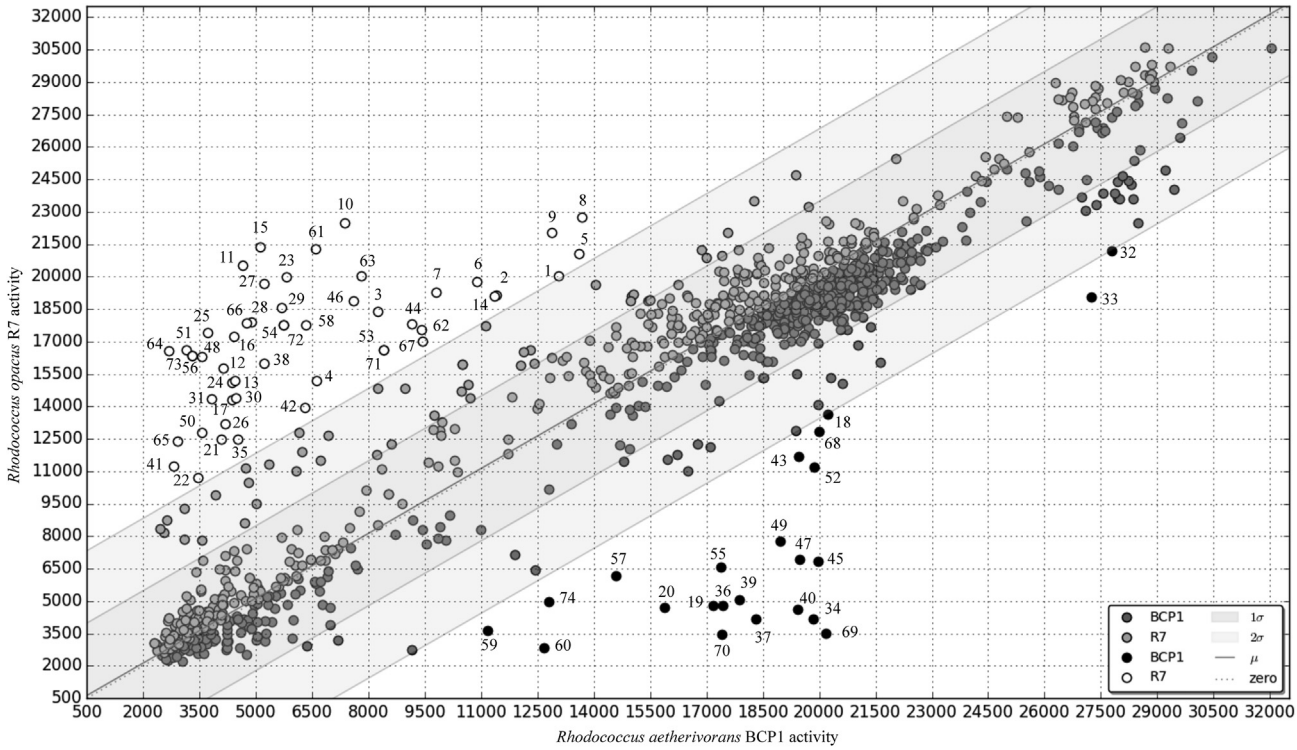


Fig. 1. Scatter plot comparing the metabolic activities of *R. opacus* R7 and *R. aetherivorans* BCP1 from PM plates 9–20. The metabolic activities of BCP1 and R7 are represented as dots on a 2D scatterplot. Plot bands represent mean (μ_A), standard deviation (σ_A) and double standard deviation ($2\sigma_A$) from the mean of the activity difference (ΔA). The dashed line (named as "zero" in the legend) correlates the points corresponding to compounds inducing the same metabolic response in the two strains. The points located outside of the double standard deviation band represent compounds inducing significantly different metabolic responses in BCP1 and R7. In particular, dots below the bands represent PM compounds inducing higher metabolic activity in BCP1 compared to R7 (full black dots in the legend). Circles above the bands indicate compounds inducing higher metabolic activity in R7 compared to BCP1 (white dots in the legend). Light gray and dark gray dots, reported in the legend, represent compounds not inducing significantly different metabolic responses in the two strains. Numbers in the plot represent the PM chemicals listed in Table S1.

independent localization of *ectD* in BCP1 and R7 with respect to *ectABC* might be related to different mechanisms generating hydroxyectoine instead of ectoine. In line with this, production of ectoine or hydroxyectoine by *R. opacus* PD630 under osmotic stress was reported to be affected by the carbon source [14].

Differences were observed between BCP1 and R7 in CDSs predicted to be involved in biosynthesis, transport and catabolism of glycine betaine (betaine), which is considered one of the main compatible solutes (Table S2). Betaine is synthesized through the activity of a choline dehydrogenase (BetA) and a betaine aldehyde dehydrogenase (BetB) [15]. Among the 9 CDSs encoding BetA in R7, 2 *betA* genes are co-localized with *betB* (see Panel A in Fig. 2). In both cases, a CDS encoding a urea-carboxylase-related amino acid permease (UctT) is placed between the two genes. One *betA-uctT-betB* gene cluster is included in a chromosomal region putatively involved in choline uptake and catabolism through the pathway choline → glycine betaine → dimethylglycine → sarcosine → glycine. Indeed, in addition to *betA* and *betB* genes, this region includes CDSs encoding: i) choline uptake protein BetT, ii) a large subunit of a phenylpropionate dioxygenase that shows similarity (35% aa identity) to the GbcA protein of *Pseudomonas aeruginosa* converting glycine betaine to dimethylglycine [16], iii) a dimethylglycine oxidase that oxidizes the dimethylglycine to sarcosine (DmgO), and iv) a heterotetrameric sarcosine oxidase (*soxBDAG* gene cluster) that catalyzes oxidative demethylation of sarcosine to glycine (Panel A in Fig. 2). Interestingly, this chromosomal region also includes CDSs predicted to encode enzymes involved in tetrahydrofolate-dependent assimilation of methyl groups related to the ability to utilize betaine as carbon and energy source in *Arthrobacter* spp. [17]. The lack of a spatial association between *betA, betB* and *betT* genes in BCP1 and the absence of a *soxBDAG* gene cluster in BCP1 support the difference between the two strains in terms of osmolyte tolerance. The salt tolerance of R7 can be further supported by the presence in R7 genome of CDSs encoding both a proline/glycine betaine transporter ProP and a multi-component proline permease ProU (composed of the three components, ProV, ProX and Prox), the latter absent in BCP1.

Both R7 and BCP1 possess CDSs coding for components of Trk and Kdp K^+ transport systems that allow efflux of K^+ during the osmo-adaptation process [18]. In both strains, a *kdpABCDE* operon codes for the three components of a membrane-associated P-type ATPase (KdpA-C), the osmosensitive K^+ channel histidine kinase KdpD and the response regulator KdpE. In both BCP1 and R7, the Trk system is

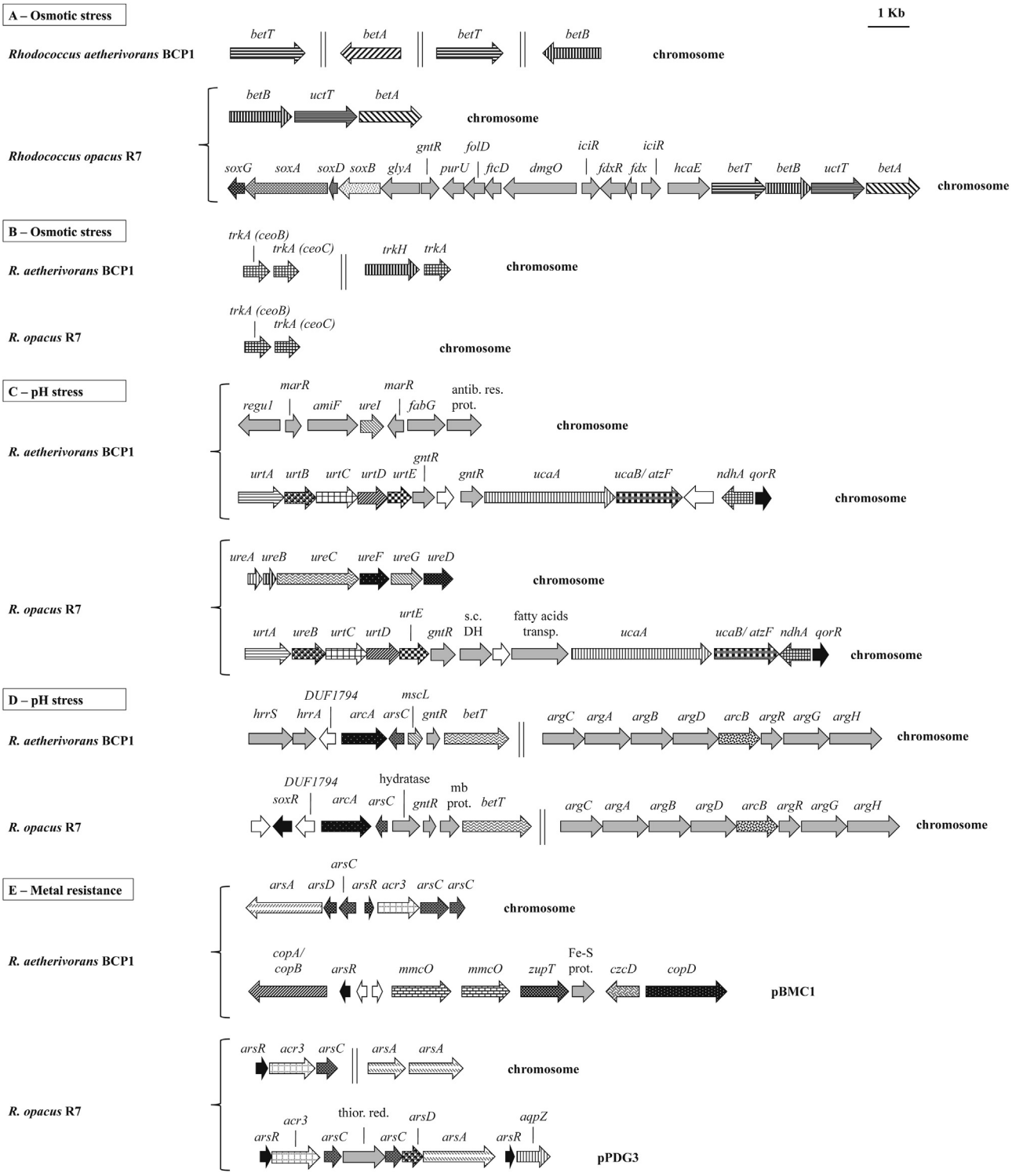


Fig. 2. Representative gene clusters involved in the *R. opacus* R7 and *R. aetherivorans* BCP1 stress responses whose presence or organization differentiates the two strains (except for *arcB* and the flanking genes). Genes are visualized as arrows. The predicted protein products of the genes are listed in Tables S2–S5, except for the following genes (represented in light gray): *hcaE*, phenylpropionate dioxygenase, large subunit; *iciR*, IclR family transcriptional regulator; *fdx*, ferredoxin; *fdxR*, ferredoxin reductase; *dmgO*, dimethylglycine oxidase; *ftcD*, formiminotetrahydrofolate cyclodeaminase; *folD*, methylenetetrahydrofolate dehydrogenase; *purU*, formyltetrahydrofolate deformylase; *gntR*, GntR family transcriptional regulator; *glyA*, serine hydroxymethyltransferase; *regu1*, putative nitrile hydratase regulator clustered with urea transport; *marR*, MarR family transcriptional regulator; *amiF*, formamidase; *fabG*, 3-oxoacyl-(acyl-carrier-protein) reductase; antib. res. prot., probable antibiotic resistance protein; *hrrA*, hemoglobin-dependent two component system response regulator HrrA; *hrrS*, sensory histidine kinase HrrS; DUF1794, protein including a DUF1794 domain with unknown function; *argC*, N-acetyl-gamma-glutamyl-phosphate reductase; *argA*, glutamate N-acetyl-transferase; *argB*, acetylglutamate kinase; *argD*, acetylornithine aminotransferase; *argR*, arginine pathway regulatory protein; *argG*, argininosuccinate synthase; *argH*, argininosuccinate lyase; s.c. DH, short chain dehydrogenase; mb prot., membrane protein; Fe-S prot., iron-sulfur protein; thior. red., thioredoxin reductase. Genes encoding hypothetical proteins are represented in white. Vertical lines are reported between gene clusters not close to each other.

composed of two TrkA proteins highly similar to CeoB and CeoC (70–75% aa identity) of *Mycobacterium tuberculosis* [18]. In BCP1, the association of a *trkH* gene with an additional *trkA* gene suggests a functional association between integral membrane protein TrkH and regulatory protein TrkA (Panel B in Fig. 2). Trk systems are considered secondary transporters, as the contribution of TrkH to K^+ transport is limited compared to other K^+ transport systems, e.g. Kpd [18].

3.2. Tolerance to acid and alkaline stress

R. aetherivorans BCP1 and *R. opacus* R7 showed metabolic activities at pH values ranging from 5 to 10, suggesting the presence of mechanisms to counteract both acidic and alkaline environments. Fig. 1 shows that, unlike BCP1, R7 was metabolically active at pH 4.5 in the presence of several amino acids such as L-methionine, L-phenilalanine, L-valine, L-homoserine, L-norvaline and α-amino-N-butyric acid (circles numbered from 21 to 31 in Fig. 1).

CDSs coding for transporters are involved in mechanisms of resistance to pH stress, i.e. multisubunit F_1F_0-ATPase and other monovalent cation/proton antiporters (Table S3). Other gene products supporting bacterial survival at acidic environment are the amino acid decarboxylase systems that control cellular pH by consuming H^+ [19]. CDSs encoding arginine, lysine and glutamate decarboxylases were found in each genome, as well as one CDS encoding a glutamate/γ-aminobutyrate antiporter (composing the GAD system with glutamate decarboxylase).

Differences between BCP1 and R7 were observed in genes linked to urea degradation (Table S3). Two distinct enzymes, urease and urea amidolyase, are known to degrade urea into carbonic acid and ammonia that raises the pH and counteracts acid stress [20]. Only R7 genome possesses a complete urease gene cassette encoding three structural proteins (UreA, UreB and UreC), along with CDSs coding for three accessory proteins (UreF, UreG, UreD) (Panel C in Fig. 2). This might be related to the capacity of R7 to resist to pH 4.5, unlike BCP1. The property of BCP1 to tolerate increasing concentrations of urea is unclear, as BCP1 does not present a urease gene cassette. Conversely, it shows both urea carboxylase (UcaA) and an allophanate hydrolase (UcaB) (composing the urea amidolyase activity, Panel C Fig. 2) along with a specific urea channel UreI that is missing in R7 and is not clustered with urea-decomposition-related genes (Panel C Fig. 2). The *ure*I gene codes for a protein that mediates uptake of urea and it co-localizes with a formamidase (AmiF), leading to production of ammonia from amides [21].

Ammonia is also produced from fermentative degradation of L-arginine through the activity of three enzymes belonging to the arginine deiminase (ADI) pathway: arginine deiminase (ADI or ArcA), ornithine carbamoyltransferase (OTC or ArcB), and carbamate kinase (CK or ArcC). Despite the fact that the ADI pathway is apparently confined to anaerobic or facultative anaerobic bacteria organisms, in BCP1 and R7 as well as in other aerobic *Actinomycetales*, CDSs encoding arginine deiminase and ornithine carbamoyltransferase are present, which are not clustered into the same chromosomal region (Panel D in Fig. 2, Table S3). This is in contrast with the cluster organization typically found on the bacterial chromosome [21]. A previous work suggested that the lack of spatial association between *arcA* and *arcB* could be related to the arginase activity of OTC revealed in *Streptomyces* and *Nocardia* strains [22]. The presence of a bifunctional OTC in *Actinomycetales* involved in both biosynthesis and catabolism of arginine was therefore suggested [21]. Interestingly, in both BCP1 and R7, the *arcB* gene is flanked by genes involved in arginine biosynthesis (*arg* genes) within a chromosomal region well conserved among Actinomycetales. Conversely, the CDSs flanking the *arcA* gene are only partially conserved. In both BCP1 and R7 as well as in other *Rhodococcus* spp., the arginine deiminase coding gene (*arcA*) is clustered with CDSs encoding a protein containing the DUF1794 domain with unknown function, a choline uptake protein BetT, an arsenate reductase ArsC and a GntR transcriptional regulator. In BCP1, *arcA* is also associated with a CDS coding for an MscL mechano-sensitive channel and a two-component regulatory system. In R7, *arcA* co-localizes with a CDS encoding a SoxR transcriptional regulator (Panel D Fig. 2).

3.3. Resistance to antibiotics, metals and other toxic compounds

In PM analysis, *R. aetherivorans* BCP1 and *R. opacus* R7 showed the ability to counteract a wide range of toxic compounds and antibiotics belonging to different classes. Interestingly, in spite of the significant diversity in genome size, a similar resistance/sensitivity pattern was seen in the two strains (around 4% of the conditions tested in PM11-PM20 induced different metabolic responses in the two strains; they are represented by circles or dots from 34 to 74 in Fig. 1). BCP1 and R7 showed resistance to antibiotics of the following classes: quinolones, fluoroquinolones, glycopeptides and sulfonamides. They were also resistant to all the tested doses of polymyxin B, phosphomycin, carbenicillin and nitrofurantoin. Conversely, BCP1 and R7 were sensitive to macrolides, rifampicin, rifamycin, vancomycin, puromycin, phenethicillin, oxacillin, fusidic acid, blasticidin S, chelerythrine and novobiocin. Several genetic determinants were found in both genomes that can contribute to antibiotic resistance mechanisms (Table S4). Fluoroquinolone resistance may be related to *gyrA* and *gyrB* genes, while resistance to β-lactams can be ascribed to CDSs encoding β-lactamases, penicillin-binding proteins and metal-dependent hydrolase of the β-lactamase superfamily III. CDSs encoding GCN5-related N-acetyltransferases may support aminoglycoside antibiotic resistance [23]. Further, CDSs coding for proteins involved in antibiotic resistance in the two strains included a vancomycin resistance protein, quaternary ammonium compound resistance protein SugE and several transport systems (e.g. multidrug transporters and efflux pumps) (Table S4). Some transcriptional regulators are predicted to activate antibiotic resistance mechanisms. In particular, one CDS encoding a WhiB family transcriptional regulator is included in a chromosomal region conserved in

BCP1, R7 and in *M. tuberculosis*. In *Mycobacterium* spp., this transcriptional regulator was reported to activate mechanisms of resistance to diverse antibiotic classes [24].

Table 1 reports the antimicrobials of various kinds (from PM11 to PM20) able to discriminate between the two strains in PM experiments on the basis of the IC_{50} value. Among these, the higher resistance of R7 to penicillin G could be related to the presence of a CDS encoding the penicillin G acylase precursor that is absent in BCP1. Higher metabolic activities were shown in BCP1 as compared to R7 in the presence of ethionamide, which targets the enoyl-acyl reductase InhA involved in mycolic acid biosynthesis [25]. In *M. tuberculosis*, the activation of ethionamide occurs via the reaction catalyzed by the monooxygenase EthA. BCP1 and R7 possess CDSs encoding InhA-like enoyl-acyl reductases and EthA-like flavin-binding monooxygenases (Table S4). The difference in ethionamide sensitivity between R7 and BCP1 might be due to different EthA regulatory mechanisms occurring in the two strains and/or to different protein sequences.

Both *R. aetherivorans* BCP1 and *R. opacus* R7 showed wide resistance to metalloids and some transition metals [12]. In particular, the strains showed tolerance to the tested concentrations of alkali metals (Li^+, Ce^+), metalloids (TeO_3^{2-}, SeO_3^{2-}, SiO_3^{2-}, AsO_2^-, AsO_4^{3-}) and some transition metals (Cd^{2+}, Co^{2+}, Cu^{2+}, Cr^{3+}, $Cr_2O_7^{2-}$, Fe^{3+}, Mn^{2+}, WO_4^{2+}, Zn^+). The metal ions toxic for both R7 and BCP1 were Ni^+, VO_4^{3-} (*o*-vanadate), and VO_3^- (*m*-vanadate). A significant difference between the two strains was observed in the resistance to thallium acetate (Tl^+) (Table 1, dots from 69 to 70 in Fig. 1). Tl^+ resistance was previously linked to the presence of the native plasmid pHG204 (190 kbp) in *Rhodococcus* sp. strain MR22 [26].

CDSs predicted to be involved in mechanisms of metal resistance are copper chaperones, copper-resistance proteins, multicopper oxidases, heavy-metal-translocating P-type ATPases, Mg and Co efflux proteins, Co/Zn/Cd resistance protein CzcD and several transcriptional regulators. Tellurium (Te) resistance genes (*terA*, *terD*) were found in both genomes, along with arsenic-related genes involved in bacterial arsenic resistance and transformation. Multiple copies of *arsR*, *arsB* and *arsC* are present in both genomes. In BCP1, one gene cluster *arsCDA* has a chromosomal localization and it is associated with *arsR*, *acr3* and two additional *arsC* genes, while other clustered CDSs coding for uncharacterized metal-resistance proteins are localized on plasmid pBMC1 (Panel E in Fig. 2). In R7, one gene cluster *arsCDAR* has a plasmid position (pPDG3) and it is associated with two additional *arsC*, one *acr3*, one *arsR* and a CDS encoding an aquaporin Z (Panel E in Fig. 2). In particular, the *arsC* and *acr3* genes have

Table 1
List and IC_{50} values of the PM11-20 inhibitors/toxicants giving different metabolic response in *R. aetherivorans* BCP1 and *R. opacus* R7.[a,b]

Substrate/Chemical	Panels (wells)	Mode of action	IC_{50}	
			BCP1	R7
Antibiotics				
Paromomycin	PM12 (C1–C4)	Protein synthesis, 30S ribosomal subunit, aminoglycoside	1.70	0.64
Apramycin	PM20 (A5–A8)	Protein synthesis, aminoglycoside	>4.40	3.63
Amikacin	PM11 (A1–A4)	Wall, lactam	>4.40	3.56
Ethionamide	PM17 (B9–B12)	Inhibits mycolic acid synthesis	>4.40	4.12
Hygromycin B	PM17 (B5–B8)	Protein synthesis, aminoglycoside	>4.40	3.59
Neomycin	PM11 (F9–F12)	Protein synthesis, 30S ribosomal subunit, aminoglycoside	3.51	2.59
Carbenicillin	PM14 (G5–G8)	Wall, lactam	3.41	4.40
Penicillin G	PM12 (A1–A4)	Wall, lactam	1.76	4.24
Gentamicin	PM11 (G5–G8)	Protein synthesis, 30S ribosomal subunit, aminoglycoside	3.63	4.36
Membrane permeability and transport				
Hexachlorophene	PM20 (E9–E12)	Membrane, electron transport	3.14	1.36
2-Polyphenol	PM18 (H5–H8)	Membrane	4.13	4.38
Niaproof	PM17 (E1–E4)	Membrane, detergent, anionic	1.30	0.61
Toxic cations				
Thallium(I) acetate	PM13 (F9–D12)	Toxic cation	2.49	0.60
Respiration/Ionophores/Uncouplers				
Pentachlorophenol	PM18 (C9–C12)	Respiration, ionophore, H+	1.58	2.53
Oxidizing agents				
Plumbagin	PM18 (H9–H12)	Oxidizing agent	4.10	4.15
Chelator, lipophilic				
Fusaric acid	PM14 (B5–B8)	Chelator, lipophilic	>4.40	3.63
1-Hydroxy-Pyridine-2-thione	PM14 (C5–C8)	Chelator, lipophilic	3.55	>4.40
5,7-Dichloro-8-hydroxyquinoline	PM15 (C1–C4)	Chelator, lipophilic	3.46	4.36
5,7-Dichloro-8-hydroxyquinaldine	PM15 (B9–B12)	Chelator, lipophilic	2.52	3.67

[a] IC_{50} values are expressed in well units and are defined as the well or fraction of a well at which a particular per-well parameter (*i.e.* the area of the curve) is at half of its maximal value over a concentration series. For each compound, IC_{50} values range between a minimum of 0.60 (no metabolic activity in any of the wells) and a maximum of 4.40 (optimal growth in all the wells).

[b] IC_{50} is reported only for compounds with which the difference between the areas of the kinetic curves of BCP1 and R7 strains was over 10,000 units in at least one of the concentrations tested, or over 7000 units in at least two of the concentrations tested (only for hexachlorophene).

been suggested to play a key role in resistance to arsenic [27]. In both BCP1 and R7, some scattered and clustered genes related to heavy metal resistance are found on plasmids (some of them are shown in Fig. 2), representing part of the genetic pool involved in horizontal gene transfer (HGT) events. Chromosomal duplication and HGT mechanisms of heavy-metal resistance genes have major importance in fitness of bacteria colonizing polluted habitats [28]. In particular, analysis of the distribution and organization of *ars*-like genes in microorganisms isolated from arsenic-polluted soils indicated that HGT played an important role in spreading arsenate and arsenite resistance ability under selective pressure [28,29]. In line with this, native plasmids of several *Rhodococcus* strains were reported to mediate heavy metal resistance such as arsenate, arsenite, cadmium, and thallium [8,9].

Lastly, CDSs involved in oxidative stress response and other detoxification processes can support the ability of the two strains to resist oxidizing agents, DNA synthesis inhibitors, folate antagonists, metal chelators and other toxicants (Table S5). Superoxide dismutase and catalase are central antioxidant enzymatic scavengers. In particular, three and six genes were annotated as catalase KatE in BCP1 and R7, respectively, and one CDS in each strain was predicted to encode the heme-containing enzyme catalase-peroxidase KatG. CDSs coding for superoxide dismutases belonging to Cu/Zn and Mn families were found in both strains. Both genomes possess CDSs encoding transcriptional regulators involved in the stress response (SoxR, Fur, Zur and Rex), as well as CDSs encoding glutathione peroxidases, ferroxidases, and proteins involved in biosynthesis of mycothiol. This low-molecular-weight thiol is typically produced by Actinobacteria for protection against the hazards of aerobic metabolism [30]. The CDSs encoding NADPH:quinone reductases in BCP1 and R7 show high similarity (>70% aa identity) with the NADPH:quinone reductase overexpressed in *Rhodococcus jostii* RHA1 during carbon starvation [5]. Interestingly, in R7, BCP1 and RHA1 genomes, the position of a CDS coding for a redox-sensing transcriptional regulator QorR is conserved upstream of the NADPH:quinone reductase coding gene, as well as the proximity of CDSs encoding allophanate hydrolase and urea carboxylase. Additional stress-response-related CDSs in BCP1 and R7 encode polyols transporters, sulfate and thiosulfate import proteins CysA, Nudix proteins, DedA, D-Tyrosyl-tRNA deacylase (Table S5) and carbon starvation protein A. Similarly to the high number of universal stress protein family genes (USPs) found in the RHA1 genome [5], R7 and BCP1 possess 14 and 11 USP genes, respectively. Only in BCP1, a CDS encoding a phytochrome-like two-component sensor histidine kinase was found. Interestingly, this CDS was flanked by a gene predicted to code for serine phosphatase RsbU that activates the σ^B factor in response to various stress conditions [31].

The limited knowledge of the genetic basis supporting oxidative stress response, as well as other detoxification processes in *Rhodococcus* spp., hampers the interpretation of metabolic differences between the two strains in the presence of some antimicrobials reported in Table 1, Table S1 and Fig. 1. However, in addition to genetic traits predicted to play a role in antimicrobial resistance mechanisms, a fundamental contribution to the wide resistance ability of the two strains can be ascribed to the peculiar features of the mycolic-acid-containing cell wall typical of *Rhodococcus* spp., as well as to phenomena of intracellular sequestration, adsorption of metals/antimicrobials to cell wall and binding to carotenoid pigments [10,32,33]. In line with this, some differences in resistance/sensitivity to antimicrobials shown by the two strains might also be related to differences in cell membrane lipids, mycolic acid composition and cell pigment production.

In summary, the results taken together show that *R. opacus* R7 has a higher resistance than *R. aetherivorans* BCP1 to acidic pHs and to a few osmolytes (tested in PM9-10); conversely, in spite of the considerable difference in genome sizes, the two strains have similar behaviors in counteracting the toxicity of metals, antibiotics, oxidative agents and several other toxic compounds (at the concentrations supplied in PM11-20). The latter properties might not be fully related to the presence of specific genetic determinants, although the present poor knowledge about the genetic basis of survival mechanisms in *Actinobacteria*, error-prone information given by the automatic genome annotation process and the actual role of non-genetic factors in *Rhodococcus* spp. do not allow a firm conclusion on this point. Indeed, the difference in genome size could be due to genetic redundancy, but might also be related to other metabolic features not tested in the present study. The conclusion is that, from the point of view of environmental remediation technologies and transformation of toxic compounds, the two strains are expected to have similar behavior in response to stressors that can occur in contaminated soils. Further studies will be necessary to reveal the role of cell membrane composition, pigment and biomolecule production in relation to the environmental persistence of *R. aetherivorans* BCP1 and *R. opacus* R7.

Acknowledgments

This work was supported by the "Vasco e GC Rossi – Microbial Biofilm" grant A.10.N4.RICER.797 and by the RFO-2012-14 grant, University of Bologna. Bioinformatics analysis was supported by the Italian Ministry of Education and Research through the Flagship "InterOmics" (PB05) and HIRMA (RBAP11YS7K) Projects. The authors thank F. Decorosi and G. Spini for help with phenotype microarray experiments.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.resmic.2016.06.008.

References

[1] Martínková L, Uhnáková B, Pátek M, Nešvera J, Křen V. Biodegradation potential of the genus *Rhodococcus*. Environ Int 2009;35:162–77.

[2] LeBlanc JC, Gonçalves ER, Mohn WW. Global response to desiccation stress in the soil actinomycete *Rhodococcus jostii* RHA1. Appl Environ Microbiol 2008;74:2627–36.
[3] De Carvalho CCR. Adaptation of *Rhodococcus erythropolis* cells for growth and bioremediation under extreme conditions. Res Microbiol 2012;163:125–36.
[4] Bequer Urbano S, Albarracín VH, Ordoñez OF, Farías ME, Alvarez HM. Lipid storage in high-altitude Andean Lakes extremophiles and its mobilization under stress conditions in *Rhodococcus* sp. A5, a UV-resistant actinobacterium. Extremophiles 2013;17:217–27.
[5] Patrauchan MA, Miyazawa D, LeBlanc JC, Aiga C, Florizone C, Dosanjh M, et al. Proteomic analysis of survival of *Rhodococcus jostii* RHA1 during carbon starvation. Appl Environ Microbiol 2012;78:6714–25.
[6] Alvarez HM, Silva RA, Cesari AC, Zamit AL, Peressutti SR, Reichelt R, et al. Physiological and morphological responses of the soil bacterium *Rhodococcus opacus* strain PD630 to water stress. FEMS Microbiol Ecol 2004;50:75–86.
[7] Alvarez HM, Steinbuchel A. Physiology, biochemistry and molecular biology of triacylglycerol accumulation by *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus*. Microbiology monographs series. Heidelberg: Springer Verlag; 2010. p. 263–90.
[8] Dabbs ER, Sole GJ. Plasmid-borne resistance to arsenate, arsen- ite, cadmium, and chloramphenicol in a *Rhodococcus* species. Mol Gen Genet 1988;211:148–54.
[9] Desomer J, Dhaese P, van Montagu M. Conjugative transfer of cadmium resistance plasmids in *Rhodococcus fascians* strains. J Bacteriol 1988;170:2401–5.
[10] Ivshina IB, Kuyukina MS, Kostina LV. Adaptive mechanisms of nonspecific resistance to heavy metal ions in alkanotrophic actinobacteria. Russ J Ecol 2013;44:123–30.
[11] Cisek AA, Rzewuska M, Witkowski L, Binek M. Antimicrobial resistance in *Rhodococcus equi*. Acta Biochim Pol 2014;61:633–8.
[12] Orro A, Cappelletti M, D'Ursi P, Milanesi L, Di Canito A, Zampolli J, et al. Genome and Phenotype Microarray analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: genetic determinants and metabolic abilities with environmental relevance. PLoS One 2015;10:e0139467.
[13] Viti C, Decorosi F, Tatti E, Giovannetti L. Characterization of chromate-resistant and -reducing bacteria by traditional means and by a high-throughput phenomic technique for bioremediation purposes. Biotechnol Prog 2007;23:553–9.
[14] Lamark T, Kaasen I, Eshoo MW, Falkenberg P, McDougall J, Strøm AR. DNA sequence and analysis of the *bet* genes encoding the osmoregulatory choline-glycine betaine pathway of *Escherichia coli*. Mol Microbiol 1991;5:1049–64.
[15] Wargo MJ, Szwergold BS, Hogan DA. Identification of two gene clusters and a transcriptional regulator required for *Pseudomonas aeruginosa* glycine betaine catabolism. J Bacteriol 2008;190:2690–9.
[16] Meskys R, Harris RJ, Casaite V, Basran J, Scrutton NS. Organization of the genes involved in dimethylglycine and sarcosine degradation in *Arthrobacter* spp.: implications for glycine betaine catabolism. Eur J Biochem 2001;268:3390–8.
[17] Cholo MC, van Rensburg EJ, Osman AG, Anderson R. Expression of the genes encoding the Trk and Kdp potassium transport systems of *Mycobacterium tuberculosis* during growth in vitro. BioMed Res Int 2015;2015. Article ID 608682.
[18] Cotter PD, Hill C. Surviving the acid test: responses of gram-positive bacteria to low pH. Microbiol Mol Biol Rev 2003;67:429–53.
[19] Kanamori T, Kanou N, Atomi H, Imanaka T. Enzymatic characterization of a prokaryotic urea carboxylase. J Bacteriol 2004;186:2532–9.
[20] Van Vliet AHM, Kuipers EJ, Stoof J, Poppelaars SW, Kusters JG. Acid-responsive gene induction of ammonia-producing enzymes in *Helicobacter pylori* is mediated via a metal-responsive repressor cascade. Infect Immun 2004;72:766–73.
[21] Zúñiga M, Pérez G, González-Candelas F. Evolution of arginine deiminase (ADI) pathway genes. Mol Phylogenet Evol 2002;25:429–44.
[22] De la Fuente JL, Martin JF, Liras P. New type of hexameric ornithine carbamoyltransferase with arginase activity in the cephamycin producers *Streptomyces clavuligerus* and *Nocardia lactamdurans*. Biochem J 1996;320:173–9.
[23] McGarvey KM, Queitsch K, Fields S. Wide variation in antibiotic resistance proteins identified by functional metagenomic screening of a soil DNA library. Appl Environ Microbiol 2012;78:1708–14.
[24] Burian J, Yim G, Hsing M, Axerio-Cilies P, Cherkasov A, Spiegelman GB. The mycobacterial antibiotic resistance determinant WhiB7 acts as a transcriptional activator by binding the primary sigma factor SigA (RpoV). Nucleic Acids Res 2013;41:10062–76.
[25] Morlock GP, Metchock B, Sikes D, Crawford JT, Cooksey RC. *ethA, inhA*, and *katG* loci of ethionamide-resistant clinical *Mycobacterium tuberculosis* isolates. Antimicrob Agents Chemother 2003;47:3799–805.
[26] Kalkus J, Dorrie C, Fischer D, Reh M, Schlegel HG. The giant linear plasmid pHG207 from *Rhodococcus* sp. encoding hydrogen autotrophy: characterization of the plasmid and its termini. J Gen Microbiol 1993;139:2055–65.
[27] Li X, Zhang L, Wang G. Genomic evidence reveals the extreme diversity and wide distribution of the arsenic-related genes in *Burkholderiales*. PLoS One 2014;9:e92236.
[28] Villegas-Torres MF, Bedoya-Reina OC. Horizontal *arsC* gene transfer among microorganisms isolated from arsenic polluted soil. Int Biodeterior Biodegrad 2011;65:147–52.
[29] Cai L, Liu G, Rensing C, Wang G. Genes involved in arsenic transformation and resistance associated with different levels of arsenic-contaminated soils. BMC Microbiol 2009;9:4.
[30] Garg A, Soni B, Singh Paliya B, Verma S, Jadaun V. Low-molecular-weight thiols: glutathione (GSH), mycothiol (MSH) potential antioxidant compound from actinobacteria. J App Pharm Sci 2013;3:117–20.
[31] Shin J-H, Brody MS, Price CW. Physical and antibiotic stresses require activation of the RsbU phosphatase to induce the general stress response in *Listeria monocytogenes*. Microbiology 2010;156:2660–9.
[32] Kuyukina MS, Ivshina IB, Rychkova MI, Chumakov OB. Effect of cell lipid composition on the formation of nonspecific antibiotic resistance in alkanotrophic rhodococci. Microbiology 2000;69:51–7.
[33] Schmidt A, Haferburg G, Sineriz M, Merten D, Büchel G, Kothe E. Heavy metal resistance mechanisms in actinobacteria for survival in AMD contaminated soils. Chem Erde Geochem 2005;65:131–44.

Di Canito et al. BMC Genomics (2018) 19:587
https://doi.org/10.1186/s12864-018-4965-6

BMC Genomics

RESEARCH ARTICLE Open Access

Genome-based analysis for the identification of genes involved in *o*-xylene degradation in *Rhodococcus opacus* R7

Alessandra Di Canito[1], Jessica Zampolli[1], Alessandro Orro[2], Pasqualina D'Ursi[2], Luciano Milanesi[2], Guido Sello[3], Alexander Steinbüchel[4,5] and Patrizia Di Gennaro[1*]

Abstract

Background: Bacteria belonging to the *Rhodococcus* genus play an important role in the degradation of many contaminants, including methylbenzenes. These bacteria, widely distributed in the environment, are known to be a powerhouse of numerous degradation functions, due to their ability to metabolize a wide range of organic molecules including aliphatic, aromatic, polycyclic aromatic compounds (PAHs), phenols, and nitriles. In accordance with their immense catabolic diversity, *Rhodococcus* spp. possess large and complex genomes, which contain a multiplicity of catabolic genes, a high genetic redundancy of biosynthetic pathways and a sophisticated regulatory network. The present study aimed to identify genes involved in the *o*-xylene degradation in *R. opacus* strain R7 through a genome-based approach.

Results: Using genome-based analysis we identified all the sequences in the R7 genome annotated as dioxygenases or monooxygenases/hydroxylases and clustered them into two different trees. The *akb*, *phe* and *prm* sequences were selected as genes encoding respectively for dioxygenases, phenol hydroxylases and monooxygenases and their putative involvement in *o*-xylene oxidation was evaluated. The involvement of the *akb* genes in *o*-xylene oxidation was demonstrated by RT-PCR/qPCR experiments after growth on *o*-xylene and by the selection of the R7–50 leaky mutant. Although the *akb* genes are specifically activated for *o*-xylene degradation, metabolic intermediates of the pathway suggested potential alternative oxidation steps, possibly through monooxygenation. This led us to further investigate the role of the *prm* and the *phe* genes. Results showed that these genes were transcribed in a constitutive manner, and that the activity of the Prm monooxygenase was able to transform *o*-xylene slowly in intermediates as 3,4-dimethylphenol and 2-methylbenzylalcohol. Moreover, the expression level of *phe* genes, homologous to the *phe* genes of *Rhodococcus* spp. 1CP and UPV-1 with a 90% identity, could explain their role in the further oxidation of *o*-xylene and R7 growth on dimethylphenols.

Conclusions: These results suggest that R7 strain is able to degrade *o*-xylene by the Akb dioxygenase system leading to the production of the corresponding dihydrodiol. Likewise, the redundancy of sequences encoding for several monooxygenases/phenol hydroxylases, supports the involvement of other oxygenases converging in the *o*-xylene degradation pathway in R7 strain.

Keywords: *Rhodococcus*, Microbial genomics, *o*-xylene degradation, *Gene clusters*, Contaminated soil

* Correspondence: patrizia.digennaro@unimib.it
[1]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milan, Italy
Full list of author information is available at the end of the article

Background

Methylbenzenes are pollutants of great relevance for their toxic properties and their wide spread in environment, commonly present in crude petroleum and in various industrial processes [1]. Different methylbenzenes, including the three xylene isomers, can be degraded by several bacterial strains, with a degradation pathway that depends on the position of the methyl groups on the aromatic ring [2]. These bacteria can be divided into two groups: i) microorganisms that can degrade both *m*- and *p*-xylene; and ii) microorganisms that can only degrade the *ortho* isomer. The two degradation pathways are rarely found simultaneously in the same microorganism.

Although bacteria capable of degrading *m*-xylene or *p*-xylene under aerobic conditions and their corresponding catabolic pathways have been well documented [3–5], bacteria capable of degrading *o*-xylene are not so common and their metabolic pathways have been poorly investigated [6–10]. Three *o*-xylene degradation pathways (leading to their corresponding catechols) have been reported [1, 8, 11–13] (Fig. 1). The first pathway is initiated by oxidation of the methyl group of *o*-xylene to form 2-methylbenzylalcohol, subsequently metabolized to form 3-methylcatechol as reported in the case of *Rhodococcus* sp. strain B3 [8] (Fig. 1, route A). The second degradation pathway of *o*-xylene is initiated by a ring-hydroxylating 2,3-dioxygenase leading to the 3,4-dimethylcatechol in *Rhodococcus* sp. strain DK17 (Fig. 1, route B) [10]. While the third pathway is initiated by a ring-hydroxylating monooxygenase at the different position 3, or 4, as performed by the Toluene *o*-xylene Monooxygenase (ToMo) in *P. stutzeri* strain OX1 (Fig. 1, routes C1 and C2), leading to the formation of 3,4-dimethylcatechol [1].

One of the main roles in the degradation of many contaminants, including methylbenzenes, is played by bacteria belonging to the *Rhodococcus* genus. These bacteria, widely distributed in the environment, are characterized as a powerhouse of numerous degradation functions, since they are able to metabolize a wide range of organic molecules including aliphatic, aromatic and polycyclic aromatic compounds (PAHs), phenols, and nitriles [14]. In accordance with their immense catabolic diversity, *Rhodococcus* spp. are characterized to possess large and complex genomes, which contain a multiplicity of catabolic genes, a high gene redundancy of biosynthetic pathways and a sophisticated regulatory network [15]. Many of them also possess a variety of large linear plasmids and smaller circular plasmids that contribute to and also explain the immense repertoire of catabolic abilities [16]. The most known example is represented by the genome of *Rhodococcus jostii* strain RHA1 [17], isolated for its ability to aerobically degrade polychlorinated biphenyls (PCBs) [18], and also able to utilize a wide range of compounds as sole carbon and energy source.

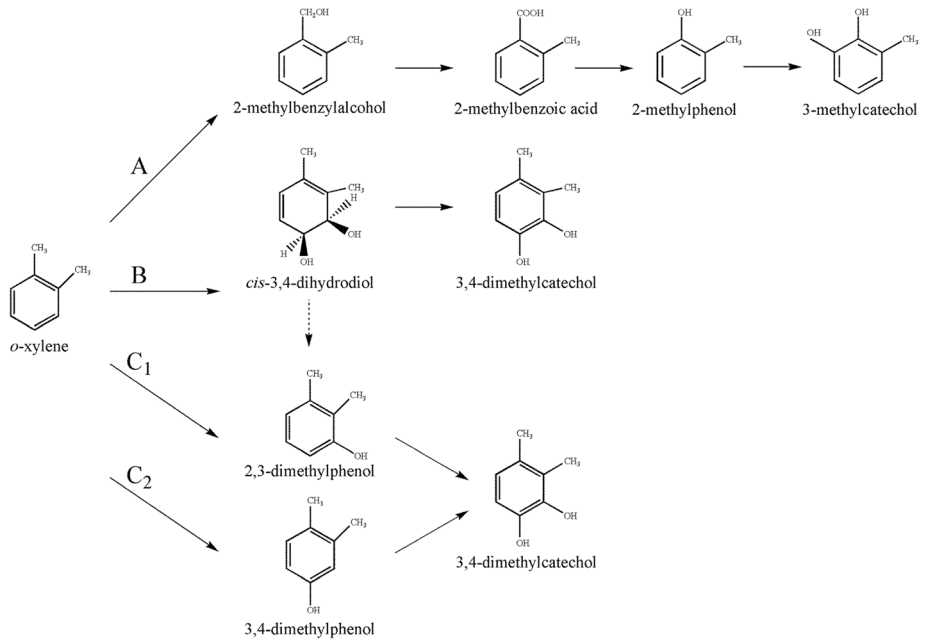


Fig. 1 Overview of different *o*-xylene degradation pathways in bacteria. Monooxygenation of the methyl group and subsequent oxidations in *Rhodococcus* sp. strain B3 (route A); dioxygenation and subsequent dehydrogenation of the aromatic ring in *Rhodococcus* sp. strain DK17 (route B); two successive monooxygenation reactions of the aromatic ring in *P. stutzeri* strain OX1 (route C1 and C2). Dashed arrow indicates a spontaneous dehydration

Analyses of the 9.7 Mb large genome of RHA1 provided the evidence of catabolic pathway redundancy and horizontal gene transfer events [17].

To date, several gene clusters involved in the degradation of multiple aromatic compounds have been identified from genome analysis of several *Rhodococcus* spp. strains, including genes for biphenyl [19], isopropylbenzene and ethylbenzene [20] and methylbenzenes [2]. However, few not in depth genetic studies have been reported regarding the abilities of *Rhodococcus* strains to degrade *o*-xylene. The only data regarding the genes involved in *o*-xylene degradation in bacteria belonging to *Rhodococcus* genus derives from the identification of the *akb* genes in *Rhodococcus* sp. strain DK17 [10]. This strain is able to grow on *o*-xylene (and toluene, ethylbenzene, isopropylbenzene) through a multicomponent *o*-xylene dioxygenase [10, 12]. The DK17 *o*-xylene dioxygenase is described to perform a ring-oxidizing pathway leading to the 3,4-dimethylcathechol formation either by a dioxygenation or two monooxygenations, which can introduce two oxygen atoms successively. Thus, a deeper analysis concerning *o*-xylene degradation in *Rhodococcus* is necessary to understand which genes and enzymes could be involved in this metabolism.

In this context, the metabolically versatile *Rhodococcus opacus* R7 [21], known for its ability to grow on naphthalene, several long- and medium-chain *n*-alkanes, and aromatic hydrocarbons belonging to BTEX group (benzene, toluene, ethylbenzene and xylenes) [22, 23], and also able to grow on *o*-xylene, can be used to add information of the metabolism of this compound. The whole genome of R7 strain was completely sequenced and it revealed to possess multiple genes for the degradation of a large set of aliphatic, aromatic and PAHs compounds [24]. Moreover, the genome analysis revealed the presence, beside the chromosome, of five plasmids (pPDG1, pPDG2, pPDG3, pPDG4, pPDG5) that provided the evidence of high catabolic pathway redundancy.

Through a genome-based analysis, the present work aimed to identify genes and molecular mechanisms involved in *o*-xylene degradation in *R. opacus* R7. Based on the previous identification of 2,3-dimethylphenol (2,3-DMP) and 3,4-dimethylphenol (3,4-DMP) in the R7 culture medium and the fact that these intermediates were metabolized by R7 strain, when supplied as the sole carbon and energy source, we identified these compounds as intermediates of R7 *o*-xylene degradation pathway [21, 22]. However, literature data suggested that the formation of dimethylphenol could be attributed to the dehydration of dihydrodiol deriving from the dioxygenation activity when *o*-xylene is supplied to *Rhodococcus* sp. strain DK17 [2]. For this reason, we searched and identified in the R7 genome all the sequences annotated as dioxygenases or monooxygenases/hydroxylases and clustered them into two different trees in order to select the oxygenases putatively involved in the *o*-xylene oxidation. Moreover, we demonstrated that the selected genes were involved in *o*-xylene degradation in *R. opacus* R7 by RT-PCR/qPCR, mutant analysis, cloning and expression experiments, thus revealing the complexity of this metabolic network.

Results

R7 genome sequence analysis

The formation of 2,3-dimethylphenol and 3,4-dimethylphenol (and the oxidation to 2-methylbenzylalcohol) in the R7 *o*-xylene degradation pathway [21, 22] suggests the involvement of monooxygenases able to oxidize *o*-xylene. However, the only *o*-xylene degradation pathway described in literature for bacteria of *Rhodococcus* genus is through the dioxygenase system of *Rhodococcus* sp. DK17 leading to the corresponding dihydrodiol that could be dehydrated to DMPs [2]. Therefore, we hypothesized that in the case of *R. opacus* strain R7 the formation of the identified intermediates could be explained by the involvement of different oxygenase systems for *o*-xylene degradation. For this reason, first we identified all the sequences annotated as dioxygenases or monooxygenases/hydroxylases in the R7 genome and clustered them into two different trees in order to select all the oxygenases putatively involved in *o*-xylene oxidation.

Analysis and clusterization of dioxygenases

A preliminary genome RAST annotation of R7 allowed the identification of 83 potential dioxygenases. Among those, only 57 were selected as catalytic subunits of R7 dioxygenases. In order to cluster the R7 catalytic subunits, 22 reference sequences putatively involved in the aromatic compound degradation were considered. The generated tree reveals that these amino acid sequences are divided into eight clades (Fig. 2). All the sequences are listed in Additional file 1: Table S1, including the reference sequences with the relative source strain and all the sequences belonging to R7 strain.

Clade number 3 includes dioxygenases putatively involved in the upper pathways of BTEX compounds and polycyclic aromatic hydrocarbons (Fig. 2, blue box), while the other seven clades include all the dioxygenases putatively involved in the peripheral pathways of different aromatic compounds (Fig. 2).

Among the sequences in clade number 3 of the dioxygenase-tree (Fig. 2, extended clade), the catalytic subunit of ethylbenzene dioxygenase (EtbAa1) of *R. jostii* RHA1 [25] and the *o*-xylene dioxygenase (AkbA1a) of *Rhodococcus* sp. DK17 [10], are shown to cluster near the only homologous dioxygenase sequence (AkbA1a) of R7. Multiple alignments of AkbA1a of R7 with proteins belonging to Bacterial Rieske non-heme iron oxygenases

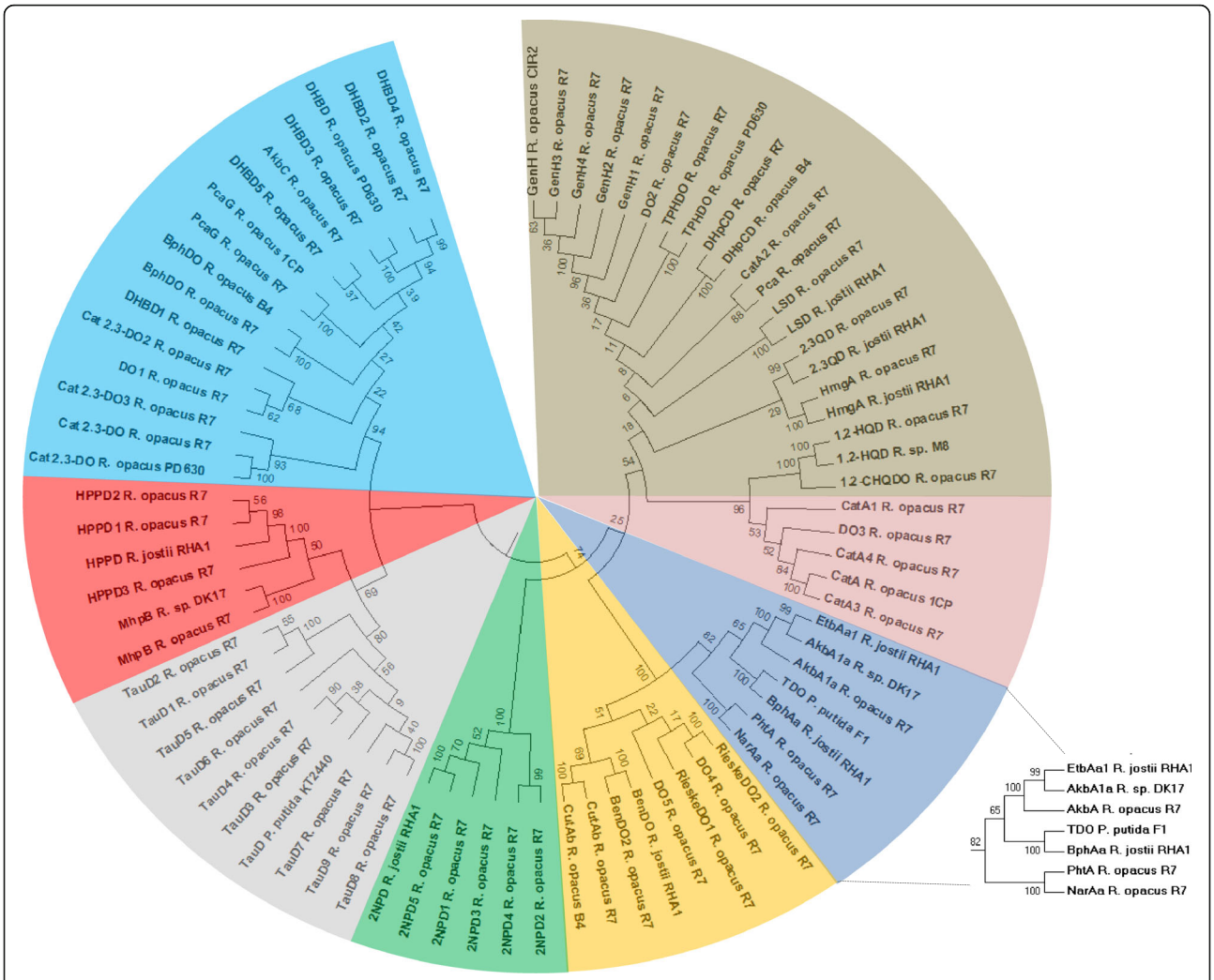


Fig. 2 Phylogenetic tree of *R. opacus* R7 proteins containing the catalytic subunit of dioxygenases. The tree was constructed aligning selected reference protein sequences with R7 sequences. Numbers represent the bootstrap values on the branches calculated for maximum likelihood (ML) method selected from the package MEGA version 6 with 100 bootstraps. Color scheme for dioxygenases: brown, clade 1; pink, clade 2; blue, clade 3; yellow, clade 4; green, clade 5; grey, clade 6; red, clade 7; light blue, clade 8. The extended clade includes the AkbA1a of *R. opacus* R7 homologous to the AkbA1a of *R.* sp. DK17 involved in the oxidation of *o*-xylene. Abbreviation of dioxygenase names are reported in Additional file 1: Table S1

reveals the coordination of the center iron-sulfur (Fe-S) (CxH - CxxH) with the amino acids that coordinate the iron atom of the active site (H - H - D). Moreover, the *akbA1a* gene encoding for the AkbA1a dioxygenase shows a nucleotide identity around 90% with the *etbAa1* of *R. jostii* RHA1 and *akbA1a* of *Rhodococcus* sp. DK17. For these features, the AkbA1a was taken into consideration for its involvement in *o*-xylene catabolism.

The remaining clades of the tree analysis include several dioxygenases involved in the mechanism of aromatic ring cleavage (e. g. gentisate 1,2-dioxygenases, catechol 1,2-dioxygenases, catechol 2,3-dioxygenases, homogentisate dioxygenase and protocatechuate dioxygenase), that are not taken into consideration.

Analysis and clusterization of monooxygenases/hydroxylases

From all the sequences derived from the whole genome of R7, the attention was also focused on putative sequences annotated as monooxygenases/hydroxylases. Then, a multiple amino acidic sequence alignment and a clusterization analysis were performed to predict protein functions using characterized monooxygenases/hydroxylases from different bacteria as reference (Fig. 3). All the sequences are listed in Additional file 2: Table S2, including the reference sequences with the relative source strain and all the sequences belonging to the R7 strain. The sequences obtained by this analysis are clustered into 10 clades. Clade number 1 includes the phenol hydroxylases P164, P165, P166, P167, P149 and P150 used as reference

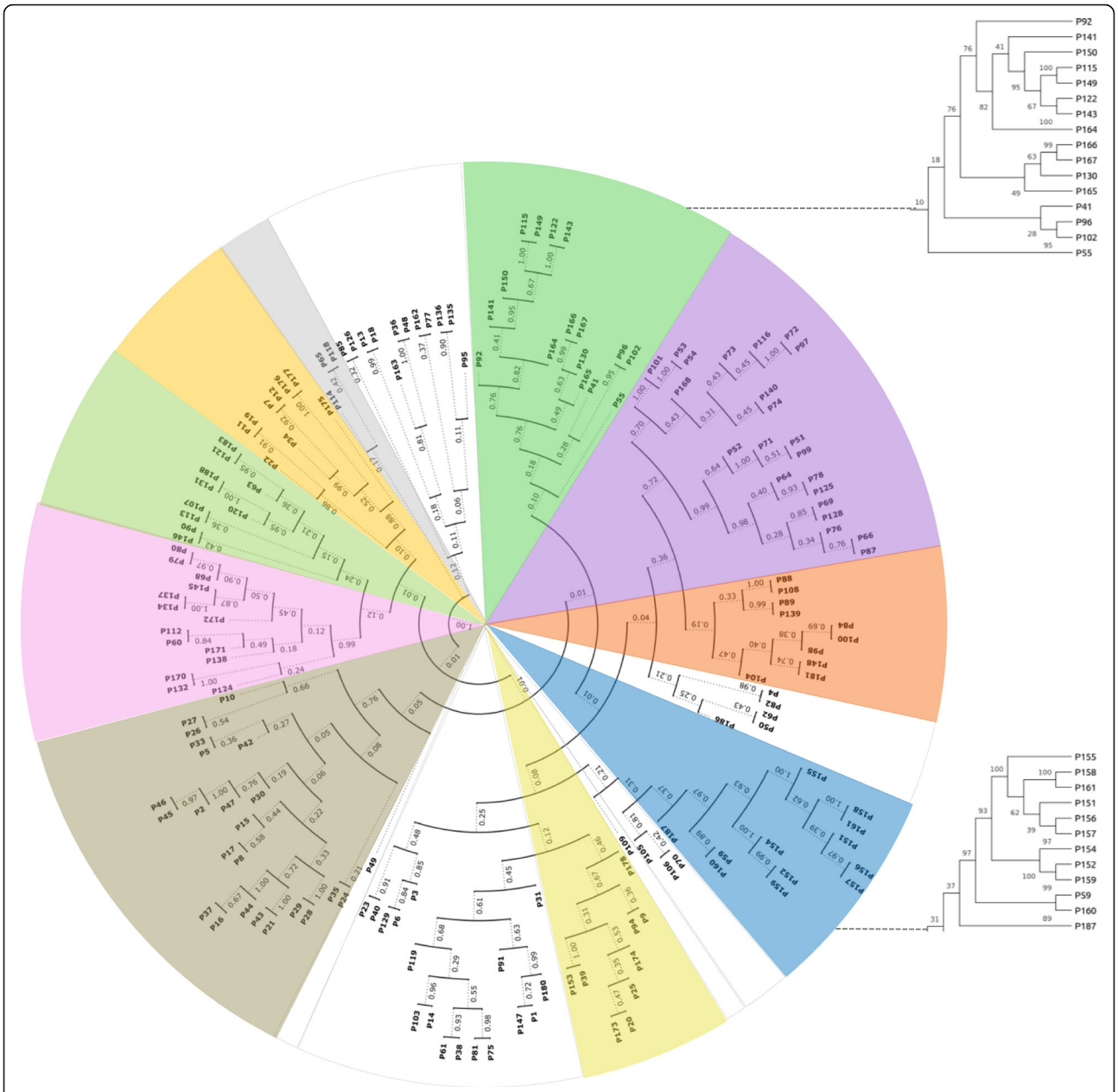


Fig. 3 Phylogenetic tree of *R. opacus* R7 proteins containing the catalytic subunit of hydroxylases and monooxygenases. The tree was constructed aligning selected reference protein sequences with R7 sequences. Numbers represent the bootstrap values on the branches calculated for maximum likelihood (ML) method selected from the package MEGA version 6 with 100 bootstraps. Distinct clades are labeled with different colors: green, clade 1; purple, clade 2; red, clade 3; blue, clade 4; yellow, clade 5; brown, clade 6; pink, clade 7; light green, clade 8; orange, clade 9; grey, clade 10. White clades surrounded by grey lines correspond to not classified proteins: clade of not classified protein 1–6. The extended clade 1 includes the R7 proteins PheA1a, PheA2a, PheA3a; the extended clade 4 includes the R7 PrmA protein. The tree image was built with the ETE Toolkit using the circular plot function, with equal branch lengths and labeling each node with bootstrap support. Abbreviation of hydroxylase and monooxygenase names are reported in Additional file 2: Table S2

sequences. Amongst these sequences, solely the P149 of *R. opacus* 1CP [26] was found to be similar to P115 (PheA1a) (98%), P122 (PheA2a) (92%), and P143 (PheA3a) (92%) of *R. opacus* R7. Accordingly, from this clade we selected the P115, P122 and P143 sequences (encoded by the *pheA1a* gene, *pheA2a* gene, and *pheA3a* gene, respectively) for further molecular analysis.

In clade number 4, only the P59 (PrmA) (encoded by the *prmA* gene) of R7, annotated as *alpha* chain methane monooxygenase component, is shown to cluster with all

the phenol hydroxylases/monooxygenases used as reference sequences. Among the reference sequences, the TouA component (P151) of Toluene *o*-xylene Monooxygenase (ToMo) and the PhN component (P152) of Phenol Hydroxylase (PH) from *P. stutzeri* OX1, were mainly considered because they are the most described in literature for *o*-xylene oxidation [27–30]. In fact, comparing the amino acid sequences of the TouA component (P151) of ToMo, the PhN component (P152) of PH, and of PrmA (P59), the most residues of the catalytic site of the three proteins were found to be conserved. Thus, P59 was selected to investigate its involvement in *o*-xylene degradation.

R7 proteins of clades number 2, 3 and 5, 6 including cyclohexanone monooxygenases [31], FAD-dependent monooxygenases [32], salicylate hydroxylases [33], and putative P450 hydroxylases respectively [34], were excluded from further molecular analysis in this paper, as little is known regarding their putative involvement in *o*-xylene oxidation.

Regarding the other clades number 7, 8, 9, 10 including, respectively, putative nitrilotriacetate monooxygenases [35], alkanesulfonate monooxygenases [36], other putative monooxygenases/hydroxylases, and ubiquinone monooxygenases [37], we excluded their involvement in the hydroxylation/monooxygenation of *o*-xylene, because their function seems to be very distant from our search.

Moreover, the proteins of the non-classified clades 1–6 were excluded because they were lacking in reference sequences.

Involvement of the *akb* genes in *o*-xylene degradation by RT-PCR experiments

Analyses of the R7 genome sequences evidenced the presence of the *akbA1a* gene in the *akb* gene cluster allocated on the megaplasmid pPDG5 (Table 1). This gene, coding for a large subunit dioxygenase component (AkbA1a), clustered with the following: the *akbA2a* gene coding for a small subunit dioxygenase component (AkbA2a), the *akbA3* gene for a ferredoxin component (AkbA3), the *HP* sequence for an hypothetical protein (HP) of unknown function, the *akbA4* gene for a reductase component (AkbA4), and the *akbB* gene coding for a dihydrodiol dehydrogenase (AkbB) (Fig. 4, panel a). Downstream (in the opposite direction) of these sequences, we found two sequences homologous (near the 80%) to the *akbS* and *akbT* sequences encoding for the sensor and regulator elements of DK17 strain, potentially involved in the regulatory mechanism.

Moreover, we found a second group of genes (*akbCDEF* genes) coding for complete *meta*-cleavage enzymes of the lower pathway allocated on the pPDG2 plasmid, including: a *meta*-cleavage dioxygenase (AkbC), a *meta*-cleavage hydrolase (AkbD), an hydratase component (AkbE), and an aldolase (AkbF), respectively.

The involvement of the *akb* genes in the *o*-xylene degradation of R7 strain was analyzed by RT-PCR experiments. For this, RT-PCR were performed with RNA derived from *R. opacus* R7 cells grown in presence of *o*-xylene, or 2,3-DMP, or toluene, or malate as control. Separate cDNA synthesis reactions were performed and cDNA was then amplified with primer pairs used to amplify the target genes. The target genes were *akbA1A2a* coding for the small and the large components of the dioxygenase, or *akbB* coding for the dihydrodiol dehydrogenase, or *akbC* coding for the *meta*-cleavage dioxygenase. RT-PCR analysis showed that the *akbA1A2a*, *akbB,* and *akbC* genes were transcribed in R7 cells after growth on *o*-xylene (or toluene) (Fig. 5, panel a). These results indicate that *o*-xylene induced the transcription of the *akb* genes, suggesting the involvement of a dioxygenation route for R7 *o*-xylene degradation. But, as described above, the analyses of intermediates revealed that 2,3-dimethylphenol and 3,4-dimethylphenol were non-inducers of the pathway. Indeed, RT-PCR experiments with the same *akbA1A2a*, *akbB,* and *akbC* genes, after growth in presence of 2,3-dimethylphenol and 3,4-dimethylphenol, did not show any amplification. This propelled us in the direction to search in the R7 genome for other sequences encoding for monooxygenases/hydroxylases and to demonstrate their subsequent involvement in alternative pathways for *o*-xylene degradation leading to dimethylphenols.

Involvement of the *akb* genes in *o*-xylene degradation by the identification of *R. opacus* R7 mutants in this cluster

Random mutagenesis performed after electroporation of R7 cells with the pTNR vector generated mutants of R7 unable to growth on *o*-xylene.

We investigated the growth phenotypes and substrate transforming capabilities of R7 mutants by the transposon insertion detection. Among the single transposed mutant, the clearest phenotype was observed in the R7–50 mutant strain, in which the mutation was constituted by the insertion of the transposon in the *akbS* gene (Fig. 4, dashed box). This strain was considered a leaky mutant for the growth on *o*-xylene, as it is reported in Fig. 6 (panel a growth on malate, panel B growth on *o*-xylene) in comparison to the R7 wild type strain. In fact, Fig. 6 displays a lower rate of growth for the mutant in respect to the wild type strain when grown on *o*-xylene, while there is a similar trend when both strains are grown on malate.

These data are in accordance with what was observed by Kim et al. [38] when the ATP-binding motif of the sensor *akbS* gene was mutated in the DK17 strain. The mutation in the *akbS* gene allowed the incapacity of DK17 strain to grow well on *o*-xylene. So, our results

Table 1 List of the identified genes

Gene name	Protein name	Homologous protein	Bacterium	aa Identity	Function	Accession number	Reference
akbB (813 bp)	AkbB	AkbB	*R.* sp. DK17	85%	Dihydrodiol dehydrogenase	AII11489.1	Kim et al., 2004
akbA4 (1263 bp)	AkbA4	AkbA4	*R.* sp. DK17	81%	Ferredoxin reductase	AII11490.1	Kim et al., 2004
akbA3 (150 bp)	AkbA3	AkbA3	*R.* sp. DK17	69%	Ethylbenzene dioxygenase ferredoxin	CP008952.1	Kim et al., 2004
akbA2a (549 bp)	AkbA2a	AkbA2a	*R.* sp. DK17	84%	Ethylbenzene dioxygenase small subunit	AII11492.1	Kim et al., 2004
akbA1a (1323 bp)	AkbA1a	AkbA1a	*R.* sp. DK17	92%	Ethylbenzene dioxygenase large subunit	AII11493.1	Kim et al., 2004
akbS (4812 bp)	AkbS	AkbS	*R.* sp. DK17	76%	Sensor kinase	AII11494.1	Kim et al., 2004
akbT (534 bp)	AkbT	AkbT	*R.* sp. DK17	86%	Response regulator	AII11495.1	Kim et al., 2004
akbF (762 bp)	AkbF	AkbF	*R.* sp. DK17	63%	4-Hydroxy-2-oxovalerate aldolase	AII11049.1	Kim et al., 2004
akbE (900 bp)	AkbE	AkbE	*R.* sp. DK17	64%	2-Hydroxypenta-2,4-dienoate hydratase	AII11050.1	Kim et al., 2004
akbD (858 bp)	AkbD	AkbD	*R.* sp. DK17	67%	2-Hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	AII11051.1	Kim et al., 2004
akbC (900 bp)	AkbC	AkbC	*R.* sp. DK17	87%	2,3-Dihydroxybiphenyl 1,2-dioxygenase	AII11058.1	Kim et al., 2004
pheA1a (1617 bp)	PheA1a	PheA1 (1) PheA1 (2) PheA1 (3)	*R. opacus* 1CP	83% 93% 98%	Phenol hydroxylase	AII08653.1	Gröning et al., 2014
pheA1b (573 bp)	PheA1b	PheA2 (1) PheA2 (2) PheA2 (3)	*R. opacus* 1CP	73% 100% 74%	Phenol hydroxylase reductase component	AII08654.1	Gröning et al., 2014
pheA2a (1617 bp)	PheA2a	PheA1 (1) PheA1 (2) PheA1 (3)	*R. opacus* 1CP	83% 100% 92%	Phenol hydroxylase	AII08806.1	Gröning et al., 2014
pheA2b (561 bp)	PheA2b	PheA2 (1) PheA2 (2) PheA2 (3)	*R. opacus* 1CP	69% 75% 94%	Phenol hydroxylase reductase component	AII08807.1	Gröning et al., 2014
pheA3a (1617 bp)	PheA3a	PheA1 (1) PheA1 (2) PheA1 (3)	*R. opacus* 1CP	83% 97% 91%	Phenol hydroxylase	AII10865.1	Gröning et al., 2014
pheA3b (654 bp)	PheA3b	PheA2 (1) PheA2 (2) PheA2 (3)	*R. opacus* 1CP	61% 64% 58%	Phenol hydroxylase reductase component	CP008949.1	Gröning et al., 2014
prmA (1635 bp)	PrmA	PrmA	*R. jostii* RHA1	97%	Methane monooxygenase component A *alpha* chain	AII03499.1	Sharp et al., 2007
prmC (1044 bp)	PrmC	PrmC	*R. jostii* RHA1	94%	Methane monooxygenase component C	AII03498.1	Sharp et al., 2007
prmB (1107 bp)	PrmB	PrmB	*R. jostii* RHA1	97%	Methane monooxygenase component A *beta* chain	AII03497.1	Sharp et al., 2007
prmD (342 bp)	PrmD	PrmD	*R. jostii* RHA1	98%	Methane monooxygenase regulatory protein	AII03496.1	Sharp et al., 2007

indicate that *akbS* is necessary for the growth on *o*-xylene also in the R7 strain. Moreover, these data confirm the results of RT-PCR and suggest the kind of regulation involved in the *o*-xylene degradation pathway in the R7 strain. Indeed, this degradation process is likely mediated by the system sensor-regulator AkbS-AkbT through the binding of *o*-xylene.

Identification of the involvement of the *phe* genes and the *prm* genes in *o*-xylene degradation by RT-PCR experiments

Based on the identified sequences from genome analysis and the previous metabolic intermediates of R7 *o*-xylene degradation pathway, we analyzed the involvement of some putative sequences encoding for monooxygenases/phenol hydroxylases in this pathway. In particular, we selected the sequences deriving from clade number 1 (called *phe* sequences) and the sequences from clade number 4 (called *prm* sequences) (Table 1). We identified a first *phe* sequence (*pheA1a*) encoding for the monooxygenase PheA1a (P115) that showed a nucleotide identity of 98% with the sequences of the *pheA1(3)* gene of *R. opacus* 1CP, involved in the phenol hydroxylation [26]. In the R7 genome, this gene (*pheA1a*) clustered with another gene (*pheA1b*) encoding for a phenol hydroxylase-reductase component and other open reading frames (ORFs) encoding for unknown functions (Fig. 4, panel b). From the same group of sequences, we also selected other two sequences, *pheA2a* (PheA2a) (P122) homologous to the *pheA1(2)* gene

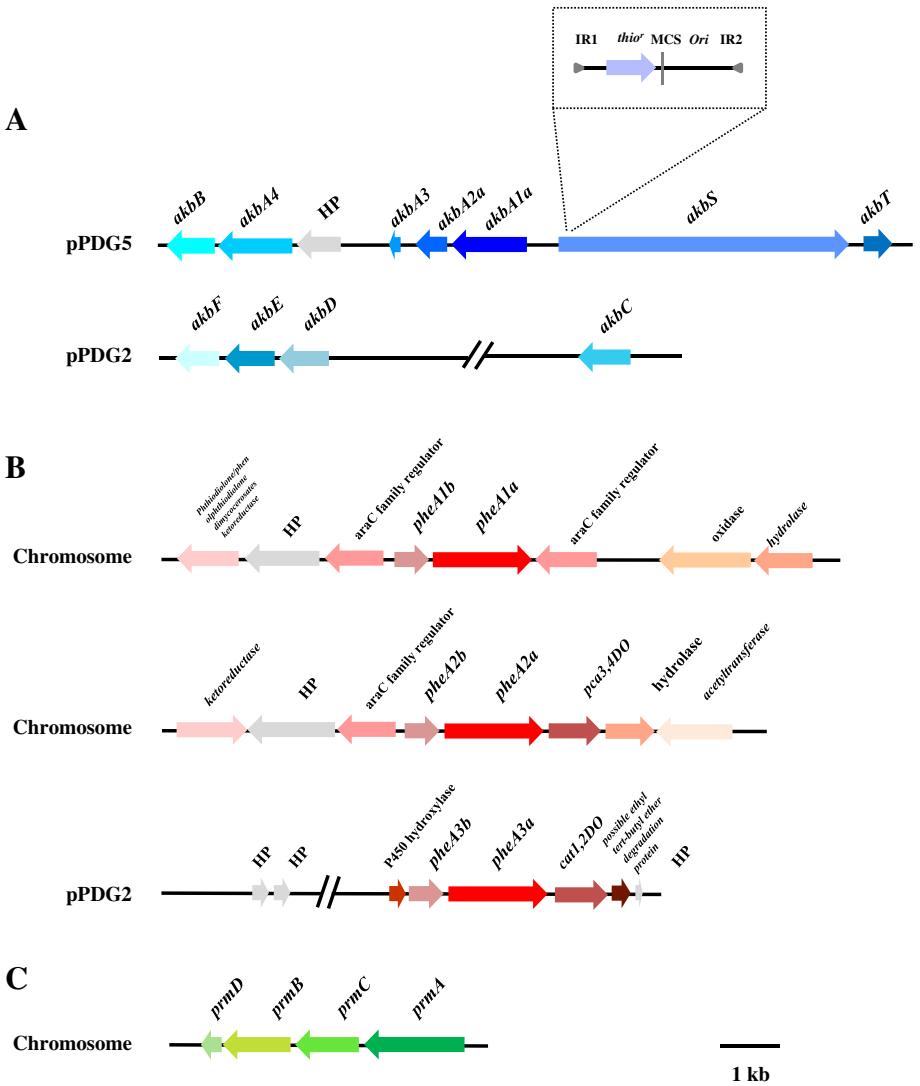


Fig. 4 Genetic organization of the *akb, phe,* and *prm* genes in *R. opacus* R7. **a** On pPDG2 plasmid: *akbA1a,* large subunit of *o*-xylene dioxygenase; *akbA2a,* small subunit of *o*-xylene dioxygenase; *akbA3*, ferredoxin component; *akbA4*, reductase component; *akbB*, dihydrodiol dehydrogenase; *akbS*, sensor kinase; and *akbT*, response regulator. On pPDG5 plasmid: *akbC*, *meta*-cleavage dioxygenase, *akbD*, *meta*-cleavage hydrolase product, *akbE*, hydratase and *akbF*, aldolase. Dashed box reported the IS*1415* insertion deriving from the transposon mutagenesis and dashed lines localized the insertion element within the *akbS* gene of R7. **b** Three *phe* gene clusters: *pheA1a* (P115), phenol hydroxylase, and *pheA1b*, phenol hydroxylase-reductase component located on the chromosome; *pheA2a* (P122), phenol hydroxylase, and *pheA2b*, phenol hydroxylase-reductase component located on the chromosome; *pheA3a* (P143), phenol hydroxylase, and *pheA3b*, phenol hydroxylase-reductase component located on the pPDG2 plasmid. **c** *prm* gene cluster located on the chromosome: *prmA* (P59), large hydroxylase subunit of a monooxygenase, *prmC*, small hydroxylase subunit of a monooxygenase, *prmB*, reductase component, and *prmD*, regulatory coupling protein. Genes with unknown or hypothetical functions were reported as HP. Identified genes (listed in Table 1) and their orientation are shown by arrow

(99%) of *R. opacus* 1CP and *pheA3a* (PheA3a) (P143) homologous to the *pheA1(2)* gene (97%) of *R. opacus* 1CP. The *pheA2a* gene clustered with the *pheA2b* gene coding for a phenol hydroxylase-reductase component and a *pca* 3,4-dioxygenase; while the *pheA3a* gene clustered with the *pheA3b* gene coding for a phenol hydroxylase-reductase component and a catechol 1,2 dioxygenase. The last *phe* gene cluster was found allocated on the pPDG2 plasmid.

We performed RT-PCR experiments on these identified sequences after growth of R7 cells in presence of *o*-xylene, or toluene, or 2,3-DMP, or malate as control. Separate cDNA synthesis reactions were performed and cDNA was then amplified with primer pairs used to amplify the target genes. RT-PCR analysis showed that in R7 cells grown on *o*-xylene (or on toluene), both the *pheA1a* gene (P115) and the *pheA3a* gene (P143) were amplified as well as on malate (Fig. 5, panel b). Among

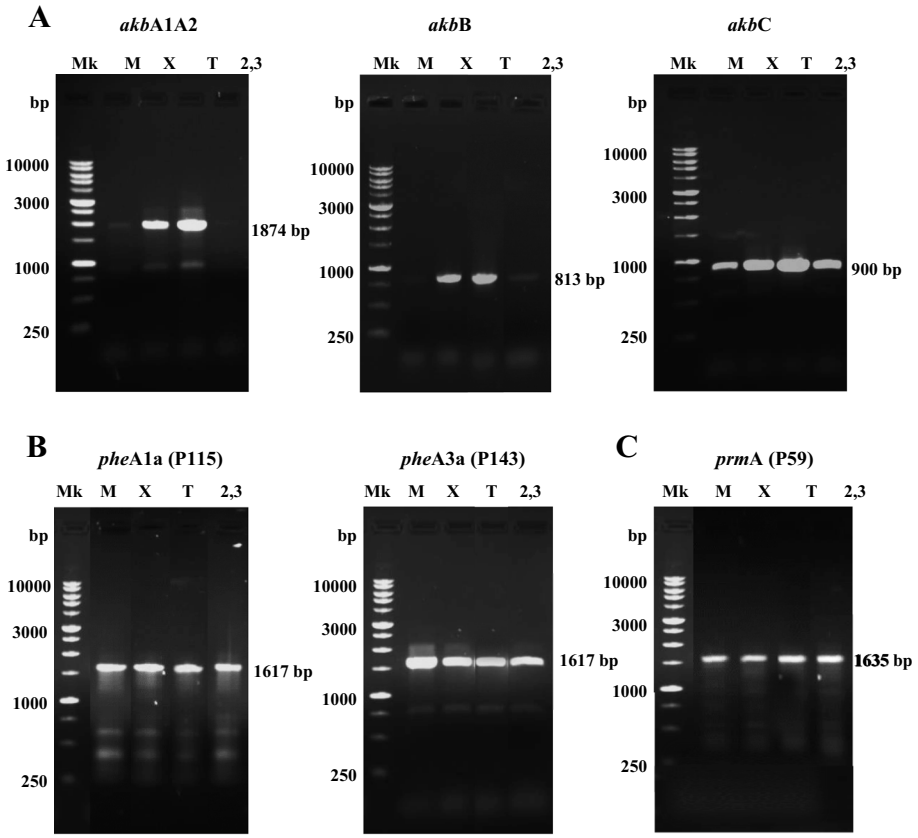


Fig. 5 RT-PCR experiments from RNA extracted from *R. opacus* R7. RT-PCR experiments with *akbA1A2*, *akbB* and *akbC* genes (**a**); RT-PCR experiments with *pheA1a*, *pheA3a* (**b**), and with *prmA* genes (**c**). M 100 to 10,000-bp markers, M growth on mineral medium M9 and malate used as control, X growth on *o*-xylene, T growth on toluene, 2,3 growth on 2,3-dimethylphenol

the identified *phe* sequences, the *pheA2a* sequence (P122) was not tested because we decided to test only the main representatives *pheA1a* (P115) and *pheA3a* (P143) as they are allocated on the chromosome and on the pPDG2 plasmid, respectively.

Moreover, as R7 strain was also able to oxidize *o*-xylene leading to the corresponding 2-methylbenzylalcohol, and data on RHA1 strain indicated the presence of genes up-regulated on propane coding for components of ethylbenzene dioxygenase [39], we decided to include the *prm* genes in the analysis (Fig. 4, panel c). The *prm* genes were found in a cluster constituted by the *prmA,C,B,D* genes, allocated on the chromosome with a percentage of amino acid identity near the 90% with the corresponding gene products of RHA1 strain. The *prmA* gene and *prmC* gene coded the large hydroxylase and the small hydroxylase subunits of a monooxygenase (annotated as propane monooxygenase), as *prmB* for

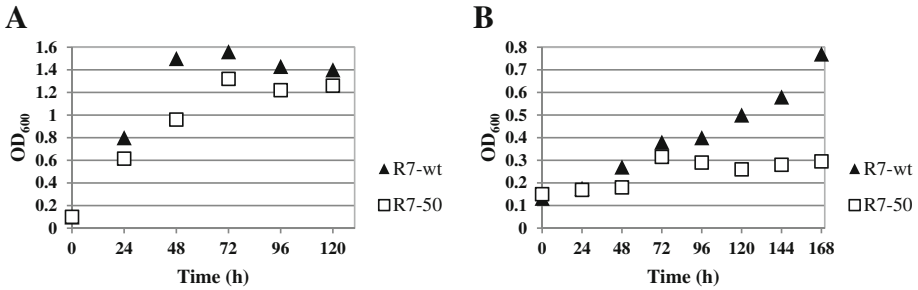


Fig. 6 Growth curves on *o*-xylene in wild type *R. opacus* R7 and R7–50 mutant strain. Cells were grown in presence of M9 mineral medium and malate (**a**) or M9 mineral medium supplied in an atmosphere saturated with *o*-xylene (**b**)

the reductase component and *prmD* for the regulatory coupling protein, respectively.

Also concerning these genes, we tested their involvement in *o*-xylene degradation by expression of the *prmA* gene (P59) in presence of the same substrates reported above for *phe* genes (Fig. 5, panel c). This gene was expressed in the presence of *o*-xylene as well as on malate, toluene and DMPs. These results indicated that the *prm* genes were amplified similarly to the *phe* genes, suggesting that they could work even when the strain was in absence of the hydrocarbon or phenols. Moreover, the gene redundancy of several monooxygenases/phenol hydroxylases supported the hypothesis of alternative pathways for *o*-xylene degradation in R7 strain. At the same time, the amplification of the PrmA (P59) could explain the formation of the corresponding 2-methylbenzyalcohol.

Quantitative real-time RT-PCR (qPCR) analysis

Quantitative real-time reverse transcription-PCR (qPCR) experiments were performed to quantify the levels of transcription of *akbA1a* (AkbA1a), *prmA* (P59) and *pheA1a* (P115) genes of R7 strain, representative of the selected catalytic subunit of different oxygenase systems putatively involved in *o*-xylene oxidation. qPCR experiments were performed after growth of R7 cells in presence of *o*-xylene, toluene and 2,3-DMP or malate as control. The values of transcription after R7 malate-grown cells were used as a basal level for comparison with the quantities determined with the substrates of interest. The level of *akbA1a* gene was approximately 19 ± 7.5-fold higher after growth on *o*-xylene (with a similar trend on toluene) than on malate. On the other hand, this analysis confirms that *prmA* and *pheA1a* gene transcription levels increased much less, which probably reflects their constitutive expression. In fact, the transcription levels of *akbA1a* gene in respect to *prmA* and *pheA1a* genes after growth on *o*-xylene (and on toluene), were found to be significantly different, with respective values of 0.23 ± 0.04 and 0.44 ± 0.11 (Table 2). A different trend was observed for the expression of the *prmA* and the *pheA1a* genes (*akbA1a* gene was not tested as it was not amplified in RT-PCR) after growth on 2,3-DMP. In this case, results showed an increase of the *pheA1a* transcription levels 5.18 ± 0.91-fold higher after growth on the corresponding dimethylphenol, whereas *prmA* was not induced.

These results demonstrated that *o*-xylene was able to activate mainly the transcription of the *akbA1a* gene whilst a very low level of the other two genes during the aerobic growth of R7 cells on *o*-xylene. Meanwhile, in presence of 2,3-DMP a higher level of expression of phenol hydroxylase was observed.

Involvement of the *prm* genes by cloning and expression of the activity in *R. erythropolis* AP

In order to evaluate the role of the *prmACBD* gene cluster in the *o*-xylene metabolism, the region of 4.3 kb was isolated from R7 genomic DNA as *Nde*I/*Hind*III fragment. The PCR product was cloned into the shuttle-vector *E. coli-Rhodococcus* pTipQC2 to obtain pTipQC2-*prmACBD*-R7.

The recombinant plasmid pTipQC2-*prmACBD*-R7 was isolated from *E. coli* DH5a and transferred by electroporation into *Rhodococcus erythropolis* AP, which was not able to use the *o*-xylene as only carbon and energy source. The *prmACBD* gene cluster was expressed under the inducible thiostrepton promoter (P*tipA*) through experiments with resting cells of *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) exposed to *o*-xylene to identify the metabolites. The activity of the recombinant strain was compared to the activity of wild type AP strain treated in the same conditions as control. *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) cells, which were pre-grown on LB and washed in mineral medium M9, were exposed to *o*-xylene dissolved in isoctan in a biphasic system. The water phase was analyzed at different incubation times by reverse-HPLC analysis; 3,4-dimethylphenol and 2-methylbenzylalcohol were identified by comparison with reference compounds (standard mixture) (Fig. 7). These compounds were observed in the first 2 h of exposure, then they were progressively metabolized and disappeared after 6 h. It was not possible to confirm the formation of the 2,3-dimethylphenol. None of these metabolites was identified in the wild type host strain. These results suggested that the *prmACBD* gene cluster could have a role within the *o*-xylene metabolism, in particular in the first step of oxidation.

Discussion

The genome-based analysis of *R. opacus* strain R7 revealed a considerable multiplicity of genes potentially involved in *o*-xylene catabolism. Although much is known about the

Table 2 qPCR analysis. Relative gene expression of *R. opacus* R7 grown on *o*-xylene, toluene and 2,3-dimethylphenol (2,3-DMP)

Substrate	Normalized genes amount relative to malate condition ($2^{-\Delta\Delta Ct}$)					
	akbA1a		*prmA*		*pheA1a*	
o-xylene	19.14	±7.55	0.23	±0.04	0.44	±0.11
toluene	15.14	±6.56	1.53	±0.39	2.14	±0.65
2,3-DMP	–	–	0.04	±0.01	5.18	±0.91

Values are means of three replicates ± standard deviation

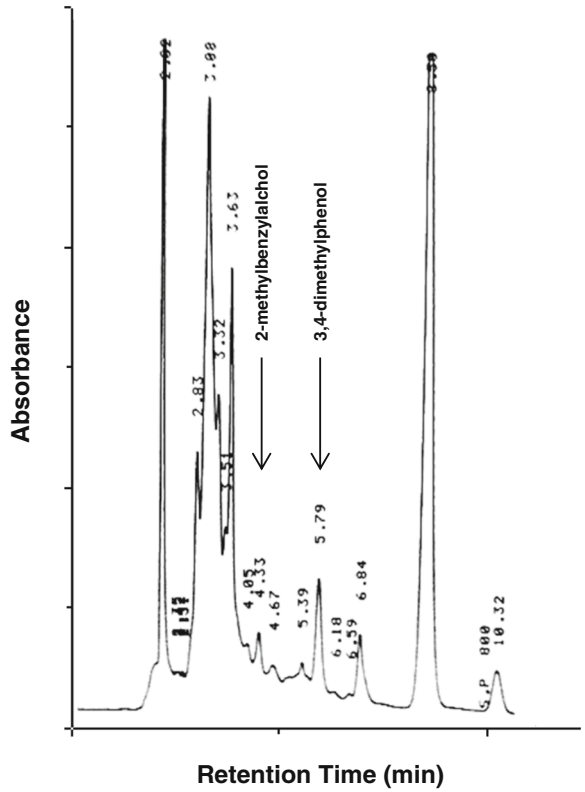


Fig. 7 Reverse-Phase HPLC analysis of culture broth of *R. erythropolis* AP (pTipQC2-*prmACBD*-R7). The analysis was performed after exposition to *o*-xylene of the strain on the water phase of the biphasic system with isoctane. The retention times were 4.33 min for 2-methylbenzylalcohol, 5.79 min for 3,4-dimethylphenol

ability of *Rhodococcus* strains to grow on toluene and ethylbenzene [39–41], little is known about the catabolism of *o*-xylene in bacteria belonging to the *Rhodococcus* genus [11]. *R. opacus* R7 is a strain isolated for its ability to grow on *o*-xylene as the only carbon and energy source. The strain is able to grow on *o*-xylene but not on *m*- and *p*-xylene. The inability of R7 strain to grow on the latter two compounds reinforced the hypothesis that the xylenes are metabolized at least by two different pathways [12]. Moreover, we have previously identified [22] the 2,3- and 3,4-dimethylphenols, as the main intermediates in the culture medium of R7 exposed to *o*-xylene, which are used by the strain as the only carbon and energy source, and not the corresponding dihydrodiol. Otherwise, in literature is reported by Kim et al. [10] that *o*-xylene is oxidized to the corresponding dihydrodiol. Moreover, Kim and co-authors also reported the direct formation of dimethylphenols in presence of *m*- and *p*-xylenes by *Rhodococcus* sp. strain DK17. This suggests that alternative oxidation mechanisms of xylenes are possible [11]. Whether this is through the action of a dioxygenase, forming a dihydrodiol, which dehydrates to a phenolic intermediate, or through the action of a monooxygenase which can directly hydroxylate the aromatic ring (or a combination of the two steps), it remains to be investigated.

In this context, a genome-based approach was used to better understand the R7 peculiar *o*-xylene pathway. Consequently, we decided to investigate the role of some selected genes and to demonstrate their involvement in this catabolism (Fig. 8). As a first step we analyzed and clustered all the R7 genome oxygenase sequences generating two phylogenetic trees (Figs. 2 and 3).

From the dioxygenase tree analysis, we selected the AkbA1a dioxygenase coded by the *akbA1a* gene (included in the *akb* gene cluster), whose sequences are 90% homologous to the sequence of the DK17 strain. In this paper, we demonstrate that the *akb* genes are induced by the presence of *o*-xylene supplied as the only carbon and energy source, both by RT-PCR/qPCR experiments and by selection of the R7–50 leaky mutant on *o*-xylene. RT-PCR analysis showed that *o*-xylene activated the transcription of the *akbA1A2* genes coding for a *o*-xylene dioxygenase and the *akbB* gene coding for a dihydrodiol dehydrogenase, suggesting the dioxygenation route for the *o*-xylene oxidation. These data confirmed what was reported for DK17 strain [10].

However, these data are apparently in disagreement with what we observed in the R7 metabolic analysis [21], because we identified the 2,3- and 3,4-DMPs and no literature data supports enough that DMPs could derive from the dehydration of the corresponding dihydrodiol. Moreover, R7 strain was also able to grow on the corresponding 2,3- and 3,4-DMPs as the only carbon and energy source, suggesting an alternative pathway for the *o*-xylene oxidation through a monooxygenation. Since the focus of the present work was the identification of genes involved in the initial oxidation of *o*-xylene, we wanted to verify the formation mechanism of 2,3- and 3,4-DMPs from *o*-xylene. To support this hypothesis, we analyzed all the R7 monooxygenases/phenol hydroxylases sequences (Fig. 3). As a result, on the basis of sequence identities with other *o*-xylene monooxygenases/phenol hydroxylases, genome sequences of other bacteria, and comparison of the protein catalytic site, we found putative sequences that could be involved in *o*-xylene degradation. In clade number 4, only one R7 protein sequence (PrmA) (P59) showed a significant amino acid identity with respect to reference *o*-xylene monooxygenases (like Toluene *o*-xylene Monooxygenase, ToMo). As the ToMo is the best known monooxygenase able to oxidize *o*-xylene with the formation of the corresponding DMPs, we hypothesized that PrmA could also be involved in this monooxygenation/hydroxylation. Thus, we investigated the role of the *prmA* gene in the *o*-xylene (or toluene) degradation by RT-PCR and quantitative real time RT-PCR experiments on R7 cells grown in presence of *o*-xylene. Results showed that

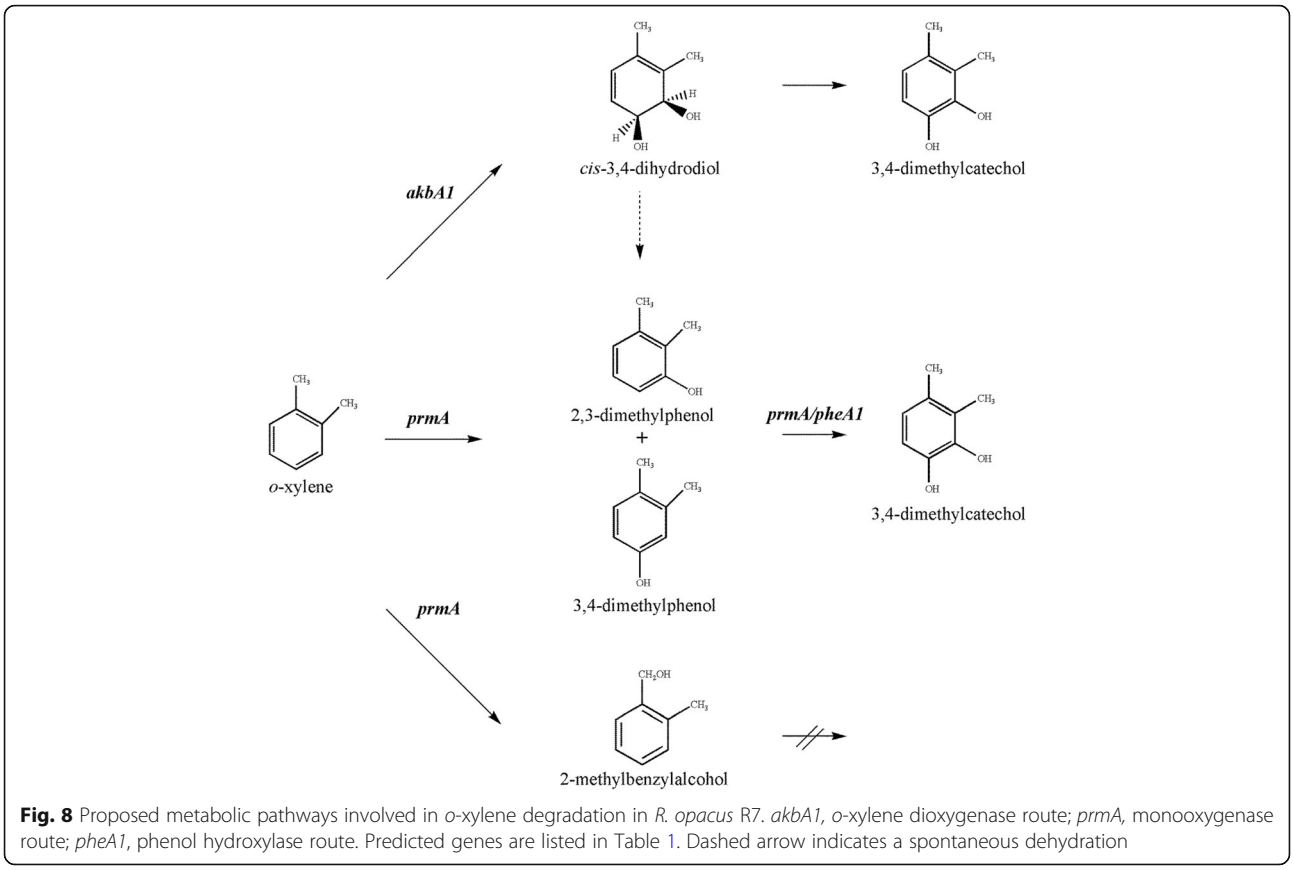


Fig. 8 Proposed metabolic pathways involved in *o*-xylene degradation in *R. opacus* R7. *akbA1*, *o*-xylene dioxygenase route; *prmA*, monooxygenase route; *pheA1*, phenol hydroxylase route. Predicted genes are listed in Table 1. Dashed arrow indicates a spontaneous dehydration

although the levels of the *prmA* gene increased little with respect to the growth on malate, it could be a first indication of the involvement of PrmA in this pathway. Then, the activity of the PrmACBD multicomponent monooxygenase was examined after the cloning of the corresponding genes in another *Rhodococcus* strain unable to use *o*-xylene like *R. erythropolis* AP. Results evidenced that *o*-xylene was slowly transformed in 3,4-DMP and 2-methylbenzylalcohol, that could be then metabolized by other monooxygenases/phenol hydroxylases. Indeed, R7 strain is able to grow also in presence of the DMPs when supplied as sole carbon and energy source. To support these data we have also looked for R7 monooxygenases/phenol hydroxylases. From R7 genome analysis, we identified three sets of two component phenol hydroxylases, constituted of an oxygenase component and a reductase component; two sets were allocated on the chromosome on two different regions, and the third one on the pPDG2 plasmid of R7, respectively. In this case, *phe* genes were selected for RT-PCR experiments to evaluate their involvement in *o*-xylene oxidation. In all the growth conditions utilized we observed an amplification of the corresponding genes. The *pheA1a* encoding for the PheA1a (P115) was also tested in the quantitative real time RT-PCR on R7 cells grown in presence of *o*-xylene, toluene, 2,3-DMP and malate. Results showed a significant increase of *pheA1a* gene transcriptional levels during the growth on 2,3-DMP in respect to the growth on malate. This suggested that the *phe* genes could be involved in the second step of *o*-xylene degradation.

R7 strain showed a substrate versatility in respect to different substituted phenols, including 2,3-DMP and 3,4-DMP. This substrate versatility could likely be the result of gene redundancy and the presence of several phenol hydroxylase (iso)enzymes. These data are in accordance with literature, where it is reported that: three phenol hydroxylases in *R. opacus* B4, four in *R. opacus* M213, five in *R. opacus* PD630 and four in the reference strain *R. jostii* RHA1 have evident activities and expression profiles for this class of enzymes in these bacteria [26].

Thus, considering such metabolic diversity of R7 strain, we would deduce that, although the *akb* genes are the specific activated genes for *o*-xylene degradation, other genes such as *prm* genes can induce an increase of levels of phenols that can converge towards the phenol oxidation route.

The co-activation of multiple oxygenases could contribute to such strategy in these kinds of bacteria particularly resistant to environmental stress. Indeed, it has been demonstrated [42] that large genome with multiple

broad-specificity catabolic enzymes such as those reported in RHA1 strain could have a competitive advantage in environmental changing soil conditions.

Conclusions

In conclusion, in this paper we demonstrate that *R. opacus* R7 is able to degrade *o*-xylene by the activation of the *akb* genes leading to the production of the corresponding dihydrodiol. Likewise, the redundancy of sequences encoding for several monooxygenases/phenol hydroxylases, supports the involvement of other genes that can induce an increase of levels of phenols that can converge towards the phenol oxidation route.

The activation of multiple converging oxygenase systems represents a strategy in bacteria of *Rhodococcus* genus to degrade a wide range of recalcitrant compounds and to persist in severely contaminated environments.

Methods

Bacterial strain and growth conditions

The bacterial strain used in this study is *R. opacus* R7, isolated for its ability to grow on naphthalene and *o*-xylene as previously described [21] (deposited to the Institute Pasteur Collection, CIP identification number 107348). The strain was grown at 30 °C in M9 mineral medium [43], supplemented with the following carbon sources as only carbon and energy source: *o*-xylene, toluene (final concentration of 1 g/l) or 2,3-dimethylphenol, 3,4-dimethylphenol (final concentration of 5–10 mM), or 2-dimethylbenzylalcohol, 3-dimethylbenzylalcohol or malate (final concentration of 10 mM). The *R. opacus* R7 growth on *o*-xylene, toluene, took place on M9 mineral medium in an atmosphere saturated with these aromatic compounds in a sealed system. The mutant R7–50 strain used in this paper was grown in the same conditions utilized for the wild type R7 strain.

Rhodococcus erythropolis AP, isolated in our laboratory (CIP 110799) for its ability to grow on diesel fuel, was maintained on M9 mineral medium in a saturated atmosphere of diesel fuel at 30 °C.

Bioinformatic analysis: Nucleotide sequence determination and protein sequence analysis

The preliminary annotation of *R. opacus* R7 genome sequences was performed using the RAST (Rapid Annotation using Subsystem Technology) service [44].

BLASTn tool [45] of NCBI pipeline was used to determine nucleotide sequence homology and to make manual curation.

R. opacus R7 putative gene clusters for *o*-xylene catabolism were identified on chromosome and megaplasmids using BLAST tool and Clustal Omega [46].

R. opacus R7 protein sequences were preliminary annotated using the RAST that allowed to identify potential monooxygenases/hydroxylases and dioxygenases using text string searching.

These sequences annotated as monooxygenases/hydroxylases and dioxygenases were aligned separately against PDB (RCSB Protein Data Bank) database to identify reference sequences. Reference proteins were selected on the basis of the highest similarity or literature data. If no match was identify against PDB database, the same procedure was applied using BLASTp of NCBI pipeline.

Afterwards, the identified reference sequences were aligned against R7 genome using the NCBI pipeline in order to verify to have considered all R7 putative monooxygenases/hydroxylases and dioxygenases.

The retrieved sequences were aligned using the multiple sequence alignment (MSA) tool of Clustal Omega program using the default parameters (neighbour joining method, the Gonnet transition matrix, gap opening penalty of 6 bits, maintain gaps with an extension of 1 bit, used bed-like clustering during subsequent iterations, and zero number of combined iterations).

For each group of oxygenases, the MSA was used for the cluster analysis inferred using the maximum likelihood (ML) method selected from the package MEGA version 6 [47]. The following parameters were used: JTT matrix, used all sites and gamma distribution of mutation rates with gamma optimized to 2. As a test of inferred phylogeny, 100 bootstrap replicates were used.

The resulting groups allowed to define two different trees, one for all the dioxygenases and one for all the monooxygenases/hydroxylases of R7 showing clades with putative functions identified by InterPro/UniProt databases.

Preparation, analysis, and DNA manipulation

Total DNA from *R. opacus* R7 was extracted according to method reported by Di Gennaro et al. [22]. The extract was precipitated by 0.1 volume of 3 M sodium acetate and after centrifugation, the DNA was isolated and purified. Standard methods of DNA manipulation were used in this work [43]. For the recovery and purification of DNA fragments from agarose, Extraction Kits by Nachery and Nagel (Fisher Scientific, Germany) were used. Amplification of fragment containing genes target was achieved by PCR performed using primers designed ad hoc (Additional file 3: Table S3 and Additional file 4: Table S4) to amplify the sequences of interest.

RNA extraction and RT-PCR, quantitative real-time RT-PCR (qPCR)

Total RNA was extracted from bacterial cultures of *R. opacus* R7 (100 ml) grown at 30 °C on M9 mineral medium supplemented with different substrates supplied (as described above) as the only carbon and energy source: *o*-xylene, toluene at the concentration of 1 g/l, 2,3-dimethylphenol and 3,4-dimethylphenol at

the concentration of 5–10 mM and 10 mM malate used as reference.

RNA extraction protocol was performed using the RNA-Total RNA Mobio Isolation Kit (Qiagen Italia, Italy) according to the manufacturer's instructions and at the end the DNase treatment was performed. Reverse transcription was performed with iScript cDNA Synthesis kit (BIO-RAD, Italy) to obtain the corresponding cDNAs. For the cDNA synthesis 200 ng of total RNA was reverse-transcribed as follows: after denaturation for 5 min at 25 °C, reverse transcription was performed for 1 h at 42 °C and then 5 min of elongation at 85 °C.

RT-PCR experiments were performed by amplification of the cDNA samples, each in 25-μl PCR volume containing 2 μl of the reverse-transcribed RNA samples.

Amplifications of the *akbA1A2*, *akbB*, *akbC*, *prmA* (P59), *pheA1a* (P115), *pheA3a* (P143) genes and 16S rDNA were performed using 2.5 U/μl of Long Range DNA Rabbit Polymerase (Eppendorf, Germany).

Thermo cycling conditions were as follows: 3 min at 95 °C, 95 °C for 30 s, specific T_m for 45 s, 72 °C for 4 min, for 35 cycles; and 72 °C for 3 min. Amplification of 16S rDNA was performed using the universal bacterial primers 27f and 1495r as described in Di Gennaro et al. [21]. The internal housekeeping gene (16S rDNA) was used as reference to evaluate relative differences in the integrity of individual RNA samples.

Quantitative real-time Reverse Transcriptase-PCR (qPCR) analyses were performed on the same samples used for RT-PCR. The reverse-transcribed samples were amplified using the StepOnePlus Real-Time PCR System (Applied Biosystem, Italy). Each 10-μl qPCR volume contained 4.4 μl of the reverse-transcribed RNA samples, 5 μl of PowerUp SYBR Green Master Mix (Applied Biosystem, Thermo Scientific, Italy), and 300 nM of each primer, listed in Additional file 3: Table S3. Thermocycling conditions were as follows: 30 s at 95 °C, followed by 40 cycles of 5 s at 95 °C, 10 s at 60 °C and 45 s at 72 °C and one cycle 15 s at 95 °C, 1 min at 60 °C and 15 s at 60 °C. Expression of the housekeeping gene, 16S rDNA, was used as reference gene to normalize tested genes in *R. opacus* R7. The ΔΔCt method with 16S rDNA as reference gene was used to determine relative abundance of target transcripts in respect to malate as control. Data are expressed as mean ± standard deviation derived from at least three independent experiments.

In order to exclude DNA contamination, negative controls were performed by omitting the reverse transcriptase in RT-PCR experiments, which were conducted with the same temperature program and the same primer sets for 35 cycles of amplification.

The primers used in the RT-PCR analysis and qPCR are described in Additional file 3: Table S3.

Mutant preparation

Transposon-induced mutagenesis in R. opacus R7 using IS1415 (pTNR-TA vector)

Plasmid pTNR-TA [48] was transferred into *R. opacus* strain R7 by electroporation as described by Treadway et al. [49], using a Gene Pulser II (BIO-RAD, Italy) set at 2.50 kV, 600 Ω, 25 μF in presence of maximum 1 μg DNA in a 2 mm-gap electro-cuvette (BIO-RAD, Italy). Afterwards, the electroporation mixture was suspended in 2.5 ml LB and it was incubated for 4 h at 30 °C under shaking. Cells were plated on Luria-Bertani (LB) supplemented with 12.5 μg/ml thiostrepton and they were grown at 30 °C for 5 days to select thiostrepton-resistant cells. Transposon-induced mutants were transferred to M9 mineral medium agar plates with 12.5 μg/ml thiostrepton and 10 mM malate.

The transposon-induced mutants obtained were tested on M9 mineral medium agar plates with the following carbon sources as the only carbon and energy source at the final concentration of 1 g/l: *o*-xylene; toluene; 2,3-dimethylphenol and 3,4-dimethylphenol (5–10 mM); 2 dimethylbenzylalcohol.

Analysis of pTNR-TA insertion sites

The genomic DNA of each transposon-induced mutant was extracted and the Two-Step gene walking PCR method was applied [50]. Insertions of IS1415 into the genomes of these mutants were confirmed by PCR using primers reported in Additional file 4: Table S4. Genomic DNA of the wild type strain was used as a negative control. Homology searches of the interrupted DNA sequences from mutants were conducted by BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) [45].

Construction of the recombinant strain *R. erythropolis* AP (pTipQC2-*prmACBD*-R7)

The *prmACBD* gene cluster was ligated as *Nde*I/*Hind*III fragment into a shuttle-vector *E. coli-Rhodococcus*, pTipQC2 [51]. The ligation mixture was used to transform *E. coli* DH5α by electroporation with standard procedures [52] and the recombinant clones were selected on LB agar supplemented with ampicillin (100 μg/ml) at 37 °C. Ampicillin-resistant clones were selected and the recombinant plasmid (pTipQC2-*prmACBD*-R7) was isolated. The same recombinant plasmid was used to transform *R. erythropolis* strain AP by electroporation according to Zampolli et al. [23]. Immediately after electroporation, 2.5 ml recovery broth (LB medium with 1.8% sucrose) were added and cells were incubated at 30 °C for 4 h. Cells were plated on LB supplemented with chloramphenicol 50 μg/ml and grown at 30 °C for 3–4 days. Recombinant strain *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) was used for bioconversion experiments in presence of *o*-xylene to

evaluate the activity of the *prmACBD* system in comparison to the activity of the wild type strain.

Bioconversion experiments of the recombinant strain *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) in presence of *o*-xylene

Cells of *R. erythropolis* AP (pTipQC2-*prmACBD*-R7) were grown in 500-mL Erlenmeyer flasks containing 100 mL of LB at 30 °C. When the culture reached the $O.D._{600}$ of 0.6, they were induced with thiostrepton. After overnight growth, the cells were collected by centrifugation for 10 min at 20,000 g, washed and re-suspended in the mineral medium M9 to have final $O.D._{600}$ of 1.

Bioconversion experiments were performed in a biphasic system with *o*-xylene (1 g/l) dissolved in isoctan at 20% of the final volume of the culture. At fixed times, 2 h, 4 h, 6 h, and 24 h, a flask was sacrificed to determine the metabolites produced during growth on *o*-xylene. It occurred by HPLC analysis after acidification with H_2SO_4. After settling of the suspension, the aqueous phase was drawn from the organic phase, stripped under a gentle stream of nitrogen, resuspended in order to concentrate each culture 1:20 and filtered with 0.45 μm filters for HPLC analysis.

Analytical methods

HPLC analyses were performed with a Waters 600E delivery system (Waters Corporation, Milford, MA, USA) equipped with a Waters 486 UV-Vis detector and a Waters reverse-phase μBondapak 3.9 × 300 mm C18 column. The mobile phase was acetonitrile:water (50:50) at a flow rate of 1 mL/min in isocratic conditions. The detection was carried out at 254 nm and the retention times of the identified peaks were compared with those of authentic compounds. Co-elution experiments were also performed in which the culture broth samples were supplemented with authentic compounds. In these conditions, the retention times were 4.33 min for 2-methylbenzylalcohol, 6.39 min for 2,3-dimethylphenol, and 5.79 min for 3,4-dimethylphenol.

Additional files

Additional file 1: Table S1. Sequences used to generate dioxygenase tree. (DOCX 25 kb)

Additional file 2: Table S2. Sequences used to generate monooxygenase tree. (DOCX 44 kb)

Additional file 3: Table S3. List of utilized oligos. (DOCX 15 kb)

Additional file 4: Table S4. List of utilized oligos for transposon identification. (DOCX 13 kb)

Abbreviations

2,3-DMP: 2,3-dimethyl phenol; 3,4-DMP: 3,4-dimethyl phenol; BLAST: Basic Local Alignment Search Tool; BTEX: benzene, toluene, ethylbenzene and xylenes; cDNA: Complementary Deoxyribonucleic Acid; DMPs: dimethyl phenols; HP: hypothetical protein; HPLC: High-Performance Liquid Chromatography; LB: Luria-Bertani; NCBI: National Center for Biotechnology Information; O.D.600: Optical Density at 600 nm; ORFs: Open Reading Frames; PAHs: polycyclic aromatic hydrocarbons; PCBs: polychlorinated biphenyls; PDB: RCSB Protein Data Bank; qPCR: quantitative real time RT-PCR; RAST: Rapid Annotation using Subsystem Technology; RT-PCR: Reverse Transcriptase Polymerase Chain Reaction; ToMo/PH: Toluene *o*-xylene Monooxygenase/Phenol Hydroxylase

Acknowledgments

We would like to thank the Department of Organic Chemistry of the University of Milano for the support in the analytical instrument procedures. We thank also our students Matteo Uggeri and Delia Valsecchi.

Funding

We thanks CNR for supporting our research with the Project Interomics and the Italian Ministry of University.

Availability of data and materials

The datasets analysed in this study are available at DDBJ/EMBL/GenBank under the Accession Numbers CP008947, CP008948, CP008949, CP008950, CP008951, CP008952 for *R. opacus* R7.
All data generated during this study are included in this article and its additional files.

Authors' contributions

PDG and LM planned the experiments. ADC, JZ, performed molecular experiments, PDU and AO performed bioinformatic analysis, GS performed chemical analysis. GS and AS analysed and interpreted the data. PDG, JZ and GS wrote the manuscript. All the authors reviewed the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

All procedures performed in this study were compliance with Ethical Standards. This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

Not Applicable.

Competing interests

The authors declare that they have competing interests.

Publisher's Note

Author details

[1]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milan, Italy. [2]ITB, CNR, via Fratelli Cervi 19, 20133 Segrate, Milan, Italy. [3]Department of Chemistry, University of Milano, via Golgi 19, 20133 Milan, Italy. [4]Department of Molecular Microbiology and Biotechnology, Westfälische Wilhelms-Universität Münster, Münster, Germany. [5]Environmental Sciences Department, King Abdulaziz University, Jeddah, Saudi Arabia.

Received: 26 March 2018 Accepted: 30 July 2018
Published online: 06 August 2018

References

1. Barbieri P, Palladino L, Di Gennaro P, Galli E. Alternative pathways for *o*-xylene or *m*-xylene and *p*-xylene degradation in *Pseudomonas stutzeri* strain. Biodegradation. 1993;4:71–80.
2. Kim D, Kim YS, Kim SK, Kim SW, Zylstra GJ, Kim YM, et al. Monocyclic aromatic hydrocarbon degradation by *Rhodococcus* sp. strain DK17. Appl Environ Microbiol. 2002;6:3270–8.
3. Galli E, Barbieri P, Bestetti G. Potential of *Pseudomonas* in the degradation of methylbenzenes. In: Galli E, Silver S, Witholt E, editors. *Pseudomonas*: molecular biology and biotechnology. Washington, DC: American Society for Microbiology; 1992. p. 268–76.

4. Harayama S, Kok M, Neidle EL. Functional and evolutionary relationships among diverse oxygenases. Annu Rev Microbiol. 1992;46:565–601.
5. Williams PA, Sayers JR. The evolution of pathways for aromatic hydrocarbon oxidation in *Pseudomonas*. Biodegradation. 1994;5:195–217.
6. Jindrovà E, Chocovà M, Demnerovà K, Brenner V. Bacterial aerobic degradation of benzene, toluene, ethylbenzene and xylene. Folia Microbiol. 2002;47:83–93.
7. Schraa G, Bethe BM, van Neerven ARW, Van Den Tweel WJJ, Van Der Wende E, Zehnder AJB. Degradation of 1,2-dimethylbenzene by *Corynebacterium* strain C125. Antoine van Leewenhock. 1987;53:159–70.
8. Bickerdike SR, Holt RA, Stephens GM. Evidence for metabolism of *o*-xylene by simultaneous ring and methyl group oxidation in a new soil isolate. Microbiology. 1997;143:2321–9.
9. Vardar G, Wood TK. Protein engineering of toluene-*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1 for synthesizing 4-methylresorcinol, methylhydroquinone, and pyrogallol. Appl Environ Microbiol. 2004;70:3253–62.
10. Kim D, Chae JC, Zylstra GJ, Kim YS, Kim SK, Nam MH, et al. Identification of a novel dioxygenase involved in metabolism of *o*-xylene, toluene, and ethylbenzene by *Rhodococcus* sp. strain DK17. Appl Environ Microbiol. 2004; 70:7086–92.
11. Kim D, Choi KY, Yoo M, Choi JN, Lee CH, Zylstra GJ, et al. Benzylic and aryl hydroxylations of *m*-xylene by *o*-xylene dioxygenase from *Rhodococcus* sp. strain DK17. Appl Microbiol Biotechnol. 2010;86:1841–7.
12. Kim D, Kim Y-S, Jung JW, Zylstra J, Kim YM, Kim S-K, et al. Regioselective oxidation of xylene isomers by *Rhodococcus* sp. strain DK17. FEMS Microbiol Lett. 2003;223:211–4.
13. Kukor JJ, Olsen RH. Molecular cloning, characterization and regulation of a *Pseudomonas pikettii* PKO1 gene encoding phenol hydroxylase and expression of the gene in *Pseudomonas aeruginosa* PA01c. J Bacteriol. 1990; 172:4624–30.
14. Martìnkovà L, Uhnàkovà B, Pàtek M, Nesvera J, Kren V. Biodegradation potential of the genus *Rhodococcus*. Environ Int. 2009;35:162–77.
15. Alvarez HM. Central metabolism of the species of the genus *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus*. Berlin Heidelberg: Springer; 2010. p. 91–108.
16. Larkin MJ, Kulakov LA, Allen CCR. Genomes and plasmids in *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus*. Berlin, Heidelberg: Springer; 2010. p. 73–90.
17. McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, et al. The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci U S A. 2006;103:15582–7.
18. Seto M, Masai E, Ida M, Hatta T, Kimbara K, Fukuda M, et al. Multiple polychlorinated biphenyl transformation systems in the gram-positive bacterium *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol. 1995;61:4510–3.
19. Sakai M, Miyauchi K, Kato N, Masai E, Fukuda M. 2-Hydroxypenta-2,4-dienoate metabolic pathway genes in a strong polychlorinated biphenyl degrader, *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol. 2003;69:427–33.
20. Kesseler M, Dabbs ER, Averhoff B, Gottschalk G. Studies on the isopropylbenzene 2,3-dioxygenase and the 3-isopropylcatechol 2,3-dioxygenase genes encoded by the linear plasmid of *Rhodococcus erythropolis* BD2. Microbiology. 1996;142:3241–51.
21. Di Gennaro P, Rescalli E, Galli E, Sello G, Bestetti G. Characterization of *Rhodococcus opacus* R7, a strain able to degradate naphthalene and *o*-xylene isolated from polycyclic aromatic hydrocarbon-contaminated soil. Res Microbiol. 2001;152:641–51.
22. Di Gennaro P, Terreni P, Masi G, Botti S, De Ferra F, Bestetti G. Identification and characterization of genes involved in naphthalene degradation in *Rhodococcus opacus* R7. Appl Microbiol Biotechnol. 2010;87:297–308.
23. Zampolli J, Collina E, Lasagni M, Di Gennaro P. Biodegradation of variable-chain-length *n*-alkanes in *Rhodococcus opacus* R7 and the involvement of an alkane hydroxylase system in the metabolism. AMB Express. 2014;4:73.
24. Orro A, Cappelletti M, D'Ursi P, Milanesi L, Di Canito A, Zampolli J, et al. Genome and phenotype microarray analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: genetic determinants and metabolic abilities with environmental relevance. Plos One. 2015; https://doi.org/10.1371/journal.pone.0139467.
25. Gonçalves ER, Hara H, Miyazawa D, Davies JE, Eltid LD, Mohn WW. Transcriptomic assessment of isozymes in the biphenyl pathway of *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol. 2006;72:6183–93.
26. Gröning JAD, Eulberg D, Tischler D, Kaschabek SR, Schlömann M. Gene redundancy of two-component (chloro)phenol hydroxylases in *Rhodococcus opacus* 1CP. FEMS Microbiol Lett. 2014;361:68–75.
27. Bertoni G, Martino M, Galli E, Barbieri P. Analysis of the gene cluster encoding toluene/*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1. Appl Environ Microbiol. 1998;64:3626–32.
28. Cafaro V, Notomista E, Capasso P, Di Donato A. Regiospecificity of two multicomponent monooxygenases from *Pseudomonas stutzeri* OX1: molecular basis for catabolic adaptation of this microorganism to methylated aromatic compounds. Appl Environ Microbiol. 2005;71:4736–43.
29. Notomista E, Lahm A, Di Donato A, Tramontano A. Evolution of bacterial and archaeal multicomponent monooxygenases. J Mol Evol. 2003;56:435–45.
30. Leahy JG, Batchelor PJ, Morcomb SM. Evolution of the soluble diiron monooxygenases. FEMS Microbiol Rev. 2003;27:449–79.
31. Yachnin BJ, McEvoy MB, MacCuish RJD, Morley KL, Lau PCK, Berghuis AM. Lactone-bound structures of cyclohexanone monooxygenase provide insight into the stereochemistry of catalysis. ACS Chem Biol. 2014;9:2843–51.
32. Jensen CN, Ali ST, Allen MJ, Grogan G. Exploring nicotinamide cofactor promiscuity in NAD(P)H-dependent flavin containing monooxygenases (FMOs) using natural variation within the phosphate binding loop. Structure and activity of FMOs from *Cellvibrio* sp. BR and *Pseudomonas stutzeri* NF13. J Mol Catal B Enzym. 2014;109:191–8.
33. Montersino S, van Berkel WJH. Functional annotation and characterization of 3-hydroxybenzoate 6-hydroxylase from *Rhodococcus jostii* RHA1. Biochim Biophys Acta. 1824;2012:433–42.
34. Bernhardt R, Urlacher VB. Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. Appl Microbiol Biotechnol. 2014;98:6185–203.
35. Uetz T, Schneider R, Mario S, Egli T. Purification and characterization of a two-component monooxygenase that hydroxylates nitrilotriacetate from "*Chelatobacter*" strain ATCC 29600. J Bacteriol. 1992;174:1179–88.
36. Eichhorn E, Davey CA, Sargent DF, Leisinger T, Richmond TJ. Crystal structure of *Escherichia coli* alkanesulfonate monooxygenase SsuD. J Mol Biol. 2002;324:457–68.
37. Gin P, Hsu AY, Rothman SC, Jonassen T, Lee PT, Tzagoloff A, Clarke CF. The *Saccharomyces cerevisiae* COQ6 gene encodes a mitochondrial flavin-dependent monooxygenase required for coenzyme Q biosynthesis. J Biol Chem. 2003;278:5308–25316.
38. Kim D, Chae JC, Zylstra GJ, Sohn HY, Kwon GS, Kim E. Identification of two-component regulatory genes involved in *o*-xylene degradation by *Rhodococcus* sp. strain DK17. J Microbiol. 2005;43:49–53.
39. Sharp JO, Sales CM, LeBlanc JC, Liu J, Wood TK, Eltis LD, et al. An inducible propane monooxygenase is responsible for *N*-nitrosodimethylamine degradation by *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol. 2007; 73:6930–8.
40. Patrauchan MA, Florizone C, Eapen S, Gòmez-Gil L, Sethuaman B, Fukuda M, et al. Roles of ring-hydroxylating dioxygenases in styrene and benzene catabolism in *Rhodococcus jostii* RHA1. J Bacteriol. 2007;190:37–47.
41. Pathak A, Chauhan A, Blom J, Indest KJ, Jung CM, Stothard P, et al. Comparative genomics and metabolic analysis reveals peculiar characteristics of *Rhodococcus opacus* strain M213 particularly for naphthalene degradation. PLoS One. 2016;17:1–32.
42. Iino T, Wang Y, Miyauchi K, Kasai D, Masai E, Fujii T, et al. Specific gene responses of *Rhodococcus jostii* RHA1 during growth in soil. Appl Environ Microbiol. 2012;78:6954–62.
43. Maniatis T, Fritsch EF, Sambrook J. Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory, Cold Spring Harbor; 1982.
44. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. BMC Genomics. 2008;9:75.
45. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
46. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22:4673–80.
47. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30:2725–9.
48. Sallam KI, Mitani Y, Tamura T. Construction of random transposition mutagenesis system in *Rhodococcus erythropolis* using IS1415. J Biotechnol. 2006;121:13–22.

49. Treadway SL, Yanagimachi KS, Lankenau E, Lessard PA, Stephanopoulos G, Sinskey AJ. Isolation and characterization of indene bioconversion genes from *Rhodococcus* strain I24. Appl Microbiol Biotechnol. 1999;51:786–93.
50. Pilhofer M, Bauer AP, Schrallhammer M, Richter L, Ludwig W, Schleifer K-H, et al. Characterization of bacterial operons consisting of two tubulins and a kinesin-like gene by the novel two-step gene walking method. Nucleic Acids Res. 2007; https://doi.org/10.1093/nar/gkm836.
51. Nakashima N, Tamura T. A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35°C. Biotechnol Bioeng. 2004;86:136–48.
52. Sambrook J, Russell DW. Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory, Cold Spring Harbor; 1989.

Applied Microbiology and Biotechnology
https://doi.org/10.1007/s00253-018-9539-7

MINI-REVIEW

Genome analysis and -omics approaches provide new insights into the biodegradation potential of *Rhodococcus*

Jessica Zampolli[1] · Zahraa Zeaiter[1] · Alessandra Di Canito[1] · Patrizia Di Gennaro[1]

Received: 5 October 2018 / Revised: 21 November 2018 / Accepted: 22 November 2018

Abstract

The past few years observed a breakthrough of genome sequences of bacteria of *Rhodococcus* genus with significant biodegradation abilities. Invaluable knowledge from genome data and their functional analysis can be applied to develop and design strategies for attenuating damages caused by hydrocarbon contamination. With the advent of high-throughput -omic technologies, it is currently possible to utilize the functional properties of diverse catabolic genes, analyze an entire system at the level of molecule (DNA, RNA, protein, and metabolite), simultaneously predict and construct catabolic degradation pathways. In this review, the genes involved in the biodegradation of hydrocarbons and several emerging plasticizer compounds in *Rhodococcus* strains are described in detail (aliphatic, aromatics, PAH, phthalate, polyethylene, and polyisoprene). The metabolic biodegradation networks predicted from omics-derived data along with the catabolic enzymes exploited in diverse biotechnological and bioremediation applications are characterized.

Keywords *Rhodococcus* · -omics · Biodegradation · Recalcitrant compounds · Gene cluster

Introduction

Rhodococcus genus is a heterogeneous group of microorganisms taxonomically associated with the *Actinobacteria* phylum, one of the largest within the bacteria domain (Ludwig et al. 2012). This phylum represents Gram-positive bacteria with high G-C content, often widespread in aquatic and terrestrial ecosystems. Bacteria of this group exhibit remarkable catabolic versatility and an array of unique enzymatic capabilities that reveal their environmental and biotechnological importance (Barka et al. 2016). For instance, they can degrade a large number of organic and xenobiotic compounds often toxic and recalcitrant categorized in different groups such as aliphatic, aromatic, and polyaromatic hydrocarbons and heterocyclic, halogenated, and nitro-substituted compounds (Martínková et al. 2009). Additional attributes of rhodococci are the modification of steroid, the enantio-selective synthesis, the production of amides from nitriles, the conversion of many plant secondary metabolites found in soil and rhizosphere, such as alkaloids, terpenes, and sterols, and the elimination of sulfur from coal and petroleum products (Larkin et al. 2006; Van Der Geize and Dijkhuizen 2004).

The genus of *Rhodococcus* represents an important reservoir of physiological and functional diversity. Their metabolites have a biotechnological and industrial significance. Indeed, they can produce carotenoids, wax esters, oils, biosurfactants, bioflocculation agents, and acrylamide (Jones and Goodfellow 2010). For instance, *Rhodococcus opacus* PD630 was shown to accumulate lipids in its cytosol due to its natural ability to tolerate and to utilize phenolic compounds, which emerges this strain as a promising microbial host for lignocellulose conversion into value-added products (Yoneda et al. 2016).

Moreover, rhodococci were characterized for their ability to thrive harsh environmental conditions and for their biodegradative potential with respect to very toxic and recalcitrant compounds. In this regard, *Rhodococcus jostii* RHA1 was assessed for its resistance to desiccation and nutrient starvation (LeBlanc et al. 2008; Patrauchan et al. 2012) by evaluating genes' expression through a proteomic approach while *Rhodococcus erythropolis* DCL14 was studied for its ability to tolerate various concentrations of water miscible (ethanol, butanol, and dimethylformamide, up to 50% *v*/*v*) and water immiscible solvents (dodecane, bis(2-ethylhexyl) phthalate, and toluene, up to 5% *v*/*v*) (De Carvalho et al. 2004).

✉ Patrizia Di Gennaro
patrizia.digennaro@unimib.it

[1] Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milan, Italy

Published online: 15 December 2018

The recent awareness that *Rhodococcus* genus represents a promising potential in bioremediation, biotransformations, and biocatalysis applications is prompting research and interest from the scientific community in many countries of the world. Indeed, the number of publications and patents on rhodococci has increased during the last several years; in particular, research on the biotechnological exploitation of rhodococci has intensified significantly within the last 37 years, as illustrated by citation analysis (Fig. 1). Genome sequencing is the milestone in unveiling the bioremediation and biotechnological potential in bacteria. In fact, since the advent of genome sequencing two decades ago (Fleischmann et al. 1995), the technical and biological knowledge have increased. Nowadays, sequencing of bacterial genomes is a standard procedure, and the information from tens of thousands of bacterial genomes has had a major impact on our views of the bacterial world (Land et al. 2015) (Table 1). In regard to *Rhodococcus* genus, several public and private genome projects involving *Rhodococcus* members are now in progress due to their genome complexity, demonstrated by a multiplicity of catabolic genes and a high genetic redundancy of biosynthetic pathways often connected to a sophisticated regulatory network. Thus, combining functional genome studies with biochemical and physiological knowledge could enhance the exploitation of rhodococcal biotechnological use (Alvarez 2010).

This review intends to report new insights into the biodegradation potential of *Rhodococcus* genus based on genome analyses and "-omics" approaches. It summarizes the main characteristics of members of *Rhodococcus* genus and presents a collection of catabolic functions of different *Rhodococcus* species involved in the degradation and/or use of several toxic compounds, from hydrocarbons to several emerging contaminants as plasticizer compounds.

Genomics of *Rhodococcus* and "-omics studies"

This review presents an update collection of *Rhodococcus* genes and metabolic pathways involved in the biodegradation of main group of toxic compounds basing on "-omic" approaches. The suffix "-omics" mainly indicates studies undertaken on a genome-wide scale, including the genome itself (genomic), RNA transcription (transcriptomic), protein (proteomic), and metabolic products (metabolomic), relying on bioinformatic data.

Particularly, bacterial genomic information could be exploited on at least two levels, to (i) elucidate gene function of unknown enzymes and to (ii) understand the metabolic network of strains endowed with a broad catabolic diversity (Vilchez-Vargas et al. 2010). Hence, genomic analyses constitute an immense source for discovering and exploiting novel biocatalysts.

To date, among the bacterial genome completely sequenced and listed at https://www.ncbi.nlm.nih.gov/genome/browse/?report=2, 236 genome sequences belong to *Rhodococcus* sp. strains, of which 82 are fully sequenced. Accordingly, almost 80 publications on genome sequencing projects of *Rhodococcus* strains were released in the last four years in eminent journals, *Genome Announcement*. Furthermore, transcriptomic, proteomic, and metabolomic projects are in continuous increase.

Omics era elucidated the biotechnological potential of *Rhodococcus* genus in bioremediation and ecological applications. Since the first completely sequenced genome of *R. jostii* RHA1, *Rhodococcus* spp. showed to possess complex genomes that are among the largest available sequenced genomes in prokaryotes. In particular, RHA1 9.7 Mb genome is arranged in four linear replicons: one chromosome and three linear megaplasmids pRHL1, pRHL2, and pRHL3 (McLeod et al. 2006).

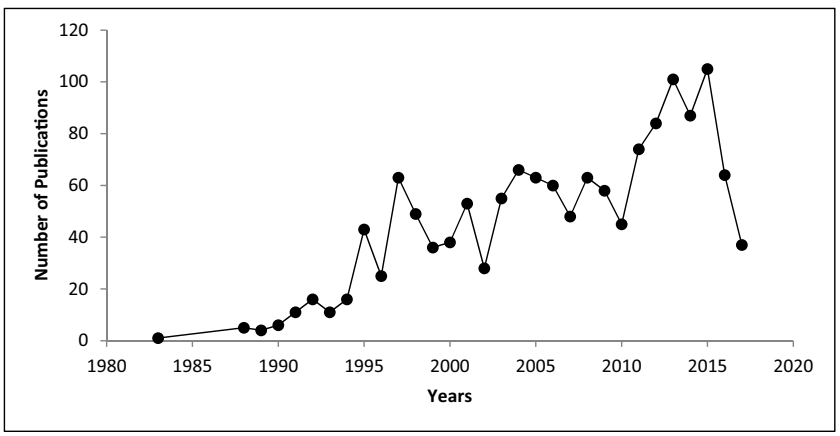


Fig. 1 Number of publications on *Rhodococcus* strains

Table 1 Omics approaches and biodegradative capabilities of different organic compounds reported in members of *Rhodococcus* genus

Compound	Strain	Type of -omic approach	Reference
Aliphatic hydrocarbons			
n-Hexadecane Commercial diesel oil Artificial mixture	*R. erythropolis* PR4	Genome comparative analysis and transcriptomic analysis	Sekine et al. 2006; Laczi et al. 2015
Medium- and long-chain alkanes	*R. aetherivorans* BCP1 and *R. opacus* R7	Genome sequencing and comparative analyses	Orro et al. 2015
Short-chain alkanes	*R. aetherivorans* BCP1	Transcriptional analysis and proteomic approach	Cappelletti et al. 2015
Sterol	*R. jostii* RHA1	Transcriptomic analysis	Rosłoniec et al. 2013; Kulig et al. 2015
Monoaromatic hydrocarbons			
BTEX	*R. aetherivorans* BCP1 and *R. opacus* R7	Genome sequencing and comparative analyses	Orro et al. 2015
	R. jostii DK17	Genome sequencing and comparative analysis	Yoo et al. 2012; Kim et al. 2002
o-Xylene	*R. opacus* R7	Genome analysis	Di Canito et al. 2018
Phenol	*R. opacus* PD630	Transcriptomic analysis (RNA-Seq)	Yoneda et al. 2016
2,4-DNT and 2,6-DNT	*R. pyridinivorans* NT2	Proteomic analysis	Kundu et al. 2016
Polycyclic aromatic hydrocarbons			
PCBs	*R. jostii* RHA1	Genomic and catabolic reconstruction; microarray hybridization	McLeod et al. 2006; Goncalves et al. 2006
Naphthalene	*R. opacus* M213	Genomic and metabolite analysis	Pathak et al. 2016
PAH	*R. aetherivorans* BCP1 and *R. opacus* R7	Genome sequencing and comparative analyses	Orro et al. 2015
Plasticizer compounds			
Phthalates	*R. jostii* RHA1	Transcriptomic	Hara et al. 2007
	R. jostii DK17		Choi et al. 2005
Isoprene	*R.* sp. AD45	Genomic	Crombie et al. 2015
Polyethylene	*R. ruber* C209	Transcriptomic	Gravouil et al. 2017
DTDB	*R. erythropolis* MI2	Genomic and proteomic	Khairy et al. 2016

Other *Rhodococcus* strains including *Rhodococcus opacus* R7 (Di Gennaro et al. 2014), *Rhodococcus opacus* M213 (Pathak et al. 2016), and *R. opacus* PD630 (Holder et al. 2011) have been also described for their large genomes (10.1, 9.19, and 9.17 Mbp, respectively) and their numerous plasmids (5 in R7, 2 in M213, and even 9 in PD630).

Moreover, it has been reported that a significant component (6%) of the biotransformations with potential biotechnological application (primarily xenobiotic reactions) recorded in the Minnesota biocatalysis/biodegradation database (http://http://umbbd.msi.umn.edu/) are assigned to *Rhodococcus* spp. This biotransformation activity in *Rhodococcus* genus is second only to *Pseudomonas* and it is achieved through numerous catabolic pathways.

In fact, a large number of enzymatic classes have been predicted to be involved in the degradation of aromatic compounds in RHA1; indeed genomic analyses revealed 203 oxygenases, of which 86 dioxygenases and 88 putative flavoprotein monooxygenases. Additionally, 50 hydroxylases, of which 28 putative cytochrome P450 hydroxylases putatively involved in aromatic and steroids compounds, have been also reported (McLeod et al. 2006).

The immense catabolic diversity shown by the members of this genus lies hidden in the multiplicity of catabolic genes, the high genetic redundancy of biosynthetic pathways, and the sophisticated regulatory networks of their genomes (Alvarez 2010). In this context, *Rhodococcus erythropolis* PR4 and *Rhodococcus* sp. DK17 showed a broad degradative abilities towards aliphatic and monoaromatic hydrocarbons, respectively (Laczi et al. 2015; Kim et al. 2002), while *R. opacus* R7 is capable of degrading the latter two as well as polycyclic aromatic hydrocarbons (Di Gennaro et al. 2010; Zampolli et al. 2014; Orro et al. 2015; Di Canito et al. 2018). Recently, -omic studies ascribed other metabolic traits to *Rhodococcus* members, such as the degradation of PCBs and steroids (chlorate and cholesterol) in *R. jostii* RHA1,

R. erythropolis PR4, *Rhodococcus opacus* B4, and *Rhodococcus aetherivorans* I24 (Goncalves et al. 2006; Swain et al. 2012; Alvarez 2010; Puglisi et al. 2010), the metabolism of isoprene (Crombie et al. 2015), short-chain alkanes, and chloroform (Ciavarelli et al. 2012; Cappelletti et al. 2015), and herbicides (Fang et al. 2016) in *Rhodococcus* sp. AD45, *Rhodococcus aetherivorans* BCP1, and *Rhodococcus* sp. MET, respectively.

Besides, intriguing degradative capabilities towards natural and synthetic polymers, organic sulfur compound as 4,4-dithiodibutyric acid (DTDB), and biodesulfurization of petroleum oil have been reported in *Rhodococcus rhodochrous* RPK1, *Rhodococcus erythropolis* MI2, and *Rhodococcus erythropolis* XP (Watcharakul et al. 2016; Gravouil et al. 2017; Khairy et al. 2016; Tao et al. 2011).

The huge repertoire of catabolic abilities of this genus could be also explained by horizontal gene transfer and gene duplication phenomena. In particular, the high frequency of recombination (homologous and illegitimate recombination) may trigger the flexibility of *Rhodococcus* genomes to easily acquire new functions (Larkin et al. 2006; Larkin et al. 2010). Indeed, catabolic genes, often identified on linear plasmids, have been found to contribute to degradation pathways together with genes located on the chromosome (Goncalves et al. 2006).

It is worth to mention that many *Rhodococcus* functional traits were unveiled combining omics data with genome comparative approaches providing insights in terms of phenotypes, metabolic capacities, and cellular response to different stress conditions. This has also contributed to elucidate phylogeny and evolutionary concepts and their relation.

Biodegradation of hydrocarbon compounds

Hydrocarbon compounds are widespread in the environment and originate from natural sources and anthropogenic activities. Large amounts of aromatic compounds were derived from decaying plant material (e.g., from lignin), soil weathering processes, and volcano emissions. Moreover, anthropogenic sources, in particular spillage of petroleum products, discharge of industrial effluents, and transport accidents, release significant amount of hydrocarbons into the environment (Martínková et al. 2009).

Mechanisms underpinning biodegradation of hydrocarbons have been described in several bacterial strains (Pérez-Pantoja et al. 2009; Juarkar et al. 2010). In particular, the aerobic biodegradation of hydrocarbon compounds has been well elucidated in several *Rhodococcus* strains (Field and Sierra-Alvarez 2004; Alvarez 2010). It was well established that aromatic compounds are catabolized through many upper and/or lower pathways leading to few central intermediates (catechol, protocatechuate, gentisate) (Fig. 2). The generated metabolites are then degraded to compounds involved in the TCA cycle through central pathways. The biodegradation of each class of hydrocarbons requires a specific enzyme due to the high diversity of the molecular structures of these compounds (Abbasian et al. 2016).

Aliphatic hydrocarbons

Rhodococcus strains are able to oxidize *n*-alkanes. Basically, the first step of oxidation is mainly performed by an alkane1-monooxygenase encoded by the *alkB* gene. This gene has been identified in well-known members of the *Rhodococcus* genus (e.g., *R. erythropolis*, *R. ruber*, *R. opacus*, *R. equi*, and *R. jostii*) and can be employed as a marker of this metabolism (Táncsics et al. 2014).

A multiplicity of *alkB* is reported in several *Rhodococcus* strains; indeed, it was identified at least four homologous alkane monooxygenase genes (*alkB1*, *alkB2*, *alkB3*, and *alkB4*) in two different rhodococcal strains (*Rhodococcus* strains Q15 and NRRLB-16531) (Whyte et al. 2002) and two different alkane hydroxylase systems (*alkBa* and *alkBb*) in *Rhodococcus ruber* SP2B (Amouric et al. 2010). Although the multiplicity of *alkB* in the genome of *Rhodococcus* spp. is frequently observed, this phenomenon is not universal for all members of the genus (Táncsics et al. 2014). Indeed, only one copy of *alkB* gene was identified in the genomes of *R. jostii* RHA1, (McLeod et al. 2006), *R. opacus* B4 (Sameshima et al. 2008), *R. aetherinovorans* BCP1 (Cappelletti et al. 2011; Cappelletti et al. 2013), and *R. opacus* R7 (Zampolli et al. 2014).

In addition, other alkane hydroxylase systems have been described in *Rhodococcus*, including (i) P450 cytochrome alkane monooxygenase systems (CYP153) that hydroxylate C_5–C_{16} *n*-alkanes (van Beilen et al. 2006) and (ii) short-alkane monooxygenases that are involved in the oxidation of C_1–C_6 *n*-alkanes (Cappelletti et al. 2011). Cytochromes P450 (CYPs) are hemoproteins that oxidize a large number of chemical compounds, such as xenobiotics, antibiotics, steroids, terpenes, alkanes, fatty acids, and alkaloids in aerobiosis (Bernhardt and Urlacher 2014; Larkin et al. 2006). In fact, *Actinobacteria* display several P450 cytochromes distributed over 220 families. For instance, 25, 26, and 45 P450 cytochromes were reported in *R. jostii* RHA1 (McLeod et al. 2006), *R. opacus* R7 (Di Gennaro et al. 2014), and *R. aetherinovorans* BCP1, respectively (Cappelletti et al. 2013). Moreover, a *cyp125* gene was identified in *R. jostii* RHA1 genome. This gene is highly upregulated during growth on cholesterol and encodes a CYP125 that catalyzes the hydroxylation of the C26 atom of sterols through side-chain oxidation leading to the C26-carboxylic acid formation (Rosłoniec et al. 2013). Transcriptomic analysis of this strain suggested that the CYP257A1 (novel cytochrome P450 enzyme family) (Rosłoniec et al. 2009) is involved in sterol

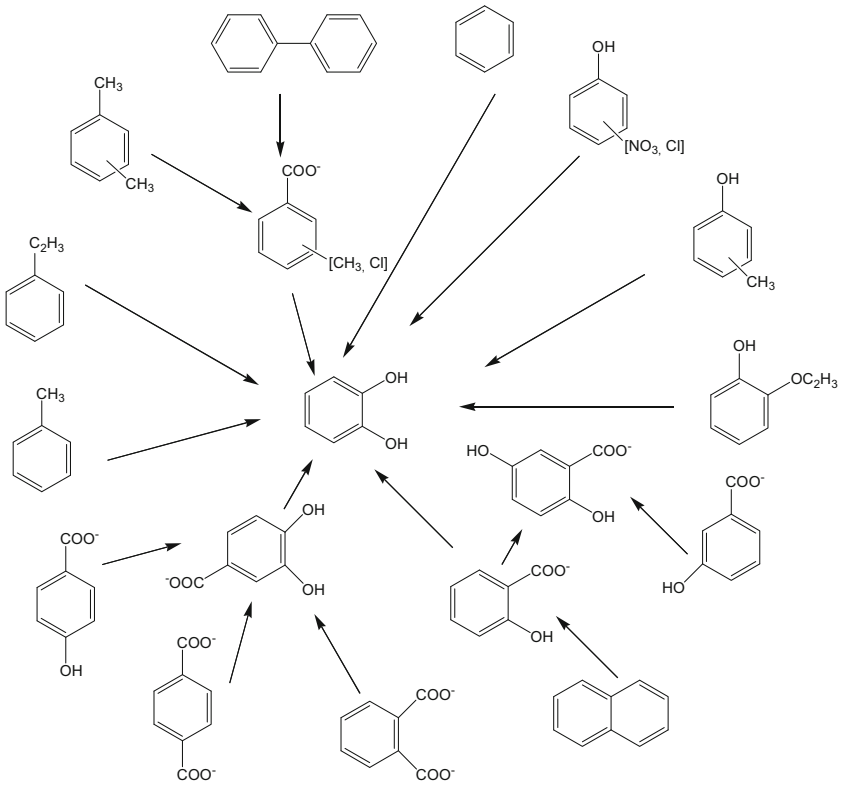


Fig. 2 Scheme of biodegradation pathways for aromatic hydrocarbon compounds in *Rhodococcus*

metabolism, while protein analysis showed that it *N*-demethylates the alkaloid substrate, i.e., dextromethorphan (Kulig et al. 2015). Another cytochrome P450 able to hydroxylate octane was described in *Corynebacterium* sp. 7E1C (in recent literature known as *Rhodococcus rhodochrous*) (Cardini and Jurtshuk 1970).

Besides, aerobic bacterial metabolism of gaseous and short-chain *n*-alkanes (ranged from ethane to *n*-butane) was reported in *Rhodococcus* genus. For instance, *R. aetherivorans* BCP1 was described for its ability to grow on gaseous *n*-alkanes as inducing condition for the co-metabolism with low chlorinated compounds. This strain presents two *prm* gene clusters: the first, *prm* gene cluster (*prmA,C,B,D* genes), is located on the chromosome and implicated in the degradation of propane and butane, and the second, *smo* gene cluster (soluble di-iron monooxygenase) (*smoA,B,D,C* genes), is positioned on the pBMC2 plasmid and involved in the degradation of short-chain *n*-alkanes (Cappelletti et al. 2015).

Aromatic hydrocarbons

The degradation of aromatic hydrocarbons employs the conversion of such substrates into a number of central intermediates, which are consequently subjected to ring cleavage. The reactions of the pathways consist in activation of the aromatic ring through oxygenases generating di- or hydroxylated intermediates. The subsequent ring cleavage of these intermediates can be catalyzed by either intradiol or extradiol dioxygenases (Martínková et al. 2009).

The BTEX system, comprising benzene, toluene, ethylbenzene, and xylene isomers, is listed as priority pollutants. Among BTEX-degrading bacteria, *Rhodococcus* sp. strain DK17, capable of degrading *o*-xylene and toluene, has been studied (Kim et al. 2002, 2010). The sequenced genome of this strain revealed that *akb* gene cluster is involved in the metabolism of these hydrocarbons. This cluster is composed of two genes encoding for the large subunit of an oxygenase (designated *akbA1a* and *akbA1b*) followed respectively by other two genes encoding for the small subunit (designated *akbA2a* and *akbA2b*). Within this cluster, three other genes *akbA3*, *akbA4*, and *akbB* encoding respectively for a reductase (AkbA4), a ferredoxin component (AkbA3), and a dehydrogenase (AkbB) were identified. Besides, in the same genomic region, *akbCDEF* genes encoding proteins involved in the ring cleavage (lower pathway) were reported. The latter genes are capable to hydroxylate aromatic compounds (i.e., ethylbenzene) to 2,3- and 3,4-*cis*-dihydrodiols and able to perform benzylic and aryl hydroxylations on *m*-xylene (Kim et al. 2002).

Another studied BTEX-degrading bacteria is *R. jostii* RHA1 (Seto et al. 1995). This strain efficiently assimilates ethylbenzene (ETB), isopropylbenzene (IPB), and biphenyl (BPH) via a common pathway. Transcriptomic analysis revealed that three gene clusters *bph*, *etb1*, and *etb2*, identified on large linear plasmids, are induced by BPH, ETB, and IPB,

which indicates that all these hydrocarbons are catabolized by the same enzymes and the genes are controlled by a common regulatory system.

Basically, these compounds are converted to benzoate and 2-hydroxypenta-2,4-dienoate (HPD) in four successive steps catalyzed by (i) four-component biphenyl dioxygenase (ferredoxin reductase, ferredoxin, and two-subunit terminal oxygenase coded by *bphA1A2A3A4*, *etbAa1Ab1*, and *etbAa2Ab2Ac*; (ii) dihydrodiol dehydrogenase (*bphB1* and *bphB2*); (iii) 2,3-dihydroxybiphenyl dioxygenase (*bphC1* and *etbC*); and finally 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoatehydrolase (*bphD1*). ETB degradation results in propionate and HPD. HPD is further metabolized to pyruvate and acetyl coenzyme A (CoA) through the lower biphenyl pathway divided into three steps (encoded by *bphGFE* in *R. jostii* RHA1) (Martínková et al. 2009; Goncalves et al. 2006) while benzoate is metabolized to succinate and acetyl-CoA through the benzoate pathway.

It was found that *bphA1A2A3A4* and *etbAa1Ab1* genes are under the control of a two-component system encoded by *bphS* and *bphT* genes, which encode the sensor histidine kinase and response regulator, respectively. Indeed, the transcriptional induction of the *bphA1* promoter by biphenyl, benzene, alkylbenzenes, and chlorinated benzenes requires *bphS1T1* genes. In a further study, Takeda et al. showed that the inducing-substrate spectrum of BphS1 includes substrates of BphS2. However, the only difference between *bphS1T1* and *bphS2T2* systems is that BPH is an inducing substrate only for the *bphS1T1* (Takeda et al. 2004). Therefore, they suggested that in the presence of a single-ring aromatic compound, such as ethylbenzene, both *bphS1* and *bphS2* are responsible for the transcriptional activation of degradation genes in RHA1.

Recently, a particular BTEX degradation pathway was found in *R. opacus* R7. This strain was found to metabolize *o*-xylene with the same dioxygenase system previously identified in DK17 (*akb*) and in RHA1 (*etb*) strains (allocated on two plasmids, pPDG5 and pPDG2). Furthermore, R7 was shown to possess numerous monooxygenases/phenol hydroxylases included in the *o*-xylene degradation pathway, highlighting the redundancy of oxygenases genes in this strain (Di Canito et al. 2018).

Phenols

Phenolic compounds (e.g., cresols), released and accumulated in the environment, pose serious health hazards to living organisms (Michałowicz and Duda 2007).

A wide range of phenols are degraded by *Rhodococcus* strains. The degradation begins by the conversion of phenol compound to catechol by phenol hydroxylase, which is further metabolized through *ortho*- or *meta*-cleavage. Basically, catechol 1,2-dioxygenase, an *ortho*-cleaving enzyme, and catechol 2,3-dioxygenase, a *meta*-cleaving enzyme, are proteins involved in the β-ketoadipate pathway in which the bacterial degradation of catechol to central metabolic intermediates occurs (Guzik et al. 2011; Szőköl et al. 2014; Zídková et al. 2013). In *Rhodococcus erythropolis* CCM2595, the phenol hydroxylase enzyme was described; it consists of two subunits encoded by the *pheA1* and *pheA2* genes clustered with the *pheR* gene, coding for an AraC-type transcriptional regulator (activator) (Zídková et al. 2013). These genes were found to be induced by phenol and other aromatic compounds such as *p*-chlorophenol, *p*-nitrophenol, resorcinol, and *p*-cresol (Fialova et al. 2003). Moreover, this gene cluster was found in *R. jostii* RHA1 in two copies, on the chromosome and on the plasmid pRHL1 (McLeod et al. 2006).

Other *Rhodococcus* strains with phenol degradation ability were reported. For instance, *Rhodococcus erythropolis* UPV-1 is efficiently capable of degrading PAHs, phenol, and a mixture of *o*-, *m*-, and *p*-cresols (Irvine et al. 2000). In addition, *Rhodococcus opacus* 1CP is able of degrading *p*-cresols through 4-methylcatechol and 3-methyl-*cis*,*cis*-muconate via *ortho*-pathway. This strain possesses also another catechol-1,2dioxygenase activity, in addition to dioxygenases with specificities for 4-chlorophenol and *p*-toluate (Kolomytseva et al. 2007).

Interestingly, an important biotechnological trait attributed to *Rhodococcus* genus is the capability of accumulating lipids to high extent in their cytosol, up to 80% in *R. opacus* PD630 (Holder et al. 2011). Recent study employing omics techniques has shown a positive correlation between phenolic compound degradation and lipid production in *Rhodococcus rhodochrous*. This could explain the capacity of *Rhodococcus* strains to utilize recalcitrant compounds including complex polyaromatic structure (Shields-Menard et al. 2017).

Polyaromatic hydrocarbons

A wide range of PAH compounds are catabolized by bacteria of *Rhodococcus* genus, i.e., fluoranthene, pyrene, and chrysene (Xu-Xiang et al. 2006). Basically, PAH degradation initiates in the cytosol by the action of intracellular dioxygenases. Subsequently, PAHs are oxidized to *cis*-dihydrodiols, which in turn are re-oxidized to aromatic dihydroxy compounds (catechols) and eventually channeled through the *ortho* or *meta*-cleavage pathways (Cerniglia 1984; Smith 1990). PAH metabolic pathway has been investigated in many *Rhodococcus* strains by growing the bacterium on naphthalene (the simplest and most soluble PAH). The latter is considered a model compound to investigate the ability of bacteria to degrade PAHs. For instance, *R. opacus* R7 degrades naphthalene through the dioxygenation of the aromatic ring, via 1,2-dihydro-1,2-dihydroxynaphthalene, that is further oxidized to salicylate and gentisate (Di Gennaro et al. 2010). In this pathway, two genes, *narAa* and *narAb*, encoding

respectively for the large (NarAa) and the small (NarAb) components of the naphthalene dioxygenase are involved. These genes are clustered with *rub1*, *rub2*, *rub1bis*, *narB*, and *orf* (*orf7*) genes, encoding for three rubredoxins, a naphthalene dihydrodiol dehydrogenase, and a protein of unknown function, respectively. In addition, two regulatory genes (*narR1*, *narR2*) were described in R7 encoding for putative regulatory proteins NarR1 and NarR2, belonging to GntR and NtrC families, accordingly. Nonetheless, no LysR-type regulatory gene was reported. The final aromatic ring cleavage in this strain proceeds through the gentisate pathway, encoded by *gen* and *sal* genes (Di Gennaro et al. 2010). Moreover, the lower pathway of naphthalene degradation was found in two copies, the first is on pPDG4 plasmid and the second on pPDG1 plasmid, which lacks *genL* gene (Orro et al. 2015). Besides, salicylate metabolism in R7 involves *genC*, *genB*, and *genA* genes located upstream the genes associated with gentisate degradation. In particular, *genA*, *genB*, *genC*, *genH*, *genI*, and *genL* encode for salicylate CoA ligase, salicylate CoA synthetase, salicylate hydroxylase, gentisate 1,2-dioxygenase, 3-maleylpyruvate isomerase, and a protein of unknown function, respectively.

Naphthalene degradation pathway has been also described in *Rhodococcus* sp. NCIMB12038. In particular, *nar* region is involved in this degradation; it is composed of *narA* and *narB* genes which are transcribed in several units. Contrariwise, these genes are transcribed as single units in P200 and P400 strains (Kulakov et al. 2005).

Other genetic determinants that likely participate in naphthalene (and *o*-xylene) degradation were reported by Martínková et al. (2009). For instance, in *Rhodococcus opacus* TKN14 (Kulakov et al. 2000), *nid* gene cluster was identified; it is composed of *nidAB*, *nidE*, and *nidF* genes encoding for the subunits of oxygenase, rubredoxin, and auxiliary protein, respectively (Maruyama et al. 2005). *nid* genes were found to be induced by *o*-xylene and are supposed to be involved in the degradation pathways of a wide range of aromatic hydrocarbons in *Rhodococcus* species.

Thanks to -omics approach a peculiar naphthalene metabolism has been described in *R. opacus* M213, which occurs via *o*-phthalate and salicylate pathways. Conversely to naphthalene pathways previously described in other rhodococci, which occur via gentisate pathway in NCIMB 12038, B4 and R7 and through catechol pathway in P200 and P400 strains. The metabolism of naphthalene in M213 occurs via the cyclization of 2-hydroxycinnamic acid or 2-carboxycinnamic acid and eventually oxidizing into phthalic anhydride or phthalic acid. It is particularly interesting that several genes associated with this peculiar biodegradation pathway were found on genome islands (GEIs), such as small and large subunits of naphthalene dioxygenase (NDO), *cis*-naphthalene dihydrodiol dehydrogenase gene, and a putative naphthalene degradation regulatory protein (Pathak et al. 2016).

Degradation of plasticizer compounds

Plasticizers and their derivatives are among the prime examples of emerging compounds. A large amount of these compounds derives from anthropogenic sources, in particular discharge of industrial activities (Gravouil et al. 2017; Koutny et al. 2006). In this paragraph, we focused on -omic studies related to the degradation of phthalate, phthalate esters (PEs), polyisoprene, and 4,4-dithiodibutyric acid (DTDB) in *Rhodococcus* strains (Table 2).

Phthalate

After decades of global industrial use as plasticizers, phthalate esters are now recognized as ubiquitous toxicologically significant environmental pollutants (Hara et al. 2007).

Among strains of *Rhodococcus* genus, *R. jostii* RHA1 is able to grow on a variety of monoalkyl PEs, including methyl, butyl, hexyl, 2-ethylhexyl phthalates, and terephthalate (Hara et al. 2010). This strain is able to degrade phthalate (PTH), terephthalate, and 4-hydroxybenzoate via protocatechuate *ortho*-cleavage pathway (Patrauchan et al. 2005). Nevertheless, it can also degrade terephthalate via the catechol branch of the 3-oxoadipate pathway (Hara et al. 2007).

Transcriptomic analysis revealed the metabolic pathway of PTH in RHA1. Two identical functional copies of *pad* genes

Table 2 Enzymes and/or pathways involved in biodegradative metabolism of representative plasticizer compounds identified in *Rhodococcus* members

Degradation substrate	Enzyme and/or pathway	Strain name	Reference
Phthalates	*pad*, *pat* genes	*R. jostii* RHA1	Hara et al. 2007
		R. sp. DK17	Choi et al. 2005
Polyethylene	Alkane pathway (*alk*) Laccase	*R. ruber* C208	Gravouil et al. 2017
Isoprene	Isoprene pathway (*iso*)	*R.* sp. AD45	Crombie et al. 2015
Polyisoprene	*lcp* genes	*R. rhodochrous* RPK1	Watcharakul et al. 2016
DTDB	*nox* gene	*R. erythropolis* MI2	Khairy et al. 2016

carried on two large linear plasmids, pRHL1 and pRHL2, are employed. These genes encode a ring-hydroxylating 3,4-dioxygenase system (PadAaAbAcAd), a dihydrodiol dehydrogenase (PadB), a decarboxylase (PadC), and a regulatory protein (PadR). Moreover, close to *pad* genes, *pat* genes were identified, which encode an ABC transport system (PatABCD) and an ester hydrolase (PatE) (Patrauchan et al. 2012) and are highly upregulated in presence of phthalate rather than terephthalate. Interestingly, in a further study involving *patB* gene knockout, ATP-binding cassette (ABC) transporter was identified necessary for the uptake of phthalate and monoalkyl phthalate esters in RHA1. Moreover, in the same study, it was demonstrated that PatE which is a member of monoalkyl PE hydrolase family is able to specifically hydrolyze monoalkyl PEs to phthalate, rather than terephthalate or other PEs (Hara et al. 2010). In addition, other genes named *tpa* gene were identified to be putatively involved in terephthalate (TPA) uptake and degradation. The *tpa* genes encode a transport protein (TpaK), the large and small subunits plus the reductase of a ring-hydroxylating 1,2-dioxygenase system (TpaAaAbB), a dihydrodiol dehydrogenase (TpaC), and a regulatory protein (TpaR). Interestingly, the nearly identical regions of pRHL1 and pRHL2 were also identified to nearly identical regions that are duplicated in pDK2 and pDK3 in *Rhodococcus* sp. strain DK17 (Choi et al. 2005), which have also been shown to encode PTH and TPA degradation proteins.

Polyethylene

Polyethylene (PE) is a notable synthetic polymer used worldwide, particularly in the packaging industry. PE is stable and tends to accumulate in natural environment, representing by itself the majority of all plastic wastes (Gravouil et al. 2017). The degradation of PE in nature is a complex procedure in which hydrocarbon chains are oxidized into short aliphatic fragments (e.g., complex mixture of alkanes, alkenes, ketones, aldehydes, alcohols, carboxylic acids); subsequently, these fragments are mineralized by specific microorganisms from the environment (Koutny et al. 2006).

Species of *Rhodococcus* genus have been associated with PE degradation in several studies (Gravouil et al. 2017; Santo et al. 2013). For instance, Sivan and colleagues reported their ability to grow as biofilm on PE-oxidized films. Nevertheless, the molecular mechanism of PE consumption has not been revealed (Sivan et al. 2006).

Recent study based on transcriptomics and lipidomics approaches pointed out the consumption pathways of polyethylene degradation in *Rhodococcus ruber* C208. In this context, non-oxidized short-chain PE and oxidized PE samples resulting from abiotic treatment induce the pathways related to alkane degradation and β-oxidation of fatty acids in *R. ruber* C208. This observation revealed these pathways as being the bulk for the consumption of PE. The induction of these pathways could be explained by the presence of short aliphatic fragments in the samples, which could be recognized as natural inducers (Gravouil et al. 2017).

Abiotic treatment of PE (thermo and photo-oxidization) releases mixture of fragments displaying a wide range of lengths and oxidation levels. Fragments of low molecular weight are more likely to be consumed than PE films by *Rhodococcus* species.

Santo and colleagues have reported the role of laccase enzyme in catalyzing the biotic oxidation of PE and consequently in increasing the amount of metabolized short fragments. Indeed, the authors identified a copper-binding laccase in *R. ruber* C208, which is able, when overproduced, to induce a reduction of the PE Mw and Mn of 20% and 15%, respectively (Santo et al. 2013). In addition, Gravouil and colleagues identified three close homologs to laccase in the RNA library of *R. ruber* C208, but interestingly, none of them were shown to be significantly upregulated in the presence of the PE samples, regardless their oxidation levels (Gravouil et al. 2017).

Polyisoprene

Natural rubber is synthetized in large amount by many plants and is used in the production of tires, sealings, and latex gloves. The main component of rubber is the poly(*cis*1,4-isoprene) (Sharkey 1996). Nonetheless, the economic importance of rubber and the enormous amount of rubber waste materials released into the environment, only a limited number of studies have investigated its biodegradation in bacteria.

Van Hylckama Vlieg and colleagues revealed the genes involved in isoprene utilization in *R.* sp. AD45. In this strain, the consumption of isoprene includes its oxidation to epoxide, conjugation with glutathione, and dehydrogenation steps. The authors identified the presence of 10 genes, where two of them encode enzymes involved in isoprene degradation: a glutathione S-transferase with activity towards 1,2-epoxy-2-methyl-3-butene (*isoI*) and a 1-hydroxy-2-glutathionyl-2-methyl-3-butene dehydrogenase (*isoH*) (Van Hylckama Vlieg et al. 2000). Another glutathione S-transferase gene (*isoJ*) was reported. The latter acts with 1-chloro-2,4-dinitrobenzene and 3,4-dichloro-1-nitrobenzene, but not with 1,2-epoxy-2-methyl-3-butene. In addition, the authors identified downstream of *isoJ*, six genes (*isoABCDEF*) that encode for a putative alkene monooxygenase, and they figured out an amino acid sequence encoded by an additional gene (*isoG*) which shows high homology with respect to α-methylacyl-coenzyme A racemase (Van Hylckama Vlieg et al. 2000).

Subsequently, Crombie and colleagues have identified the complete set of inducible genes responsible for isoprene degradation. The authors reported the genome of *R.* sp. AD45. They found that genes involved in isoprene metabolism are concentrated in a small region on a megaplasmid, containing a relatively large number of transposases (Crombie et al. 2015). Their

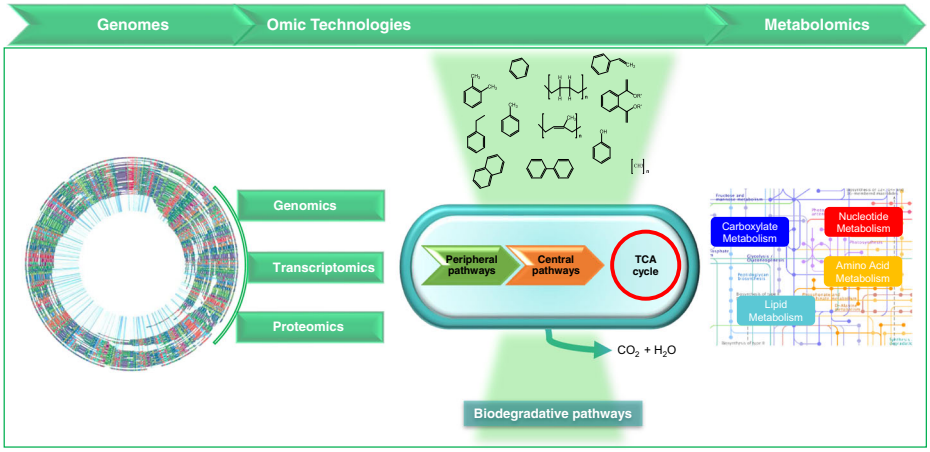


Fig. 3 Schematic representation of relationship between omics technologies and *Rhodococcus* biodegradative pathways

analyses showed high level of transcription of 22 contiguous genes when induced by isoprene or epoxyisoprene, suggesting their involvement in isoprene metabolism, including the monooxygenase, glutathione transferase, dehydrogenase (IsoH), and glutathione biosynthesis genes. In addition, the authors showed that only two aldehyde dehydrogenases, a disulfide reductase, and hypothetical protein are induced by isoprene, suggesting that this cluster may contain all the genes specific to this metabolism. Finally, they demonstrated that the previous identified genes were only induced by isoprene in a strain able to oxidize isoprene to epoxyisoprene, demonstrating that isoprene itself was not the inducing molecule. Hence, they suggested that *isoGHIJABCDEF* are co-transcribed as an operon with a promoter upstream of *isoG* (Crombie et al. 2015).

Watcharakul and colleagues identified a gene coding for the latex-clearing protein (*lcp*) in *Rhodococcus rhodochrous* strain RPK1, which is putatively involved in the initial oxidative attack on the polyisoprene polymer. This gene was also found in many *Actinobacteria* and all known rubber-degrading *Actinobacteria* for which the genome sequences are available (Nanthini et al. 2015; Rose et al. 2005). In addition, a hypothetical *lcp* gene in the genomes of *Rhodococcus rhodochrous* strain MTCC11081, *Rhodococcus* sp. MK3027, and *Rhodococcus* sp. ARG-BN062 was identified. Particularly, the predicted Lcp amino acid sequence of this gene includes the DUF2236 domain which constitutes the central part of most Lcp proteins. Moreover, by isolating and characterizing the Lcp protein of *R. rhodochrous* RPK1, the authors pointed out the presence of different Lcp subgroups which vary among them by metal ion dispositions and spectroscopic properties (Watcharakul et al. 2016).

4,4-Dithiodibutyric acid

4,4-Dithiodibutyric acid (DTDB) is a synthetic organic sulfur compound employed in diverse biochemical fields (Imbernon et al. 2015; Jang and Keng 2006). However, the biotechnological significance of this compound is its application in studies for the improvement of polythioester (PTE) synthesis. Although DTDB is non-toxic compound, its biodegradation results in the formation of very harmful metabolite: the poly(4-mercaptobutyric acid) (4MB), which is a PTE precursor substrate (Khairy et al. 2016). Hence, investigating DTDB catabolism in bacteria is necessary for the generation of metabolically engineered PTE production strains. *R. erythropolis* strain MI2 has been investigated for its rare ability to use DTDB as sole carbon source and electron donor for aerobic growth (Khairy et al. 2016). Genome and proteome analysis of this strain has clarified the DTDB degradation by elucidating the 4MB catabolism pathway. Basically, DTDB is cleaved by the action of Nox (RERY_03780) to 4MB, which is then oxidized forming 4-oxo-4-sulfanylbutyric acid via the LLMMI2 F420-dependent enzyme (RERY_05640) encoded by one of the 126 monooxygenase-encoding genes found in MI2 genome. The ultimate step is supposed to be a desulfurization reaction. Thus, 4-oxo-4-sulfanylbutyric could either be degraded by SQRMI2 (RERY_02710) or the putative desulfhydrase (RERY_06500) which forms succinic acid and volatile hydrogen sulfide. H_2S could be oxidized by SQR and then reacts with a free thiol group forming a sulfurhydrate-containing compound. The latter may be oxidized forming thiosulfite by protein encoded by RERY_02720, which is then oxidized by the putative rhodanese-related sulfurtransferase (RERY_02740), giving sulfite and succinic acid as final products. Sulfite could be oxidized to sulfate prior to be transported outside the cell through the transmembrane exporter protein (RERY_02750), while succinic acid is most probably utilized for growth (Khairy et al. 2016). More studies are needed to verify the postulated enzymatic reactions.

Conclusions

In the last decades, -omics data have increased our understanding of the biological aspects, genetics, and metabolic pathways that could be exploited for biotechnological and bioremediation

applications (Khairy et al. 2016; Juwarkar et al. 2010). Traditional approaches of genetics and biochemistry have been applied to *Rhodococcus* bacteria. Nonetheless, understanding the mechanisms and pathways of hydrocarbons biodegradation have been limited. Nowadays, omics techniques with advanced high-throughput analytical technologies have a special impact in the wake of predicting and constructing metabolic pathways and since the possibility to exploit constructed pathways and gene functions in biodegradation and bacterial biotechnological potential (Fig. 3) (Pathak et al. 2016; Yoneda et al. 2016). With this aim, combination of molecular tools and markers could be useful in bioremediation processes.

Recent achievements of combining genomic, proteomic, and bioinformatics approaches provided insights not only on gene functions, degradation pathways, and molecular mechanisms, but also in the genome-wide gene expression of *Rhodococcus* bacteria in various environmental conditions (Kim et al. 2018). However, further researches (i.e., cell physiology and regulatory studies, environmental stress response) are required in the field of hydrocarbon degradation in *Rhodococcus*. The analysis of generated data from currently available sequenced genomes along with genetics and molecular biology approaches will allow to investigate in depth both biodegradation properties and biotechnological potential of these bacteria, which have been recognized as a untapped source of genetic and functional diversity.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in this study were in compliance with ethical standards. This article does not contain any studies with human participants or animals performed by any of the authors.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Abbasian F, Palanisami T, Megharaj M, Naidu R, Lockington R, Ramadass K (2016) Microbial diversity and hydrocarbon degrading gene capacity of a crude oil field soil as determined by metagenomics analysis. Biotechnol Prog 32:638–648

Alvarez HM (2010) Central metabolism of the species of the genus *Rhodococcus*. In: Alvarez HM (ed) Biology of *Rhodococcus*. Springer-Verlag, Berlin Heidelber, pp 91–108

Amouric A, Quéméneur M, Grossi V, Liebgott PP, Auria R, Casalot L (2010) Identification of different alkane hydroxylase systems in *Rhodococcus ruber* strain SP2B, an hexane-degrading actinomycete. J Appl Microbiol 108:1903–1916

Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Klenk H-P, Clément C, Ouhdouch Y, van Wezel GP (2016) Taxonomy, physiology, and natural products of *Actinobacteria*. Microbiol Mol Biol Rev 80:1–43

Bernhardt R, Urlacher VB (2014) Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations. Appl Microbiol Biotechnol 98:6185–6203

Cappelletti M, Fedi S, Frascari D, Ohtake H, Turner RJ, Zannoni D (2011) Analyses of both the *alkB* gene transcriptional start site and *alkB* promoter-inducing properties of *Rhodococcus* sp. strain BCP1 grown on n-alkanes. Appl Environ Microbiol 77:1619–1627

Cappelletti M, Di Gennaro P, D'Ursi P, Orro A, Mezzelani A, Landini M, Fedi S, Frascari D, Presentato A, Zannoni D, Milanesi L (2013) Genome sequence of *Rhodococcus* sp. strain BCP1, a biodegrader of alkanes and chlorinated compounds. Genome Announc 1: e00657–e00613

Cappelletti M, Presentato A, Milazzo G, Turner RJ, Fedi S, Frascari D, Zannoni D (2015) Growth of *Rhodococcus* sp. strain BCP1 on gaseous n-alkanes: new metabolic insights and transcriptional analysis of two soluble di-iron monooxygenase genes. Front Microbiol https://doi.org/10.3389/fmicb.2015.00393

Cardini G, Jurtshuk P (1970) The enzymatic hydroxylation of *n*-octane by *Corynebacterium* sp. strain 7E1C. J Biol Chem 245:2789–2796

Cerniglia CE (1984) Microbial metabolism of polycyclic aromatic hydrocarbons. Adv Appl Microbiol 30:31–71

Choi KY, Kim D, Sul WJ, Chae J-C, Zylstra GJ, Kim YM, Kim E (2005) Molecular and biochemical analysis of phthalate and terephthalate degradation by *Rhodococcus* sp. strain DK17. FEMS Microbiol Lett 252:207–213

Ciavarelli R, Cappelletti M, Fedi S, Pinelli D, Frascari D (2012) Chloroform aerobic cometabolism by butane-growing *Rhodococcus aetherivorans* BCP1 in continuous-flow biofilm reactors. Bioprocess Biosyst Eng 35:667–681

Crombie AT, El Khawand M, Rhodius VA, Fengler KA, Miller MC, Whited GM, Mcgenity TJ, Murrell JC (2015) Regulation of plasmid-encoded isoprene metabolism in *Rhodococcus*, a representative of an important link in the global isoprene cycle. Environ Microbiol 17:3314–3329

De Carvalho CCCR, Da Cruz AARL, Pons MN, Pinheiro HMRV, Cabral JMS, Da Fonseca MMR, Ferreira BS, Fernandes P (2004) *Mycobacterium* sp., *Rhodococcus erythropolis*, and *Pseudomonas putida* behavior in the presence of organic solvents. Microsc Res Tech 64:215–222

Di Canito A, Zampolli J, Orro A, D'Ursi P, Milanesi L, Sello G, Steinbüchel A, Di Gennaro P (2018) Genome-based analysis for the identification of genes involved in *o*-xylene degradation in *Rhodococcus opacus* R7. BMC Genomics 19:587

Di Gennaro P, Terreni P, Masi G, Botti S, De Ferra F, Bestetti G (2010) Identification and characterization of genes involved in naphthalene degradation in *Rhodococcus opacus* R7. Appl Microbiol Biotechnol 87:297–308

Di Gennaro P, Zampolli J, Presti I, Cappelletti M, D'Ursi P, Orro A, Mezzelani A, Milanesi L (2014) Genome sequence of *Rhodococcus opacus* strain R7, a biodegrader of mono- and polycyclic aromatic hydrocarbons. Genome Announc 2:e00827

Fang Y, Du Y, Hu L, Xu J, Long Y, Shen D (2016) Effects of sulfur-metabolizing bacterial community diversity on H_2S emission behavior in landfills with different operation modes. Biodegradation 27: 237–246

Fialova A, Cejkova A, Masak J, Jirku V (2003) Comparison of yeast (*Candida maltosa*) and bacterial (*Rhodococcus erythropolis*) phenol hydroxylase activity and its properties in the phenolic compounds biodegradation. Commun Agric Appl Biol Sci 68:155–158

Field JA, Sierra-Alvarez R (2004) Biodegradability of chlorinated solvents and related chlorinated aliphatic compounds. Rev Environ Sci Biotechnol 3:185–254

Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, McKenney K, Sutton G, FitzHugh W, Fields C, Gocayne JD, Scott J, Shirley R, Liu LI, Glodek A, Kelley JM, Weidman JF,

Phillips CA, Spriggs T, Hedblom E, Cotton MD, Utterback TR, Hanna MC, Nguyen DT, Saudek DM, Brandon RC, Fine LD, Fritchman JL, Fuhrmann JL, Geoghagen NSM, Gnehm CL, McDonald LA, Small KV, Fraser CM, Smith HO, Venter JC (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496–512

Goncalves ER, Hara H, Miyazawa D, Davies JE, Eltis LD, Mohn WW (2006) Transcriptomic assessment of isozymes in the biphenyl pathway of *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol 72: 6183–6193

Gravouil K, Ferru-Clément R, Colas S, Helye R, Kadri L, Bourdeau L, Moumen B, Mercier A, Ferreira T (2017) Transcriptomics and lipidomics of the environmental strain *Rhodococcus ruber* point out consumption pathways and potential metabolic bottlenecks for polyethylene degradation. Environ Sci Technol 51:5172–5181

Guzik U, Greń I, Hupert-Kocurek K, Wojcieszyńska D (2011) Catechol 1,2-dioxygenase from the new aromatic compounds-degrading *Pseudomonas putida* strain N6. Int Biodeterior Biodegrad 65:504–512

Hara H, Eltis LD, Davies JE, Mohn WW (2007) Transcriptomic analysis reveals a bifurcated terephthalate degradation pathway in *Rhodococcus* sp. strain RHA1. J Bacteriol 189:1641–1647

Hara H, Stewart GR, Mohn WW (2010) Involvement of a novel ABC transporter and monoalkyl phthalate ester hydrolase in phthalate ester catabolism by *Rhodococcus jostii* RHA1. Appl Environ Microbiol 76:1516–1523

Holder JW, Ulrich JC, DeBono AC, Godfrey PA, Desjardins CA, Zucker J, Zeng Q, Leach ALB, Ghiviriga I, Dancel C, Abeel T, Gevers D, Kodira CD, Desany B, Affourtit JP, Birren BW, Sinskey AJ (2011) Comparative and functional genomics of *Rhodococcus opacus* PD630 for biofuels development. PLoS Genet 7:e1002219

Imbernon L, Oikonomou EK, Norvez S, Leibler L (2015) Chemically crosslinked yet reprocess able epoxidized natural rubber via thermo-activated disulfide rearrangements. Polym Chem 6:4271–4278

Irvine VA, Kulakov LA, Larkin MJ (2000) The diversity of extradiol dioxygenase "edo" genes in cresol degrading rhodococci from a creosote-contaminated site that express a wide range of degradative abilities. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 78: 341–352

Jang L, Keng H (2006) Development and characterization of as a monolayer for protein chips. Sens Mater 18:367–380

Jones A, Goodfellow M (2010) Genus II. *Rhodococcus* (Zopf 1891) emend Goodfellow et al. 1998. In: Bergey's Manual of Systematic Bacteriology, vol 4, 2nd edn. Springer, Berlin, pp 1–65

Juwarkar AA, Singh SK, Mudhoo A (2010) A comprehensive overview of elements in bioremediation. Rev Environ Sci Biotechnol 9:215–288

Khairy H, Meinert C, Wübbeler JH, Poehlein A, Daniel R, Voigt B, Riedel K, Steinbüchel A (2016) Genome and proteome analysis of *Rhodococcus erythropolis* MI2: elucidation of the 4,4′-dithiodibutyric acid catabolism. PLoS One 11:e0167539

Kim D, Kim Y-S, Kim S-K, Kim SW, Zylstra GJ, Kim YM, Kim E (2002) Monocyclic aromatic hydrocarbon degradation by *Rhodococcus* sp. strain DK17. Appl Environ Microbiol 68:3270–3278

Kim SH, Han HY, Lee YJ, Kim CW, Yang JW (2010) Effect of electrokinetic remediation on indigenous microbial activity and community within diesel contaminated soil. Sci Total Environ 408:3162–3168

Kim D, Choi KY, Yoo M, Zylstra GJ, Kim E (2018) Biotechnological potential of *Rhodococcus* biodegradative pathways. J Microbiol Biotechnol 28:1037–1051

Kolomytseva MP, Baskunov BP, Golovleva LA (2007) Intradiol pathway of *para*-cresol conversion by *Rhodococcus opacus* 1CP. Biotechnol J 2:886–893

Koutny M, Sancelme M, Dabin C, Pichon N, Delort AM, Lemaire J (2006) Acquired biodegradability of polyethylenes containing pro-oxidant additives. Polym Degrad Stab 91:1495–1503

Kulakov LA, Allen CCR, Lipscomb DA, Larkin MJ (2000) Cloning and characterization of a novel *cis*-naphthalene dihydrodiol dehydrogenase gene (*narB*) from *Rhodococcus* sp. NCIMB12038. FEMS Microbiol Lett 182:327–331

Kulakov LA, Chen S, Allen CCR, Larkin MJ (2005) Web-type evolution of *Rhodococcus* gene clusters associated with utilization of naphthalene. Appl Environ Microbiol 71:1754–1764

Kulig JK, Spandolf C, Hyde R, Ruzzini AC, Eltis LD, Grönberg G, Hayes MA, Grogan G (2015) A P450 fusion library of heme domains from *Rhodococcus jostii* RHA1 and its evaluation for the biotransformation of drug molecules. Bioorganic Med Chem 23: 5603–5609

Laczi K, Kis Á, Horváth B, Maróti G, Hegedüs B, Perei K, Rákhely G (2015) Metabolic responses of *Rhodococcus erythropolis* PR4 grown on diesel oil and various hydrocarbons. Appl Microbiol Biotechnol 99:9745–9759

Land M, Hauser L, Jun SR, Nookaew I, Leuze MR, Ahn TH, Karpinets T, Lund O, Kora G, Wassenaar T, Poudel S, Ussery DW (2015) Insights from 20 years of bacterial genome sequencing. Funct Integr Genomics 15:141–161

Larkin MJ, Kulakov LA, Allen CC (2006) Biodegradation by members of the genus *Rhodococcus*: biochemistry, physiology, and genetic adaptation. Adv Appl Microbiol 59:1–29

Larkin MJ, Kulakov LA, Allen CC (2010) Genomes and plasmids in *Rhodococcus*. In: Alvarez HM (ed) Biology of *Rhodococcus*. Springer, Berlin, pp 73–90

LeBlanc JC, Gonçalves ER, Mohn WW (2008) Global response to desiccation stress in the soil actinomycete *Rhodococcus jostii* RHA1. Appl Environ Microbiol 74:2627–2636

Ludwig W, Euzéby J, Schumann P, Buss HJ, Trujillo ME, Kämpfer P, Whiteman WB (2012) Road map of the phylum *Actinobacteria*. In: Goodfellow M, Kämpfer P, Busse HJ, Trujillo ME, Suzuki KI, Ludwig W, Whitman WB (eds) Bergey's manual of systematic bacteriology. Springer-Verlag, New York, pp 1–28

Martínková L, Uhnáková B, Pátek M, Nešvera J, Křen V (2009) Biodegradation potential of the genus *Rhodococcus*. Environ Int 35:162–177

Maruyama T, Ishikura M, Taki H, Shindo K, Kasai H, Haga M, Inomata Y, Misawa N (2005) Isolation and characterization of *o*-xylene oxygenase genes from *Rhodococcus opacus* TKN14. Appl Environ Microbiol 71:7705–7715

McLeod MP, Warren RL, Hsiao WWL, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJM, Holt R, Brinkman FSL, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA 103:15582–15587

Michałowicz J, Duda W (2007) Phenols-sources and toxicity. Polish J Environ Stud 6:347–362

Nanthini J, Chia KH, Thottathil GP, Taylor TD, Kondo S, Najimudin N, Baybayane P, Singh S, Sudesh K (2015) Complete genome sequence of *Streptomyces* sp. strain CFMR 7, a natural rubber degrading actinomycete isolated from Penang, Malaysia. J Biotechnol 214:47–48

Orro A, Cappelletti M, D'Ursi P, Milanesi L, Di Canito A, Zampolli J, Collina E, Decorosi F, Viti C, Fedi S, Presentato A, Zannoni D, Di Gennaro P (2015) Genome and phenotype microarray analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: genetic determinants and metabolic abilities with environmental relevance. PLoS One 10:e0139467

Pathak A, Chauhan A, Blom J, Indest KJ, Jung CM, Stothard P, Bera G, Green SJ, Ogram A (2016) Comparative genomics and metabolic analysis reveals peculiar characteristics of *Rhodococcus opacus*

strain M213 particularly for naphthalene degradation. PLoS One 11: e0161032

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

Patrauchan MA, Miyazawa D, LeBlanc JC, Aiga C, Florizone C, Dosanjh M, Davies J, Eltis LD, Mohn WW (2012) Proteomic analysis of survival of *Rhodococcus jostii* RHA1 during carbon starvation. Appl Environ Microbiol 78:6714–6725

Pérez-Pantoja D, Donoso R, Junca H, González D, Pieper H (2009) Phylogenomics of aerobic bacterial degradation of aromatics. Handbook of hydrocarbon and lipid microbiology, In, pp 1355–1397

Puglisi E, Cahill MJ, Lessard PA, Capri E, Sinskey AJ, Archer JAC, Boccazzi P (2010) Transcriptional response of *Rhodococcus aetherivorans* I24 to polychlorinated biphenyl-contaminated sediments. Microb Ecology 60:505–515

Rose K, Tenberge KB, Steinbüchel A (2005) Identification and characterization of genes from *Streptomyces* sp. strain K30 responsible for clear zone formation on natural rubber latex and poly(cis-1,4-isoprene) rubber degradation. Biomacromolecules 6:180–188

Rosłoniec KZ, Wilbrink MH, Capyk JK, Mohn WW, Ostendorf M, Van Der Geize R, Dijkhuizen L, Eltis LD (2009) Cytochrome P450 125 (CYP125) catalyses C26-hydroxylation to initiate sterol side-chain degradation in *Rhodococcus jostii* RHA1. Mol Microbiol 74:1031–1043

Rosłoniec KZ, van der Geize R, Dijkhuizen L (2013) CYP257A1 of *Rhodococcus jostii* strain RHA1 represents a novel cytochrome P450 enzyme family with demethylase activity and a putative physiological role in sterol metabolism. Dissertation, University of Groningen

Sameshima Y, Honda K, Kato J, Omasa T, Ohtake H (2008) Expression of *Rhodococcus opacus alkB* genes in anhydrous organic solvents. J Biosci Bioeng 106:199–203

Santo M, Weitsman R, Sivan A (2013) The role of the copper-binding enzyme-laccase-in the biodegradation of polyethylene by the actinomycete *Rhodococcus ruber*. Int Biodeterior Biodegrad 84:204–210

Sekine M, Tanikawa S, Omata S, Saito M, Fujisawa T, Tsukatani N, Tajima T, Sekigawa T, Kosugi H, Matsuo Y, Nishiko R, Imamura K, Ito M, Narita H, Tago S, Fujita N, Harayama S (2006) Sequence analysis of three plasmids harboured in *Rhodococcus erythropolis* strain PR4. Environ Microbiol 8 (2):334–346

Seto M, Kimbara K, Shimura M, Hatta T, Fukuda M, Yano K (1995) A novel transformation of polychlorinated biphenyls by *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol 61:3353–3358

Sharkey TD (1996) Isoprene synthesis by plants and animals. Endeavour 20:74–78

Shields-Menard SA, AmirSadeghi M, Green M, Womack E, Sparks DL, Blake J, Edelmann M, Ding X, Sukhbaatar B, Hernandez R, Donaldson JR, French T (2017) The effects of model aromatic lignin compounds on growth and lipid accumulation of *Rhodococcus rhodochrous*. Int Biodeterior Biodegrad 121:79–90

Sivan A, Szanto M, Pavlov V (2006) Biofilm development of the polyethylene-degrading bacterium *Rhodococcus ruber*. Appl Microbiol Biotechnol 72:346–352

Smith MR (1990) The biodegradation of aromatic hydrocarbons by bacteria. Biodegradation 1:191–206

Swain K, Casabon I, Eltis LD, Mohn WW (2012) Two transporters essential for reassimilation of novel cholate metabolites by *Rhodococcus jostii* RHA1. J Bacteriol 194:6720–6727

Szőköl J, Rucká L, Šimčíková M, Halada P, Nešvera J, Pátek M (2014) Induction and carbon catabolite repression of phenol degradation genes in *Rhodococcus erythropolis* and *Rhodococcus jostii*. Appl Microbiol Biotechnol 98:8267–8279

Takeda H, Yamada A, Miyauchi K, Masai E, Fukuda M (2004) Characterization of transcriptional regulatory genes for biphenyl degradation in *Rhodococcus* sp. strain RHA1. J Bacteriol 186: 2134–2146

Táncsics A, Benedek T, Farkas M, Máthé I, Márialigeti K, Szoboszlay S, Kukolya J, Kriszt B (2014) Sequence analysis of 16S rRNA, *gyrB* and *catA* genes and DNA-DNA hybridization reveal that *Rhodococcus jialingiae* is a later synonym of *Rhodococcus qingshengii*. Int J Syst Evol Microbiol 64:298–301

Tao F, Zhao P, Li Q, Su F, Yu B, Ma C, Tang H, Tai C, Wu G, Xu P (2011) Genome sequence of *Rhodococcus erythropolis* XP, a biodesulfurizing bacterium with industrial potential. J Bacteriol 193:6422–6423

Van Beilen JB, Funhoff EG, Van Loon A, Just A, Kaysser L, Bouza M, Holtackers R, Röthlisberger M, Li Z, Witholt B (2006) Cytochrome P450 alkane hydroxylases of the CYP153 family are common in alkane-degrading eubacteria lacking integral membrane alkane hydroxylases. Appl Environ Microbiol 72:59–65

Van Der Geize R, Dijkhuizen L (2004) Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. Curr Opin Microbiol 7:255–261

Van Hylckama Vlieg JET, Leemhuis H, LutjeSpelberg JH, Janssen DB (2000) Characterization of the gene cluster involved in isoprene metabolism in *Rhodococcus* sp. strain AD45. J Bacteriol 182: 1956–1963

Vilchez-Vargas R, Junca H, Pieper DH (2010) Metabolic networks, microbial ecology and "omics" technologies: towards understanding in situ biodegradation processes. Environ Microbiol 12:3089–3104

Watcharakul S, Röther W, Birke J, Umsakul K, Hodgson B, Jendrossek D (2016) Biochemical and spectroscopic characterization of purified latex clearing protein (Lcp) from newly isolated rubber degrading *Rhodococcus rhodochrous* strain RPK1 reveals novel properties of Lcp. BMC Microbiol 16:92

Whyte LG, Smits THM, Labbé D, Witholt B, Greer CW, Van Beilen JB (2002) Gene cloning and characterization of multiple alkane hydroxylase systems in *Rhodococcus* strains Q15 and NRRL B-16531. Appl Environ Microbiol 68:5933–5942

Xu-Xiang Z, Shu-Pei C, Cheng-Jun Z, Shi-Lei S (2006) Microbial PAH-degradation in soil: degradation pathways and contributing factors. Pedosphere 16:555–565

Yoneda A, Henson WR, Goldner NK, Park KJ, Forsberg KJ, Kim SJ, Pesesky MW, Foston M, Dantas G, Moon TS (2016) Comparative transcriptomics elucidates adaptive phenol tolerance and utilization in lipid-accumulating *Rhodococcus opacus* PD630. Nucleic Acids Res 44:2240–2254

Yoo M, Kim D, Choi KY, Chae JC, Zylstra GJ, Kim E (2012) Draft genome sequence and comparative analysis of the superb aromatic-hydrocarbon degrader *Rhodococcus* sp. strain DK17. J Bacteriol 194(16):4440

Zampolli J, Collina E, Lasagni M, Di Gennaro P (2014) Biodegradation of variable-chain-length *n*-alkanes in *Rhodococcus opacus* R7 and the involvement of an alkane hydroxylase system in the metabolism. AMB Express 4:73

Zídková L, Szoköl J, Rucká L, Pátek M, Nešvera J (2013) Biodegradation of phenol using recombinant plasmid-carrying *Rhodococcus erythropolis* strains. Int Biodeterior Biodegrad 84:179–184