Dipartimento di / Department of

Dipartimento di Scienze dell'Ambiente e della Terra (DISAT)

Dottorato di Ricerca in / PhD program Scienze Ciclo / Cycle XXIX

Curriculum in (se presente / if it is) Scienze Ambientali

TITOLO TESI / THESIS TITLE

Assessment of Microbial Degradative Functions for the Evaluation of the Environmental Quality of Hydrocarbon Contaminated Soils

Cognome / Surname Zampolli Nome / Name Jessica

Matricola / Registration number 708661

Tutore / Tutor: D.ssa Collina Elena Maria

Cotutore / Co-tutor: D.ssa Di Gennaro Patrizia

Coordinatore / Coordinator: Prof.ssa Frezzotti Maria Luce

ANNO ACCADEMICO / ACADEMIC YEAR 2015/2016

Index

Chapter 1

1. Introduction and Aim

The purpose of this chapter is to give a comprehensive overview of soil quality concept, the complexity of soil system related to bioremediation approach and recent advances in molecular, ''-omics" and in genetic engineering technologies. The present PhD thesis focuses on the analysis of metabolic capabilities and genetic features of *Rhodococcus opacus* R7 in comparison to reference strains such as *Rhodococcus* sp. BCP1 and *Rhodococcus jostii* RHA1 to identify new genomic sequences and gene clusters involved in hydrocarbon metabolisms. These new tools are useful to assess environmental quality of contaminated soils.

1.1 Introduction

1.1.1 Preface

Soil contamination has been one of the major issues of the last decades, because of the increase of anthropogenic pollution. Environmental contamination is widespread due to the great number of facilities and processes, affecting human health, water resources, ecosystems and other receptors [1]. Petroleum products including gasoline, diesel or lubricants, can be released to the environment through accidents, managed spills, or as unintended by-products of industrial, commercial or private actions causing local and diffuse pollution to the environment. Petroleum mixtures can be affected by air, water and organisms, thereby changing the location and their composition in soil, water or air [2]. Annually, around 35 million metric tons of petroleum enters the sea [3].

It is now widely recognized that contaminated land is a potential threat to human health, and its continual discovery over recent years has led to international efforts to remedy many of these sites, 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 [4].

The toxicity of petroleum hydrocarbons to microorganisms, plants, animals and humans led to the development of many bioassays for petroleum pollution which depend on low concentrations (5-100mg/l) of crude oil or petroleum fractions killing or inhibiting the growth of microalgae and juvenile forms of marine animals. The toxic effects of hydrocarbons on terrestrial higher plants and their use as weedkillers have been ascribed to the oil dissolving the lipid portion of the cytoplasmic membrane, thus allowing cell contents to escape [3]. 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 identifying and determining the behavior and fate of organic compounds in natural ecosystems [5].

1.1.2 Soil quality

Soil quality represents an integral value of the compositional structures and natural functions of soil in relation to soil use and environmental conditions on site. In a biological and biochemical way, soil represents the most important complex interface for the global interchange of matter and energy. Among the indigenous soil components, different organisms and especially microorganisms play a key role in ecologically important biogeochemical processes [6]. It is well documented that human industrial activities strongly affect biological systems and, more in particular, the soil quality [7, 8]. For decades soils have increasingly been subjected to pollution and other adverse side-effects of different activities which negatively affect soil organisms. In particular, processes involving petroleum extraction, transport or processing [2] lead to extensive release in the environment of hydrocarbons mainly containing saturated hydrocarbons, polycyclic aromatic hydrocarbons and nitrogen-sulfur-oxygen compounds. A European Environment Agency (EEA) study showed that about 34% of contaminated sites in Europe were affected with mineral oil in 2006. This value increases to 53% including common petroleum substances like polycyclic aromatic hydrocarbons (PAH) and volatile aromatic hydrocarbons: benzene, toluene, ethylbenzene and xylenes (BTEX) [9].

A proper risk assessment is essential for the understanding and management of the soil polluted by hydrocarbons. Regarding regulation policies, several national regulations define two generic soil quality standards (SQSs). The first SQS divides soils between clean and slightly contaminated soils [10, 11]. This SQS has been defined differently in the national policies, i.e. background value (BV) in The Netherlands (VROM, 2012), Reference Generic Level (RGL) in Spain (Presidency Ministry, 2005), threshold contamination concentration (CSC) in Italy (Italian Ministry of Environment, 2006) or threshold value in Finland (Finish Ministry of the Environment, 2007). The second SQS establishes differences between slightly and seriously contaminated soils. This concentration can be defined as a generic value, like in The Netherlands with the intervention value (IV) and in Spain with the RGL multiplied by 100 times. In other European countries the SQS must be site-specific established such as in Italy with the risk contamination standard (CSR). In France, some generic soil concentrations are established, although they cannot be used as soil screening values to perform further detailed

site-specific risk assessment [12, 13]. In Germany, a few substances present regulation values, the different soil levels are risk based, and formulae for calculating the values are available [14].

Under stress conditions caused by adverse anthropogenic effects such as spread of chemical pollutants, the development and biochemical activities of soil microorganisms undergo several alterations. [15]. To prevent negative ecological consequences, microbiologically-related parameters should be involved in the determination of soil quality, such as: microbial biomass, composition of microbial communities (N_2-fixing bacteria, Actinomycetes, Pseudomonads), biochemical process-linked activities (respiration (CO_2 release), nitrification/denitrification, dehydrogenase activity).

1.1.3 Hydrocarbons in the environment

n-Alkanes

Hydrocarbons are energy-rich organic compounds consisting of carbon and hydrogen atoms and they can be saturated (only single C-H bonds) or unsaturated (one or more double or triple C-H bonds). Alkanes are saturated hydrocarbons with the general formula C_nH_{2n+2}; they can be linear (*n*-alkanes), cyclic (cyclo-alkanes), with formula C_nH_{2n}, or branched (iso-alkanes). Those having between one and four carbon atoms (methane to butane) are gaseous at ambient temperature. Larger molecules are liquid (>pentane C_5) or solid (>octadecane C_{18}) (the sub-index indicates the number of carbon atoms of the alkane molecule). Alkanes can constitute up to 50% of crude oil, depending on the oil source, but they are also produced by many living organisms such as plants, green algae, bacteria or animals. This probably explains why alkanes are present at low concentrations in most soil and water environments [16].

Monoaromatic hydrocarbons

Benzene, toluene, ethylbenzene and xylenes (BTEX) are volatile simple aromatic hydrocarbons commonly present in crude petroleum and petroleum products such as gasoline. They are also produced in Tg (i.e. megatons) per year as bulk chemicals for industrial use as solvents and starting materials for the manufacture of pesticides, plastics and synthetic fibers. BTEX are considered to be one of the major causes of environmental pollution because of widespread occurrence due to leakage from underground petroleum storage tanks and spills at petroleum production wells, refineries, pipelines and distribution terminals. BTEX compounds are included in European priority lists and are frequently observed as groundwater pollutants due to their widespread use in

petroleum industry or as solvents [17]. Benzene, toluene, and xylenes (BTX) are more soluble than other components of gasoline such as alkanes and polyaromatic hydrocarbons. Although immediate lethal toxicity is rare, the sublethal toxicity of these substances may cause a wide range of effects on living organisms, including humans, from subtle biochemical and physiological disturbances to severe toxic and mutagenic effects on development and reproduction [18, 19]. Benzene is of the most concern because of its association with the development of leukemia in humans [20].

Despite the toxicity and persistence, many microorganisms are able to transform or mineralize these compounds and use them as a sole source of carbon and energy [17].

Mineralization of BTEX has been reported under aerobic conditions as well as in the presence of other electron acceptors. However, diverse pathways are assumed to be responsible for a productive breakdown of BTEX components under aerobic conditions.

Polycyclic aromatic hydrocarbons

Polycyclic aromatic hydrocarbons (PAHs) are composed of fused, aromatic rings whose biochemical persistence arises from dense clouds of π-electrons on both sides of the ring structures, making them resistant to nucleophilic attack [21]. Polycyclic aromatic hydrocarbons (PAHs) are an ubiquitous class of hydrophobic organic compounds that are formed and released as a result of the incomplete combustion of organic material. Anthropogenic sources such as road traffic, petroleum sludge, asphalt, wood preservative wastes and combustion of fossil fuels, predominate [22]; anyway also natural sources such as volcanic eruptions and forest fires contribute to the total PAHs burden. PAHs consist of fused benzene rings in linear, angular or clustered arrangements. They contain by definition only carbon and hydrogen atoms. However, sulfur, nitrogen and oxygen atoms may readily substitute in the benzene rings to form heterocyclic aromatic compounds, which are commonly grouped with the PAHs. On the basis of their diffusion and toxicity, 16 PAHs compounds have been identified as priority pollutants by the United States Environmental Protection Agency (US-EPA) (Figure 1.1) [23].

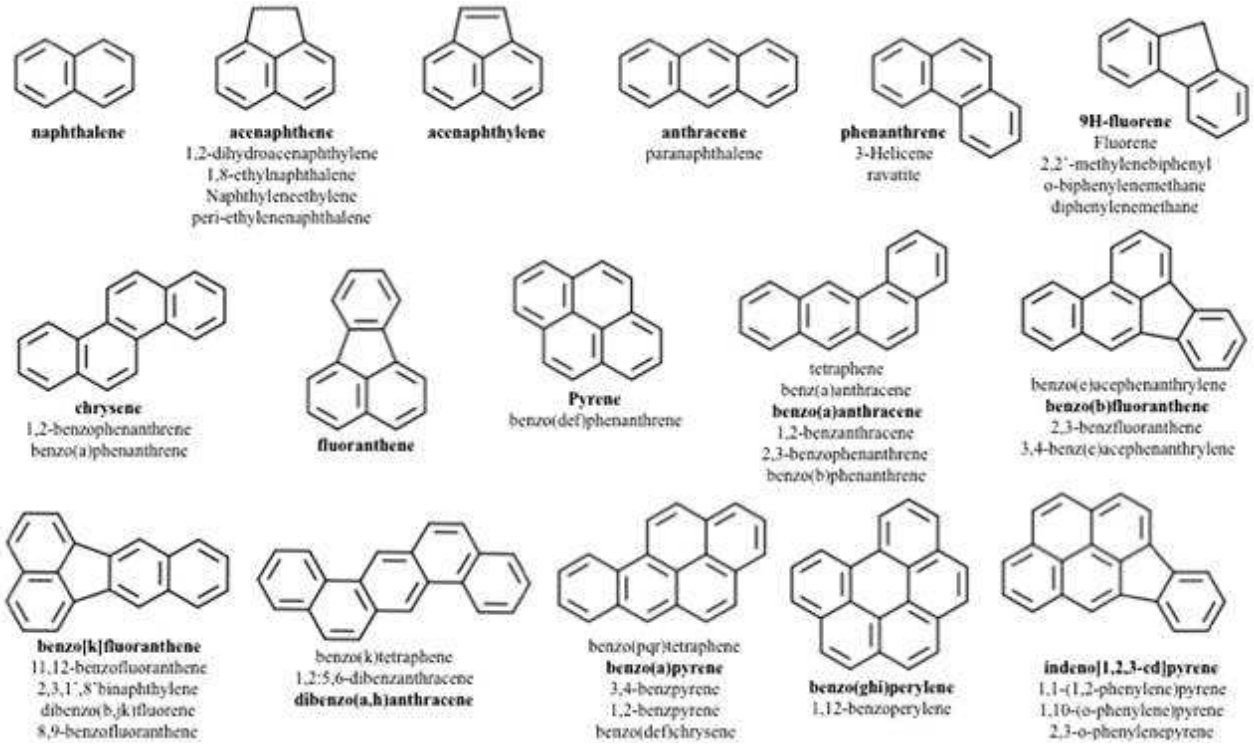

Figure 1.1 - Chemical structures of the 16 priority PAH compounds identified by US-EPA [23].

Generally, PAHs are lipophilic compounds that show high affinity with organic matter, indeed they are strongly sorbed to the organic matter in soil. Thus, they are relatively unavailable to degradation processes. PAHs can therefore remain in soil for many centuries, posing a long-term threat to the environment, despite partial degradation, volatilization and leaching of low molecular weight PAHs. Sorption generally increases with the number of benzene rings in PAH-molecule, since this implies higher lipophilicity [24].

Many PAHs have toxic, mutagenic and/or carcinogenic properties [25, 26]. Since PAHs are highly lipid soluble, they are readily absorbed by the gastrointestinal tract of mammals. They rapidly diffuse in a wide variety of tissues with a marked tendency for the localization in body fats. Metabolism of PAHs occurs via the cytochrome P450-mediated mixed function oxydase system with oxidation or hydroxylation as the first step [27]. The resulting epoxides or phenols might get detoxified in a reaction with the production of glucoronides, sulfates or glutathione conjugates. Some of the epoxides might be metabolized into dihydrodiols which, in turn, could undergo conjugation to form soluble detoxification products or be oxidized to diolepoxides. Carcinogenicity has been demonstrated for some of these epoxides [25]. For these reasons, PAHs are considered environmental pollutants that can have a detrimental effect on the flora and fauna of affected habitats, resulting in the uptake and accumulation of toxic chemicals in food chains and, in some

instances, in serious health problems and/or genetic defects in humans. Naphthalene, the first member of the PAH group, is a common micropollutant in potable water. The toxicity of naphthalene has been well documented and cataractogenic activity has been reported in laboratory animals [25, 26]. Naphthalene binds covalently to molecules in liver, kidney and lung tissues, thereby enhancing its toxicity; it is also an inhibitor of mitochondrial respiration [28]. Acute naphthalene poisoning in humans can lead to haemolytic anaemia and nephrotoxicity. In addition, dermal and ophthalmological changes have been observed in workers occupationally exposed to naphthalene. Phenanthrene is known to be a photosensitizer of human skin, a mild allergen and mutagenic to bacterial systems under specific conditions [26]. It is a weak inducer of sister chromatid exchanges and a potent inhibitor of gap junctional intercellular communication. Equivocal results for tumour initiation have been obtained with skinpainting studies in mice. However, the toxicity of benzo(a)pyrene, benzo(a)anthracene, benzo(b)fluoranthene, benzo(k)fluranthene, dibenz(a,h)anthracene and indeno(1,2,3-c,d)pyrene has been studied and there is sufficient experimental evidence to show that they are carcinogenic [26, 29, 30].

1.1.4 Environmental bioremediation

Bioremediation is defined as the process whereby organic compounds are biologically degraded under controlled conditions to an innocuous state, or to levels below concentration limits established by regulatory authorities. Bioremediation procedures offer an economical and effective possibility to degrade or render harmless various contaminants using natural biological activity [4]. In this strategy microbial communities are used for their ability to metabolize recalcitrant compounds. Microbial bioremediation strategies can be either applied *ex situ* or *in situ* in order to restore a contaminated environment. *Ex situ* treatment involves removal of contaminants at an off-site separate treatment facility, usually a bioreactor or an effluent treatment plant (ETP). These *ex situ* approaches provide optimal conditions and constant monitoring unlike the *in situ* bioremediation approaches that do not require the excavation of contaminated soil [4]. Conversely, *in situ* treatment involves bioremediation of the contaminated site either by natural attenuation (progressive removal of contaminant), intentional biostimulation of indigenous microbial communities or by bioaugmentation which includes deliberate inoculation of laboratory grown microbial strains having exceptional ability to degrade pollutants at the contaminated sites. Biostimulation provides the addition of nutrients, oxygen or other electron donors and acceptors to increase the population or activity of naturally occurring microorganisms available for

bioremediation. Bioaugmentation is the addition of single microorganisms or microbial communities that can biotransform or biodegrade contaminants [31]. It is more usual that the remediation of contaminated area is obtained by inoculating microbial communities instead of a single microorganism culture. Microbial consortia are composed by different bacterial species that cooperate so that each strain carries out one single step of the entire biodegradation process. Success of *ex situ or in situ* bioremediation relies heavily on the relative abundance, structure, catabolic versatility and biotic/abiotic interactions of the microbial communities (aerobic/anaerobic) that are indigenously present, amended or stimulated at contaminated sites, industrial wastewater/effluent treatment plants or within biofilms of the bioreactors [32].

Since contaminant compounds are transformed by living organisms through reactions that take place as a part of their metabolic processes, environmental biotechnology focuses on the development of techniques to optimize environmental parameters to allow these metabolic pathways to proceed faster and, therefore, to improve the biodegradation.

Soil is a chemically complex environment, due in part to the wide range of compounds produced by plants. Further complexity is introduced by decomposition and soil biogenesis. Many compounds of plant origin are chemically bound to soil, although they are likely present at concentrations that individually do not support bacterial growth. A potentially successful competitive strategy for bacteria in this environment is to use the diverse compounds simultaneously. This strategy, rather than rapid response to transient nutrient sources, may underlie selective pressure for large genomes having numerous paralogous genes [33]. In the current scenario, several microbial bioremediation strategies are sought after as an indispensable, ecofriendly and cost-effective solution towards restoring the polluted ecosystems. Although it will not always be suitable, however, as the range of contaminants on which it is effective is limited, the time scales involved are relatively long, and the residual contaminant levels achievable may not always be appropriate.

1.1.5 Hydrocarbon bioremedation

Many microorganisms (bacteria, filamentous fungi and yeasts) can degrade different kind of hydrocarbons, using them as carbon source (Table 1.1).

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

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

The microbial degradation of a hydrocarbon typically occurs because microorganisms can benefit from the use of the contaminant compound as an electron donor and carbon source to support growth. The redox reactions result in the flow of electrons from the substrate to a terminal electron acceptor (e.g. an oxidant such as O_2) and the release of energy that is used to support cell synthesis [34, 35]. In the absence of O_2, a large number of alternative terminal electron acceptors (NO^{3-}, Fe^{3+}, $Fe(OH)_3$, SO_4^{2-}) can be utilized to support anoxic respiration. 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 recalcitrant compounds are often bacteria belonging to following genera: *Pseudomonas*, *Alcaligenes*, *Flavobacterium, Arthrobacter, Nocardia, Rhodococcus* and *Corynebacterium* [36] (Table 1.1). Indeed, this kind 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 more widespread and more versatile [37]. Some strains, many of which were *Pseudomonas* sp., could produce biosurfactant

with proper nutrients available. In fact, biosurfactants, usually including rhamnolipid, saponin, and alkyl polyglucosides could overcome the difficulties in oil-cell surface contact to accelerate the biodegradation of crude oil and its production by decreasing the surface tension of medium. Among different genera of microorganisms, the genus *Rhodococcus* 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 [38]. *Rhodococcus* is a metabolically versatile genus of actinobacteria frequently found in the environment. The biotechnological importance of rhodococci derives from their "lifestyles"; these heterotrophs commonly occur in soil where they degrade a wide range of organic compounds. Their assimilatory abilities have been attributed to their diversity of enzymatic activities as well as their mycolic acids, proposed to facilitate the uptake of hydrophobic compounds [39]. 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. They have the ability to accumulate triacylglycerol, to produce biosurfactants and to adapt to organic solvents. They 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) [40]. Consistent with the immense catabolic diversity of rhodococci, they often possess large genomes, like 9.7 Mb of genome sequences of *Rhodococcus jostii* RHA1 [33] distributed between a chromosome and three mega linear plasmids that act as storage for multiple copies of many degradative genes. An important feature is their ability to continue biodegradation under potentially adverse conditions such as low temperatures, which is important for effective bioremediation in very cold climates [41, 42].

n-Alkanes bioremediation

As alkanes are apolar molecules with a low chemical reactivity, their metabolism by microorganisms poses challenges related to their low water solubility, their tendency to accumulate in cell membranes, and the energy needed to activate the molecule. However, several microorganisms, both aerobic and anaerobic, can use diverse alkanes as a source of carbon and energy [16]. Various alkane degraders are bacteria that have a very versatile metabolism, so that they can use as carbon source many other compounds in addition to alkanes [43, 44]. Most frequently, alkanes are not preferred growth substrates for these bacteria, which will rather utilize other compounds before

turning to alkanes. On the other hand, some bacterial species are highly specialized in degrading hydrocarbons. They are called hydrocarbonoclastic bacteria and play a key role in the removal of hydrocarbons from polluted environments [16].

So far some alkane-degradative systems of a small number of Gram-negative bacteria have been well characterized, such as those of *Alkanivorax, Pseudomonas and Acinetobacter*. Much less is known about the alkane-degradative systems of Gram-positive bacteria. Homologs of *alkB* gene were amplified from some *Rhodococcus* spieces [45].

Monoaromatic hydrocarbon bioremediation

Large amounts of natural aromatic compounds (molecules containing one or more benzene rings) in the biosphere originate from decaying plant material (e.g. from lignin, a complex phenolic polymer). In addition, xenobiotic aromatic compounds are released due to leaks and spills from various industrial processes, and thus lead to hazardous environmental contaminations. Major sources of aromatic pollutants are the chemical and pharmaceutical industries, intensive agriculture, pulp and paper bleaching, mining and fossil fuels. Moreover, many of these synthetic aromatic compounds persist in the environment due to their recalcitrant nature. The abilities of microorganisms to degrade natural aromatic compounds (e.g., aromatic amino acids and amines, hydroxylated benzoates and phenylcarboxylic acids) are widespread among microbial species. Xenobiotic aromatic compounds are degraded by a smaller group of bacteria, particularly by strains of the Gram-negative bacteria of *Pseudomonas, Sphingomonas, Acinetobacter, Ralstonia* and *Burkholderia* and the Gram-positive genus *Rhodococcus*.

Various strains of rhodococci degrade benzene and its derivatives (toluene, ethylbenzene, xylenes, biphenyl), polycyclic and heterocyclic aromatic compounds, phenolic compounds, aromatic acids, halogenated aromatics, including polychlorinated biphenyls (PCBs), amino- and nitro-derivatives of aromatic compounds (e.g., aniline and nitrophenol), ethers and pesticides and desulfurize coal and petroleum products [46, 41]. All five BTX compounds (including the three xylene isomers) have been found to be biodegradable under aerobic and anaerobic conditions. Aerobic degradation of toluene, *p*-xylene, and *m*-xylene has been shown to be genetically encoded by the TOL plasmid. Benzene and only recently *o*-xylene have been shown to be degraded by pure cultures of aerobic bacteria [20].

Polycyclic aromatic hydrocarbon bioremediation

Microorganisms can degrade polycyclic aromatic hydrocarbons (PAHs) by either metabolic or cometabolic reactions. Co-metabolism is defined as the oxidation of non-growth substrates during the growth of an organism on another carbon or energy source. Co-metabolism is important for the degradation of both PAHs mixture and high molecular weight (HMW) PAHs (4-6 aromatic rings). On the other hand, PAHs with 2- to 4- benzene rings have been described for years to act as growth substrate for microbes, mainly bacteria. Although metabolism of PAHs by bacteria in anaerobic conditions has been reported [47], attention has been mostly paid to the aerobic metabolism. A variety of bacteria capable of using these compounds have been investigated so far. The initial step of the aerobic metabolism of PAHs is biologically slow and metabolically expensive. It usually occurs usually occurs via incorporation of molecular oxygen into the aromatic nucleus, by a multicomponent dioxygenase enzyme system which gives rise to a cisdihydrodiol. This compound is then aromatised again through a cis-diol dehydrogenase to yield a dihydroxylated intermediate. The latter substrate can be cleaved by dioxygenase activity, via *ortho* or *meta*-cleavage, leading to a linear compound entering the tricarboxylic acid cycle (TCA) [24].

1.1.6 Cultivation-indipendent approach

Advent of molecular technologies over the past few decades has led environmental microbiologists to recognize microbial communities as an imperative ecological parameter in monitoring polluted sites either by detecting community shifts in response to pollution or their resilience towards anthropogenic disturbances.

Successful execution of versatile bioremediation strategies requires a thorough understanding of factors governing the growth, metabolism, dynamics and functions of indigenous microbial communities at contaminated sites. Recent innovative breakthroughs in genotypic profiling, ultrafast genome pyrosequencing, metagenomics, metatranscriptomics, metaproteomics and metabolomics along with bioinformatics tools have provided crucial insights of microbial communities and their mechanisms in bioremediation of environmental pollutants. Moreover, advances in these technologies have significantly improved the process of efficacy determination and implementation of microbial bioremediation strategies [32].

It has been estimated that only 0.1-1% of the total soil bacterial population can be cultured by applying standard cultivation techniques and moreover, 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 [48].

Catabolic (or functional) genes encode for specific enzymes in catabolic 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 rapid development of molecular techniques in recent years provides insights into microbial processes, including determination of microbial diversity and activity *in situ* at the polluted site, straightforward screenings for particular gene diversity, gene quantification, whole-genome sequencing of bacterial isolates and of DNA and mRNA from total communities. 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 [49]. 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 [50] and evolved as the so-called GeoChip [51]. A further approach for the detection of catabolic functions related to aerobic aromatic biodegradation described the use of oligoarrays that specifically targeted Rieske non-heme iron oxygenases or monooxygenases [52].

1. 2 Scope of the thesis work

The aim of this project was to obtain molecular tools to use as "marker" sequences for the environmental quality evaluation of contaminated soils. The genomic characterization of bacteria belonging to *Rhodococcus* genus allowed to identify catabolic functions involved in the degradation of the more diffused environmental contaminants. The project can be divided into three main goals:

1. isolation of interesting bacteria belonging to *Rhodococcus* genus and sequencing of the whole-genome. Characterization of novel metabolic potentials in hydrocarbon-degradative bacteria belonging to *Rhodococcus* genus using Phenotype Microarray analysis;

2. genome analysis for the characterization of different genes and metabolic pathways involved in the degradation of different hydrocarbons and identification of "marker" sequences;
3. on the basis of the identified catabolic functions, assessment of contaminated soil in microcosms experiments.
The final purpose of this project is to identify degradative capabilities to evaluate the depuration potential of an hydrocarbon contaminated soil using a culture-independent approach.

1.3 References

[1] Burgess LC (2013) Organic pollutants in soil. In: Brevik EC, Burgess LC (Eds), Soils and Human Health. CRC Press 83-106.

[2] Pinedo J, Ibáñez R, Lijzen JPA, Irabien Á (2013) Assessment of soil pollution based on total petroleum hydrocarbons and individual oil substances. J Environm Managem 130:72-79.

[3] Rosenberg E, EZ Ron (1996) Bioremediation of petroleum contamination, in Bioremediation: principles and applications, R.L.C.a.D.L. Crawford, Editor. 1996, Cambridge university press: Cambridge.

[4] Vidali M (2001) Bioremediation. An overview. Pure Appl. Chem. 73:1163-1172.

[5] Alexander M (1981) Biodegradation of chemicals of environmental concern. Science 211:132-8.

[6] Filip Z (2002) International approach to assessing soil quality by ecologically-related biological parameters. Agricult Ecosyst Environ 88:169-174.

[7] Avidano L, Gamalero E, Cossa GP, Carraro E (2005) Characterization of soil health in an italian polluted site by using microorganisms as bioindicators. Appl Soil Ecol 30:21-33.

[8] Filip Z (1995) Einfluss chemischer Kontaminanten (insbesondere Schwermetalle) auf die Bodenorganismen und ihre ökologisch bedeutenden Aktivitäten UWSF-Z. Umweltchem. Ökotox. 7:92-102.

[9] EEA (European Environment Agency) (2011) Overview of contaminants affecting soil and groundwater in Europe (http://www.eea.europa.eu/data-and-maps/figures/overview-of-contaminants-affecting-soil-and-groundwater-in-europe).

[10] Swartjes FA, Rutgers M, Lijzen JP, Janssen PJ, Otte PF, Wintersen A, Brand E, Posthuma L (2012) State of the art of contaminated site management in the Netherlands: policy framework and risk assessment tools. Sci Total Environ 427-428:1-10.

[11] VROM (Dutch Ministry of Housing, Spatial Planning and the Environment) (2012) Soil Remediation Circular 2009. Staatscourant 3 April 2012, Nr. 6563. Ministry of Housing, Spatial Planning and the Environment, The Hague.

[12] Carlon, C, (2007) Derivation methods of soil screening values in europe. EUR 22805-EN. In: A Review and Evaluation of National Procedures towards Harmonization. European Commission, Joint Research Centre, Ispra, Italy.

[13] MATE (French Ministry of Environment), 2002. Management of (Potentially) Polluted Sites. Preliminary View, the Initial Diagnostic, Simplified Risk Assessment. Version 2. BRGM Editions, March 2000, Annex 5 December 2002. MATE, La Défense, France (in French).

[14] BMU (Federal Ministry for the Environment, Nature Conservation and Nuclear Safety) (1999) Federal Soil Protection and Contaminated Sites Ordinance (BBodSchV). Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, Bonn, Germany.

[15] Reible D, Demnerova K (2001) Innovative approaches to the on-site assessment and remediation of contaminated sites. Kluwer Academic Publishers. Prague, Czech Republic.

[16] Rojo F (2009) Degradation of alkanes by bacteria. Environ Microbiol 11:2477-2490.

[17] Jindrova E, Chocová M, Demnerová K, Brenner V (2002) Bacterial aerobic degradation of benzene, toluene, ethylbenzene and xylene. Folia Microbiol 47:83-93.

[18] Díaz E (2004) Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. Int Microbiol 7:173-180.

[19] Andreoni V, Gianfreda L (2007) Bioremediation and monitoring of aromatic-polluted habitats. Appl Microbiol Biotechnol 76:287-308.

[20] Evans PJ, Mang DT and Young LY (1991) Degradation of toluene and *m*-xylene and transformation of *o*-xylene by denitrifying enrichment cultures. Appl Environ Microbiol 57:450-454.

[21] Johnsen AR, Wick LY, Harms H (2004) Principles of microbial PAH-degradation in soil. Environmental Pollution 133:71-84.

[22] Cooke M, Dennis AJ (1983) Polycyclic aromatic hydrocarbons: formation, metabolism and measurement. Battelle Press Inc, Columbus OH, USA.

[23] Keith LH and Telliard WA (1979) ES&T Special Report: Priority pollutants: I-a perspective view. Environ Sci Technol 13:416-423.

[24] Zocca C. Study of polycyclic aromatic hydrocarbon-transforming microbial communities from contaminated sites. PhD thesis in Environmental and Industrial Biotechnology XVII cycle.

[25] Goldman R, Enewold L, Pellizzari E, Beach JB, Bowman ED, Krishnan SS and Shields PG (2001) Smoking increase carcinogenic polycyclic aromatic hydrocarbons in human lung tissue. Cancer Res 61:6367-6371.

[26] Mastrangelo G, Fadda E and Marzia V (1996) Polycyclic aromatic hydrocarbons and cancer in man. Environ Health Perspect 104:1166-1170.

[27] Stegeman JJ, Schlezinger JJ, Craddock JE and Tillitt DE (2001) Cytochrome P450 1A expression in mid water fishes: potential effects of chemical contaminants in remote oceanic zones. Environ Sci Technol 35:54-62.

[28] Falahatpisheh MH, Donnelly KC and Ramos KS (2001) Antagonistic interactions among nephrotoxic polycyclic aromatic hydrocarbons. J Toxicol Environ Health 62:543-560.

[29] Liu K, Han W, Pan WP and Riley JT (2001) Polycyclic Aromatic Hydrocarbon (PAH) emissions from a coal fired pilot FBC system. J Hazard Mater 84:175-188.

[30] Sram RJ, Binkova B, Rossner P, Rubes J, Topinka J and Dejmek J (1999) Adverse reproductive outcomes from exposure to environmental mutagens. Mutat Res 428:203-215.

[31] Leung M (2004) Bioremediation: techniques for cleaning up a mess. Bio Teach J 2:18-22.

[32] Desai C, Pathak H, Madamwar D (2010) Advances in molecular and ''-omics" technologies to gauge microbial communities and bioremediation at xenobiotic/anthropogen contaminated sites. Biores Technol 101:1558-1569.

[33] McLeod MP, Warren RL, Hsiao WWL, Araki N, Myhre M, Fernandes C, Eltis LD et al. (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci 103:15582-7.

[34] Field J and Sierra-Alvarez R (2004) Biodegradability of chlorinated solvents and related chlorinated aliphatic compounds. Rev Environm Science Biotechnol 3:185-254.

[35] Tiedje J (1993) Bioremediation from an ecological perspective. In: *In situ* bioremediation: when does it work?, Washington, DC: National Academy Press.

[36] Wagner-Döbler I, Bennasar A, Vancanneyt M, Störmpl C, Brümmer I, Eichner C, et al. (1998) Microcosm enrichment of biphenyl-degrading microbial communities from soils and sediments. Appl Environ Microbiol 64:3014-3022.

[37] Harayama S, Rekik M, Wubbolts M, Rose K, Leppik RA and Timmis K (1989) Characterization of five genes in the upper pathway operon of TOL plasmid pWW0 from *Pseudomonas putida* and identification of the gene products. J Bacteriol 171:5048-5055.

[38] Warhurst AM and Fewson CA (1994) Biotransformations catalyzed by the genus *Rhodococcus*. Crit Rev Biotechnol 1:29-73.

[39] Gurtler V, Mayall BC, Seviour R (2004) Can whole genome analysis refine the taxonomy of the genus *Rhodococcus*? FEMS Microbiol Rev 28:377-403.

[40]de Carvalho CC, da Cruz AA, Pons MN, Pinheiro HM, Cabral JM, da Fonseca MM, et al. (2004) *Mycobacterium* sp., *Rhodococcus erythropolis*, and *Pseudomonas putida* behavior in the presence of organic solvents. Microsc Res Tech 64:215-222.

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

[42] 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.

[43] Margesin R, Labbé D, Schinner F, Greer CW and Whyte LG (2003) Characterization of hydrocarbon-degrading microbial populations in contaminated and pristine alpine soils. Appl Environ Microbiol 69:3085-3092.

[44]Harayama S, Kasai Y, Hara A (2004) Microbial communities in oil-contaminated seawater. Curr Opin Biotechnol 15:205-214.

[45] 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.

[46] van der Geize R, Dijkhuizen L (2004) Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. Curr Opin Microbiol 7:255-61.

[47] Rockne KJ, Chee-Sanford JC, Sanford RA, Hedlund BP, Staley JT and Strand SE, (2000) Anaerobic naphthalene degradation by microbial pure cultures under nitrate-reducing conditions. Appl Environ Microbiol 66:1595-1601.

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

[49] Jørgensen KS (2008) Advances in monitoring of catabolic genes during bioremediation. Indian J Microbiol 48:152-155.

[50] 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.

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

[52] Iwai, S, Kurisu, F, Urakawa, H, Yagi, O, Kasuga, I, and Furumai H (2008) Development of an oligonucleotide microarray to detect di- and monooxygenase genes for benzene degradation in soil. FEMS Microbiol Lett 285:111-121.

Chapter 2

2. Sequenced Genome Analysis

This chapter focuses on the isolation of interesting bacteria belonging to *Rhodococcus* genus for their metabolic capabilities and phenotypic features. The whole-genome sequencing and the genome analysis of *Rhodococcus opacus* R7 were performed. R7 genome was analyzed through its comparison to reference strain genomes, such as *Rhodococcus jostii* RHA1 and *Rhodococcus* sp. BCP1 and literature retrievals. The genomic characterization led to a complete description of main genomic features of R7 strain that make *R. opacus* R7 sequences interesting to identify marker sequences to use in the assessment of environmental quality of a contaminated soil.

2.1 Introduction

The *Rhodococcus* genus comprises of Gram-positive, non-motile, non-sporulating, aerobic bacteria, with a high G+C content and mycolic acid-containing cell wall. This genus was firstly proposed by Zopf and revised by Tsukamura [1] and Goodfellow and Alderson [2]; it nowadays contains nearly 60 recognized species [3, 4]. Members of *Rhodoccocus* genus are widely distributed in soil, water and marine sediments [5]. Some of them have also evolved for pathogenicity in humans, animals such as *R. equi* that can cause pneumonia in foals [6], and plants (i.e. *R. fascians*) [7]. 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 [3]. *Rhodococcus* species are particularly interesting for their structural cell envelope, genomes and plasmids, along with metabolic and catabolic abilities. The cell envelopes of rhodococci are dominated by the presence of large branched chain fatty acids, the mycolic acids and this bacteria often present arabinogalactan (AG) 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 [3].

Compounds that are metabolically transformed or degraded by these bacteria are aliphatic and aromatic hydrocarbons, oxygenates, halogenated compounds, including polychlorinated biphenyls, nitro-aromatics, heterocyclic compounds, nitriles and various herbicides [5]. *Rhodococcus* spp. are able to perform steroid modifications, enantio-selective synthesis, production of amides from nitriles, and to convert plant secondary metabolites found in soil and rhizosphere, such as alkaloids, terpenes, and sterols [8]. Various *Rhodococcus* strains are also efficient at removing sulphur from coal and petroleum products [9, 10].

In line with the immense catabolic diversity shown by the members of this genus, *Rhodococcus* spp. are characterized to possess large and complex genomes, which contain a multiplicity of catabolic genes, a high genetic redundancy of biosynthetic pathways and a sophisticated regulatory network [11]. 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 [3]. Up to date, some *Rhodococcus* genomes have been sequenced with different level of completeness [12]. The first genome completely sequenced belongs to *Rhodococcus jostii* RHA1 strain. This strain was isolated from lindane-contaminated soil for its exceptional ability to aerobically degrade polychlorinated biphenyls (PCBs) [13]. It was further described to utilize a wide range of aromatic compounds, carbohydrates, nitriles and steroids as sole carbon and energy sources. Analyses of the 9.7 Mb RHA1 large genome provided the evidence of catabolic pathway redundancy and horizontal gene transfer events [14]. The RHA1 genome is distributed between a chromosome and three mega linear plasmids that act as storage for multiple copies of many degradative genes (pRHL1 1100 kb, pRHL2 450 kb and pRHL3 330 kb). The plasmid genes encode important catabolic capabilities, including apparently redundant biphenyl and alkyl benzene pathways that cometabolize PCBs. The smallest plasmid, pRHL3, is an actinomycete invertron containing large terminal inverted repeats. Recently, *Rhodococcus opacus* PD630, *Rhodococcus opacus* B4 and *Rhodococcus* sp. BCP1 strains have been completely sequenced. *R. opacus* PD630 was well studied for the production and accumulation of energy-rich triacylglycerols and *R. opacus* B4 for the tolerance to organic solvents [15, 16]. *Rhodococcus* sp. BCP1 was isolated from an aerobic butane-utilizing consortium able to co-metabolize chloroform, vinyl chloride, and trichloroethylene [17, 18, 19]. BCP1 was also able to catabolize a wide range of aliphatic, alicyclic, and carboxylated alkanes [20, 21]. *Rhodococcus* sp. BCP1 genome consisted of one chromosome and two plasmids (pBMC1 and pBMC2 of 120373 bp and 10329 bp, respectively), giving a total genome size of 6.2 Mb with a G-C content of 70.4%. A total of 6206 open reading frames (ORFs) and 58 RNAs genes (8 rRNAs and 50 tRNAs) were annotated [21].

2.2 Materials and Methods

2.2.1 Genome sequencing and annotation

The genome sequencing of *R. opacus* R7 was performed using a 454 sequencing technology (Roche GS FLX Titanium). *R. opacus* R7 genome sequencing 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 [22] NCBI pipeline was used to annotate genome sequences and to make manual curation. A more complete metabolic reconstruction has been performed with Genomes KEGG.

2.2.2 Whole-genome alignments

The whole genome of *R. opacus* R7 has been compared with the genomes of *Rhodococcus* sp. BCP1, *R. opacus* PD630, *R. opacus* B4 *R. jostii* RHA1 and *R. pyridinivorans* SB3094, to find out similar and conserved regions between them and to highlight genes and potential unique functions of the genome under analysis. For a global alignment and visualization of the comparison of all six genomes, the Mauve tool (2.3 Version) has been used [23] that interactively shows the relative positions of sequence regions (colored blocks) that are found in more than one genome. A more accurate analysis has been performed using the Last program [24] to obtain local pairwise all-vs-all alignments of all chromosome and plasmid sequences of the six genomes. With a post-process overlapping all the regions, the common regions are subtracted from the dataset to obtain a final list of regions that can be considered unique for each genome (present in one genome and not aligned in none of the other five). Only unique regions longer than 300 bp are considered. The genomic diversity in this group of bacteria is also represented in the diagram of Venn. This type of diagram represents the number of predicted protein-coding genes unique or shared amongst the genomes under investigation.

2.2.3 Metabolic network reconstruction

The preliminary approach to reconstruct metabolic capabilities of *R. opacus* R7 was to build a metabolic map using Pathway Tools software version 18.5. It is a comprehensive biology system software for management, analysis, and visualization of integrated collections of genome, pathway,

and regulatory data. It supports creation, curation, dissemination and Web-publishing of organism specific databases, called Pathway/Genome Databases (PGDBs), that integrates many types of data. It performs computational inferences, including prediction of metabolic pathways [25]. This software takes as input an annotated genome and compares the encoded proteins with the proteins in the MetaCyc and EcoCyc databases and maps them in metabolic pathways from BioCyc database. MetaCyc consists of a set of 2260 pathways from 2600 different organisms and their constituent enzymes, EcoCyc describes the genome and the biochemical machinery of *Escherichia coli* K-12 MG1655 and BioCyc consists of a collection of 5500 Pathway/Genome Databases (PGDBs). If the enzymes are found in the genome, the pathway is scored as present. The deduced metabolic network tends to have more pathways than actually present in the bacterium, but it supports well as a foundation element for inference [26].

An accurate comparative analysis of the molecular functions has been performed starting from the annotations computed with RAST in which each ORF function was annotated in a manually curated hierarchical taxonomy based on categories, subcategories, subsystems and roles. Categories contain more specific subcategories and subsystems which represent the actual function (a particular cellular basal metabolism or regulation mechanism) of the gene in which it plays a specific role.

The enzymes of the subsystems represented in the RAST annotation have been mapped to the pathways of the Pathway Tools database. These pathways and the corresponding set of reactions present in the two genomes are compared in order to find differences between them and to quantify, for each pathway, how many components (enzymes) can be found in the genome annotations.

Pathway Tools and RAST have a different annotation system that does not permit to have a complete description of the pathway in terms of presence of their components (reactions or enzymes) in the genome. For this reason the results of the two annotation systems have been integrated together.

2.2.4 Phylogenomic analyses

In order to compare the R7 genome with other *Rhodococcus* genus members, a phylogenetic analysis of 29 different strains was performed. The dataset contains bacteria belonging to the following species: *erythropolis* (2), *rhodochrous* (2), *opacus* (4) (including R7), *wratislaviensis* (2),

jostii (2), *rhodnii* (1), pyridinivorans (2), *ruber* (2) and other less representative species (see results for more details).

It was built the concatemer of four marker genes of the 28 strains, that were previously identified as conserved and informative for the bacteria classification: *16S rRNA* gene, *secY* gene, *rpoC* gene and *rpsA* gene [27]. The sequences were aligned with the Muscle sequence alignment program [28] using default parameters and the phylogenetic tree was built with PhyML program [29] that iteratively search for the optimal tree with a Nearest Neighbor Interchange algorithm evaluating Maximum Likelihood probability at each step starting from 5 random generated trees.

2.3 Results and Discussion

2.3.1 Genome features of *Rhodococcus opacus* R7

The genome of *Rhodococcus opacus* R7 was completely sequenced [30]. R7 genome was arranged in a chromosome of 8466345 bp and in 5 plasmids (pDG1 of 656443 bp, pDG2 of 426388 bp, pDG3 of 352342 bp, pDG4 of 191359 bp, and pDG5 of 25175 bp), giving a total genome size of 10.1 Mb with a G-C content of 66.0%. RAST analysis of R7 genome sequence identified a total of 9602 open reading frames (ORFs) and 62 RNAs genes (9 rRNAs and 53 tRNAs). By RAST analysis, the annotated ORFs of *R. opacus* R7 were classified into categories as reported in Table 2.1.

R7 genome showed 457 subsystems and subsystem categories representing the metabolism of carbohydrates, amino acids and cofactors, vitamins, prosthetic groups, or pigments, are the most abundant, and they account for 1236, 995, or 665 proteins, respectively. 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 amongst the 134 annotated are predicted to catalyze the oxidation of organic compounds with industrial and environmental relevance, such as linear alkanes, cyclic ketones, aromatic compounds (e.g., benzoate, catechol, gentisate, salycilate, and byphenil), aminopolycarboxilic acids, nitroalkanes, and phenylalkanoic acids. Forty-five ORFs encode cytochrome P450 monooxygenases that catalyzeregio- and stereospecific oxidation of a vast number of substrates.

Table 2.1 - Number of functions in each of the RAST category.

Category	*R. opacus* R7
Cofactors, Vitamins, Prosthetic Groups, Pigments	665
Cell Wall and Capsule	111
Virulence, Disease and Defense	124
Potassiummetabolism	17
Miscellaneous	115
Phages, Prophages, Transposable elements, Plasmids	0
Membrane Transport	108
Ironacquisition and metabolism	16
RNA metabolism	106
Nucleosides and Nucleotides	156
ProteinMetabolism	257
Cell Division and Cell Cycle	35
Motility and Chemotaxis	5
Regulation and Cell signaling	75
SecondaryMetabolism	13
DNA metabolism	103
Fatty Acids, Lipids and Isoprenoids	745
NitrogenMetabolism	71
Dormancy and Sporulation	3
Respiration	221
Stress Response	162
Metabolism of AromaticCompounds	267
Amino Acids and Derivatives	995
SulphurMetabolism	126
PhosphorusMetabolism	35
Carbohydrates	1236
Total	**5767**

2.3.2 Metabolic network reconstruction

The metabolic network reconstruction of *R. opacus* R7 genome was performed using pathway tools software 18.5. The metabolic reconstruction resulted in a model containing 1735 metabolic reactions: 322 pathways, 1836 enzymatic reactions, 10 transport reactions, 2023 enzymes, 13 transporters were assigned by this software analysis. The identified pathways were divided in classes, as described in Table 2.2 and categorized into the RAST categories (Table 2.1).

RAST annotation assigned 26 functional categories, 457 subsystems and 6089 functional roles for *R. opacus* R7. Although the integration of the information of pathway tools and RAST allowed to obtain an improved pathway reconstruction, only around 50% of the pathways have all their enzymes represented in the annotation. This percentage increases up to 70% when pathways, represented by all the enzymes except one, are considered.

Table 2.2 - Numbers of metabolic pathways deriving from metabolic reconstruction of *R. opacus* R7 genome.

Pathway Class	Number of Pathways
Activation/inactivation/interconversion	3
Biosynthesis	220
Degradation/utilization/assimilation	171
Detoxification	5
Generation of precursor metabolites and energy	25
Metabolic clusters	8
Superpathways	82
TOT	322

2.3.3 Whole-genome comparison

Whole genome sequence comparison of R. *opacus* R7 with a set of five other reference genomes (*Rhodococcus* sp. BCP1, *R. jostii* RHA1, *R. opacus* PD630, *R. opacus* B4 and *R. pyridinivorans* SB3094) was performed using the Mauve program, in the Mauve 2.3 Version and using the Last Software (Figure 2.1).

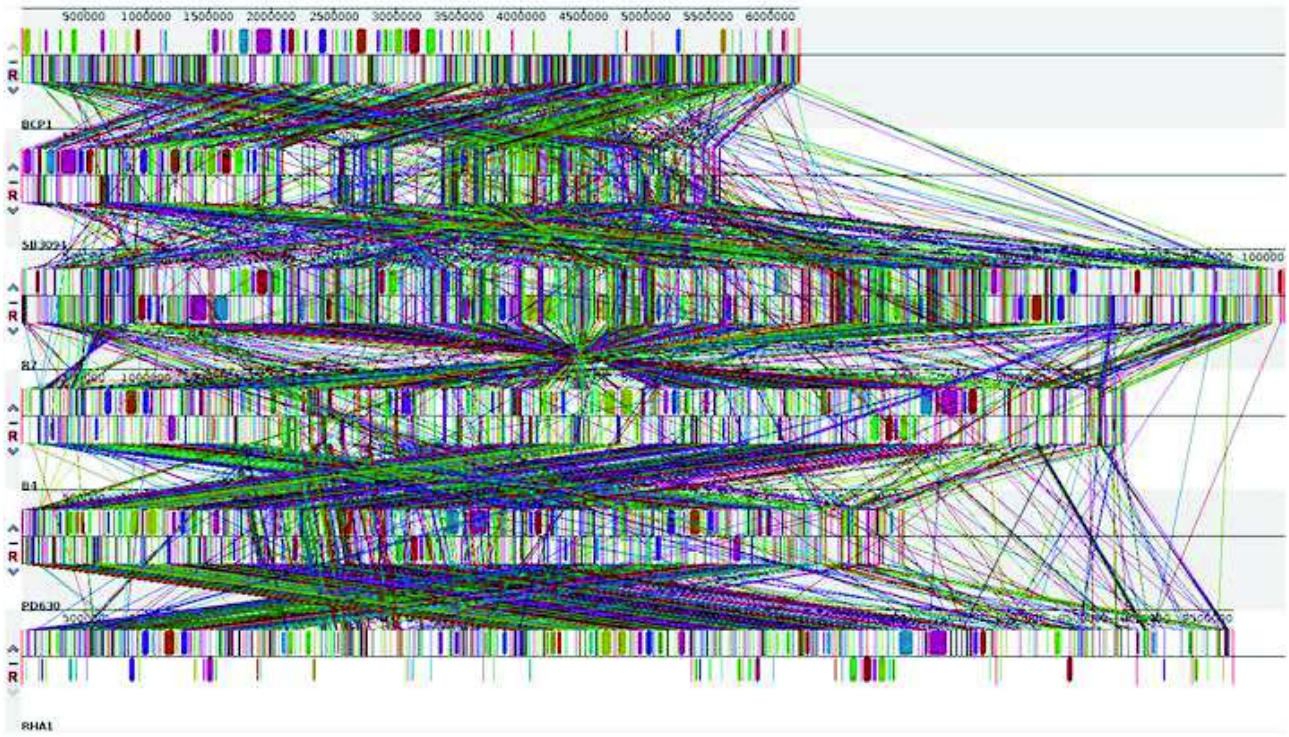

Figure 2.1 - Whole genome sequence comparison of *R. opacus* R7 with a set of five other reference genomes (*R. jostii* RHA1, *Rhodococcus* sp. BCP1, *R. opacus* PD630, *R. opacus* B4, *R. pyridinivorans* SB3094.

Aligned segments of pair of genomes have been collected in a dataset, filtered to remove regions <300 bp in length and grouped by genome. Similarity score between one reference genome and each other genome has been calculated considering the total amount of bases shared between the reference genome and each other genome under analysis.

Results showed that *R. opacus* R7 and *Rhodococcus* sp. BCP1 shared a total of 52% (R7 as reference) of similarity, calculated on all chromosomes and plasmids (Table 2.3).

Table 2.3 - Similarity scores for *Rhodococcus* genomes under analysis. Values in the matrix represent the percentage of bases shared in regions longer than 300 bp.

	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%	-

Considering the genome comparison amongst the *R. opacus* members, they share a range of similarity of 70%-90% percentage. A lower percentage of similarity resulted from the comparison of either *R. pyridinivorans* SB3094 or *Rhodococcus*sp. BCP1 with each *R. opacus* strain (around 50% and 40%, respectively). The analysis of relative positions of the sequences shared amongst the genomes (chromosome and plasmids) showed that they mainly do not maintain the same order (Table 2.4).

Table 2.4 - Numbers of similar regions that maintain the same order in each pair of genomes respect the total number of similar regions in R7 strain.

	R. sp. BCP1	*R. opacus* R7	*R. jostii* RHA1	*R. opacus* PD630	*R. opacus* B4	*R. pyridinivorans* SB3094
R. sp. BCP1	-	286/25919 (1.10%)	475/23532 (2.02%)	400/17911 (2.23%)	481/22904 (2.10%)	642/12397 (5.18%)
R. opacus R7	-	-	123/40390 (0.30%)	121/30862 (0.39%)	129/39122 (0.33%)	261/21603 (1.21%)
R. jostii RHA1	-	-	-	220/26720 (0.82%)	233/33364 (0.70%)	553/20422 (2.71%)
R. opacus PD630	-	-	-	-	195/25465 (0.77%)	462/15651 (2.95%)
R. opacus B4	-	-	-	-	-	546/19708 (2.77%)
R. pyridinivorans SB3094	-	-	-	-	-	-

The percentage of the regions of each *Rhodococcus* strain shared amongst all the genomes under analysis, ranged between 37% and 50% values, indicating a high level of genetic variability amongst *Rhodococcus* genus strains. The genomic diversity in this group of bacteria is represented in the diagram of Venn shown in Figure 2.2.

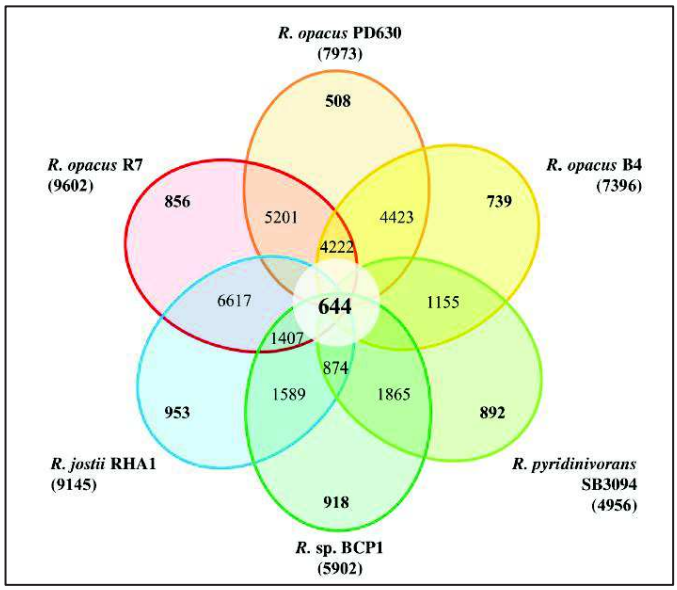

Figure 2.2 - Genomic comparison of *R. opacus* R7 with other *Rhodococcus* strains.

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 *Rhodococcus* sp. BCP1 and *R. opacus* R7 are located on the chromosomes; this result suggests that these regions could be characteristic of the *Rhodococcus* genus.

2.3.4 Unique and non-unique genomes features

Genomic comparison of *R. opacus* R7 with five other strains belonging to *Rhodococcus* genus allowed to identify unique genomic regions in its genomes (Table 2.5). The genomic divergences were investigated in the complete and well characterized sequenced genomes of *R. jostii* RHA1, *R.* sp. BCP1, *R. opacus* PD630, *R. opacus* B4 and *R. pyridinivorans* SB3094. The comparison showed that they are highly conserved except for the unique regions (longer than 300 bp); the majority of them including annotated genes.

Table 2.5 - Unique regions larger than 4 kb identified in R7 strain.

Region n.	Predicted key functions	Size (bp)	Unique gene number
1	Amino acids biosynthesis	4373	3
2	Aromatic compounds degradation	5323	4
3	Transcriptional regulators	4004	4
4	Phage region	5208	2
5	Unknown function	8845	2
6	Unknown function	7095	6
7	Phage region	6060	6
8	Polysaccharides biosynthesis	7648	3
9	Amino acids biosynthesis	4256	2
10	Unknown function	4289	2
11	Unknown function	5176	4
12	Amino Acids biosynthesis	4051	3
13	Phosphonate metabolism	4131	4
14	Unknown function	5775	4
15	Monosaccharides metabolism	5707	3
16	Fatty acids metabolism	4195	2
17	Antibiotic resistance; Folate Biosynthesis	4702	6
18	Monosaccharides metabolism	7569	6
19	Aromatic compounds degradation	4104	7
20	Amino acids biosynthesis	4662	4
21	Membrane transport systems	7072	5
22	Amino acids biosynthesis	4935	4
23	Unknown function	6723	6
24	Unknown function	6764	4
25	Phage region	20558	16
26	DNA recombination/repair	5651	2

27	Unknown function	7525	1
28	Unknown function	6665	9
29	Carbohydrates Metabolism	4823	5
30	Phage region	14443	16
31	Antibiotic resistance	4233	4
32	ABC Transporters	6368	5
33	Amino acids biosynthesis	4221	4
34	Unknown function	4522	4
35	Amino acids metabolism	5411	4
36	Aliphatic hydrocarbon degradation	5491	3
37	Flavonoid metabolism	4180	4
38	Carbohydrates metabolism; Aromatic compounds degradation	5293	7
39	Unknown function	6728	9
40	Aliphatic hydrocarbon degradation	4132	1
41	Unknown function	4476	6
42	Stress response	4006	4
43	Unknown function	9977	13
44	Unknown function	4324	1
45	Oxidoreductase	4899	6
46	Mobile element	4390	1
47	Serine/threonine protein kinase	7470	10
48	Unknown function	7688	4
49	Unknown function; Phage region	8499	9

The six strains have different amount of unique regions in the entire genomes as reported in Table 2.6. In particular, *Rhodococcus* sp. BCP1 showed the highest content of unique regions (16.3%) followed by *R. jostii* RHA1 (11.41%) and *R. opacus* R7 (11.32%). Within the R7 unique regions, a 40% of the annotated genes were predicted to code for hypothetical proteins.

Tabel 2.6 - Uniqueness characteristics of the six compared *Rhodococcus* spp. strains.

Uniqueness characteristics	*R. opacus* R7	*R.* sp. BCP1	*R. jostii* RHA1	*R. opacus* PD630	*R. opacus* B4	*R. pyridinivorans* SB3094
Lenght of unique regions (bp)	1,145,011	1,015,869	1,106,630	650,025	912,915	545,841
Unique regions (%)	11.32	16.3	11.41	7.1	10.33	10.44
ORFs in unique regions	1202	1266	1247	575	1058	1197
ORFs of hypothetical proteins in unique regions (%)	39.77	30.88	48.44	42.78	41.30	46.45
Not Annotated (%)	29.03	29.46	24.78	32.52	27.32	27.57

Moreover, around 30% of these regions was found to be non-coding sequence. Genes encoding several enzymatic classes predicted to be involved in the organic compounds metabolism were identified in unique regions in *R. opacus* R7, including oxygenases, hydroxylases, dehydrogenases, hydrolases, oxidoreductases, ligases, isomerases, aldolases, ion-transporting ATPases and cytochromes P450. The presence of these genes suggests different or specific metabolic capabilities of these *Rhodococcus* strains. In particular, in *R. opacus* R7 unique regions, two alkanal monooxygenases and one alkane-1-monooxygenase (annotated in RAST as alkanal monooxygenase), one cyclohexanone monooxygenase, and one pentachlorophenol monooxygenase were found. Other genes encoding several enzymatic classes are also located in unique regions such as kinases, peptidases, reductases, permeases, synthetases, esterases, cyclases, transcriptional regulators and membrane proteins. Interestingly, mobile elements as transposases, phage and mobile element proteins were found, suggesting the occurrence of recombination events. The ability to counteract stress conditions is strictly related to the ability to degrade a wide range of organic compounds and concerning this aspect, several genes coding for proteins involved in resistance to oxidative stress, metals and antibiotics were found in these regions (Table 2.5). The unique regions contained some heavy metal-resistance genes (i.e. lead-cadmium-zinc-mercury transporting ATPases, cobalt-zinc-cadmium resistance protein, metal-binding enzymes), stress-resistence genes (i.e. catalases, universal stress proteins, chaperons), and antibiotic-resistance genes (i.e. drug/metabolites transporters, beta-lactamase related proteins, penicillin-binding proteins).

2.3.5 Taxonomic classification of *Rhodococcus opacus* R7

The first taxonomical study of phylogeny of *Rhodococcus opacus* R7 strain was developed and a multi-locus sequence analysis (MLSA) maximum likelihood (ML) tree was constructed based on four marker genes (*16SrRNA*, *secY*, *rpoC* and *rpsA*) that were previously identified as conserved and informative for the bacteria classification [27]. The phylogenetic tree based on sequence alignments with reference strains of *Rhodococcus* genus showed that *R. opacus* R7 is correlated to *Rhodococcus opacus* and *Rhodococcus wratislaviensis* species in a clade that also includes *R. jostii* RHA1 (Figure 2.3). Conversely, *R. opacus* R7 doesn't belong to the same clade of *Rhodococcus* sp. BCP1 and *R. aetherivorans* strains including also *R. ruber* species. In conclusion, it can be confirmed that *Rhodococcus opacus* R7 is taxonomically related to *Rhodococcus opacus* species.

Figure 2.3 – Phylogenetic Tree. Phylogenetic analysis of *R. opacus* R7 based on sequence alignments with reference strains of *Rhodococcus* genus.

2.4 Conclusions

The *R. opacus* R7 genome characterization revealed a variety of genetic traits that likely contributing to have a complete description of main features of R7 strain that make the strain genome interesting to identify marker sequences to use in the assessment of environmental quality of a contaminated soil. Furthermore, this analysis showed some interesting genetic features related to hydrocarbon metabolisms.

2.5 References

[1] Tsukamura M (1974) Necessity for differentiation of rhodochrous group from tubercle bacilli. Jpn J Microbiol. 18:94-7.

[2] Goodfellow M, Alderson G (1977) The actinomycete-genus *Rhodococcus*: a home for the "rhodochrous" complex. J Gen Microbiol 100:99-122.

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

[4] http://www.bacterio.net/rhodococcus.html.

[5] Martínková L, Uhnáková B, Pátek M, Nesvera J, Kren V (2009) Biodegradation potential of the genus *Rhodococcus*. Environ Int 3:162-77.

[6] Letek M, González P, MacArthur I, Rodríguez H, Freeman TC, Valero-Rello A, et al. (2010) The genome of a pathogenic *Rhodococcus*: cooptive virulence underpinned by key gene acquisitions. PLoS Genet 6: e1001145.

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

[8] Larkin MJ, Kulakov LA, Allen CCR (2006) Biodegradation by members of the genus *Rhodococcus*: biochemistry, physiology, and genetic adaptation. Adv Appl Microbiol 59:1-29.

[9] van der Geize R, Dijkhuizen L (2004) Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. Curr Opin Microbiol 7:255-61.

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

[11] Alvarez HM (2010) Central metabolism of the species of the genus *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus*. Springer-Verlag Berlin Heidelber 91-108.

[12] http://www.ncbi.nlm.nih.gov/genome/Rhodococcus

[13] 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 61:4510-3.

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

[15] 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.

[16] Holder JW, Ulrich JC, DeBono AC, Godfrey PA, Desjardins CA, Zucker J, et al. (2011) Comparative and functional genomics of *Rhodococcus opacus* PD630 for biofuels development. PLoS Genet 7:e1002219.

[17] Frascari D, Pinelli D, Nocentini M, Fedi S, Pii Y, Zannoni D (2006) Chloroform degradation by butane-grown cells of *Rhodococcus aetherovorans* BCP1. Appl Microbiol Biotechnol 73:421-8.

[18] Cappelletti M, Frascari D, Zannoni D, Fedi S (2012) Microbial degradation of chloroform. Appl Microbiol and Biotechnol 96:1395-1409.

[19] 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 monooxygenase genes. Front Microbiol 6:393.

[20] 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 77:1619-27.

[21] Cappelletti M, Di Gennaro P, D'Ursi P, Orro A, Mezzelani A, Landini M, et al. (2013) Genome sequence of *Rhodococcus* sp. strain BCP1, a biodegrader of alkanes and chlorinated compounds. Genome Announc 1:e00657–13.

[22] 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.

[23] Darling AE, Mau B, Perna NT (2010) ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE 5:e11147.

[24] Kiełbasa SM, Wan R, Sato K, Horton P, Frith MC (2011) Adaptive seeds tame genomic sequence comparison. Genome Res 21:487-493.

[25] Paley SM, Latendresse M, Karp PD (2012) Regulatory network operations in the Pathway Tools software. BMC Bioinformatics 13:243.

[26] Wang Y, Lilburn TG (2009) Biological resource centers and systems biology. Bioscience 59:113-125.

[27] Adékambi T, Butler RW, Hanrahan F, Delcher AL, Drancourt M, Shinnick TM (2011) Core gene set as the basis of multilocus sequence analysis of the subclass Actinobacteridae. PLoS ONE 6:e14792.

[28] Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl Acids Res 32:1792-1797.

[29] Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307-21.

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

Chapter 3

3. Phenotype Microarray Analysis

The aim of this chapter is to analyze the wide metabolic potential of bacteria belonging to *Rhodococcus* genus. The characterization of novel metabolic abilities of hydrocarbon-degradative bacteria is showed using the Phenotype Microarray analysis in presence of 41 different hydrocarbon compounds. The metabolic abilities of *R. opacus* R7 strain were investigated using this high-throughput technology based on the irreversible reduction of tetrazolium violet to formazan as a reporter of active metabolism. A software allowed then the correlation of the phenotype showed by R7 strain to the putative genotype to identify interesting sequences involved in these metabolisms. Moreover the sensitivity of R7 strain in presence of several toxic chemicals was tested to evaluate its ability to survive in a multi drug contaminated environment.

3.1 Introduction

The genome sequence only represents a snapshot of the real phenotypic capabilities of an organism, providing very few indications on other crucial aspects of the underlying life cycle such as response to environmental and genetic perturbations, fluctuations in time, gene essentiality and so on. To gain a systemic and exhaustive description of living entities, static information deriving from genome sequence is not enough and other levels of knowledge must be taken into consideration [1]. Compared to the increasing number of *Rhodococcus* spp. genomes that have been sequenced, very limited studies are available on their wide metabolic abilities and on their potentials for biodegradation activities related to genomic features. A high-throughput assessment of metabolic abilities is given by Phenotype Microarray (PM) system (Biolog), which consists of 96-well plates provided with different conditions of growth (substrates, metabolites, pH conditions, toxic chemicals) for bacteria in order to generate phenotypic data [2]. This technology has the potential to accelerate the functional characterization of genes also deriving from whole-genome sequencing. This approach has been recently used to describe the respiration, viability and growth of several environmentally and clinically relevant bacteria species like *Mycobacterium*, *Pseudomonas*, *Cronobacter* and *Sinorhizobium* [3, 4, 5, 6]. *R. jostii* RHA1 and *R. opacus* PD630 were also tested for the growth on a limited number of carbon sources using Phenotype Microarray system [7].

3.2 Materials and Methods

3.2.1 Phenotype Microarray on organic/xenobiotic compounds

Phenotype Microarray was conducted using a tetrazolium-based growth assay developed by Biolog Incorporated (Biolog, Inc., Hayward, CA). This technology uses the irreversible reduction of tetrazolium violet to formazan as a reporter of active metabolism [2]. The reduction of the dye causes the formation of a purple colour that is recorded by a charge-coupled-device camera every 15 min and provides quantitative and kinetic information about the response of the microbial cells in PM plates. All the procedures were performed as indicated by the manufacturer. The metabolic abilities of R7 strain were assayed on 41 organic/xenobiotic compounds added as the only carbon and energy source. R7 was grown at 30°C on BUG agar (Biolog), and then, the strain was picked with a sterile cotton swab from the agar surface and suspended in 1 mL of sterile mineral medium. The cell suspension was added to 15 mL of mineral medium without carbon sources until cell density of 79% transmittance (T) was reached on a Biolog turbidimeter [5]. 1% dye G (vol/vol) was added to the suspension before inoculation, while the substrate was supplied to the inoculated wells with a minimum of two wells of distance. The plates were incubated at 30°C in an OmniLog reader and were monitored automatically every 15 min. Readings were recorded for 72 h and the data were analyzed using OminoLog PM software (release OM_PM_109M). The analysis was performed at least in duplicate and the results were checked for consistency [4]. The growth was also evaluated using OD_{590} after 72 hours.

Activity curves from Phenotype Microarray experiments were analyzed. A clustering analysis has been performed with the DuctApe program [8] to group together the curves using nine characteristic curve parameters and three sigmoid-like models. Errors in the fitting confirmed that a direct calculation of activity area in this case is more accurate that a model-estimated activity based on not completely developed growth curves. Activity values are reported for the custom set of xenobiotics. The PM activities obtained with different compounds were associated with enzymes (EC code) gathered from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [9]. Enzymes not available are labeled as unknown. The activities reported in the heat maps are normalized in a scale from low activity (red color) to high activity (green color).

3.2.2 Phenotype Microarray on toxic compounds (PM11 to PM20)

Phenotype Microarray was conducted using Biolog microplates, a tetrazolium-based growth assay developed by Biolog Incorporated (Biolog, Inc., Hayward, CA). R7 and BCP1 strains were assayed on microplates PM11 to PM20 testing a wide variety of antibiotics, antiseptics, heavy metals and other inhibitors. Phenotype Microarray technology uses the irreversible reduction of tetrazolium violet to formazan as a reporter of active metabolism [2]. The reduction of the dye causes the formation of a purple colour that is recorded by a charge-coupled-device camera every 15 min and provides quantitative and kinetic information about the response of the microbial cells in PM plates. All the procedures were performed as indicated by the manufacturer. Strains were grown at 30°C on BUG agar (Biolog), and then, each strain was picked with a sterile cotton swab from the agar surface and suspended in 1 mL of sterile mineral medium. Each strain suspension was added to 15 mL of physiological solution until cell density of 79% transmittance (T) was reached on a Biolog turbidimeter [5]. The culture medium for PM from 11 to 20 was prepared with a mineral medium without carbon source. After the addition of 1% dye G (vol/vol) to the suspension, 100 µl of the mixture was inoculated into each well of the microplates. All PM microplates were incubated at 30°C in an OmniLog reader, and they were monitored automatically every 15 min. Reads were recorded for 72 h and data were analyzed using OminoLog PM software (release OM_PM_109M), which generated a time course curve for tetrazolium color developed. Each strain was analyzed at least in duplicate and the results were checked for consistency [4]. The growth was also evaluated in each well using optical density measurements at 590 nm (OD_{590}) after 72 hours.

3.3 Results and Discussion

3.3.1 Phenotype Microarray Analysis on organic/xenobiotic compounds

A high-throughput assessment of metabolic abilities was performed by Phenotype Microarray system for bacteria in order to generate phenotypic data. The Phenotype Microarray analysis allowed the assessment of metabolic abilities and different conditions of growth of *R. opacus* R7 to accelerate the functional characterization of putative genes involved in xenobiotic degradation.

R. opacus R7 metabolic activity was tested on 41 xenobiotic compounds supplied as sole carbon and energy source. The tested substrates belong to four chemical categories including: 1) aliphatic hydrocarbons and cycloalkanes, 2) BTEX and other aromatic compounds, 3) polycyclic aromatics (PAHs), 4) naphthenic acids and other carboxylic acids. Phenotype Microarray analysis was

conducted following both the tetrazolium-based metabolic activity and the growth measured as OD_{590} increment over 72 h. *R. opacus* R7 was able to grow on several hydrocarbons and their putative metabolic intermediates (Figure 3.1).

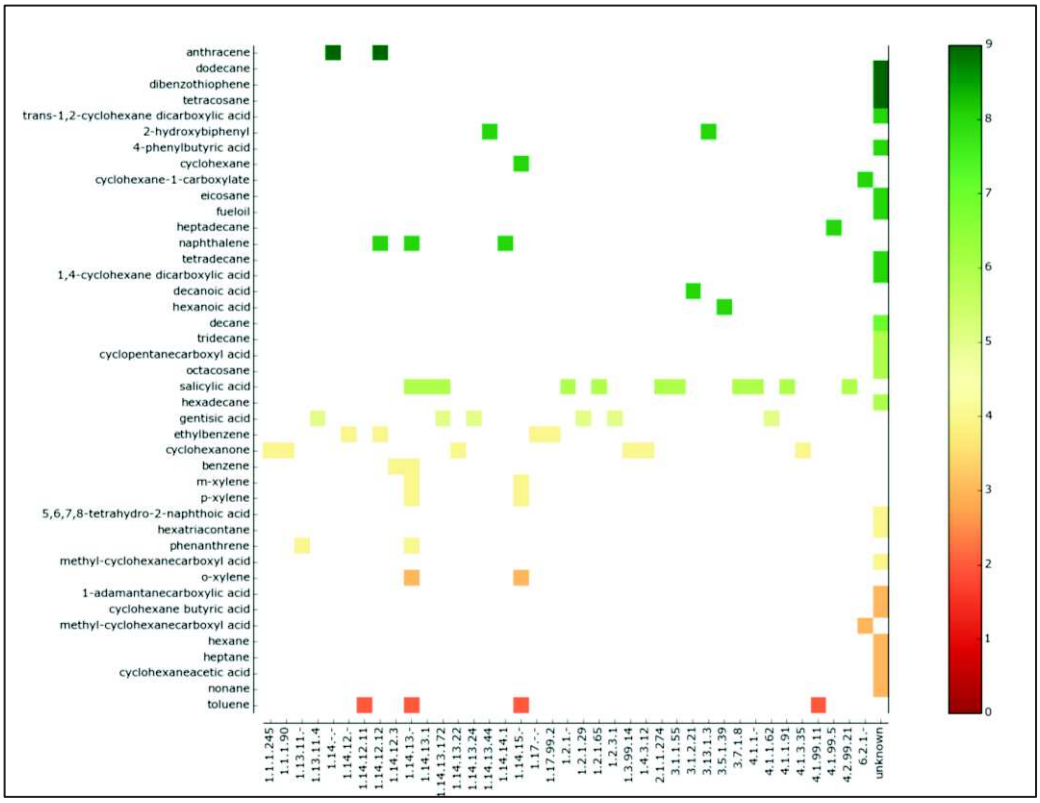

Figure 3.4 – Xenobiotics Phenotype Microarray analysis.

Based on the growth results in presence of *n*-alkanes ranging from *n*-hexane to *n*-hexatriacontane and cycloalkanes, *R. opacus* R7 had a preferential activity on *n*-dodecane, *n*-tetradecane, *n*-hexadecane and *n*-eicosane. These data confirmed previous results by Zampolli et al. (2014)[10]. Moreover, results showed that R7 strain was not able to grow in presence of odd-alkanes, except for *n*-heptadecane and R7 evidenced activity on cyclohexane and cyclohexanone. In addition, high level of activity was revealed on fuel oil, a mixture of different alkanes, including prystane and phytane. Growth analyses on aromatic hydrocarbons showed that the strain had the ability to degrade common aromatic compounds as toluene and ethylbenzene. In accordance with results reported by Di Gennaro et al. (2010) [11], R7 has the ability to grow only on the *ortho*-xylene isomer. The metabolic ability of R7 was tested also in presence of dibenzothiophene (DBT) as only carbon and energy source; the strain showed the highest activity among all compounds tested and in presence of its possible intermediate, 2-hydroxybiphenyl, R7 showed high activity.

R. opacus R7 was tested also on polycyclic aromatic hydrocarbons (PAHs) and their metabolic intermediates. Generally PAHs are metabolized by two main reactions: the peripheral one transforming a large number of polycyclic aromatic derivatives into central intermediates; and the ring cleavage reaction of intermediates to products that are subjected to the TCA cycle. Phenotype Microarray results showed that R7 has the ability to degraded PAHs compounds. The higher activity was shown on anthracene. The ability of R7 strain to grow on putative PAH degradation intermediates was also investigated. Results showed a good level of activity on gentisic acid and on salicylic acid. These experiments confirmed data from a previous work on R7 strain [11].

Among tested substrates, different naphthenic acids were supplied as only carbon and energy sources. These compounds are natural components of hydrocarbon deposits (petroleum, oil sands, bitumen, and crude oils) and highly toxic to a number of different organisms. R7 reported a significant response on the cyclopentanecarboxylic acid (CPCA) and cyclohexanecarboxylic acid (CHCA), while on methyl-cyclohexanecarboxylic acid (m-CHCA), cyclohexane butyric acid (CHBA), cyclohexaneacetic acid (CHAA), 1-adamantanecarboxylic acid (ACA) and 5,6,7,8-tetrahydro-2-naphthoic acid (THAN) R7 reported a low activity. The strain reported also a significant response to different acids that could be possible intermediates of different metabolic pathways, such as 4-phenylbutyric acid, hexanoic acids and 1,4-cyclohexane dicarboxylic acid.

The correlation between phenotype and the predicted genotype was analyzed using an adjustment of DuctApe software. The enzyme functions involved in the metabolism of 19 substrates out of the 41 tested were predicted by the software (Table 3.1). The majority of unidentified enzymes were related to the metabolism of *n*-alkanes and naphthenic acids. Indeed, *n*-alkanes and naphthenic acids were not associated to any known enzymes because of lacking in the KEGG's annotation system. Conversely, the metabolism of cycloalkanes and their possible metabolic intermediates was associated to the activity of monooxygenases and deyhdrogenases. For instance, the ability to utilize cyclohexanone was correlated to the activity of a cyclohexanone dehydrogenase (EC 1.3.99.14) and a cyclohexanone monooxygenase (EC 1.14.13.22).

The enzyme families predicted to be implicated in BTEX metabolism were: monooxygenases (EC 1.14.13- and 1.14.15-) involved in different hydrocarbon degradations; benzene/toluene dioxygenases (EC 1.14.12.3 and EC 1.14.12.11) and several putative enzymes involved in ethylbenzene metabolism, like the naphthalene dioxygenase (EC 1.14.12.12), the ethylbenzene hydroxylase (EC 1.17.99.2) and a putative dehydrogenase (EC 1.17-).

Concerning the activity on PAHs, amongst the 129 oxygenases/hydroxylases annotated for *R. opacus* R7 by RAST, 61 enzymes were predicted to be involved in mono- and polycyclic aromatic hydrocarbon metabolism. Within PAHs, the analysis by DuctApe of Phenotype Microarray showed anthracene and naphthalene to be metabolized by the same enzyme, a naphthalene dioxygenase (EC 1.14.12.12). Conversely the metabolism of phenanthrene was correlated to a PAH dioxygenase (EC 1.13.11.-). The use of gentisic acid was correlated to the gentisate 1,2-dioxygenase (EC 1.13.11.4); while the use of salicylic acid was related to salicylate 5-hydroxylase (EC 1.14.13.172). Amongst the naphthenic acids, the only substrate associated to an enzymatic class was the cyclohexanecarboxylic acid (CHCA), related to the large group of CoA synthetases (EC 6.2.1.1). Regarding this enzyme class, acetate and longer chain carboxylic acids can be converted to -CoA derivatives by CoA synthetases (EC 6.2.1.1) [7].

Table 3.1 - List of the xenobiotic compounds tested in phenotype microarray analysis and EC the numbers of the enzymes identified in R7 genome and predicted to be involved in its metabolism by DuctApe software.

Index	Chemical	EC numbers for *R. opacus* R7
1	1,4-Cyclohexane dicarboxylic acid	unknown
2	1-Adamantanecarboxylic acid	unknown
3	2-Hydroxybiphenyl	1.14.13.44; 3.13.1.3
4	4-Phenylbutyric acid	unknown
5	5,6,7,8-Tetrahydro-2-naphthoic acid	unknown
6	Benzene	1.14.12.3; 1.14.13.-
7	Cyclohexane	1.14.15.-
8	Cyclohexanecarboxylic acid	6.2.1.-
9	Cyclohexanone	1.1.1.90; 1.1.1.245; 1.3.99.14; 1.4.3.12; 1.14.13.22; 4.1.3.35
10	Cyclopentanecarboxylic acid	unknown
11	Decane	unknown
12	Dodecane	unknown
13	Eicosane	unknown
14	Ethylbenzene	1.14.12.12; 1.14.12.-; 1.17.99.2; 1.17.-.-
15	FuelOil	unknown
16	Heptadecane	4.1.99.5

17	Heptane	unknown
18	Hexadecane	unknown
19	Hexatriacontane	unknown
20	Naphthalene	1.14.12.12; 1.14.13.-; 1.14.14.1
21	Nonane	unknown
22	Octacosane	unknown
23	Tetracosane	unknown
24	Tetradecane	unknown
25	Toluene	1.14.12.11; 1.14.13.-; 1.14.15.-; 4.1.99.11
26	Tridecane	unknown
27	Anthracene	1.14.12.12; 1.14.-.-
28	Cyclohexane butyric acid	unknown
29	Cyclohexaneacetic acid	unknown
30	Decanoic acid	3.1.2.21
31	Dibenzothiophene	unknown
32	Gentisic acid	1.2.1.29; 1.2.3.1; 1.13.11.4; 1.14.13.24; 1.14.13.172; 4.1.1.62
33	Hexane	unknown
34	Hexanoic acid	3.5.1.39
35	*m*-Xylene	1.14.13.-; 1.14.15.-
36	Methyl-cyclohexanecarboxylic acid	6.2.1.-
37	*o*-Xylene	1.14.13.-; 1.14.15.-
38	*p*-Xylene	1.14.13.-; 1.14.15.-
39	Phenanthrene	1.13.11.-; 1.14.13.-
40	Salicylic acid	1.2.1.65; 1.2.1.-; 1.14.13.1; 1.14.13.172; 1.14.13.-; 2.1.1.274; 3.1.1.55; 3.7.1.8; 4.1.1.91; 4.1.1.-; 4.2.99.21
41	Trans-1,2-cyclohexane dicarboxylic acid	unknown

3.3.2 Phenotype Microarray Analysis on toxic compounds

In order to acquire a global understanding of the sensitivity of *R. opacus* R7, we used phenotype microarrays (PM) analysis, using the PM11-PM20 plates, the R7 strain was assayed on 240 chemical

agents tested at four different concentrations (Figures 3.2, 3.3, 3.4). The test was conducted in comparison to the sensitivity of *Rhodococcus* sp. BCP1. In PM analysis, *R. opacus* R7 showed the ability to counteract a wide range of toxic compounds and antibiotics belonging to different classes. R7 strain was resistant to the presence of several antibiotics (Figure 3.2 Panel AI, AII, AIII).

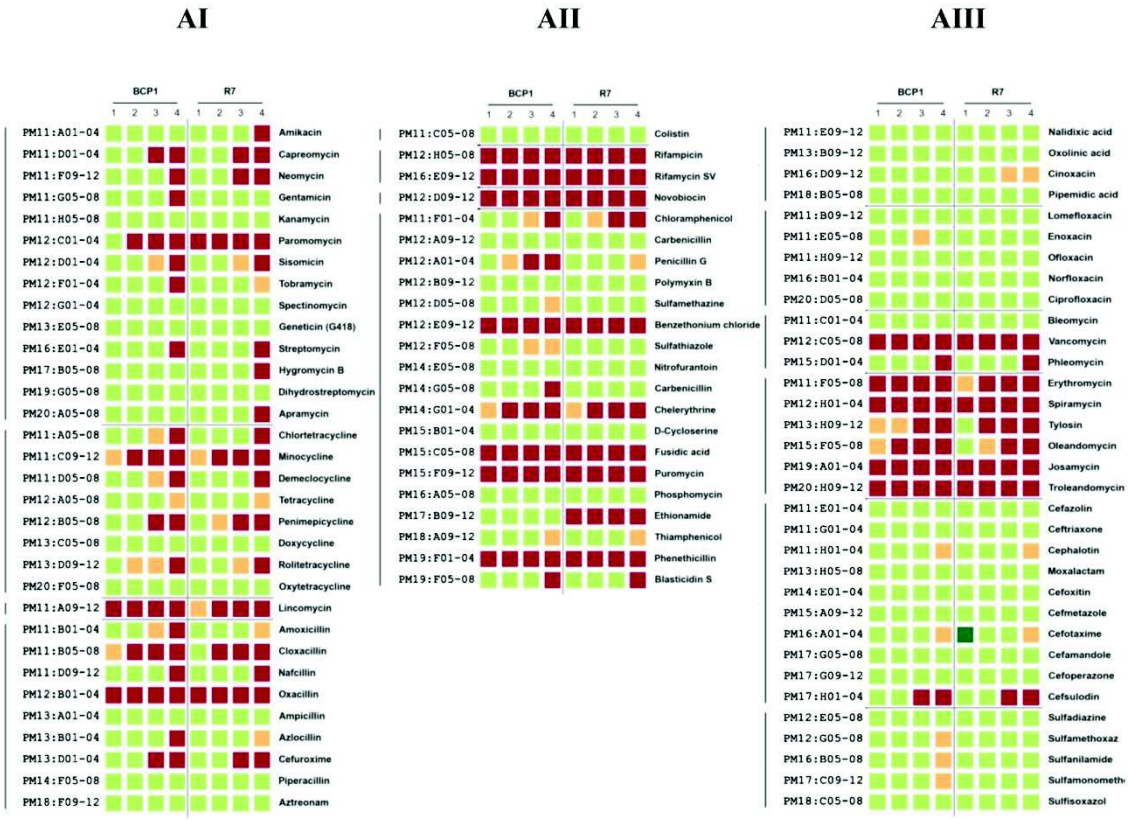

Figure 3.5 - Phenotype Microarray PM in presence of different antibiotics. Resistance differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of different antibiotics that were tested at four concentration (1, 2, 3, 4) according to Biolog procedure (AI, AII, AIII).

In particular, R7 could grow in the presence of all the tested concentrations of the antibiotics belonging to quinolone, fluoroquinolone, cephalosporin, sulfonamide, aminocoumarin (novobiocin) classes. Amongst the tested β-lactam and glycopeptide antibiotics, only oxacillin and vancomycin, significantly inhibited the growth of both the strains. Both R7 and BCP1 presented seven β-lactamase (EC 3.5.2.6) resistance genes in their genomes. Conversely, R7 showed general sensitivity to macrolide antibiotic class (i.e. erythromycin, josamycin and spiramycin) and to rifamycin group (i.e. rifampicin and rifamycin). For the tetracycline group, only minocycline was toxic for R7 strain. Amongst the aminoglycoside antibiotic, R7 was resistant to all the concentrations of kanamycin, spectinomycin, geneticin and dihydrostreptomycin and also showed high resistance to amikacin, gentamycin, sisomicin, tobramycin and hygromycin (Figure 3.2 Panel

AI). R7 did not grow on any concentration of paromomycin, while BCP1 was tolerant to the lowest concentration of this aminoglycoside antibiotic.

Amongst the other tested chemicals (preservatives, drugs etc.) (Figure 3.3 Panel AI, AII and Figure 3.4 Panel AIII, AIV), R7 was in general resistant to the amino acid hydroxamates (protein synthesis inhibitors) and to the nucleic acid analogs. Amongst the quaternary sodium salts, R7 was resistant only to dequalinium. Other compounds such as ionophores (CCCP and FCCP) and the respiration inhibitor iodonitrotetrazolium chloride inhibited the growth of the two strains even at the lowest concentration.

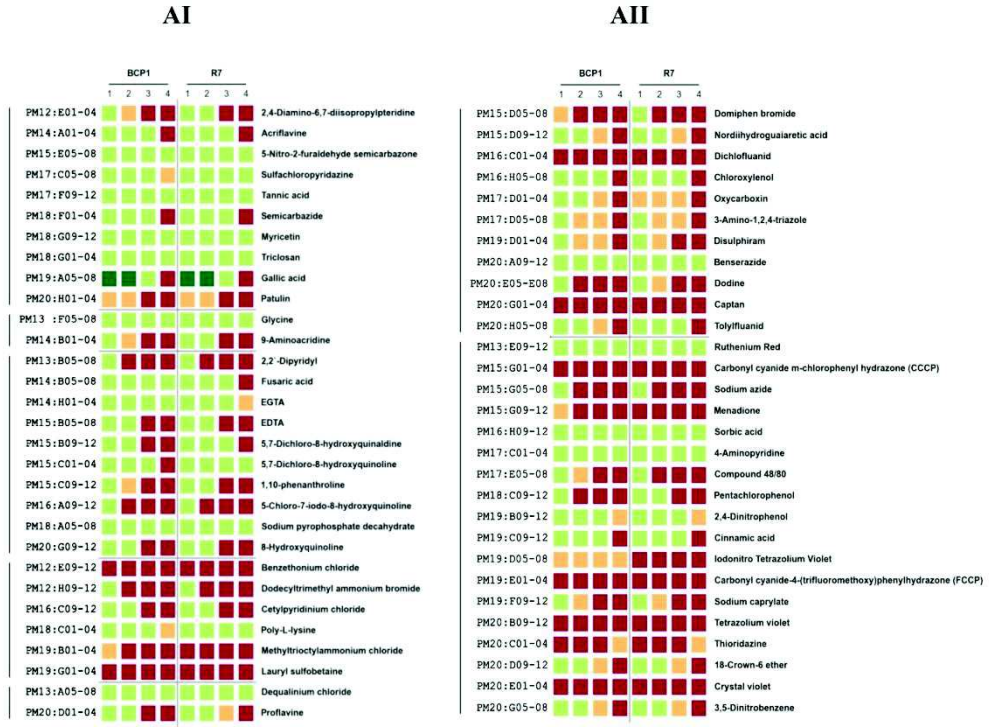

Figure 6.3 - Phenotype Microarray PM in presence of different antiseptics. Resistance differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of antiseptics (AI, AII).

Some of these phenotypic features can be attributed to the presence in its genome of a high number of genes coding for multidrug resistance proteins and transport systems (20 in R7). In particular, a higher tolerance was shown by R7 in the presence of the phenol derivative pentachlorophenol, the antifungal compound patulin and the β-lactam penicillin G. R7 showed a general tolerance to metal salts even at the highest concentrations. Exceptions include: i) the low/medium tolerance to sodium m-periodate, antimony (III) chloride and sodium caprylate (growth observed at the lowest concentration); ii) the sensitivity to vanadate sodium salts (sodium metavanadate and sodium orthovanadate) that inhibited the growth of the strain at all the tested

concentrations. Thallium acetate was the only metal that had a different effect on the two strains as it inhibited R7 growth at all the concentrations, while BCP1 showed tolerance to two out of the four concentrations of this metal salt (Figure 3.4 Panel B).

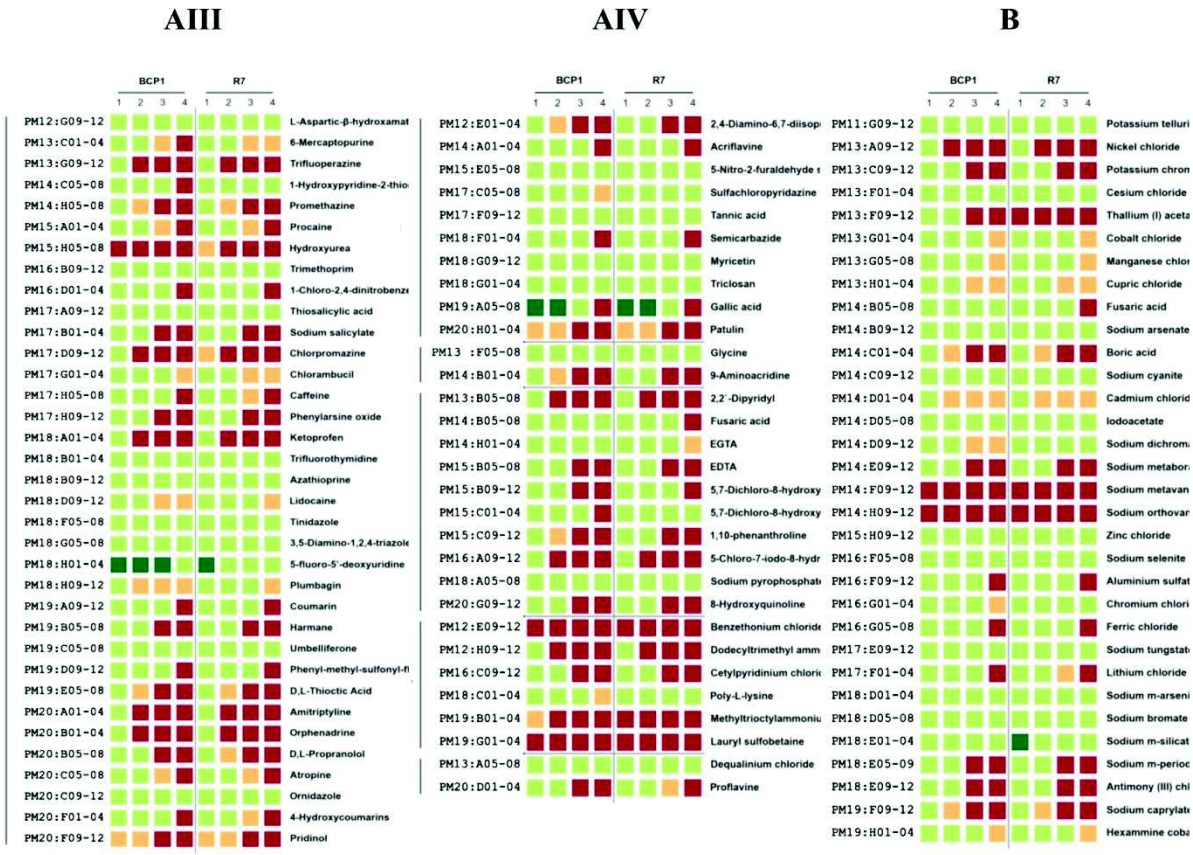

Figure 3.7 - Phenotype Microarray PM in presence of other antiseptics and metals. Resistance differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of antiseptics (AIII, AIV) and metals (B).

In the figure 3.2, 3.3, 3.4, the threshold values were established for every plate based on activity values of phenotype microarray analysis. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.

3.4 Conclusions

The aim of the present chapter was to assess the metabolic potential of *R. opacus* R7 in presence of different hydrocarbons using the Phenotype Microarray technology. Moreover, R7 sensitivity in presence of different toxic chemicals was assessed to test the ability of this strain to survive in a contaminated environment. The aim was not only to improve the knowledge about R7 degradative

capabilities, but also to correlate these metabolic abilities to the genotype. The idea was to use the data deriving from the Biolog experiments to identify main enzymatic classes involved in different metabolisms. The available enzymatic annotations appeared not enough for every kind of metabolism and all the genes that could be involved in these activities were not identified. Therefore another approach was used to enhance the information about the interesting genome sequences involved in xenobiotic metabolism. This second method is discussed in the following chapter (Chapter 4) where the phylogenetic analysis of R7 oxygenases putatively involved in xenobiotic metabolism is described.

3.5 References

[1] Fondi M and Liò P (2015) Multi -*omics* and metabolic modelling pipelines: challenges and tools for systems microbiology. Microbiological Research 171:52-64.

[2] Bochner BR (2003) New technologies to assess genotype-phenotype relationships. Nat Rev Genet 4:309-14.

[3] Yan Q, Power KA, Cooney S, Fox E, Gopinath GR, Grim CJ, et al. (2013) Complete genome sequence and phenotype microarray analysis of *Cronobacter sakazakii* SP291: a persistent isolate cultured from a powdered infant formula production facility. Front Microbiol 4:256.

[4] Biondi EG, Tatti E, Comparini D, Giuntini E, Mocali S, Giovannetti L, et al. (2009) Metabolic capacity of *Sinorhizobium* (*Ensifer*) *meliloti* strains as determined by phenotype MicroArray analysis. Appl Environm Microbiol 75:5396-404.

[5] 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 23:553-9.

[6] Khatri B, Fielder M, Jones G, Newell W, Abu-Oun M, Wheeler PR (2013) High throughput phenotypic analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* strains' metabolism using Biolog Phenotype Microarrays. PLoS ONE 8:e52673.

[7] Holder JW, Ulrich JC, DeBono AC, Godfrey PA, Desjardins CA, Zucker J, et al. (2011) Comparative and functional genomics of *Rhodococcus opacus* PD630 for biofuels development. PLoS Genet 7:e1002219.

[8] Galardini M, Mengoni A, Biondi EG, Semeraro R, Florio A, Bazzicalupo M, et al. (2014) DuctApe: a suite for the analysis and correlation of genomic and OmniLog™ Phenotype Microarray data. Genomics 103:1-10.

[9] Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucl Acids Res 8:27-30.

[10] 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.

[11] 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.

Chapter 4

4. Genome Characterization

In the previous chapter the metabolic potential of *R. opacus* R7 in presence of different hydrocarbons was evaluated using the Phenotype Microarray technology that correlated these metabolic abilities to the R7 genotype. The main enzymatic classes involved in different hydrocarbon metabolisms were identified. In the present chapter a phylogenetic analysis of R7 oxygenases putatively involved in xenobiotic metabolism was performed. This approach results in more precise and detailed information about all genes that could be involved in the hydrocarbon degradation.

4.1 Introduction

Biochemical reactions in which electrons are transferred from one molecule to another are catalyzed by a wide variety of enzymes, known as oxidoreductases or redox enzymes (EC 1.x.x.x). In order to perform these reactions, oxidoreductases often utilize (in)organic cofactors, such as flavins, metal-ions, hemes and pyrroloquinoline quinone (PQQ) for their catalytic activity. However, redox reactions can also be catalyzed by enzymes that do not require any additional cofactors. These cofactor-independent oxidoreductases typically contain several aromatic residues in their active site in order to be catalytically active [1]. Based on the type of reaction they catalyze, oxidoreductases have been divided into 22 different EC-subclasses. A more general classification was proposed by Xu [2], in which these enzymes were divided into four subgroups: (i) oxidases, (ii) peroxidases, (iii) oxygenases/hydroxylases and (iv) dehydrogenases/reductases. Oxygenases and hydroxylases catalyze the insertion of oxygen atom(s) into an organic substrate, using molecular oxygen (O_2) as oxygen donor. This subgroup consists of two enzyme types: enzymes that catalyze the insertion of a single oxygen atom, monooxygenases/hydroxylases, and enzymes that catalyze the insertion of two oxygen atoms, dioxygenases/hydroxylases. This eventually results in the oxygenation or hydroxylation of the organic substrate [1].

4.1.1 Monooxygenases

Monooxygenases are enzymes that catalyze the insertion of a single oxygen atom from O_2 into an organic substrate. In order to carry out this type of reaction, these enzymes need to activate molecular oxygen to overcome its spin-forbidden reaction with the organic substrate. This activation occurs upon donation of electrons to molecular oxygen, after which oxygenation of the

organic substrate can occur. The type of reactive oxygen-intermediate that is formed depends on which cofactor is present in the monooxygenase. In some cases, no cofactor at all is present. Several families of monooxygenases can be distinguished: the heme-dependent monooxygenases, the flavin-dependent monooxygenases, the copper-dependent monooxygenases, the non-heme iron-dependent monooxygenases, the pterin-dependent monooxygenases, other cofactor-dependent monooxygenases and cofactor-independent monooxygenases [1]. The first group of monooxygenases is also known as cytochrome P450 (CYPs) and can be found in many life forms: eukaryotes (mammals, plants and fungi) and bacteria express a wide variety of these enzymes [3]. In contrast to CYPs, genes encoding for flavin-dependent monooxygenases appear to be relatively abundant in prokaryotic genomes. Based on their amino acid sequence similarity and the available structural data, the external flavin-dependent monooxygenases were recently classified into six subclasses (Table 4.1) [4].

Table 4.1 - Classification of external flavin-dependent monooxygenases [4].

Subclass	Prototype	Reactions	Subunits	Cofactor	Coenzyme
A	*p*-hydroxybenzoate hydroxylase	Hydroxylation, epoxidation	α	FAD	NAD(P)H
B	Cyclohexanone monooxygenase	Baeyer-Villiger, *N*-oxidation	α	FAD	NADPH
C	Luciferase	Light emission, Baeyer-Villiger	α + β	–	FMN/NAD(P)H
D	*p*-hydroxyphenylacetate hydroxylase	Hydroxylation	α + β	–	FAD/NAD(P)H
E	Styrene monooxygenase	Epoxidation	α + β	–	FAD/NAD(P)H
F	Tryptophan 7-halogenase	Halogenation	α + β	–	FAD/NAD(P)H

Copper-dependent monooxygenases (EC 1.14.17.x) constitute a relatively small family of enzymes that require copper ions for hydroxylation of their substrates. Enzymes belonging to this family have been mainly found in eukaryotic organisms. Non-heme iron-dependent monooxygenases utilize two iron atoms as cofactor for their oxidative activity. These diiron-dependent enzymes are

also referred to as bacterial multicomponent monooxygenases (BMMs), catalyze hydroxylation and epoxidation reactions and consist of three components; a monooxygenase, a reductase and a small regulatory protein [5,6].

4.1.2 Dioxygenases

Dioxygenases are enzymes belonging to oxidoreductases and catalyze the introduction of two oxydrilic groups in the target molecule. Dioxygenases contain an iron atom as cofactor and they can be divided in three classes based on the position of the iron atom in the active-site.

The first class includes the non-heme iron dependent dioxygenases. The iron center plays a crucial role in enzymatic catalysis as it is normally used for co-ordination and catalytic activation of the substrate and O_2. Several dioxygenases belonging to this class are included in the cupin superfamily, which is characterized by a six-stranded-barrel fold and conserved amino acid motifs that provide the 3 His or 2- or 3His-1Glu ligande environment of a divalent metal ion [7]. Belonging to this class are the catechol dioxygenases. The catechol, an intermediate during bacterial aromatic degradation pathways, could be oxidatively cleaved by two kind of enzymes, the intradiol and the extradiol dioxygenases that require iron (III) and iron(II), respectively. The intradiol dioxygenases usually have the iron atom coordinated by two histidines, two tyrosines and one H_2O molecule in a triangular pyramid conformation [8]. The enzymes able to cleave the aromatic ring in intradiol position coordinate the iron atom by two histidines, one glutamic acid and a H_2O molecule in octahedral conformation [7]. The second class is represented by *rieske* non-heme oxygenases. Although these oxygenases mostly use NAD(P)H as an electron donor and catalyze the same oxygenation reaction, they are remarkably diverse with respect to their structure. They are multi-component enzymes of two or three protein components consisting of an electron transport chain (ETC) and an oxygenase. Oxygenase components are either homo- (α_n) or hetero-oligomers ($\alpha_n\beta_n$) and in each case, the α subunit, called large subunit, contains two conserved regions, a Rieske [2Fe-2S] center and non-heme mononuclear iron. The α subunits are known to be the catalytic components involved in the transfer of electrons to oxygen molecules. The ETC that transfers reducing equivalents from NAD(P)H to the oxygenase components consists of either a flavoprotein reductase or a flavoprotein reductase and a ferredoxin. An interaction between oxygenase and ETC components is required for the enzyme system to transfer electrons from the electron donor to the aromatic hydrocarbon electron acceptor [9].

The third class is the heme dioxygenases group, in which the iron atom is coordinated by the heme group. These enzymes catalyze the oxidation of L-tryptophan to *N*-formylkynurenine, the first and rate-limiting step in tryptophan catabolism [10].

4.2 Materials and Methods

4.2.1 Phylogenetic classification of monooxygenases and dioxygenases of *R. opacus* R7

In order to compare putative *R. opacus* R7 enzymes involved in different hydrocarbons degradation, a phylogenetic analysis of oxygenases was performed. A preliminary RAST annotation allowed to identify two main classes of enzymes belonging to R7: monooxygenases and dioxygenases. Considering only the enzymes annotated as catalytic component or subunits (usually α chain), phylogenetic cluster analysis was performed using Clustal Omega with maximum number of interactions and Tree View for graphic output. Reference sequences derived from literature and databases sequences (UniProt (UP) or PDB) were used as reference.

The nucleotide sequences of the interesting protein identified were compared with those from other *Rhodococcus* strains reported in literature and other strain references. The NCBI pipeline was used for the sequences comparison.

4.3 Results and Discussion

4.3.1 Classification of oxygenases of *Rhodococcus opacus* R7

The DuctApe software used to correlate phenotype to genetic aspects revealed no effectiveness in producing accurate information about the enzymes involved in R7 xenobiotic metabolisms. Indeed it couldn't associate each phenotype activity to every enzyme of the strain, because of lacking in annotation in database systems. In order to consider all the enzymes putatively involved in xenobiotic metabolism another method was considered.

The majority of the enzymes (EC) identified, putatively involved in the metabolism of different xenobiotics tested with Phenotype Micorarray assays, belonginged to the class of oxygenases. Therefore a preliminary phylogenetic analysis of R7 oxygenases was performed.

4.3.2 Phylogenetic classification of monooxygenases of *Rhodococcus opacus* R7

A preliminary RAST annotation identified 99 potential monooxygenases. About 65 of these are catalytic subunits and the phylogenetic analysis clusterized most part of these putative monooxygenases into six classes according to van Berkel and coworkers [4]. The analysis, performed according to the last two criteria, sequence and (if available) structural data, showed six classes of flavin-dependent monooxygenases (A, B, C, D, E, F) and a non-heme iron-dependent monooxygenase family (Figure 4.1). The classification revealed following monooxygenases categories: within A class we identified one 3-(3-hydroxyphenyl) propionate hydroxylase, several FAD binding monooxygenases similar to 2,4-dichlorophenol 6-monooxygenase of *Pseudomonas acidovorans* (UniProt); B class included several categories, including cyclopentanone and cyclohexanone monooxygenases, and some FAD binding monooxygenases similar to a probable indole-3-pyruvate monooxygenase from a *Arabidopsis thaliana* (UniProt); C class included two alkanal monooxygenases (LuxA), eight nitriloacetate monooxygenases (NTA-MO) and two alkanal sulfonate monooxygenases (SsuD); within D class only five 4-hydroxyphenylacetate 3-monooxygenases were identified; E class wss represented by a tryptophan monooxygenase similar to tryptophan 2-monooxygenase from *Pseudomonas savastanoi* (Uniprot) and F class was represented by a styrene monooxygenase.

Moreover a non-heme iron-dependent monooxygenase clade was identified that is organized in two subclasses: a membrane-bound alkane monooxygenase and one similar to a methane monooxygenase.

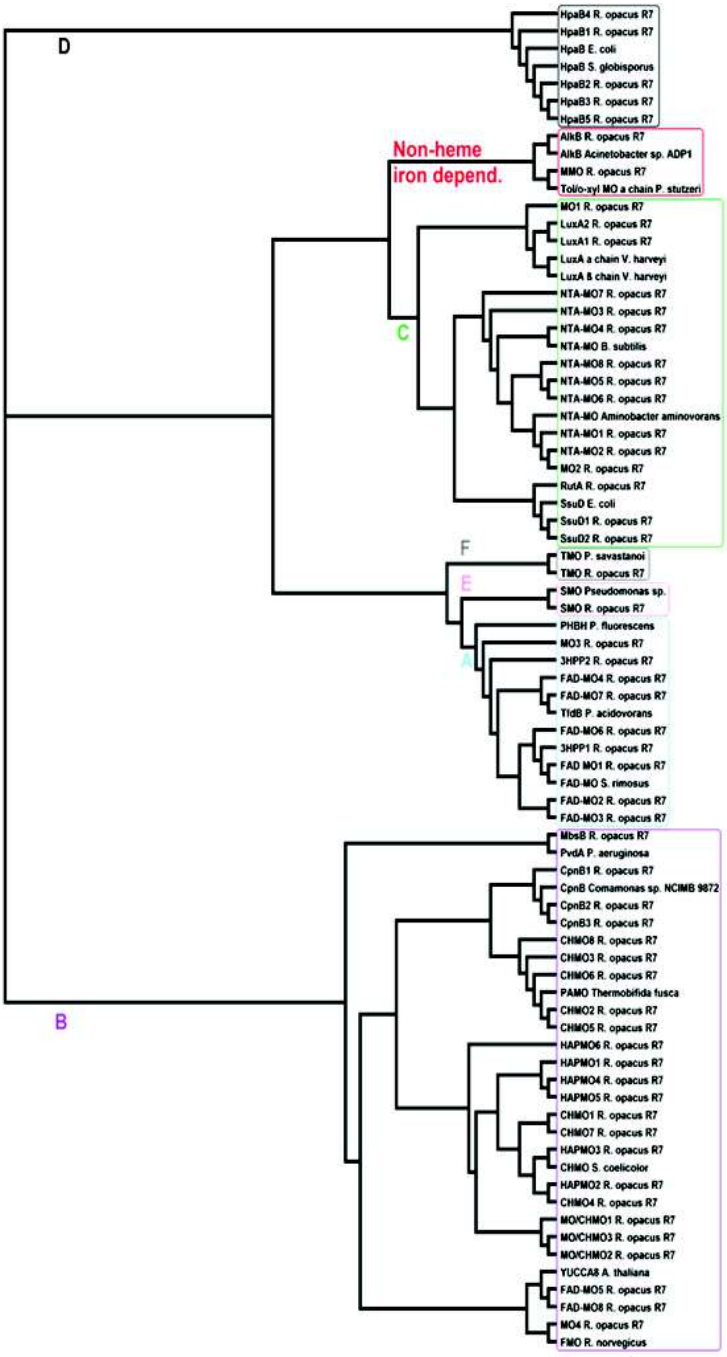

Figure 4.1 – Phylogenetic classification of *R. opacus* R7 monooxygenases protein with reference sequences.

The non-heme iron-dependent monooxygenase clade resulted interesting and a comparison of the monooxygenases included in this group was performed to evaluate how the nucleotide sequence was conserved. We used the nucleotide sequence of *alkB* gene of R7, that resulted similar to the alkB of *Acinetobacter* sp., for the comparison against NCBI database. The analysis revealed that this gene is conserved among Actinomycetes (nucleotide similarity around 99-79%), in particular among *Rhodoccoccus* species (Figure 4.2), in accordance to [11] .

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus opacus strain R7 alkB gene cluster, complete sequence	2278	2278	100%	0.0	100%	KJ573524.1
Rhodococcus opacus strain R7 sequence	2255	2255	100%	0.0	99%	CP008947.1
Rhodococcus opacus B4 DNA, complete genome	1868	1868	100%	0.0	94%	AP011115.1
Rhodococcus opacus strain 1CP, complete genome	1862	1862	100%	0.0	94%	CP009111.1
Rhodococcus jostii RHA1, complete genome	1862	1862	100%	0.0	94%	CP000431.1
Rhodococcus wratislaviensis strain IFP2016 alkane-1 monooxygenase (alkB1) gene, complete cds	1840	1840	100%	0.0	94%	JX569342.1
Rhodococcus opacus PD630, complete genome	1829	1829	100%	0.0	93%	CP003949.1
Rhodococcus wratislaviensis strain JCM 9689 alkane-1-monooxygenase (alkB) gene, partial cds	1081	1081	48%	0.0	99%	KF500475.1
Rhodococcus wratislaviensis strain V alkane 1-monooxygenase (alkB) gene, partial cds	1026	1026	52%	0.0	95%	KT248549.1
Nocardia cyriacigeorgica GUH-2 chromosome complete genome	948	948	94%	0.0	82%	FO082843.1
Rhodococcus opacus strain JCM 9703 alkane-1-monooxygenase (alkB) gene, partial cds	942	942	48%	0.0	95%	KF500464.1
Rhodococcus wratislaviensis alkB gene for alkane hydroxylase, partial cds, strain: NBRC 100605	922	922	44%	0.0	97%	LC107468.1
Rhodococcus percolatus strain JCM 10087 alkane-1-monooxygenase (alkB) gene, partial cds	920	920	48%	0.0	94%	KF500465.1
Rhodococcus koreensis strain JCM 10743 alkane-1-monooxygenase (alkB) gene, partial cds	915	915	51%	0.0	93%	KF500460.1
Rhodococcus imtechensis strain JCM 13270 alkane-1-monooxygenase (alkB) gene, partial cds	915	915	48%	0.0	94%	KF500457.1
Rhodococcus jostii strain JCM 11615 alkane-1-monooxygenase (alkB) gene, partial cds	881	881	51%	0.0	92%	KF500459.1
Nocardia nova SH22a, complete genome	880	880	92%	0.0	81%	CP006850.1
Rhodococcus sp. 1BN partial alkB gene putative alkane hydroxylase	785	785	44%	0.0	93%	AJ401611.1
Rhodococcus percolatus alkB gene for alkane hydroxylase, partial cds, strain: NBRC 100626	782	782	43%	0.0	93%	LC107465.1
Rhodococcus opacus alkB gene for alkane hydroxylase, partial cds, strain: NBRC 100624	782	782	43%	0.0	93%	LC107464.1
Rhodococcus equi 103S chromosome	771	771	90%	0.0	79%	FN563149.1
Rhodococcus opacus B4 plasmid pROB01 DNA, complete sequence	771	771	94%	0.0	79%	AP011116.1
Uncultured bacterium clone OTU 18 alkane monooxygenase-like (alkB) gene, partial sequence	739	739	41%	0.0	93%	EU307683.1

Figure 4.2 – Result of *R. opacus* R7 *alkB* alignment against NCBI database.

The nucleotide sequence of *MMO* gene of R7 (in Chapter 5 it will be named as *prmA*) resulted similar to the tol/*o*-xyl monooxygenase of *P. stutzeri*, and it was compared against NCBI database. The analysis revealed that this gene is conserved among *Rhodoccoccus* species (nucleotide identity around 96-87%) and also among *Gordonia*, *Mycobacterium*, and *Amycolatopsis* species (nucleotide identity around 84-81%) (Figure 4.3).

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus opacus strain R7 sequence	2045	2045	100%	0.0	100%	CP008947.1
Rhodococcus wratislaviensis strain IFP2016 propane monooxygenase hydroxylase beta subunit (prmAb) gene, complete cds	1807	1807	100%	0.0	96%	JX569353.1
Rhodococcus opacus strain 1CP, complete genome	1801	1801	100%	0.0	96%	CP009111.1
Rhodococcus jostii RHA1, complete genome	1796	1796	100%	0.0	96%	CP000431.1
Rhodococcus opacus PD630, complete genome	1779	1779	100%	0.0	96%	CP003949.1
Rhodococcus sp. WB1, complete genome	1258	1258	100%	0.0	87%	CP015529.1
Rhodococcus aetherivorans strain IcdP1, complete genome	1253	1253	100%	0.0	87%	CP011341.1
Amycolicicoccus subflavus DQS3-9A1, complete genome	1070	1070	100%	0.0	84%	CP002786.1
Mycobacterium sp. djl-10, complete genome	1014	1014	95%	0.0	84%	CP016640.1
Gordonia terrae strain 3612, complete genome	946	946	97%	0.0	83%	CP016594.1
Gordonia sp. KTR9, complete genome	946	946	92%	0.0	83%	CP002907.1
Mycobacterium smegmatis genome assembly NCTC8159, chromosome : 1	941	941	98%	0.0	82%	LN831039.1
Mycobacterium smegmatis strain INHR2, complete genome	935	935	98%	0.0	82%	CP009496.1
Mycobacterium smegmatis strain INHR1, complete genome	935	935	98%	0.0	82%	CP009495.1
Mycobacterium smegmatis str. MC2 155, complete genome	935	935	98%	0.0	82%	CP009494.1
Mycobacterium smegmatis str. MC2 155, complete genome	935	935	98%	0.0	82%	CP001663.1
Mycobacterium smegmatis str. MC2 155 putative thymidilate kinase (MSMEG1875) gene, complete cds	935	935	98%	0.0	82%	DQ866849.1
Mycobacterium sp. JS623, complete genome	931	931	98%	0.0	82%	CP003078.1
Mycobacterium smegmatis str. MC2 155, complete genome	931	931	98%	0.0	82%	CP000480.1
Nocardioides dokdonensis FR1436, complete genome	920	920	98%	0.0	82%	CP015079.1
Mycobacterium goodii strain X7B, complete genome	918	918	98%	0.0	82%	CP012150.1
Rhodococcus sp. WMMA185, complete genome	893	893	100%	0.0	81%	CP017014.1
Amycolatopsis methanolica 239, complete genome	893	893	97%	0.0	82%	CP009110.1

Figure 4.3 - Result of *R. opacus* R7 *prmA* alignment against NCBI database.

4.3.3 Phylogenetic classification of dioxygenases of *Rhodococcus opacus* R7

A preliminary RAST annotation identified 83 potential dioxygenases. Considering catalytic subunits and reference sequences of dioxygenases putatively involved in the aromatic compounds degradation, phylogenetic analysis revealed that most part of these putative dioxygenases clusterized into two major clades: one clade with dioxygenases putatively involved in upper pathways and another clade of dioxygenases putatively involved in peripheral pathways of organic compounds (Figure 4.4). The catalytic subunits of dioxygenases putatively involved in the degradation of BTEX compounds, polycyclic aromatic hydrocarbons, such as naphthalene, anthracene and biphenyls, clusterized in the same clade (blue box), in particular the catalytic subunit of ethylbenzene dioxygenases of *R. jostii* RHA1 and of *Rhodococcus* sp. DK17 involved in *o*-xylene degradation clusterized near R7 ethylbenzene dioxygenase (red box). Considering dioxygenases putatively involved in peripheral pathways, this analysis revealed the presence of several 1,2 gentisate dioxygenases and several 4-hydroxyphenylpyruvate dioxygenases.

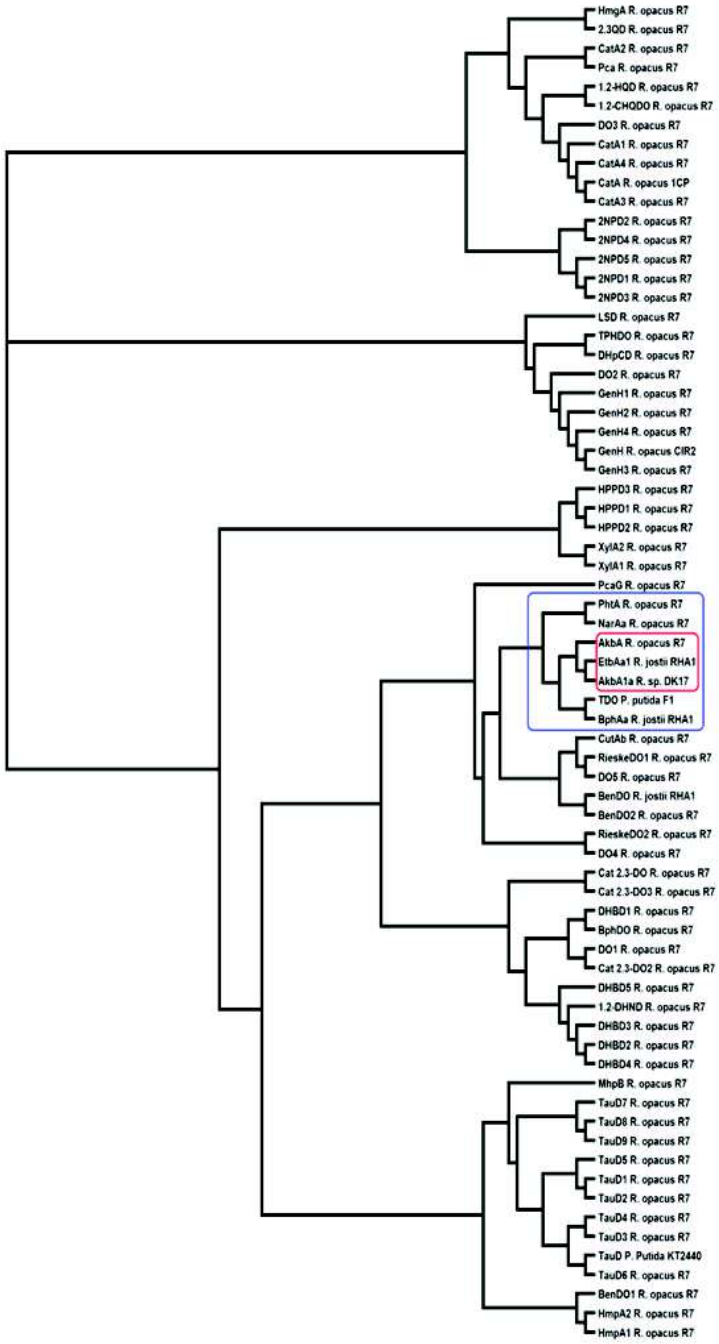

Figure 4.4 – Phylogenetic classification of *R. opacus* R7 dioxygenases protein with reference sequences.

The R7 ethylbenzene dioxygenases (AkbA) and the phthalate (PhtA) dioxygenases resulted interesting, therefore their nucleotide sequence was compared against NCBI database to evaluate the level of conservation. R7 *akbA* doesn't appear so well conserved among *Rhodococcus* species (figure 4.5), indeed in NCBI database the sequences is similar only to *etbA1 R. jostii* RHA1 and *akbA Rhodococcus* sp. DK17 with a nucleotide identity around 85%. Conversely, R7 *phtA* appears better conserved among *Rhodococcus* species (nucleotide identity around 99-79%) and *Mycobacterium* species (nucleotide identity around 79%) (figure 4.6).

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus opacus strain R7 plasmid pPDG5, complete sequence	2444	2444	100%	0.0	100%	CP008952.1
Rhodococcus sp. TFB megaplasmid pTFB1 tetralin degradation gene cluster, complete sequence; and TetR family putative regu	2438	4877	100%	0.0	99%	FJ183471.1
Rhodococcus sp. DK17 plasmid pDK2, partial sequence	1369	2733	100%	0.0	85%	AY502075.1
Rhodococcus jostii RHA1 plasmid pRHL2, complete sequence	1363	2727	100%	0.0	85%	CP000433.1
Rhodococcus sp. RHA1 ebdA1, ebdA2, ebdA3, etbD2 genes for ethylbenzene dioxygenase alpha subunit, ethylbenzene dioxyger	1363	1363	100%	0.0	85%	AB120956.1
Rhodococcus sp. RHA1 etbA1, etbA2, etbC, bphD, bphE2, bphF2 genes, comeplete cds	1363	1363	100%	0.0	85%	AB120955.1
Rhodococcus sp. RHA1 gene for aromatic ring hydroxylation dioxygenase C2, complete cds	1363	1363	100%	0.0	85%	AB048708.1
Rhodococcus sp. RHA1 gene for aromatic ring hydroxylation dioxygenase C, complete cds	1363	1363	100%	0.0	85%	AB048707.1

Figure 4.5 - Result of *R. opacus* R7 *akbA* alignment against NCBI database

Description	Max score	Total score	Query cover	E value	Ident	Accession
Rhodococcus opacus strain R7 plasmid pPDG2 sequence	2660	2660	100%	0.0	100%	CP008949.1
Rhodococcus sp. TFB PhtR (phtR), PhtAa (phtAa), PhtAb (phtAb), PhtB (phtB), PhtAc (phtAc), PhtAd (phtAd), PhtC (phtC), and PI	2654	2654	100%	0.0	99%	DQ007994.1
Microbacterium sp. J-1 phthalate dioxygenase large subunit (phtA) gene, partial cds	1144	1144	51%	0.0	95%	KT359341.1
Rhodococcus sp. DK17 plasmid pDK2 putative oph operon transcriptional regulator (ophR') gene, complete cds; oph operon,	1068	1068	96%	0.0	81%	EF494237.1
Rhodococcus jostii RHA1 plasmid pRHL2, complete sequence	1062	1062	96%	0.0	81%	CP000433.1
Rhodococcus jostii RHA1 plasmid pRHL1, complete sequence	1062	1062	96%	0.0	81%	CP000432.1
Rhodococcus sp. RHA1 plasmid pRHL2 padR2, padAa2, padAb2, padB2, padAc2, padAd2, padC2, orf11 genes, complete cds	1062	1062	96%	0.0	81%	AB154537.1
Rhodococcus sp. RHA1 plasmid pRHL1 padR1, padAa1, padAb1, padB1, padAc1, padAd1, padC1, orf8, orf9, orf10 genes, co	1062	1062	96%	0.0	81%	AB154536.1
Rhodococcus sp. RHA1 gene for aromatic ring hydroxylation dioxygenase E, complete cds	1062	1062	96%	0.0	81%	AB048709.1
Rhodococcus sp. DK17 plasmid pDK3, partial sequence	1038	1038	95%	0.0	80%	AY502076.1
Terrabacter sp. DBF63 plasmid pDBF1 genes, complete and partial cds	859	859	90%	0.0	79%	AP008980.1
Terrabacter sp. DBF63 ISTesp2 and phthalate catabolic gene cluster (phtA1, phtA2, ORF11, phtB, phtA3, phtA4, phtC, phtR), co	859	859	90%	0.0	79%	AB084235.1
Mycobacterium sp. 05-1390, complete genome	846	846	88%	0.0	79%	CP003347.1
Mycobacterium chimaera strain Mycobacterium chimaera MC045 genome assembly, chromosome: 1	845	1211	88%	0.0	79%	LT703505.1
Mycobacterium chimaera strain AH16, complete genome	845	845	88%	0.0	79%	CP012885.2
Mycobacterium indicus pranii MTCC 9506, complete genome	830	830	88%	0.0	79%	CP002275.1
Mycobacterium intracellulare MOTT-02, complete genome	824	824	88%	0.0	79%	CP003323.1
Mycobacterium intracellulare 1956, complete genome	819	819	88%	0.0	79%	CP009499.1
Mycobacterium intracellulare ATCC 13950, complete genome	819	819	88%	0.0	79%	CP003322.1
Rhodococcus sp. p52 plasmid pDF02, complete sequence	813	813	82%	0.0	79%	CP016820.1

Figure 4.6 - Result of *R. opacus* R7 *phtA* alignment against NCBI database

4.4 Conclusions

A phylogenetic analysis of *R. opacus* R7 oxygenases putatively involved in xenobiotic metabolism was performed in order to have clear information about all genes that could be involved in these metabolisms. *R. opacus* R7 monooxygenases and dioxygenases were clusterized and interesting genes were found. Thereafter, for the four interesting genes identified, *alkB*, *prmA*, *akbA*, *phtA,* the level of sequence conservation was evaluated.

4.5 References

[1] Pazmiño TDE, Winkler M, A. Glieder, Fraaije MW (2010) Monooxygenases as biocatalysts: classification, mechanistic aspects and biotechnological applications. J Biotechnol 146:9-24.

[2] Xu F (2005) Applications of oxidoreductases: recent progress. Ind Biotechnol 1:38-50.

[3] Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, et al. (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics 6:1-42.

[4] Van Berkel WJH, Kamerbeek NM, Fraaije MW (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. J Biotechnol 124:670-689.

[5] Wallar BJ, Lipscomb JD (1996) Dioxygen activation by enzymes containing binuclear non-heme iron clusters. Chem Rev 96:2625-2657.

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

[7] Fetzner S (2012) Ring-cleaving dioxygenases with a cupid fold. Appl Environm Microbiol 78:2505-2514.

[8] De Visser S, Kumar D (2011) Iron-containing enzymes versatile catalysis of hydroxylation reactions in nature. Royal Society of Chemistry.

[9] Kweon O, Kim S-J, Baek S, Chae J-C , Adjei MD, Baek D-H, et al., (2008) A new classification system for bacterial Rieske non-heme iron aromatic ring-hydroxylating oxygenases. BMC Biochemistry 9:11.

[10] Efimov I, Basran J, Thackray SJ, Handa S, Mowat CG and Raven EL (2011) Structure and reaction mechanism in the heme dioxygenases. Biochemistry 50:2717-2724.

[11] Táncsics A, Benedeka T, Szoboszlay S, Veres PG, Farkas M, Máthéc I, et al. (2015) The detection and phylogenetic analysis of the alkane 1-monooxygenase gene of members of the genus *Rhodococcus.* System Appl Microbiol 38:1-7.

Chapter 5

5. Genetic Aspects related to Hydrocarbon Degradation in *Rhodococcus opacus* R7

The aim of the present part of the thesis is to characterize different genes and metabolic pathways involved in the degradation of different hydrocarbons to identify "marker" sequences through *Rhodococcus* spp. genome analysis. In the previous chapters enzymatic classes of proteins involved in these metabolisms and interesting gene sequences were identified through clusterization of all oxygenases of *Rhodococcus opacus* R7. This chapter focuses on the identification of gene clusters involved in the degradation of different hydrocarbons, such as aliphatic, mono-aromatic and polycyclic aromatic hydrocarbons through the comparison of *R. opacus* R7 genome sequences with reference strain sequences and literature retrievals.

5.1 Introduction

The main role in soil degradation is performed in particular by bacteria belonging to the *Pseudomonas* and *Rhodococcus* genus. A common feature of the aerobic rhodococci is the presence of many types of monooxygenases and dioxygenases. It is clear that an immense gene diversity in cells that are physiologically robust could confer an advantage [1].

The number of genes that encodes for proteins with homologues in closely and distantly related species is growing rapidly. Often the homologues perform the same function. Many of these genes encode products involved in biodegradation of xenobiotic compounds [2].

5.1.1 *n*-Alkanes aerobic metabolism

Aerobic alkane degraders use O_2 as a reactant for the activation of the alkane molecule. The alkane-activating enzymes, which are monooxygenases, overcome the low chemical reactivity of the hydrocabon by generating reactive oxygen species. Oxidation of methane, which is a special case, renders methanol that is subsequently transformed to formaldehyde and then to formic acid. This compound can be converted to CO_2, or assimilated for biosynthesis of multicarbon compounds either by the ribulose monophosphate pathway, or by the serine pathway, depending on the microorganism considered. In the case of *n*-alkanes containing two or more carbon atoms, aerobic degradation usually starts by the oxidation of a terminal methyl group to render a primary alcohol, which is further oxidized to the corresponding aldehyde, and finally converted into a fatty acid. Fatty acids are conjugated to CoA and further processed by β-oxidation to generate acetyl-CoA. In some cases, both ends of the alkane molecule are oxidized through ω-hydroxylation of

fatty acids at the terminal methyl group (the ω position), rendering an ω-hydroxy fatty acid that is further converted into a dicarboxylic acid and processed by β-oxidation [3,4]. Subterminal oxidation of *n*-alkanes has also been reported [5]. The product generated, a secondary alcohol, is converted to the corresponding ketone, and then oxidized by a Baeyer-Villiger monooxygenase to render an ester. The ester is hydrolysed by an esterase, generating an alcohol and a fatty acid (Figure 5.1). Both terminal and subterminal oxidation can coexist in some microorganisms. While oxidation of fatty alcohols and fatty acids is common among microorganisms, activation of the alkane molecule requires enzymes that are much less widespread, and that can belong to different families.

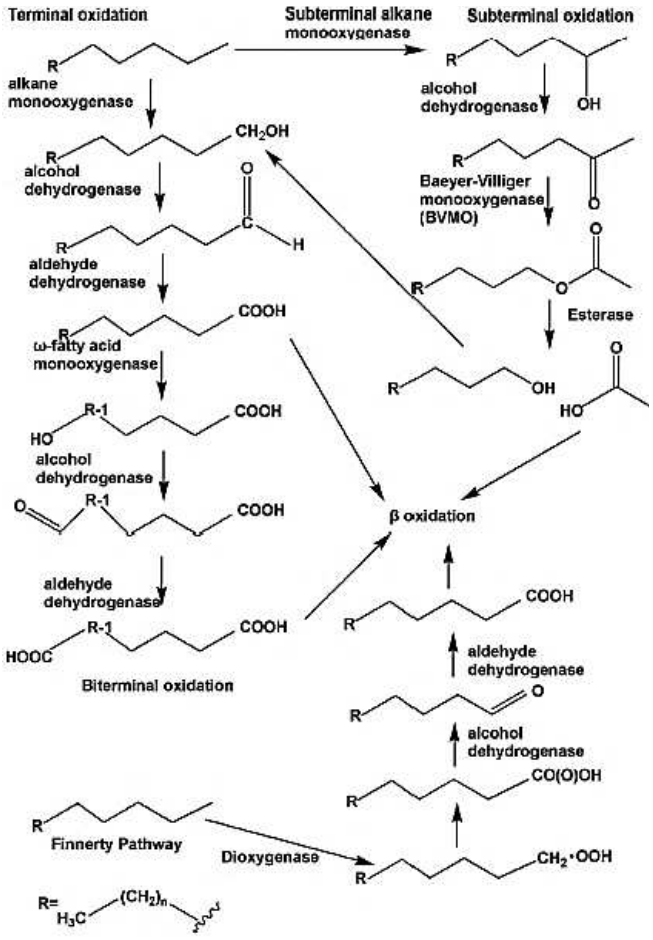

Figure 5.8 – Aerobic pathways of *n*-alkanes degradation by terminal and subterminal oxidation

5.1.2 Aromatic hydrocarbon aerobic metabolism

A wide range of natural and xenobiotic aromatic compounds are aerobically catabolized by bacteria via a large variety of upper or peripheral pathways resulting in a limited number of central intermediates. These common metabolites are further degraded through a few central pathways to

finally provide intermediates of the citrate cycle. Such arrangements for the catabolism of aromatic compounds were found in both the longer studied pseudomonads [6] and in the recently characterized *R. jostii* RHA1 [7]. According to the *R. jostii* RHA1 genome sequence, 26 peripheral pathways and 8 central pathways are involved in the catabolism of aromatic compounds [7]. Within the peripheral pathway, the aromatic compound is modified in a number of steps, including the action of monooxygenase or dioxygenase resulting in the formation of a dihydroxylated benzene ring. The main resulting metabolites are catechol and protocatechuate with hydroxyl groups at positions 1,2 and gentisate with hydroxyl groups at positions 1,4. Initial oxidative attack on BTEX converting the compound into a catechol structure and the cleavage of the catechol structure are the key steps in aerobic BTEX degradation (Figure 5.2) [8].

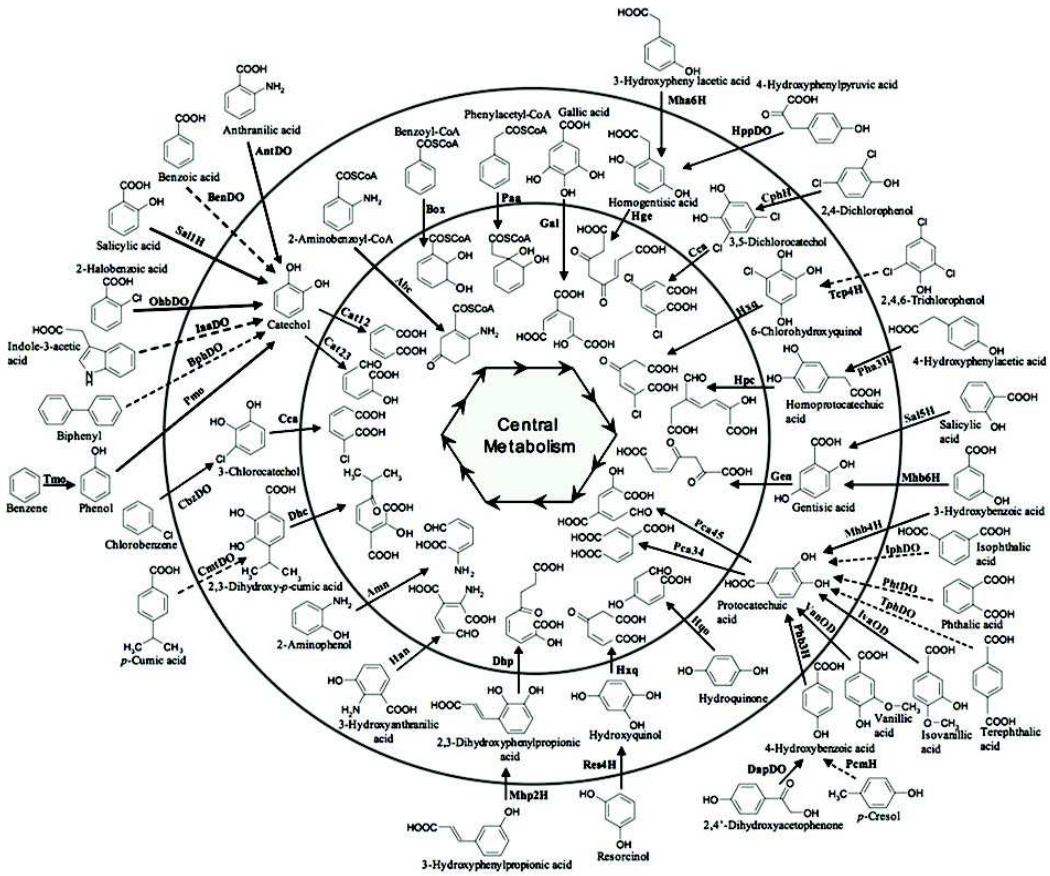

Figure 5.9 - Overview of upper and peripherial pathways for bacterial catabolism of aromatic compounds [9].

Five primary biochemical pathways have been characterized for BTEX degradation by aerobic bacteria (TOD, TOL, TOM, TBU and TMO). No single pathway can fully degrade all of the BTEX compounds individually, but most of the enzyme pathways can transform at least 3 or 4 of the compounds [10]. Based on the first step of formation of the catechols, these can be classified into

ring hydroxylation pathway and alkyl substituent oxidation pathway. In ring hydroxylation pathways (toluene dioxygenation (TOD), toluene *o*-monooxygenation (TOM), toluene *m*-monooxygenation (TBU), and toluene *p*-monooxygenation (TMO)), dioxygenases attack the aromatic rings and produce dihydroxyl compounds, while monooxygenases attack the aromatic rings and produce arene oxides as intermediates, which are highly unstable and further converted to phenols (Figure 5.3) [11].

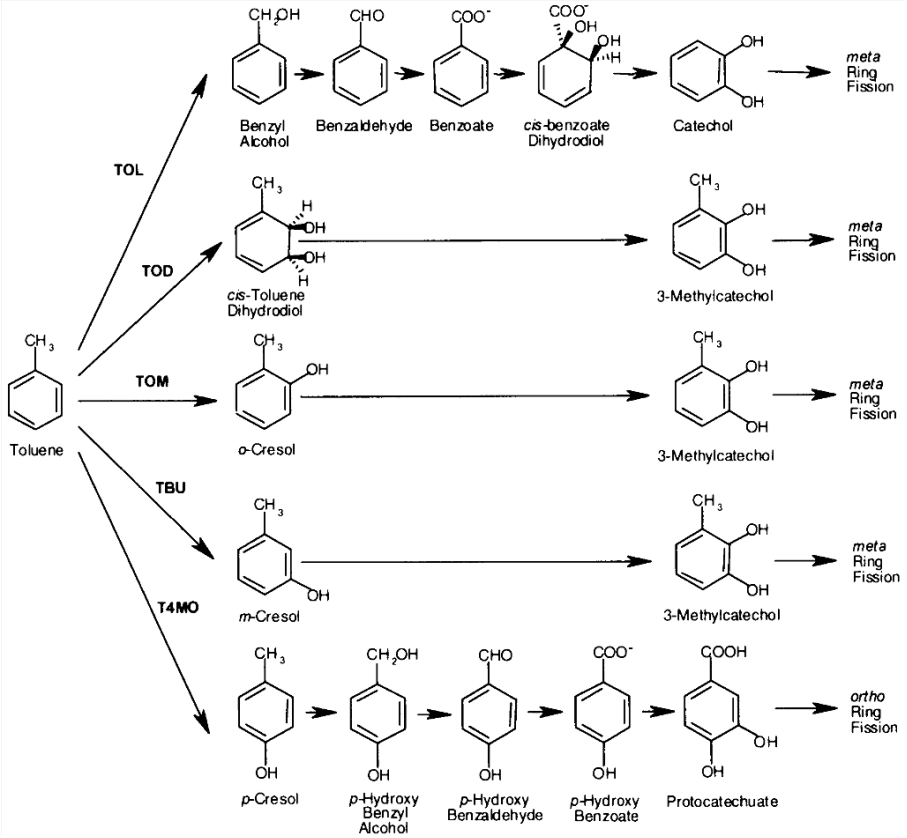

Figure 5.3 - Diversity of aerobic toluene degradation pathways.

The TOL pathway was discovered in a plasmid harbored by *Pseudomonas putida* mt-2, which initiates catabolism by oxidizing toluene at the methyl group [12]. The TOD pathway was identified in *P. putida* F1, which uses toluene dioxygenase to add two oxygen atoms to the ring [13]. The TOM pathway was found in *Burkholderia cepacia* G4 (formerly known as *P. cepacia* G4), which uses toluene *ortho*-monooxygenase in the initial attack to form *o*-cresol [14]. The TBU pathway occurs in *B. pickettii* PK01 (formerly known as *P. pickettii* PK01) which uses toluene *meta*-monooxygenase to form *m*-cresol [15]. The T4MO pathway was identified in *P. mendocina* KR1, which uses toluene *para*-monooxygenase to form *p*-cresol [16]. The expression of these pathways is not exclusive to the five archetypes mentioned above [17].

Monooxygenases are capable of performing another monooxygenation on these phenols and convert them to dihydroxyl compounds, such as catechols [18]. The alkyl substituent oxidation pathway (TOL) involves the oxidation of the alkyl substituent to produce dihydroxyl compounds. The dihydroxyl compounds produced subsequently undergo ring fission to produce intermediates of the Kreb's cycle [19].

Pseudomonas stutzeri OX1 is able to grow on a wide spectrum of aromatics, including phenol, cresols, and dimethylphenols, but also on nonhydroxylated molecules such as toluene, *o*-xylene, and benzene [20]. The degradation of the most recalcitrant compound *o*-xylene by *Pseudomonas stutzeri* OX1 was identified by Baggi et al. (1987) [21]. They observed that the monooxygenase attacked the aromatic ring (similar to TOM pathway) twice producing 3,4-dimethyl catechol, which was further cleaved through the *meta* pathway [19]. Two different monooxygenases have been found in the genome of *P. stutzeri* OX1, phenol hydroxylase (PH) [22] and toluene *o*-xylene monooxygenase (ToMO) [23,24]. ToMO is endowed with a broad spectrum of substrate specificity and with the ability to hydroxylate more than a single position of the aromatic ring in two consecutive monooxygenation reactions. Thus, ToMO is able to oxidize *o*-, *m*-, and *p*-xylene, 2,3- and 3,4-dimethylphenol, toluene, cresols, benzene, naphthalene, and styrene. Moreover, ethylbenzene, trichloroethylene, 1,1-dichloroethylene, chloroform, and tetrachloroethylene can also be oxidized by the enzyme [20].

5.1.3 Polycyclic aromatic hydrocarbon aerobic metabolism

The biodegradation of PAHs has been observed under both aerobic and anaerobic conditions. Under aerobic conditions the pathways initiate the biodegradation of PAHs by introducing both atoms of molecular oxygen into the aromatic nucleus, the reaction is catalyzed by a multicomponent dioxygenase which consists of a reductase, a ferredoxin and an iron–sulfur protein [25]. Since bacteria initiate PAH degradation by the action of intracellular dioxygenases, the PAHs must be taken up by the cells before degradation can take place. Bacteria most often oxidize PAHs to cis-dihydrodiols by incorporation of both atoms of an oxygen molecule. The cis-dihydrodiols are further oxidized, first to the aromatic dihydroxy compounds (catechols) and then channeled through the *ortho*- or *meta*-cleavage pathways [26, 27].

The low bioavailability of PAH may have prevented the evolution of suitable enzymatic pathways in soil bacteria. According to Perry (1979) [28], recalcitrant compounds generally do not serve as

growth substrates for any single microbial organism, but are thought to be oxidized in a series of steps by consortia of microbes.

Environmental bacterial isolates often degrade only a narrow range of PAHs [29] and patterns of simultaneous degradation of PAH mixtures are complex. Cometabolism and inhibition phenomena have been investigated by Bouchez et al. (1995) [29]. With mixtures of two individually degradable PAHs, either preferential degradation of one PAH or reduced degradation rates of both PAHs indicating metabolic competition were observed.

A variety of non-actinomycete bacteria, such as *Pseudomonas* spp., *Burkholderia cepacia*, *Sphingomonas* spp., *Flavobacterium* spp., *Cycloclasticus* spp., and *Stenotrophomonas* spp., have been investigated and it has been proven that they can metabolize a large number of PAHs, including naphthalene, phenanthrene, benz[a]anthracene, chrysene, fluoranthene, fluorine, anthracene, and pyrene in a soil environment.

Several actinomycetes and bacteria, such as *Mycobacterium* spp., *Gordonia* spp., and *Rhodococcus* spp., were isolated from various hydrocarbon-contaminated soils and each uses fluoranthene, pyrene, and chrysene as sole carbon and energy sources [30].

5.2 Materials and Methods

5.2.1 Bacterial strains, growth conditions and general procedures

Rhodococcus opacus R7 was grown on either Luria-Bertani LB or in a M9 mineral medium. When the strain grew on mineral medium [31], it was supplemented with naphthalene or *o*-xylene or different *n*-alkanes in an atmosphere saturated with these compounds, as the only carbon and energy source. *Rhodococcus erythropolis* AP was maintained on mineral medium in a saturated atmosphere of diesel fuel at 30°C. *Escherichia coli* DH5α was grown on LB medium at 37°C. To select *E. coli* and *Rhodococcus* trasformants, antibiotics (ampicillin at 100 μg/ml for *E. coli* and tetracycline at 25 μg/ml for *Rhodococcus*) were added in the cultural medium. Genomic DNA of *R. opacus* R7 was extracted from R7 cells by cell suspension in Tris-EDTA in presence of 2 mg of lysozime at 37°C for 30 min. After addition of SDS 10% and proteinase K (10 mg/ml) the extract was incubated for 1 h. The extract was precipitated by 0.1 volume sodium acetate 3 M and after centrifugation, the DNA was isolated and purified by phenol/chloroform extraction of the surnatant. DNA manipulation, enzymatic digests, ligation and PCR reactions were performed using standard molecular techniques [32]. For recombinant plasmids extraction NucleoSpin Plasmid Kit

by Machery and Nagel was used according to the manufacturer's instruction. Growth experiments on R7 strain were performed in presence of *n*-alkanes (1 g/l) ranging from C6 to C36 in 20 ml of M9 with OD_{600} 0.1. To determine abiotic loss, uninoculated flasks were also prepared. The flasks were incubated at 30°C for a maximum of 72 h. Every 24 h a flask was sacrified to evaluate *n*-alkanes degradation by determination of the residual *n*-alkane by gas chromatography-mass spectrometry (GC-MSD) after extraction with *n*-hexane. Bacterial growth was determined by OD_{600} measurements.

5.2.2 Identification of genetic aspects for xenobiotic degradation pathways

The degradation pathways of several hydrocarbon compounds and putative gene clusters were predicted in *R. opacus* R7 by manual protein alignment analysis. This analysis was performed using proteins previously identified in the strain or extracted from sequences of reference strains reported in databases, such as *Rhodococcus* sp. BCP1 and *R. jostii* RHA1. The identified gene clusters were partially described and compared with those from other *Rhodococcus* strains reported in literature. The NCBI pipeline and RAST server [33] were used for genome predictions and comparisons.

5.2.3 Identification and cloning of the alk gene cluster

The isolation of the complete coding region of the *alkB* gene cluster from *R. opacus* R7 was obtained from purified genomic DNA of the strain by PCR-amplification. To identify the genomic region, two primers (F1, 5′-AAGGCCATGGGGCGTTAGAGCACCGCAGCTAAT-3′ and R1, 5′-ACAGCATATGACCTAGCGGGCGGCCGCGACCCG-3′) were designed on the basis of the more conserved sequences between the *alkB* gene cluster of different bacteria belonging to the *Rhodococcus* genus deposited in Genbank DataBase. It was carried out using the following program: 95°C for 3 min; 95° C for 30 sec, 68°C for 45 sec, 72°C for 5 min, for 35 cycles; and 72°C for 3 min. DNA fragments were purified from agarose gel by the NucleoSpin Extraction II Kits by Machery and Nagel. The eluted fragment was sequenced by automated sequencing (Eurofins MWG). The 3.0 Kb fragment, containing the *alkB* gene cluster, was then amplified by PCR using the primers *Nco*I-*alk*R7 5′-AAGGCCATGGACGTGACGACGTCGGATATC-3′ and *Nde*I-*alk*R7 5′-ACAGCATATGACCTAGCGGGCGGCCGCGAC-3′ in order to generate the *Nco*I-*Nde*I ends. The 3.0 Kb fragment was cloned as PCR product into the pDrive vector (Qiagen) into *E. coli* DH5α by electroporation and the recombinant clones were selected on LB agar supplemented with

ampicillin 100 µg/ml and IPTG 1 mM and 5-bromo-4-chloro-3-indolyl-β-D galactopyranoside (X-Gal) 40 µg/ml. White colonies were isolated and plasmid (pDrive-*alk*R7) of the recombinant clones was extracted and verified with digestion *Nco*I/*Nde*I.

The *alkB* insert was ligated as *Nco*I-*Nde*I fragment into a shuttle-vector *E. coli-Rhodococcus*, pTipQT1 [34]. The ligation mixture was used to transform *E. coli* DH5α by electroporation with standard procedures [32] and the recombinant clones were selected on LB agar supplemented with ampicillin (100 µg/ml) at 37°C. The isolated recombinant plasmid (pTipQT1-*alk*R7) was used as shuttle-vector to transform others *Rhodococcus* spp. strains by electroporation. Plasmids were introduced into strains of *Rhodococcus* spp. by electroporation using a Gene Pulser II (Biorad, Italia) set at 2.50 kV, 600 Ω, 25 µF [35] in presence of about 1 µg DNA. Cells were plated on LB supplemented with tetracycline 25 µg/ml and grown at 30°C for 3-4 days. Recombinant strain *R. erythropolis* AP (pTipQT1-*alk*R7) was used for biodegradation experiments in presence of *n*-alkane C_{12} to evaluate the activity of the *alkB* system.

5.2.4 Chromatographic analyses and kinetic modeling

Residual *n*-dodecane was extracted from the inoculated and uninoculated flasks using *n*-hexane. It occurred by strong manual agitation (in test tubes), after washing of the flasks to take away any residual *n*-alkane on their wall. The suspension was centrifuged at 4000 rpm for 10 min and after 10 min of settling, 2 ml were drawn from organic phase. The organic phase extracted was diluted 1:100 for GC injection. Residual *n*-alkane was analysed by GC-MSD using a Technologies 6890N Network GC System, interfaced with 5973 Network Mass Selective Detector (MSD) (Agilent Technologies). A ZB-5MS capillary column was used (5% diphenyl-95% dimethylpolysiloxane 60 m x 0,25 mm, 0,25 µm; Alltech). Analyses were carried out in split-less injection mode using helium as carrier gas at 99,99%. The injector port was set at 250°C. The oven temperature was programmed from 60°C for 3 min, then 15°C min^{-1} to 280°C, holding this temperature for 10 min. All analyses were carried out with three replicates, and the mean values obtained are reported.

The residual *n*-dodecane concentrations as a function of time were modeled according to the Monod-type kinetic models as reported by Simkins and Alexander (1984) [36]. For first-order kinetic model, the natural logarithm of substrate concentration *vs* time was plotted and the slope of the least-squares-regression line was determined.

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

Plasmid pTNR-TA [37] was transferred into *R. opacus* strain R7 by electroporation using a Gene Pulser II (Biorad, Italia) set at 2.50 kV, 600 Ω, 25 μF [35] in presence of maximum 1 μg DNA in a 2 mm-gap electrocuvette (Biorad, Italia). 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 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 1% (wt/vol) malate 1M.

The transposon-induced mutants obtained were tested on 13 different M9 mineral medium agar plates with the following carbon sources as only carbon and energy source at final concentration of 1 g/l: naphthalene, salicylate, gentisate, naphthalene dihydrodiol, naphthol, *o*-xylene, 2,3-dimethylphenol, 3,4-dimethylphenol, toluene, *m*-cresol, benzene, phenol, 2 dimethylbenzylalcohol.

5.2.6 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 [38]. Insertions of IS*1415* into the genomes of these mutants were confirmed by PCR using primers (Table 5.1). 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) [39].

Table 5.2 – List of oligonucleotides used for Two-Step gene walking PCR amplification

Oligonucleotide name	sequence (5′ – 3′)
Walking_thio_1	GGAAAAGGACTGCTGTCGCTGCC
i-pTipThio-70-rev	CAAGGGGAAGTCGTCGCTCTCTGG
pTNR884-for	TTGGTAGCTCTTGATCCGGCAAAC
Walking_3-for	AACAACTGGCCGCCACC

5.3 Results and Discussion

Based on the genome of *R. opacus* R7 and its annotation, gene clusters involved in the degradation of different organic/xenobiotic compounds were investigated in comparison with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1. Four chemical categories of compounds were considered including: 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.

Aliphatic hydrocarbons and cycloalkanes degradation

The presence of catabolic genes involved in the growth on short-, medium- and long-chain *n*-alkanes was investigated in *R. opacus* R7 [40]. The *alkB* gene cluster has been identified as 3 kb DNA region and the sequence analysis revealed four consecutive ORFs homologues to the *alk* gene cluster components: *alkB* encoding for an alkane monooxygenase, *rubA* encoding for a rubredoxin, *rubB* encoding for a second rubredoxin, and *rubred* encoding for a rubredoxin reductase.

These data were compared to the genome sequences of *Rhodococcus* sp. BCP1 strain well characterized for the ability to grow on different *n*-alkanes, in particular on short-chain *n*-alkanes [41, 42]. BCP1 showed two copies of this gene (*alkB1* and *alkB2)* within the chromosome (Figure 5.4 panel A). 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, in addition, the *tetR* gene coding for a regulator. In BCP1 genome, an additional 1161-bp *alkB* gene (*alkB2*) was identified. This gene was not organized in operon with *rubA/B* and *rubred* genes but it was associated with genes coding for a putative esterase and a long-chain-fatty-acid-CoA ligase (EC 6.2.1.3), 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 (Table 5.2) with those coded by *alk*B cluster of R7 and with the ones of the most known bacteria belonging to the *Rhodococcus* genus [7, 43], such as *R. jostii* RAH1 or *R. opacus* PD630.

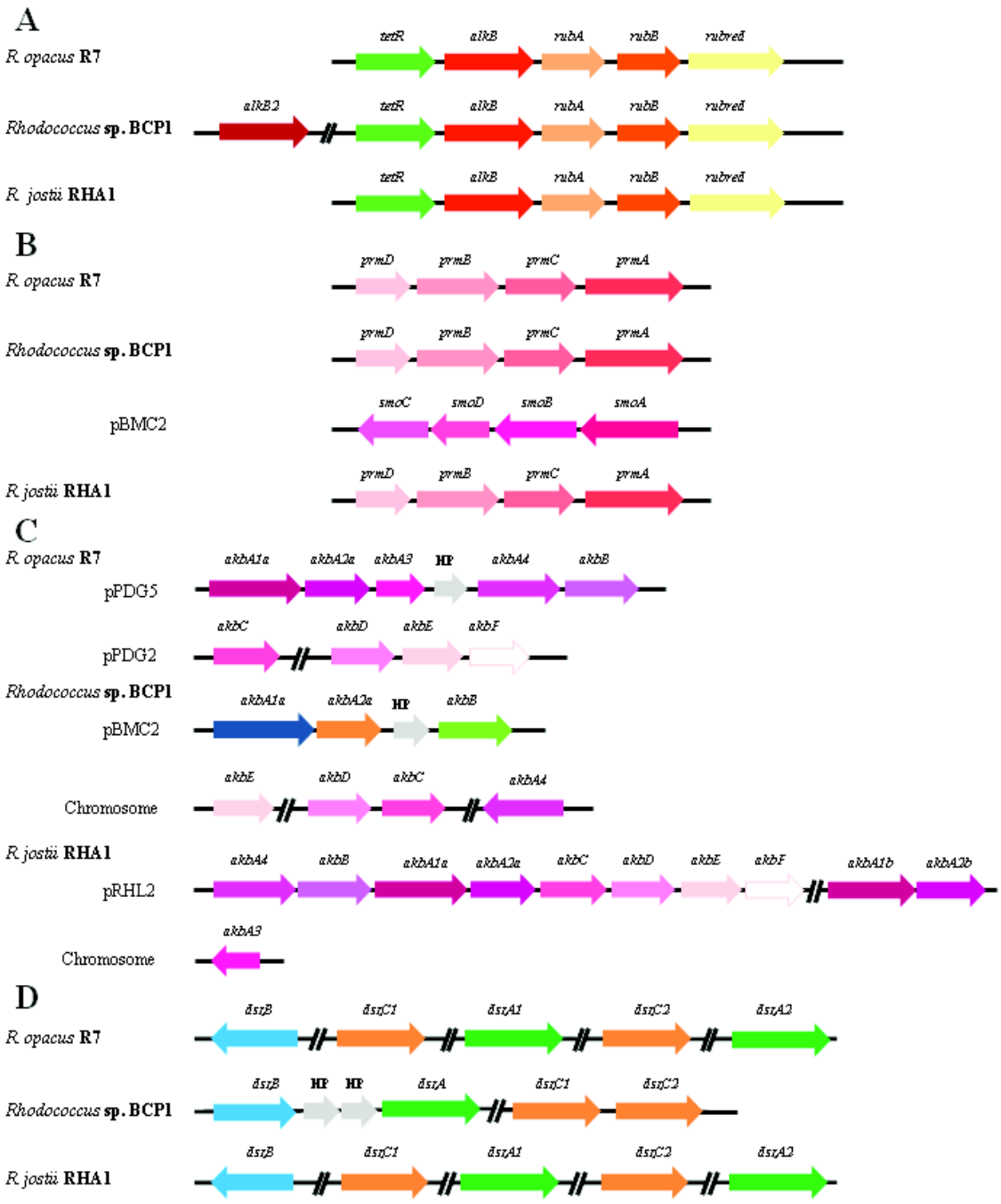

Figure 5.4 - Comparison of gene clusters from R7 genome correlated to aliphatic and aromatic hydrocarbon degradation. Comparative organization of genetic determinants for the tested aliphatic and aromatic hydrocarbons in *R. opacus* R7 with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1 as reference strains. Predicted genes and their orientation are shown by arrow. (A) *alk* gene cluster; (B) *prm* and *smo* gene cluster (*smo* gene cluster was organized in *smoA* encoding a methane monooxygenase component A, *smoB* encoding a methane monooxygenase component B, *smoD* and *smoC* encoding a methane monooxygenase regulatory protein and a methane monooxygenase component C, respectively); (C) *akb* gene cluster; (D) dsz gene cluster. When not specified, it means that genes were located on chromosome. Genes with unknown or hypothetical functions were reported as HP. Double slash indicates a distances between two genes more than 1 kb within the same plasmid or chromosome.

Two *prm* gene clusters were identified and analyzed in BCP1 (Figure 5.4 panel B). The first *prm* gene cluster, constituted by the *prmA,C,B,D* genes, 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) (*smoA,B,D,C* genes), was described to be involved in the growth of BCP1 on a wide range of short-chain *n*-alkanes [42].

Table 5.2 - Comparison of predicted genes and proteins of *alk* and *prm* gene cluster of *R. opacus* R7 and comparison with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1 homologous proteins.

Gene	Homologous protein	Function	R7 vs BPC1 (aa identity)	R7 vs RHA1 (aa identity)	Position in genome	Accession Number
alkB/alkB1	AlkB	Alkane monooxygenase	80%	91%	chrom.	AIA09965.1
alkB2	AlkB2	Alkane monooxygenase	/	/	chrom.	/
rubA	RubA	Rubredoxin	79%	98%	chrom.	AIA09966.1
rubB	RubB	Rubredoxin	84%	97%	chrom.	AIA09967.1
rubred	RubRed	Rubredoxin reductase	60%	91%	chrom.	AIA09968.1
prmA	PrmA	Methane monooxygenase component A alpha chain	98%	97%	chrom.	AII03499.1
prmC	PrmC	Methane monooxygenase component C	85%	94%	chrom.	AII03498.1
prmB	PrmB	Methane monooxygenase component A beta chain	88%	97%	chrom.	AII03497.1
prmD	PrmD	Methane monooxygenase regulatory protein	89%	98%	chrom.	AII03496.1

The comparison of R7 genome sequences with BCP1 *prm* gene cluster 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% (Table 5.2). Despite of the high sequence similarity between the two *prm* clusters, no growth of R7 in presence of gaseous *n*-alkanes was observed [40]. R7 strain showed also a considerable number of genes coding for P450 monooxygenases, 23 predicted coding sequences. We couldn't exclude that these putative enzymes are also involved in the aliphatic hydrocarbon degradation.

Aromatic hydrocarbons degradation

The ability of R7 to grow on different aromatic hydrocarbons, including BTEX compounds, was tested. The majority of the *Rhodococcus* genomes sequenced, as in the case of *R. jostii* RHA1 and *Rhodococcus* sp. DK17, were characterized for the ability to grow on BTEX and in some cases also for their genetic organization [44, 7]. In *R. jostii* RHA1, the *akb* gene cluster was identified as the cluster involved in the metabolism of alkylbenzenes. It contains two genes (designated *akbA1a* and *akbA1b*, approximately 6 kb apart from each other) encoding an oxygenase component large subunit, each one followed by genes (designated *akbA2a* and *akbA2b*) encoding 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. Comparing the *akb* gene cluster of RHA1 with R7, the same gene cluster divided in two parts was found and each part was allocated on two plasmids (Figure 5.4 panel C): *akbA1a, akbA2a, akbA3, akbA4, akbB* genes were found on the pPDG5 plasmid; while *akbCDEF* genes, putatively coding for a complete *meta*-cleavage pathway, constituted by a *meta*-cleavage dioxygenase (AkbC), a *meta*-cleavage hydrolase product (AkbD), a hydratase (AkbE), and an aldolase (AkbF), were identified on the pPDG2 plasmid with high amino acid identity (around 60%) (Table 5.3). An homologous sequence of *akb* cluster was also found on the chromosome and on the pPDG4 plasmid, but they showed a protein identity around 35%.

RAST analysis of BCP1 genome showed that coding sequences homologous to the *akb* genes were found in pBMC2 plasmid: *akbA1a* was found similar to a gene annotate as a large subunit *nah*/*bph* dioxygenase with an amino acid identity of 36%, *akbA2a* similar to a gene annotated as a biphenyl dioxygenase beta subunit (EC 1.14.12.18) with an amino acid identity of 36% and *akbB* similar to a gene annotated as a 2,3-dihydroxy-2,3-dihydro-phenylpropionate dehydrogenase (EC 1.3.1.-) with 47% protein identity.

These data indicated a redundancy of genes involved in the metabolism of the aromatic compounds in R7 strain, 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. In BCP1 only one copy of a putative cluster involved in the degradation of aromatic hydrocarbons was found.

Table 5.3 - Comparison of predicted genes and proteins of *akb* and *dsz* gene clusters of *R. opacus* R7 and comparison with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1 homologous proteins.

Gene	Homologous protein	Function	R7 vs BCP1 (aa identity)	R7 vs RHA1 (aa identity)	Position in genome	Accession Number
akbA1a	AkbA1a	Ethylbenzene dioxygenase large subunit	36%	92%	pPDG5	AII11493.1
akbA2a	AkbA2a	Ethylbenzene dioxygenase small subunit	43%	84%	pPDG5	AII11492.1
akbA3	AkbA3	Ethylbenzene dioxygenase ferredoxin	/	69%	pPDG5	CP008952.1
akbA4	AkbA4	Ferredoxin reductase	39%	81%	pPDG5	AII11490.1
akbB	AkbB	Dihydrodiol dehydrogenase	47%	85%	pPDG5	AII11489.1
akbC	AkbC	2,3-Dihydroxybiphenyl 1,2-dioxygenase	36%	87%	pPDG2	AII11058.1
akbD	AkbD	2-Hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	35%	67%	pPDG2	AII11051.1
akbE	AkbE	2-Hydroxypenta-2,4-dienoate hydratase	56%	60%	chrom.	AII11050.1
akbF	AkbF	4-Hydroxy-2-oxovalerate aldolase	/	63%	pPDG2	AII11049.1
dszA1	DszA1	Dibenzothiophene desulfurization	44%	98%	chrom.	AII08556.1

		enzyme				
dszA2	DszA2	Dibenzothiophene desulfurization enzyme	44%	97%	chromosome	AII03608.1
dszB	DszB	Possible ABC sulfonate transporter	76%	98%	chromosome	AII06125.1
dszC1	DszC1	Probable dibenzothiophene desulfurization enzyme	67%	97%	chromosome	AII08748.1
dszC2	DszC2	Probable dibenzothiophene desulfurization enzyme	84%	96%	chromosome	AII08273.1

In presence of dibenzothiophene (DBT) R7 strain showed the highest activity 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 catalyses the stepwise S-oxidation of DBT, first to dibenzothiophene 5-oxide (DBTO) and then to dibenzothiophene 5,5-dioxide (DBTO2); DszA catalyses the conversion of DBTO to 2-(2'-hydroxyphenyl) benzene sulphinate (HBPSi-) and DszB catalyses the desulphation of HBPSi- to give HBP and sulphonate [45]. A dsz gene cluster was found in R7 and BCP1 genomes (Figure 5.4 Panel D). 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% (Table 5.3). Sequence analysis revealed also two *dszC* genes (*dszC1* and *dszC2*) predicted to code for dibenzothiophene desulphurization enzymes, and not so far we found only one copy of the *dszB* gene encoding an ABC sulfonate transporter protein.

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 with a percentage reported in Table 5.3; 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.

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 [46].

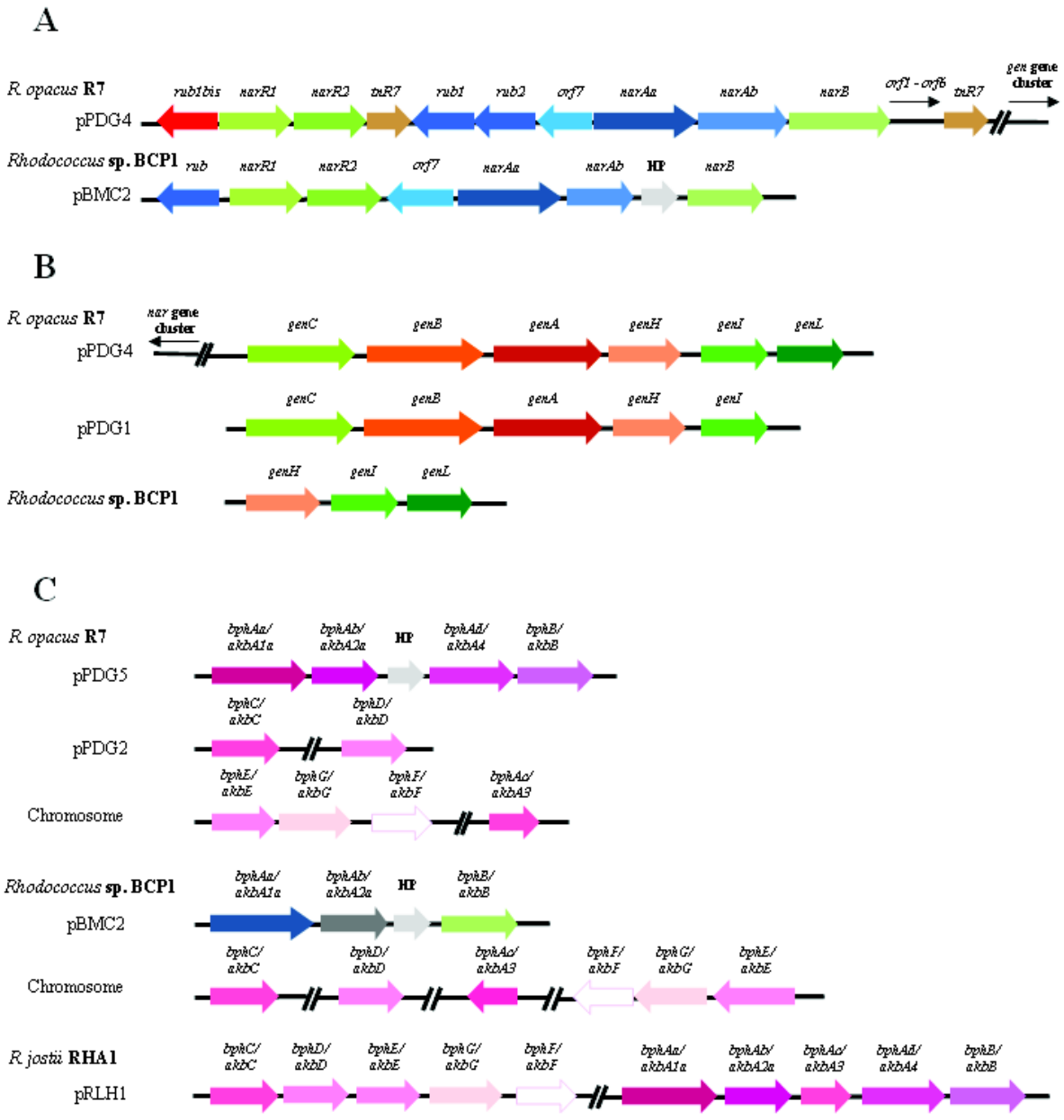

Figure 5.5 - Comparison of gene clusters from R7 genome correlated to polycyclic aromatic hydrocarbon degradation. Comparative organization of genetic determinants for naphthalene and biphenyl (as reference compounds of PAH and putative intermediates) in *R. opacus* R7 with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1 as reference strains. Predicted genes and their orientation are shown by arrow. (A) *nar* gene cluster; (B) *gen* gene cluster; (C) *bph* gene cluster organization. When not specified, it means that genes were located on chromosome. Genes with unknown or hypothetical functions were reported as HP. Double slash indicates a distances between two genes more than 1 kb within the same plasmid or chromosome.

The whole genome sequence analysis pointed out that *nar* gene cluster is located in pPDG4 plasmid in R7 strain and that the same cluster was found in pBMC2 in *Rhodococcus* sp. BCP1 (Figure 5.5 panel A). This genomic region identified in pBMC2 is the same identified as *akb* gene cluster in BCP1. In Table 5.4 protein identities are reported.

The whole genome sequence analysis revealed the presence of *orf7* within the cluster, as in R7 strain, while any of the other six CDSs identified in R7 were not found in BCP1 strain. An homologous *nar* gene cluster was not found in the reference strain *R. jostii* RHA1. The lower naphthalene catabolic pathway was previously investigated in R7 and it was hypothesized that the *gen* gene cluster was located far from the *nar* region as no amplification of the middle region was obtained. Genome sequence analysis confirmed these data (Figure 5.5 panel B).

Table 5.4 - Comparison of predicted genes and proteins of *nar* gene cluster of *R. opacus* R7 and comparison with *Rhodococcus* sp. BCP1 homologous proteins.

Gene	Homologous protein	Function	R7 vs BCP1 (aa identity)	Position in genome	Accession Number
rub1bis	Rub1bis	Rubredoxin	/	pPDG4	DQ846881
narR1	NarR1	Regulator of GntR family	87%	pPDG4	ABH01023.1
narR2	NarR2	XylR-like regulator protein	94%	pPDG4	ABH01024.1
rub1/rub	Rub1/Rub	Rubredoxin	88%	pPDG4	ABH01026.1
rub2	Rub2	Rubredoxin	/	pPDG4	ABH01027.1
orf7	Orf7	Sterol-binding domain protein/ unknown	93%	pPDG4	ABH01028.1
narAa	NarAa	Naphthalene dioxygenase large subunit	90%	pPDG4	ABH01029.1
narAb	NarAb	Naphthalene dioxygenase small subunit	88%	pPDG4	ABH01030.1
narB	NarB	Cis-naphthalene dihydrodiol dehydrogenase	96%	pPDG4	ABH01031.1

Moreover, two copies of the gene clusters involved in naphthalene lower pathway were found in R7. 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 with high protein identity (around 80%), followed by *genL* gene that showed a protein identity of 48% with the homologous genes of R7. Instead, genes involved in salicylate degradation, *genA*, *B*, *C* were

found in different regions of the BCP1 chromosome showing a lower amino acid identity (Table 5.5).

Table 5.5 - Comparison of predicted genes and proteins of *gen* gene cluster of *R. opacus* R7 and comparison with *Rhodococcus* sp. BCP1 homologous proteins.

Gene	Homologous protein	Function	R7 vs BCP1 (aa identity)	Position in genome	Accession number
genC	GenC	Salicylate hydroxylase	/	pPDG1 pPDG4	AII11448.1 AII10777.1
genB	GenB	Salicylate CoA syntethase	/	pPDG1 pPDG4	AII11449.1 AII10778.1
genA	GenA	Salicylate CoA ligase	/	pPDG1 pPDG4	AII11450.1 AII10779.1
genH	GenH	Gentisate dioxygenase	96%	pPDG1 pPDG4	AII11451.1 AII10780.1
genI	GenI	3-Maleylpyruvate Isomerase	92%	pPDG1 pPDG4	AII11452.1 AII10781.1
genL	GenL	Unknown function	89%	pPDG4	AII11453.1

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 [47, 48]. In the biphenyl metabolic pathway, biphenyl 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). Thus, 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 (Figure 5.5 panel C). We found different genes of this cluster in R7 chromosome and in pPDG2 and pPDG5 plasmids. 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 (Table 5.6).

RAST analysis of BCP1 genome identified CDSs homologous to *bph* genes in pBMC2. In particular, only large and small subunits of a dioxygenase (*bphAa* and *bphAb* genes) and a dihydrobiphenyldiol dehydrogenase (*bphB* gene) were found with an amino acid identity around 40%; this coincides with the region identified as *akb* gene cluster and *nar* gene cluster. Also in BCP1, the genes encoding the ring-cleavage enzymes were not found near the cluster within the plasmid pBMC2; however, an homologous coding sequence was found in the chromosome with low protein identity (around 30%) (Table 5.6). 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. Indeed, we can attribute the same sequence to the *akb*, *nar* and *bph* gene clusters.

Table 5.5 - Comparison of predicted genes and proteins of *bph* gene cluster of *R. opacus* R7 and comparison with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1 homologous proteins

Gene	Homologous protein	Function	R7 vs BCP1 (aa identity)	R7 vs RHA1 (aa identity)	Position in genome	Accession Number
bphAa	BphAa	Biphenyl-2,3-dioxygenase α subunit	36%	37%	pPDG5	AII11493.1
bphAb	BphAb	Biphenyl-2,3-dioxygenase β subunit	36%	35%	pPDG5	AII11492.1
bphAc	BphAc	Biphenyl-2,3-dioxygenase, ferredoxin component	33%	31%	chromosome	AII08472.1
bphAd	BphAd	Biphenyl-2,3-dioxygenase, reductase	33%	33%	pPDG5	AII11490.1
bphB	BphB	Cis-2,3-dihydrobiphenyl-2,3-	47%	78%	pPDG5	AII11489.1

		diol dehydrogenase				
bphC	BphC	2,3-Dihydroxybiphenyl-1,2-dioxygenase	36%	35%	pPDG2	AII11058.1
akbD	AkbD	2-Hydroxy-6-oxohepta-2,4-dienoate hydrolase	35%	38%	pPDG2	AII11051.1
bphE	BphE	2-Oxopent-4-enoate hydratase	59%	66%	chromosome	AII03622.1
bphF	BphF	4-Hydroxy-2-oxovalerate aldolase	87%	89%	chromosome	AII03620.1
bphG	BphG	Acetaldehyde dehydrogenase	76%	74%	chromosome	AII03621.1

Naphthenic acids degradation

Considering the ability of the strain to grow on naphthenic acids, the putative gene clusters involved in this degradation were investigated. In literature, few metabolic studies are available on the biodegradation of naphthenic acids (NAs) based on cyclohexane ring i.e. cyclohexane carboxylic acid (CHCA); while none has been focused on cyclopentane ring i.e. cyclopentane carboxylic acid (CPCA). Although no genetic information was provided in these studies [49, 50], the metabolism of CHCA was described to follow two main routes: (i) aromatization of the cycloalkane ring to produce hydroxybenzoate before ring opening [49], or (ii) activation of cycloalkane ring as CoA thioester-derivative that is further degraded through β-oxidation steps [50]. Iwaki and co-workers (2005) 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. Iwaki et al. [49] demonstrated the involvement of *pobA* product in CHCA catabolism by *Corynebacterium* strain downstream of the formation of 4-hydroxybenzoate from CHCA through several oxidation steps. One gene homologous to *pob*A was identified in R7 genomes and its product was annotated as *p*-hydroxybenzoate hydroxylase in RAST server. Similarly to what found in *Corynebacterium* strain,

*pob*A gene was flanked by a gene coding for an *Icl*R-type transcriptional regulator. The genomic region including *pob*A in R7 included also the *alkB* gene cluster, *tet*R-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* was different from those of R7 and RHA1 and included a permease coding gene and several oxidoreductases (Figure 5.6).

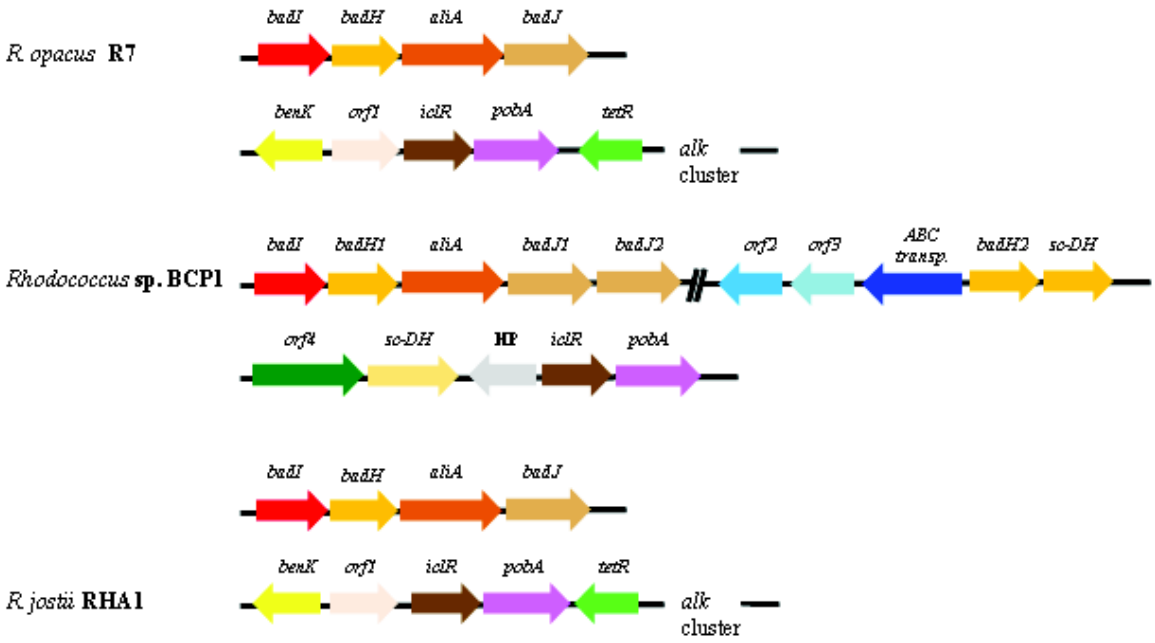

Figure 5.6 - Comparison of gene clusters from R7 genome correlated to carboxylated hydrocarbon degradations. Comparative organization of genetic determinants for naphthenic acids (as reference compounds of carboxylated hydrocarbons and putative intermediates) in R. opacus R7 with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1 as reference strain. Predicted genes and their orientation are shown by arrow. The following genes encode for: *benK*, benzoate transporter; *orf1*, O-antigen acetylase; *icl*R, transcriptional regulator *Icl*R family; *tet*R, transcriptional regulator, *tet*R family; *orf 2* and *orf 3*, permease; *ABC transp*, ABC transporter; *sc-DH*, probable short-chain dehydrogenase; *orf4*, permease. When not specified, it means that genes were located on chromosome. Genes with unknown or hypothetical functions were reported as HP. Double slash indicates a distances between two genes more than 1 kb within the same plasmid or chromosome.

The activation of the cycloalkane ring as Co-A thioester-derivative was firstly described during the anaerobic degradation of benzoic acid by *Rhodopseudomonas palustris* [51]. 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 [52]. These enzymes degrade CHCA by catalyzing the following reactions: (i) activation of cyclohexanecarboxylate as cyclohexanecarboxylate-CoA by a CoA-ligase (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 catalyses the the cyclohexane ring opening with the formation of pimelyl-CoA. On the basis of the amino acid identity percentage as compared with the 2-ketocyclohexane-carboxyl-CoA dehydrogenase (BadH) of *R. palustris*, two genes *(badH1 and badH2)* were found in BCP1 genome encoding two enzymes annotated as 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase. Only one gene homologous to *badH* 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 5.6, Table 5.6).

Table 5.6 - Comparison of predicted genes and proteins of naphthenic acids gene cluster of *R. opacus* R7 and comparison with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1 homologous proteins.

Gene	Homologous protein	Function	R7 vs BCP1 (aa identity)	R7 vs RHA1 (aa identity)	Position in genome	Accession Number
badI	BadI	Naphthoate synthase	89%	99%	chrom.	AII08541.1
badH1	BadH1	2-Hydroxycyclohexanecarboxyl-CoA dehydrogenase	83%	99%	chrom.	AII08542.1
badH2	BadH2	2-Hydroxycyclohexanecarboxyl-CoA dehydrogenase	/	/	/	/
aliA	AliA	Long-chain-fatty-acid-CoA ligase	71%	95%	chrom.	AII08543.1
badJ	BadJ	Acyl-CoA dehydrogenase	87%	99%	chrom.	AII08544.1
pobA	PobA	*p*-Hydroxybezoate hydroxylase	77%	98%	chrom.	AII08627.1

5.3.1 Genetic aspects related to aromatic peripheral pathways in *Rhodococcus opacus* R7

Considering the aromatic compounds that R7 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. Genes responsible for such catabolic pathways have been reported in several bacteria [9], including *R. jostii* RHA1. We performed a sequence comparison analysis with *Rhodococcus*. sp. BCP1 and *R. jostii* RHA1 to identify the genes predicted to be involved in these pathways in *R. opacus* R7. 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). RHA1 strain also showed several genes potentially involved in catechol catabolism. BCP1 genome presents only two catechol 1,2-dioxygenases and one catechol 2,3-dioxygenase on the chromosome. These aspects might be explained on the basis of the significant difference of their genome size. Two catechol dioxygenase genes, amongst those identified on R7 chromosome, were organized in cluster (Figure 5.7 Panel A). 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 second *cat* gene cluster identified in R7 lacked of *catC* gene; moreover, CatA2 (*catA2*) and CatB2 (*catB2*) were not found homologous to R7 genes (Table 5.7). *Rhodococcus* sp. BCP1 presented only one copy of *cat* gene cluster with the same organization of R7; BCP1 *cat* genes showed high similarity (70-90%) with those of R7 strain. The comparison of the first *cat* gene cluster of R7 with RHA1 *cat* genes revealed high similarity (96-98% aa identity), while the comparison of the second *cat* gene cluster of R7 with RHA1 *cat* genes revealed no significant similarity (Table 5.7).

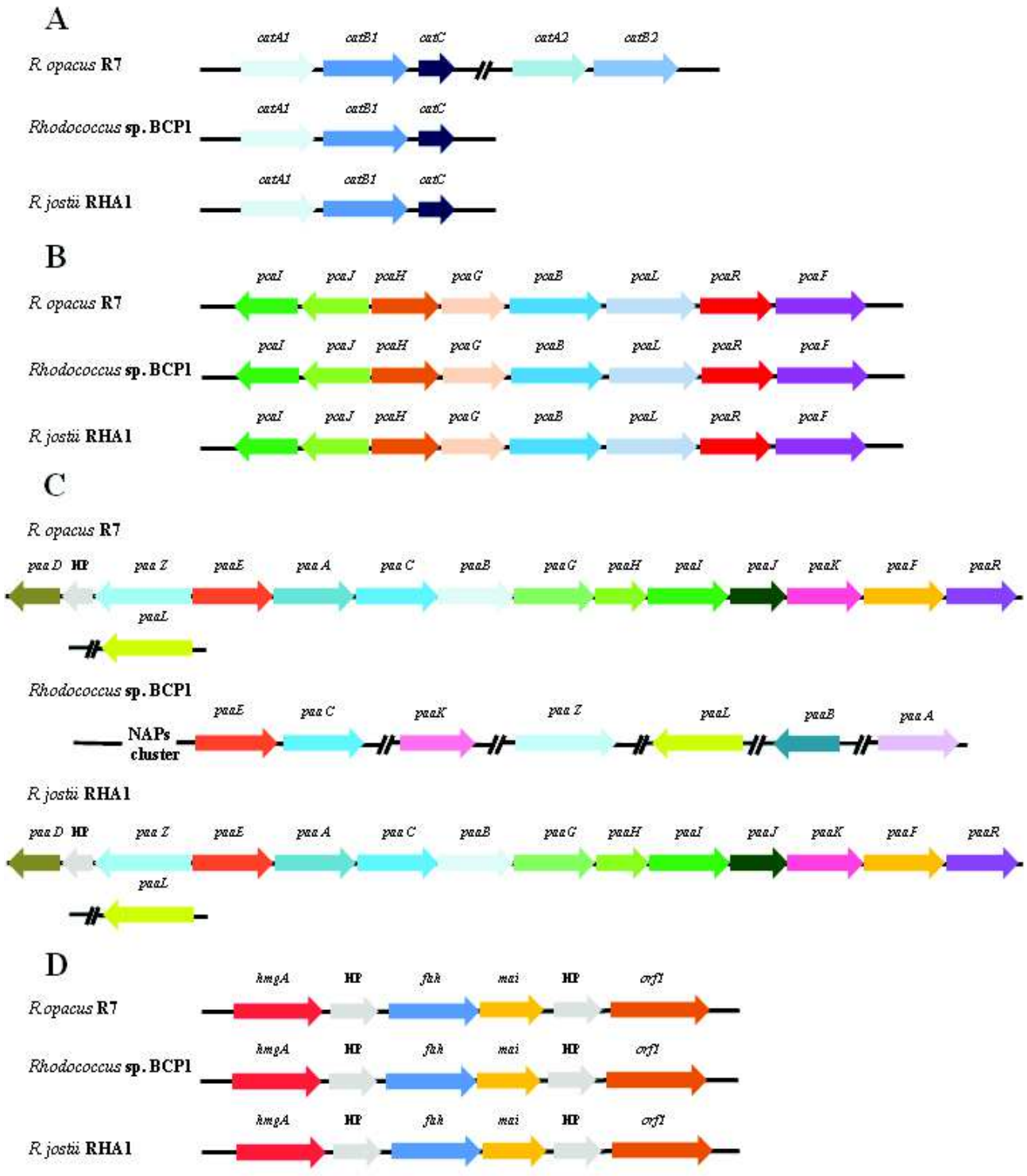

Figure 5.7 - Comparison of gene clusters from R7 genome correlated to xenobiotic peripheral pathways. Comparative organization of genetic determinants for xenobiotic peripheral pathways in *R. opacus* R7 with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1 as reference strains. Predicted genes and their orientation are shown by arrow. (A) *cat* gene cluster; (B) *pca* gene cluster; (C) *paa* gene cluster; (D) *hmg* gene cluster. When not specified, it means that genes were located on chromosome. Genes with unknown or hypothetical functions were reported as HP. Double slash indicates a distances between two genes more than 1 kb within the same plasmid or chromosome.

Both R7 and BCP1 genomes contain several genes potentially involved in protocatechuate catabolism (Figure 5.7 Panel B). 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 (Table 5.7). 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, *pcaJI* and *pcaHGBLRF*.

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 catalized by enzymes coded by the *paa* gene cluster [53]. 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 (Figure 5.7 Panel C). They also showed high percentage of similarity as reported in Table 5.7. On the contrary, the *paa* gene cluster was not conserved in *Rhodococcus* sp. BCP1 and few genes homologous to those of RHA1, are present without a co-localization.

R7 genomes analysis reported one homogentisate 1,2-dioxygenase (*hmgA*) showing 97% protein identity with RHA1 hmgA, while with BCP1 showed 85% identity the same as with RHA1 hmgA. The genetic organization of the *hmgA* flanking regions is quite similar between R7, BCP1 and RHA1, showing genes encoding an enoyl-CoA hydratase (EC 4.2.1.17), fumarylacetoacetase (EC 3.7.1.2) and a glutaryl-CoA dehydrogenase (EC 1.3.99.7). 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 5.7 Panel D; Table 5.7).

Table 5.7 - Comparison of predicted genes and proteins of *cat*, *pca*, *paa*, *hmg* clusters of *R. opacus* R7 with *Rhodococcus* sp. BCP1 and *R. jostii* RHA1 homologous proteins.

Gene	Homologous protein	Function	R7 vs BCP1 (aa identity)	R7 vs RHA1 (aa identity)	Position in genome	Accession Number
catA1	CatA1	Catechol 1,2 dioxygenase	68%	99%	chrom.	AII08813.1
catB1	CatB1	Muconate cycloisomerase	78%	98%	chrom.	AII08814.1
catC	CatC	Muconolactone isomerase	89%	96%	chrom.	AII08815.1
catA2	CatA2	Catechol 1,2 dioxygenase	/	/	chrom.	CP008947.1
catB2	CatB2	Muconate cycloisomerase	/	/	chrom.	AII05696.1
pcaI	PcaI	Succinyl-CoA 3-ketoacid-coenzyme A transferase subunit B	77%	99%	chrom.	AII09804.1
pcaJ	PcaJ	Succinyl-CoA 3-ketoacid-coenzyme A transferase subunit A	80%	100%	chrom.	AII09803.1
pcaH	PcaH	Protocatechuate 3,4-dioxygenase beta chain	71%	97%	chrom.	AII09802.1
pcaG	PcaG	Protocatechuate 3,4-dioxygenase alpha chain	57%	99%	chrom.	AII09801.1
pcaB	PcaB	3-Carboxy-cis,cis-muconate cycloisomerase	61%	98%	chrom.	AII09800.1
pcaL	PcaL	3-Oxoadipate enol-lactone hydrolase	63%	97%	chrom.	AII09799.1
pcaF	PcaF	Acetyl-CoA acetyltransferase	82%	98%	chrom.	AII09797.1
paaG	PaaG	Phenylacetate-CoA oxygenase, PaaG subunit	/	99%	chrom.	AII08339.1
paaH	PaaH	Phenylacetate-CoA oxygenase, PaaH subunit	/	100%	chrom.	AII08340.1
paaI	PaaI	Phenylacetate-CoA oxygenase, PaaI subunit	/	96%	chrom.	AII08341.1
paaJ	PaaJ	Phenylacetate-CoA oxygenase, PaaJ subunit	/	98%	chrom.	AII08342.1
paaK	PaaK	Phenylacetate-CoA oxygenase/reductase,	35%	96%	chrom.	AII08343.1

paaF	PaaF	PaaK subunit Phenylacetate-coenzyme A ligase PaaF	/	98%	chrom.	AII08344.1
paaE	PaaE	Acetyl-CoA acetyltransferase	74%	98%	chrom.	AII08335.1
paaA	PaaA	Enoyl-CoA hydratase	32%	97%	chrom.	AII08336.1
paaC	PaaC	3-Hydroxyacyl-CoA dehydrogenase	66%	95%	chrom.	AII08337.1
paaB	PaaB	Enoyl-CoA hydratase	35%	97%	chrom.	AII08338.1
paaZ	PaaZ	Aldehyde dehydrogenase	31%	97%	chrom.	AII08334.1
paaD	PaaD	Phenylacetic acid degradation protein, thioesterase	/	97%	chrom.	AII08332.1
paaL	PaaL	Acetate permease ActP (cation/acetate symporter)	42%	99%	chrom.	AII08348.1
hmgA	HmgA	Homogentisate 1,2-dioxygenase	85%	97%	chrom.	AII08874.1
fahA	FahA	Fumarylacetoacetase	79%	97%	chrom.	AII08876.1
mai	Mai	Enoyl-CoA hydratase/isomerase	79%	97%	chrom.	AII08877.1
orf1	Orf1	Long-chain-fatty-acid-CoA ligase	78%	97%	chrom.	AII08879.1

5.3.2 Functional analysis of alk gene cluster

R. opacus R7 was characterized for its ability to degrade several long- and medium-chain *n*-alkanes. The *alkB* gene cluster has been identified as 3 kb DNA region. This region has been isolated by PCR with primers designed on the basis of the alignment of conserved sequences from the *alkB* gene region from different *n*-alkane degrading *Rhodococcus* bacteria. From the sequence analysis, the covering region contained four consecutive ORFs homologues to the *alkB* gene cluster components described in 5.3 Results and Discussion paragraph, section Aliphatic hydrocarbons and cycloalkanes degradation.

The *alkB* gene cluster was isolated from R7 genomic DNA and cloned into the pDrive vector giving the plasmid pDrive-*alk*R7. The region was isolated as *Nco*I/*Nde*I fragment and cloned into the shuttle-vector *E. coli-Rhodococcus* pTipQT1. The recombinant plasmid pTipQT1-*alk*R7 was isolated from *E. coli* DH5a and transferred into *Rhodococcus erythropolis* AP. The expression of the *alkB*

gene cluster was verified under the control of the *Ptip/regulator* system with experiments with resting cells of *Rhodococcus erythropolis* AP(pTipQT1-*alk*R7) exposed to C12 were performed. The expression of the *alk* region was determined comparing the biodegradation kinetics of the recombinant strain *R. erythropolis* AP (pTipQT1-*alk*R7) with the wild type strain *R. erythropolis* AP. From kinetic analysis we can observe that the initial degradation rate was higher in the recombinant strain with respect to the wild type strain indicating a difference in the activity levels of the *alkB* system (Figure 5.8). This difference was confirmed by a statistical test for the comparison of the slopes of the regression lines at 95% significance level [40].

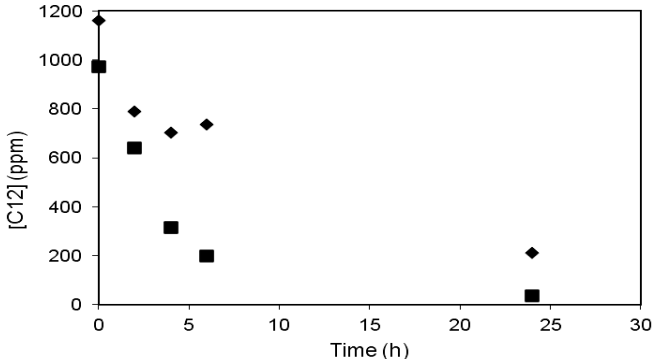

Figure 5.8 - Kinetic analyses of *n*-dodecane degradation in *R. erythropolis* AP (pTipQT1-*alk*R7) (square in the graph). Cell of the recombinant strain were grown on rich medium and then, after induction with thiostrepton, were collected and exposed to *n*-dodecane. For comparison, kinetics of *n*-dodecane degradation in *R. erythropolis* AP (triangle in the graph) without the cloned fragment expressing the *alk*B gene, is also showed.

The whole genome-sequencing analysis confirmed that the *alkB* gene cluster is the only copy of this gene present in this strain. This confirm the involvement of the *alkB* gene cluster in the *n*-alkane degradation in R7 strain.

5.3.3 Functional analysis of akb gene cluster

R. opacus R7 was characterized for its ability to degrade monoaromatic hydrocarbons as toluene, ethylbenzene and *ortho*-xylene isomer. The gene cluster analysis revealed the involvement of the *akb* gene cluster in monoaromatic hydrocarbons, including BTEX compounds. To prove the involvement of *akb* gene cluster in this metabolism, functional analyses were performed. Cloning and transforming the entire cluster to have heterologous expression in *E. coli* was difficult because literature data reported the need to change the GTG coding sequence for the expression of an alkane monooxygenase in *E. coli* [54]. Other authors were unable to show function heterologous

expression of rhodococcus genes in *Pseudomonas* or *E. coli* expression system, principally because the functional expression requires proper synthesis, correct folding, and proper assembly, and these are not always ensured for rhodococcal proteins [55]. Moreover generate deletion mutant or disrupted mutant of the entire gene cluster or part of the sequences was difficult to achieve because of the *akb* genes redundancy. For this reason *R. opacus* R7 was subjected to random transposon mutagenesis.

5.3.3.1 Random mutants analysis

R. opacus R7 was subjected to transposon mutagenesis employing the pTNR-TA vector [37, 56], which confers thiostrepton resistance on the recipient. This mutagenesis aimed to generate mutants with defective growth on 13 different carbon sources: naphthalene, salicylate, gentisate naphthalene dihydrodiol, naphthol, *o*-xylene, 2,3- dimethylphenol, 3,4-dimethylphenol, toluene, *m*-cresol, benzene, phenol, 2 dimethylbenzylalcohol.

Mutants	Criobox	NAH	SAL	GEN	Naphthol	dihydrodiol NAH	*o*-XYL	2,3DMP	3,4DMP	TOL	*m*-Cresol	BENZ	2-DMBA	PHE	MAL+Thio
M3Q10C8D	1	-	+	ND	+	+	+	+	+	+	+	+	+	+	+
M3U32C2H	2	-	+	ND	+	+	+	+	+	-	+/-	+/-	-	-	+
M3U38C7H	3	+	+	ND	+	B	+	+	+	+	+	+	+	+	+
M4V25C4A	4	-	+	p	+	+	+	-	+	+	+	+	+	+	+
M4V48C12A	5	+	-	ND	+	+	-	+	+	+/-	-	+	+	-	+
M4Z23C8B	6	-	+	ND	-	-	+	+	+	+	+	-	+	-	+
M4AF15C5G	7	+	+	ND	+	B	+	+	+	+	+	+	+	-*	+
M4AF45C11G	8	+	+	ND	+	+	+/-	-	-	+	-	-	-	-	+
M4AF56C12G	9	+	+	ND	+	-	+	+	+	-	-	+	+	-	+
M5AH49C11A	10	+	+	ND	+	B	+	+/-	+/-	+	+	+	+	-	+
M5Ai53C8B	11	+	+	ND	+/-	-	+/-	-	-	+	-	-	-	-	+
M5AL52C5C	12	+	-	ND	-	+	-	+	-	+	+	+	+	p	+
M5AQ21C2G	13	+	-	ND	-	+	+	-	+	-	-	+	-	+	+
M5AQ25C3G	14	+	+	ND	+	-	+	+	+	+	+	-	+	+	+
M5AQ35C8G	15	-	+	ND	-	+	+	+	+	+	+	+	-	-	+
M5AQ44C12G	16	+	+	ND	-	-	-	+	-	+	+	+	+	+	+
M5AQ50C1H	17	+	+	ND	-	+	-	+	-	+	+	+	+	+	+
M5AR11C5H	18	+	+	ND	+	+	+	-	+	-	+/-	+	+	+	+
M6AU26C12B	19	+	+	ND	+	B	+/-	+	+	+	+	+	+	+	+
M8(1)A15C3B	20	+	+	-	-	ND	+	+	+	+/-	+	-	+	-	+
M8(1)A18C6B	21	+	+	-	-	ND	-	-	+	+	-	+	-	-	+
M8(1)A34C10C	22	+	+	+	+	ND	+	-	+/-	-	+	+	+/-	+	+
M8(1)A45C9D	23	+	+	-	+	ND	+	-	+	+	+	-	+	-	+
M8(1)B27C3H	24	+	+	-	ND	+	-	ND	+/-	+	+	+	+	+	+
M8(1)B30C6H	25	+	+	-	ND	+	+	ND	-	+/-	+	-	+/-	-	+
M9(1)B56C8B	26	+	+	p	+	-	-	-	+	+	+	+	+	+	+
M9(1)B59C11B	27	-	+	+	-	-	+	-	-	+	+	+	-	-	+
M9(1)C7C7C	28	+	-	p	-	+	+	-	+	+	+	-	+	+	+
M9(1)C20C8D	29	+	+	-	+	+	+	+	+	-	+	-	+	+	+
M9(1)C24C12D	30	+	+	-	+	+	+	+	+	+/-	+	-	+/-	-	+
M10(2)A4C4A	31	+	+	-	+	+	-	+	+/-	+	-	+	+	+	+
M10(2)A6C6A	32	-	+	p	+	+	+	+	+	+	+	+/-	+	+	+
M10(2)B7C7F	33	+	+	-	+	+	+	+	+	+	+	-	-	-	+
M11(2)B40C4A	34	-	+		+	+	+	+	+	+	+	+	-	-	+
M11(2)C33C11D	35	+	+	-	+	-	+	+	+	+	+	-	+/-	-	+
M11(2)C35C1E	36	+	+	-	+	+	+	+	+	-	+	+	+	-	+
M12(3)A45C9D	37	-	+	+	+	B	+	+	+	+	+	+	+	-	+
M13(3)B31C7B	38	+ B	+	-	-	B	+	+	+/-	+	+	+	+	-	+
M15(4)B25C1A	39	+	+	p	+	B	+	+	+	+	+	+	+	-	+
M1A46C7A	40	+	+	-	+	B	Y	+	+	+	+	+	+	+	+
M1E38C1E	41	-	+	-	+	B	+	-	+	+	+	+	-	+	+
M1F53C7F	42	+	+	-	+	-	+	+	+	+/-	-	-	+/-	-	+

Figure 5.9 - Phenotypic characterization of 42 IS*1415* induced mutants of *R. opacus* R7 that showed a certain negative phenotype on different xenobiotics.

Among 3000 transposon-induced mutants screened, 42 showed a certain negative phenotype on different xenobiotics tested, 8 showed no growth on *o*-xylene (Figure 5.9) and 21 a weak growth on *o*-xylene (Figure 5.10) used as the sole carbon and energy source when cultivated on mineral medium agar plates.

Mutant	Criobox	NAH	SAL	NAFTOLO	NAH diDIOLO	O-XYL	2,3DMP	3,4DMP	TOL	m-CRESOLO	BENZ	2 DMBA	Phe	MAL+Thio
M3U34C4H	43	+	+	+	+	+/-	-	+	+	+	+/-	+	-	+
M5AL30C2C	44	+	+	+	-	+/-	-	+	+	-	+	+	-	+
M5AR52C10H	45	+	+	+	+	+/-	+	-	+	+	+	+	-	+
M9(1)B41C5A	46	+	+	ND	-	+/-	ND	+	+	+	+	-	-	+
M9(1)B57C9B	47	+	+	+	-	+/-	+/-	+	+	+	+	+	+	+
M10(2)A9C9A	48	+	+	+	+	+/-	-	+	+	+	+	+	-	+
M10(2)A58C10E	49	+	+	+	p+	+/-	+	+	+	+	+	+	-	+
M10(2)B26C2H	50	+	+	+	+	+/-	+	-	+	+	+	-	-	+
M11(2)C31C9D	51	+	+	+	+	+/-	-	+	+	+	+	+	-	+
M1E17C6D	52	+	+	+	-	+/-	-	+/-	+	+/-	+	+	+	+
M4AA34C8C	53	+	+	-	-	+/-	-	+	+	-	+	-	-	+
M5AL1C9B	54	+	-	+	-	+/-	+	+	+/-	+	+	+	-	+
M5AM1C6C	55	+	+	+/-	-	+/-	+	+	+/-	+	+	+	+	+
M5AO37C4E	56	+	+	+	-	+/-	+	-	+	-	+	+	-	+
M6AU26C12B	57	+	+	+	M	+/-	+	+	+	+	+	+	+	+
M10(2)A13C1B	58	+	+	+	+	+/-	+	+	+	+	+	+	-	+
M10(2)A16C4B	59	+	+	-	-	+/-	-	+	+	+	+	+	-	+
M10(2)B10C10F	60	+	+/-	-	-	+/-	-	+	+	+	+	+	-	+
M10(2)B23C11G	61	+	+/-	+	-	+/-	+	+/-	+	-	+	+	-	+
M11(2)C50C8F	62	+	+/-	-	-	+/-	-	-	+/-	-	+	+/-	-	+
M1F19C2F	63	+	+	+/-	-	+/-	-	+	+	+	+	+	+	+

Figure 5.10 - Phenotypic characterization of 21 IS*1415* induced mutants of *R. opacus* R7 that showed a weak phenotype on *o*-xylene.

The analysis of pTNR insertion sites at molecular level was performed following the two-step gene walking method [38]. The preliminary results of sequence analysis and homologies obtained from the GenBank database are summarized and graphically illustrated in table 5.8 and figure 5.11

Table 5.8 - Genotypic characterization of IS*1415* induced mutants of *R. opacus* R7 with a weak growth on *o*-xylene.

Mutant	Nucleotide identity with *R. opacus* R7	Insertion locus of IS1415
M9(1)B57C9B	99%	between *akbA1* and *akbS* genes on pDG5
M10(2)A58C10E	93%	between *akbA1* and *akbS* genes on pDG5
M10(2)B26C2H	85%	*akbS* gene on pDG5
M1E17C6D	98%	between *akbA1* and *akbS* genes on pDG5
M6AU26C12B	99%	between two mobile elements downstream *akbS* gene on pDG5

The nucleotide sequence of a putative product deduced from the pTNR-TA insertion locus in three of the mutants with a weak growth on *o*-xylene, M9(1)B57C9B, M10(2)A58C10E and M1E17C6D, exhibited respectively 99%, 93% and 98% nucleotide identity to the region between *akbA1* and *akbS* genes on pDG5 of *R. opacus* R7, which are involved in the expression of the *akb* gene cluster

(Figure 5.11). The same figure 5.11 also provides the array of the coding DNA sequence (CDS) including the transposon insertion loci. The analysis on the insertion loci in the mutant M10(2)B26C2H showed 85% nucleotide identity with *akbS* gene on pDG5 of R7; the nucleotide sequence of mutant M6AU26C12B showed 99% nucleotide identity with a region between two mobile elements downstream *akbS* gene on pDG5 of R7.

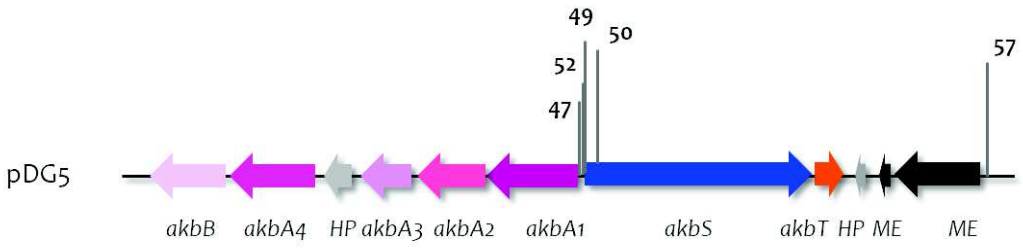

Figure 5.10 - IS*1415* insertions in the genome of *R. opacus* R7 and identification of genes adjacent to the insertion loci. The genes shown encode the following proteins: *ME*, mobile element, *HP*, hypothetical protein, *akbS*, response regulator, two-component system (sensor kinase), *akbT*, response regulator.

As reported in 5.3 Results and Discussion paragraph section Aromatic hydrocarbons degradation this gene cluster is related to the ability of R7 strain to grow on BTEX compounds and these data supported bioinformatic assumptions. Furthermore, these results suggested the involvement of the *akb* gene cluster in the *o*-xylene biodegradation.

5.4 Conclusions

This chapter aimed to identify new molecular tools, sequences, involved in the degradation of different hydrocarbons, such as aliphatic, mono-aromatic and polycyclic aromatic hydrocarbons through the comparison of *R. opacus* R7 genome sequences with reference strains, such as *Rhodococcus* sp. BCP1 and *R. jostii* RHA1. Several gene clusters involved in hydrocarbon metabolism were identified: *alk*, *prm*, *akb*, *dsz*, *nar*, *gen bph,* naphthenic acids clusters and *cat*, *pca*, *paa*, *hmg* involved in the peripheral metabolisms.

The *alk* gene cluster revealed to be the only copy within the R7 genome and therefore it was isolated and cloned in a shuttle-vector *E. coli-Rhodococcus* pTipQT1 to prove the expression of the cluster in an heterologous strain.

R. opacus R7 was subjected to transposon mutagenesis to generate mutants with defective growth on 13 different hydrocarbons to identify new sequences putatively involved in these metabolisms.

The preliminary results suggested the involvement of the *akb* gene cluster in the *o*-xylene biodegradation.

5.5 References

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

[2] Gurtler V, Mayall BC, Seviour R (2004) Can whole genome analysis refine the taxonomy of the genus *Rhodococcus*? FEMS Microbiol Rev 28:377-403.

[3] Watkinson RJ, and Morgan P (1990) Physiology of aliphatic hydrocarbon-degrading microorganisms. Biodegradation 1:79-92.

[4] Coon MJ (2005) Omega oxygenases: nonheme-iron enzymes and P450 cytochromes. Biochem Biophys Res Commun 338:378-385.

[5] Rojo F (2009) Degradation of alkanes by bacteria. Environ Microbiol 11:2477-2490.

[6] Jimenez JI, Minambres B, Garcia JL, Diaz E (2002) Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. Environ Microbiol 4:824-41.

[7] McLeod MP, Warren RL, Hsiao WWL, Araki N, Myhre M, Fernandes C, Eltis LD, et al. (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci 103:15582-7.

[8] Martínková L, Bronislava Uhnáková, Miroslav Pátek, Jan Nešvera, Vladimír Křen (2009) Biodegradation potential of the genus *Rhodococcus*. Environ International 35:162-177.

[9] Pérez-Pantoja D, Donoso R, Agulló L, Córdova M, l Seeger M, Pieper DH and González B (2012) Genomic analysis of the potential for aromatic compounds biodegradation in *Burkholderiales.* Environ Microbiol 14:1091-1117.

[10] Zylstra GJ (1994) Molecular analysis of aromatic hydrocarbon degradation. In Garte SJ (Ed.), Molecular Environmental Biology. Lewis Publishers, Boca Raton, FL 83-115.

[11] Mitchell K, Studts J, Fox B (2002) Combined participation of hydroxylase active site residues and effector protein binding in a *para* to *ortho* modulation of toluene 4-monooxygenase regiospecificity. Biochemistry 41:3176-3188.

[12] Williams PA and Murray K (1974) Metabolism of benzoate and the methylbenzoates by *Pseudomonas putida* (*arvilla*) mt-2: evidence for the existence of a TOL plasmid. J Bacteriol 120:416-423.

[13] Gibson DT, Koch JR, Schuld CL, Kallio RE (1968) Oxidative degradation of aromatic hydrocarbons by microorganisms. II. Metabolism of halogenated aromatic hydrocarbons. Biochemistry 7:3795-3802.

[14] Shields MS, Montgomery SO, Chapman PJ, Cuskey SM and Pritchard PH (1989) Novel pathway of toluene catabolism in the trichloroethylene-degrading bacterium G4. Appl Environ Microbiol 55:1624-1629.

[15] Kukor JJ and Olsen R H (1996) Catechol 2,3-dioxygenases functional in oxygen-limited (hypoxic) environments. Appl Environ Microbiol 62:1728-1740.

[16] Whited GM and Gibson DT (1991) Toluene-4-monooxygenase, a three-component enzyme system that catalyzes the oxidation of toluene to *p*-cresol in *Pseudomonas mendocina* KR1. J Bacteriol 173:3010-3016.

[17] Gülensoy N and Alvarez PJJ (1999) Diversity and correlation of specific aromatic hydrocarbon biodegradation. Biodegradation 10:331-340.

[18] Tao Y, Fishman A, Bentley W, Wood T (2004) Oxidation of benzene to phenol, catechol, and 1,2,3-trihydroxybenzene by toluene 4-monooxygenase of *Pseudomonas mendocina* KR-1 and toluene 3-monooxygenase of *Ralstonia pickettii* PKO1. Appl Environ Microbiol 70:3814-3820.

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

[20] Cafaro V, Izzo V, Scognamiglio R, Notomista E, Capasso P, Casbarra A, et al. (2004) Phenol hydroxylase and toluene/o-xylene monooxygenase from *Pseudomonas stutzeri* OX1: interplay between two enzymes. Appl Environ Microbiol 70:2211-2219.

[21] Baggi G, Barbieri P, Galli E and Tollari S (1987) Isolation of a *Pseudomonas stutzeri* strain that degrades *o*-xylene. Appl Environ Microbiol 53:2129-2132.

[22] Arenghi FL, Berlanda D, Galli E, Sello G, and Barbieri P (2001) Organization and regulation of meta cleavage pathway gene for toluene and *o*-xylene derivative degradation in *Pseudomonas stutzeri* OX1. Appl Environ Microbiol 67:3304-3308.

[23] Bertoni G, Bolognesi F, Galli E, and Barbieri P(1996) Cloning of the genes for and characterization of the early stages of toluene catabolism in *Pseudomonas stutzeri* OX1. Appl Environ Microbiol 62:3704-3711.

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

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

[26] Cerniglia CE (1984) Microbial metabolism of polycyclic aromatic hydrocarbons. Adv Appl Microbiol 30:31-71.

[27] Smith MR (1990) The biodegradation of aromatic hydrocarbons by bacteria. Biodegradation 1:191-206.

[28] Perry (1979) Microbial cooxidations involving hydrocarbons. Microbiol Rev 43:59-72.

[29] Bouchez M, Blanchet D, Vandecasteele J-P (1995) Degradation of polycyclic aromatic hydrocarbons by pure strains and by defined strain associations: inhibition phenomena and cometabolism. Appl Microbiol Biotechnol 43:154-156.

[30] Xu-Xiang Z, Shu-Pei C, Cheng-Jun Z and Shi-Lei S (2006) Microbial PAH-degradation in soil: degradation pathways and contributing factors. Pedosphere 16:555-565.

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

[32] Sambrook J, Russell DW (1989) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

[33] 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.

[34] Nakashima and 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.

[35] 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.

[36] Simkins S, Alexander M (1984) Model for mineralization kinetics with the variable of substrate concentration and population density. Appl Environ Microbiol 47:1299-1306.

[37] Sallam KI, Mitani Y, Tamura T (2006) Construction of random transposition mutagenesis system in *Rhodococcus erythropolis* using IS*1415*. J Biotechnol 121:13-22.

[38] Pilhofer M, Bauer AP, Schrallhammer M, Richter L, Ludwig W, Schleifer K-H and Petron G (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 35:e135.

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

[40] 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.

[41] 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 77:1619-27.

[42] Cappelletti M, Presentato A, Milazzo G, Turner RJ, 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 monooxygenase genes. Front Microbiol 6:393.

[43] Holder JW, Ulrich JC, DeBono AC, Godfrey PA, Desjardins CA, Zucker J, et al. (2011) Comparative and functional genomics of *Rhodococcus opacus* PD630 for biofuels development. PLoS Genet 7:e1002219.

[44] 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 86:1841-7.

[45] 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 143:2961-73.

[46] 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.

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

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

[49] 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. Microbes and Environ 20:144-150.

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

[51] Pelletier DA, Harwood CS (2000) 2-Hydroxycyclohexanecarboxyl coenzyme A dehydrogenase, an enzyme characteristic of the anaerobic benzoate degradation pathway used by *Rhodopseudomonas palustris*. J Bacteriol 182:2753-60.

[52] 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 USA. 94:6484-9.

[53] 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 187:4497-504.

[54] 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.

[55] 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.

[56] Sallam KI, Tamura N, Tamura T (2007) A multipurpose transposonbased vector system mediates protein expression in *Rhodococcus erythropolis*. Gene 386:173-182.

Chapter 6

6 Microcosms Experiments

This chapter focuses on the assessment of depuration potential of an hydrocarbon contaminated soil using a culture-independent approach in solid-phase microcosms. Soil was collected from an agricultural area and characterized, no *o*-xylene was identified. Preliminary kinetic microcosm experiments were performed to optimize the condition to evaluate the degradation rate of the *o*-xylene in presence of autochthonous bacteria in comparison to bioaugmented microcosms, in presence of *Rhodococcus opacus* R7 strain. Long-term microcosm experiments were established to monitor the catabolic functions involved in the *o*-xylene metabolism using marker sequences previously identified.

6.1 Introduction

Bioremediation processes are considered the best tool for decontamination of polluted natural ecosystems. The reclamation of polluted soils by *in situ* bioremediation is particularly attractive for both environmental and economic reasons. 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. 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)[1].

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)[2]. 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) [1].

The monitoring of introduced and indigenous strains can be achieved using probes based on conserved housekeeping genes and catabolic gene probes. Genetic databases now include not only the well-known genes coding for degradative systems of Proteobacteria but also genes for naphthalene dioxygenase and alkane hydroxylase of Gram-positive bacteria. Such probes have

widened the possibility of describing the biodiversity of degrading microbial communities in polluted environments (Cavalca et al., 2002) [3].

The BTEX compounds, comprising benzene, toluene, ethylbenzene and xylene isomers are listed as priority pollutants by the US Environment Protection Agency (EPA) and are among the top 100 chemicals on the priority list of hazardous substances published by the Agency for Toxic Substances and Disease Registry (ATSDR) [4]. These volatile organics commonly occur as soil or ground water contaminants due to spillages at petroleum production wells, refineries, pipelines, distribution terminals and leakages from fuel storage tanks. They are also present in the effluents of petroleum refineries and industries associated with the manufacture of pesticides, chemicals, detergents, varnishes and paints (Nagarajan and Loh, 2015) [5]. Conventional physico-chemical methods used for BTEX removal are either difficult to execute or expensive and invariably release secondary pollutants. On the other hand, biodegradation offers an immense range of environmentally friendly options for mineralization or transformation of the mono-aromatics into less harmful or non-hazardous compounds.

The BTEX compounds are frequently occurring groundwater contaminants. They can be biodegraded under both aerobic and anaerobic conditions and *in situ* bioremediation, either passive or active, is increasingly applied for the elimination of BTEX in groundwater. Decisions as to whether a site should be contained and monitored or actively treated are largely made on an empirical basis. Basic knowledge about the distribution, population densities and activities of BTEX degrading organisms at the polluted sites can contribute to rational decision-making. The ability to rapidly and accurately detect BTEX biodegrading bacteria and their activity in the environment is therefore of major interest. This can be done by demonstrating the occurrence of catabolic genotypes involved in BTEX degradation or their corresponding mRNA in the aquifer by employing sensitive PCR and RT-PCR detection methods (Hendrickx et al., 2006) [6].

Several metabolic routes such as TOD, TOL, TOM, TBU and TMO pathways have been characterized for aerobic removal of these mono-aromatic hydrocarbons [7] (see section 5.1.2 Aromatic hydrocarbon aerobic metabolism in Chapter 5). Most of the enzyme pathways transform at most only three or four compounds with cometabolic removal of other mono-aromatics. BTEX biodegradation has been achieved by mixed community of micro-organisms isolated from contaminated sites due to their broad enzymatic capacities and ability to support different metabolic pathways. However, conventional acclimatization has failed to produce a stable consortium to biodegrade the mono-aromatics due to the competitive behavior of the strains

following different metabolic pathways. This is because the enriched consortia, which arose by natural selection, are not spontaneous in establishing an optimum culture population. Molecular biology approaches enabled the identification of the key genes responsible for BTEX removal and subsequently facilitated the construction of recombinant strains to biodegrade the volatile organics (Lee et al., 1995; Kahraman and Geckil 2005) [8, 9]. The potential risks associated with the use of these genetically engineered strains, however, hinder their use for practical applications. Construction of a microbial consortium through mixing known species will help in overcoming the drawback caused by an isolated mixed community during BTEX biodegradation and practical difficulties in using genetically modified microorganisms (Nagarajan and Loh, 2015) [5].

6.2 Materials and Methods

The soil was collected from a non-contaminated agricultural area near Lodi. The soil was subsequently passed through a 2 mm sieve after the visible plant debris and the stones were removed. It was homogenized and stored at 4°C.

6.2.1 Characterization of soil sample

The soil pH was measured in a solution of 1M KCl and in H_2O with a Tecnavetro pH meter (soil-to-solution ratio of 1:2.5). Soil texture (sieving and sedimentation; five fractions), organic carbon, total nitrogen (Flash EA 1112 NC-Soil, Thermo-Fisher Scientific CN elemental analyzer, Pittsburgh, USA), total phosphorus (Olsen method), cation-exchange capacity (CEC)(NRCS, 2004), and field capacity were determined.

Chemical characterization

A 20 g dry soil sample was extracted using a Soxhlet BUCHI Extraction System B-811 with different cycles of extraction in dichloromethane. Then the extract was dried under gentle nitrogen flux and resuspended in 1 ml of *n*-hexane and it was analyzed in 6890N Network gaschromatograph system (ZB-5MS column, Alltech) coupled to an HP 5970 mass selective detector with quadrupole mass analyzer, 5973 Network Mass Selective Detector in scan mode (Agilent Technologies). The temperature program was 2 min at 60°C, then 6°C min^{-1} to 300°C, and 20 min at 300°C.

Microbiological characterization

The total heterotrophic and *o*-xylene degrading bacteria were evaluated using colony-forming unit counting method diluting the soil bacteria in a solution of NaCl 0.9 %. Each dilution was plated on Luria-Bertani (LB) agar medium and on M9 mineral medium (Maniatis et al., 1982) [10] supplemented with *o*-xylene (1 g/l) as the only carbon and energy source. The plates were incubated at 30°C.

6.2.2 Preparation of soil microcosms

The following spiking method was developed: 50 g of soil were divided in 6 homogeneous squares and each square was homogeneously contaminated with 2.5 ml of *o*-xylene solution 20045 µg/mL in DCM. Then the soil was mixed with a sterile spatula and it was transferred to a 250-ml glass bottle with screw cap (previously sterilized for 4 hours at 165°C in a heater). Nutrient solution (NH_4Cl solution) was added to the microcosm to re-establish the C:N:P ratio and humidity (17%). Bioaugmented microcosms were prepared adding 5 ml of a *Rhodococcus opacus* R7 culture grown until OD_{600}~ 0.8 - 1. *R. opacus* R7 was maintained in mineral M9 medium [10] supplied with naphthalene added in crystalline form (1 g/l) as the only carbon source at 30°C. Then 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. Part of the soil was sterilized (sterilized for 4 hours at 165°C in a heater the day before the experiments) and used as control. After vigorous stirring, the microcosms bottles were closed and maintained in the dark at room temperature and then sampled for analysis at fixed times. For each condition tested (see section 6.3.3 Experimental plan and Table 6.2), four microcosms were prepared and sampled at fixed times (0, 2, 3, 6, 12, 14, 16, 21, 28 days) in order to perform various determinations: 3 g of soil were sampled to measure the decrease of the contaminant, 1 g to measure the microorganisms amount and 1 g was stored at -80°C to extract subsequently DNA/RNA. The fourth microcosm was flushed with air to measure the CO_2 produced by microorganisms. The temperature was monitored during the experiments and it was around 22°C. Each replicate was independent and each experiment was conducted in duplicate.

To determine the carbon dioxide developed by microorganisms, at fixed times the microcosm gaseous phase was bubbled in a solution of sodium hydroxide (NaOH), that reacted with the CO_2 produced, then and the residual NaOH was titrated with hydrochloric acid (HCl).

At fixed times, 3 g of soil were extracted using ultrasound bath (Branson 2510) in dichloromethane as a solvent for 30 minutes. The extract was analyzed by a 6890N Network gaschromatograph

system (ZB-5MS column, Alltech) coupled to an HP 5970 mass selective detector with quadrupole mass analyzer, 5973 Network Mass Selective Detector in sim mode (Agilent Technologies).The temperature program was 4 min at 36°C, then 8°C min^{-1} to 100°C, and 5 min at 300°C.

6.2.3 DNA extraction and PCR amplification

Total DNA was extracted from 0.5 g of microcosm soil sample (stored at -80°C) using Fast DNA Spin Kit For Soil (Biomedical). Total RNA was extracted from 0.5 g of microcosm soil sample (stored at -80°C) using RNA Power Soil Total RNA Isolation Kit (MoBio) according to the manufacturer's instructions with two modifications. In the first, a second step of sample purification was performed adding 3.5 ml of phenol:chloroform:isoamyl alcohol (25:24:1), 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 second one was performed before adding the SR4 buffer: the pellet was resuspended in 5 ml sterile water, 500 µl of 5 M NaCl, and 5.5 ml of 100% isopropanol, incubated at -20°C for 30 min and then centrifuged at 2500×g for 30 min. This reduced the size of the pellet and the amount of carryover to the column.

Reverse transcription was performed with iScriptc DNA 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.

PCR-amplification was performed on the total DNA extracted from soil a sample to identify the sequences similar to *akbA1* and *akbA2* genes. Two pair of primers for each gene (Table 6. 1) were designed on the basis of the more conserved sequences between the *akb* and *prm* gene clusters of different bacteria belonging to the *Rhodococcus* genus deposited in Genbank DataBase. They were carried out using the following program: 95°C for 3 min; 95°C for 30 sec, 57°C for 45 sec, 72°C for 4 min, for 35 cycles; and 72°C for 3 min. DNA fragments were purified from agarose gel by the NucleoSpin Extraction II Kits by Machery and Nagel. The eluted fragment was sequenced by automated sequencing (Eurofins MWG).

Table 6.3 – List of oligonucleotides used for PCR amplifications

Oligonucleotide name	sequence (5′ – 3′)
akbA1-GTG-for	GTGAATCCGCAGGACGGGTGG
akbA2-TGA-rev	TCAGAGGAAGAGGTTGAGATT
akbA1-F	ATATGATCTTGGACAATGAGG
akbA1-R	ATTCTCCATATCAATCTCGGG
*prmA*f-TTG	TTGAGTAGGCAAAGCCTGACA
*prmA*r-TGA	TCAGGCGGAACTGTGCCGCC
prmA-F	AACATCTACCTGACCGTGGT
prmA-R	GCCATCAGCAGGATCGAATA

6.2.4 Experimental plan

The expression of catabolic functions was evaluated in various conditions [11, 12]: non-sterilized soil contaminated not inoculated (SNS-C-N), to determine the intrinsic degradative capacity of the indigenous soil bacterial community; non-sterilized soil not contaminated and not inoculated (SNS), to evaluate the basal respiration of the soil; contaminated soil not sterilized and inoculated with the strain *Rhodococcus opacus* R7 (SNS-C-N-I), to evaluate the effect of bioaugmentation; non-sterilized soil 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 soil not inoculated (SS-C-N) to assess the abiotic loss; sterilized and contaminated soil inoculated with R7 (SS-C-I-N), to evaluate the ability of the strain to degrade the contaminant in soil (Table 6.2).

Table 6.2 – Experimental plan of soil microcosms

Soil(S)	Nutrients (N)	Inoculum (I)	Contaminant (C)
Not sterilized (NS)	NH_4Cl	-	*o*-xylene
Not sterilized (NS)	NH_4Cl	-	-
Not sterilized (NS)	NH_4Cl	*R. opacus* R7	*o*-xylene
Not sterilized (NS)	NH_4Cl	*R. opacus* R7	-
Sterilized (SS)	NH_4Cl	-	*o*-xylene

Sterilized (SS)	NH_4Cl	*R. opacus* R7	*o*-xylene

6.3 Results and Discussion

6.3.1 Characterization of soil samples

The microcosms were prepared using soil from an agricultural area near Lodi that was characterized considering agronomic, chemical and microbiological parameters. The agronomic characterization showed that the soil had 2.5% of organic matter content, which, according to the USDA classification, is considered an average content; total nitrogen and phosphorus (assimilable phosphorus) were present in high amounts in reference to the USDA classification, respectively 2100 ppm and 52 ppm. The pH was 6.5 and its moisture content was 17%.

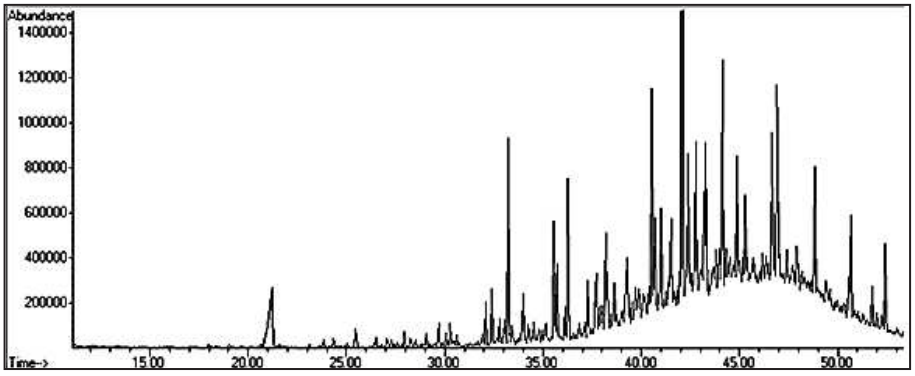

Figure 6.10 – Chromatogram of organic compounds present in a soil sample

The chemical characterization of soil showed several organic molecules (Figure 6.1) such as medium-chain *n*-alkanes as tetradecane, hexadecane and octadecane and long-chain *n*-alkanes as heptacosane; aldehydes such as *n*-decanal and carboxylic acids such as propanoic acid, pentanoic acid and hexanoic acid. Chemical analyses detected also some aromatic molecules as benzothiazole and tributyl phosphate; these compounds could have natural origins, such as *n*-alkanes, aldehydes and carboxylic acids or they could be a possible indicators of environmental pollution. A detailed chemical characterization of the organic molecules is given in Table 6.3.

Table 6.3 – Organic molecules identified with GC-MSD analysis

Name	Molecular Formula	Retention Time (min)	Match %	Abundance (%)
Hexanoic acid	$C_7H_{15}O_2$	21.21	93	100
n-Decanal	$C_{10}H_{20}O$	24.34	87	1
Benzothiazole	C_7H_5NS	25.47	88	22
Propionic acid	$C_4H_9O_2$	29.07	80	2
3-Hydroxybutyric acid	$C_4H_{10}O$	29.70	83	3
Tributylphosphate	$C_{12}H_{27}O_4P$	33.23	77	24
Pentanoic acid	$C_5H_{12}0_2$	35.53	78	14
Dodecanoic acid	$C_{12}H_{24}O_2$	36.29	90	19
Hexadecane	$C_{16}H_{34}$	35.70	93	11
Octadecane	$C_{18}H_{38}$	37.30	82	16
Tetradecane	$C_{14}H_{30}$	38.22	88	3
Heptacosane	$C_{27}H_{56}$	40.67	90	22
2-Pentadecanone	$C_{15}H_{30}O$	41.52	87	2

Microbiological characterization showed that total heterotrophic bacteria counts were 10^6 cfu/mL, while *o*-xylene degrading bacteria counts were equal to 10^4 cfu/mL.

Soil chemical analysis showed the absence of *o*-xylene. For this reason we decided to contaminate it with a compound belonging to BTEX group, in particular the experimentation was conducted in presence of *o*-xylene that is more recalcitrant than the *meta*- and *para*-isomers. Moreover, the strain *Rhodococcus opacus* R7 has the ability to degrade this isomer and the gene cluster involved in this metabolism was identified (see Chapter 3 and 5).

6.3.2 Preliminary growth kinetic experiments

Solid-phase microcosms were established to evaluate the depuration potential of an hydrocarbon contaminated soil using a culture-independent approach. Preliminary microcosms experiments were performed to evaluate the degradation rate of *o*-xylene, a compound belonging to BTEX group, by autochthonous bacteria in comparison to bioaugmented microcosms, in presence of *Rhodococcus opacus* R7 strain. At fixed times (0, 2, 3, 6 days) different parameters were determined: the decrease of the contaminant, the CO_2 production and the bacterial count.

The intrinsic *o*-xylene degradative capability of the indigenous soil bacteria was evaluated in the non-sterilized contaminated not inoculated (SNS-C-N) soil microcosm.

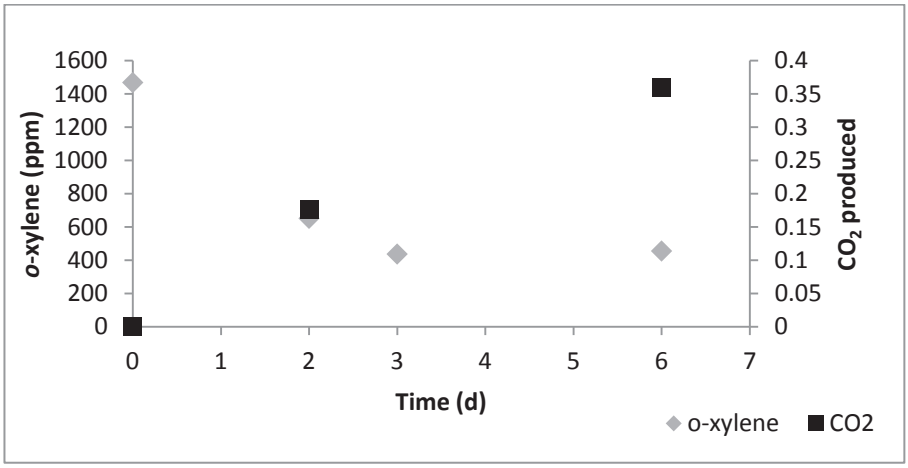

Figure 6.11 – Degradation of *o*-xylene and CO_2 produced in the non-sterilized contaminated not inoculated soil microcosm (SNS-C-N).

As showed in Figure 6.12, autochthonous bacterial population was able to degrade 63% of *o*-xylene and to produce around 0.35 mmol of CO_2 in six days. In the same incubation period *o*-xylene degrading bacteria increased slightly from 2.6 x 10^4 CFU g^{-1} to 5 x 10^4 CFU g^{-1}. Therefore, there was a high percentage of *o*-xylene degrading bacteria, but six days were not sufficient to observe the increment of autochthonous bacterial population.

The CO_2 produced in the SNS-C-N microcosms was higher than the CO_2 produced in the non-sterilized soil not contaminated and not inoculated with nutrients added (SNS-N) microcosms to evaluate the basal respiration of the soil that was around 0.18 mmol (Figure 6.3).

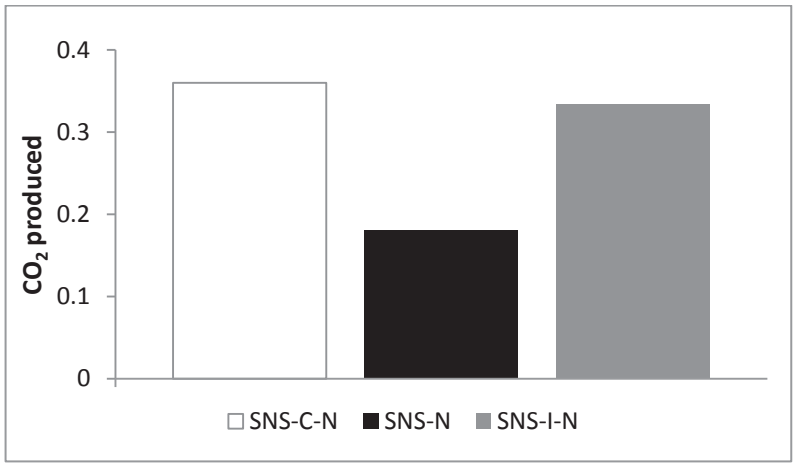

Figure 6.3 – Comparison between CO_2 produced in the non-sterilized contaminated not inoculated soil microcosm (SNS-C-N), in the non-sterilized uncontaminated not inoculated soil microcosm (SNS-N) and the non-sterilized uncontaminated inoculated soil microcosm (SNS-I-N).

Rhodococcus opacus R7 was used to evaluate the bioaugmentation efficiency in the *o*-xylene degradation, because of the phenotypic and the genotypic features of R7 strain (see Chapter 4 and 5). The *o*-xylene degradation rate measured in six days in the microcosms containing not sterilized contaminated soil with nutrients added and inoculated (SNS-C-I-N) was 73% and the CO_2 produced was around 0.45 mmol (Figure 6.4).

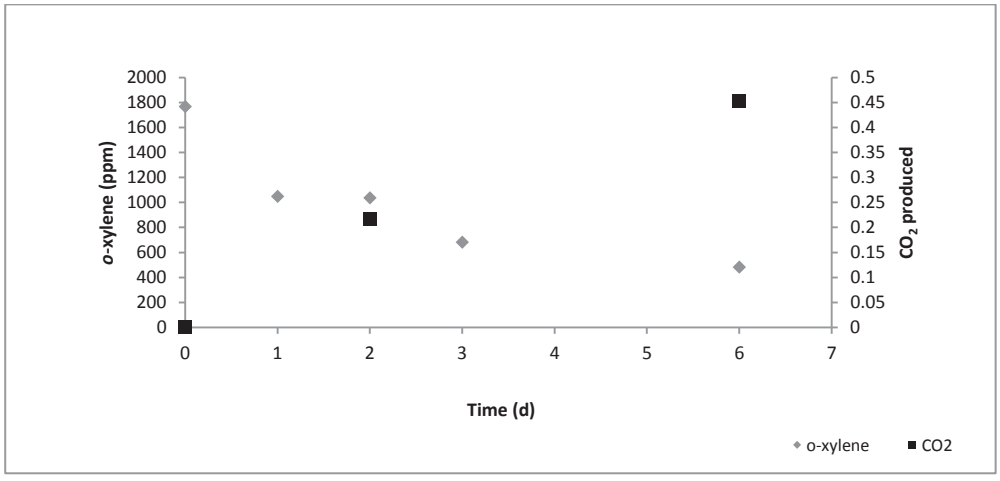

Figure 6.4 – Degradation of *o*-xylene and CO_2 produced in the non-sterilized contaminated inoculated soil microcosm (SNS-C-I-N).

In the same incubation period *o*-xylene degrading bacteria increased from 5 x 10^3 CFU g^{-1} to 5.8 x 10^4 CFU g^{-1}. We also evaluated the CO_2 produced in non-sterilized not contaminated soil inoculated with *R. opacus* R7 (SNS-I-N) to evaluate the interactions between the inoculated strain

and the autochthonous bacteria in absence of contamination. It was around 0.33 mmol that is equal to the basal respiration rate observed in SNS-C-N (Figure 6.3).

To assess the abiotic loss, microcosms with sterilized contaminated soil not inoculated with added nutrients (SS-C-N) were performed. The *o*-xylene decrease rate in these conditions was around 57% and it was compared to the one measured in microcosms with sterilized and contaminated soil inoculated with R7 and added with nutrients (SS-C-I-N), to evaluate *o*-xylene degradative ability of the strain in soil. In six days the *o*-xylene degradation rate was around 83% and the CO_2 produced was around 0.34 mmol of CO_2 (Figure 6.5).

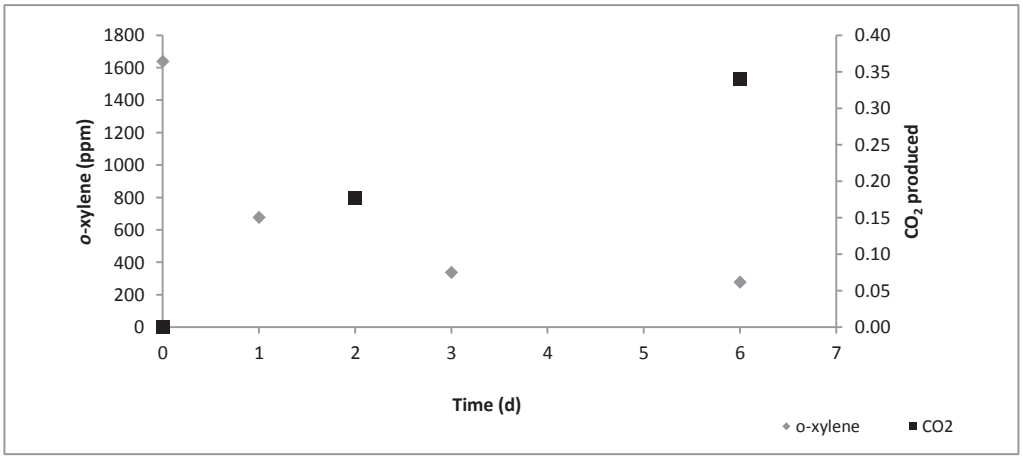

Figure 6.5 – Degradation of *o*-xylene and CO_2 produced in the sterilized contaminated inoculated soil microcosm (SS-C-I-N).

As a control, the increase of the *R. opacus R7* was observed: from 4 x 10^3 CFU g^{-1} to 2 x 10^4 CFU g^{-1}. The growth rate of soil bacteria in SNS-C-I-N and in SNS-C-N microcosms and growth rate of *R. opacus R7* in soil in six days were not significant, therefore long-term kinetic experiments were set up.

6.3.3 28 days growth kinetic experiments

Long-term experiments were set up to evaluate the biodegradation rate of *o*-xylene and the growth rate of bacteria population for 28 days. Three microcosm bottles were established for each of the following conditions: SNS-C-N-I (bottles R7-1, R7-2, R7-3) and SNS-C-N (bottles 1, 2, 3). The spiking method and inoculation procedure were the same used in the preliminary experiments (see Materials and Methods paragraph section 6.2.2 Preparation of soil microcosms). After 14 days the *o*-xylene biodegradation was 97% in the bioaugmented microcosm (R7-1), where 36 ppm of contaminant were left. In the other two bioaugmented microcosm bottles (R7-2, R7-3) the same

trend was observed and a similar final contaminant concentration was reached. Similar biodegradation rates were also observed in SNS-C-N microcosms, indeed the *o*-xylene degradation measured in bottles 1, 2, 3 was around 98%. In 14 days the bacterial counts were also evaluated in SNS-C-N-I and SNS-C-N microcosms: *o*-xylene degrading bacteria were stable during the incubation period and they were around 6 x 10^4 CFU g^{-1} and 2 x 10^4 CFU g^{-1}, respectively.

These results confirmed the preliminary kinetic measurements and suggested to extend the biodegradation time adding the contaminant in the bottles when it was over. For this reason the three bottles for each condition were re-contaminated with the same spiking procedure (Figure 6.6). After 21 days the same trend was observed for the three bottlesin each condition and the contaminant was added again (using a different method, adding pure *o*-xylene, not a DCM solution).

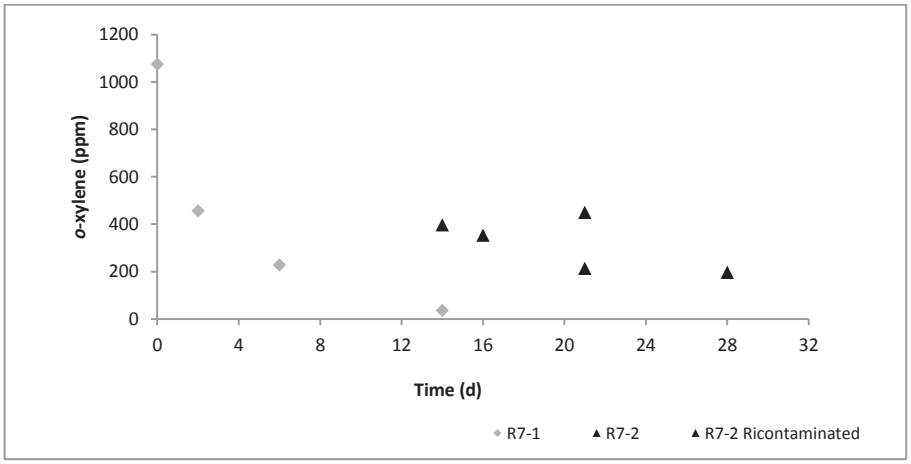

Figure 6.6 – Degradation of *o*-xylene in the not-sterilized contaminated inoculated soil microcosm (SNS-C-I-N) for a long-term kinetic experiments

With regard to the kinetics of *o*-xylene disappearance in bioaugmented microcosms, it can be noticed in figure 6.6 that the degradation rate decreased during the kinetics: low degradation rates corresponded to low *o*-xylene concentrations. Such behavior is often found for the long-term mineralization of contaminants with low water solubility [13, 14].

In bioaugmented microcosms, *o*-xylene degrading bacteria increased slightly from 6 x 10^4 CFU g^{-1} to 1 x 10^5 CFU g^{-1} in 28 days. A similar trend was observed in not bioaugmented microcosms where the *o*-xylene degrading bacteria increased slightly from 1 x 10^4 CFU g^{-1} to 3 x 10^4 CFU g^{-1} in 28 days. These results indicate that *o*-xylene biodegradation was mostly performed by the autochthonous degraders and that the *Rhodococcus opacus* R7 strain concentration was not

enough to counteract the autochthonous bacterial population. This could be attributed to the low concentration of the *R. opacus* R7 inoculum added [15] and to a high *o*-xylene degradation potential of the indigenous soil bacteria.

As shown in figure 6.7, from the several *o*-xylene enrichments of the bioaugmented microcosms it was possible to follow the *Rhodococcus opacus* R7 strain along all the growth experiments, due to the color and shape of R7 colonies growing on *o*-xylene. During 28 days, it was also observed a selection of colonies that were able to survive and degrade the *o*-xylene added.

t = 0 t = 14 d t = 28 d

Figure 6.7 - *o*-xylene enrichments on agarose plates of bacteria from soil samples of the bioaugmented microcosm. Plates showed the *o*-xylene enrichments at 0, 14, 28 days.

The total DNA/RNA was also extracted from collected oil samples to monitor the presence/absence of marker sequences after 0, 2, 21 and 28 days. The total DNA of each soil sample was extracted and PCR-amplification was performed to monitor the presence of the *akbA1* and *akbA2* genes that encode for ethylbenzene dioxygenase large subunit and ethylbenzene dioxygenase small subunit, respectively. These genes encode for the catalytic subunit of *akb* gene cluster and they are associated to BTEX degradation pathway (See section Aromatic hydrocarbons degradation in 5.3 Results and Discussion paragraph). It was shown in Chapter 4 (see 4.3.3 Phylogenetic classification of dioxygenases of *Rhodococcus opacus* R7 paragraph) that these sequences are not so well conserved among different bacteria genera, but they seem representative of *Rhodococcus* species.

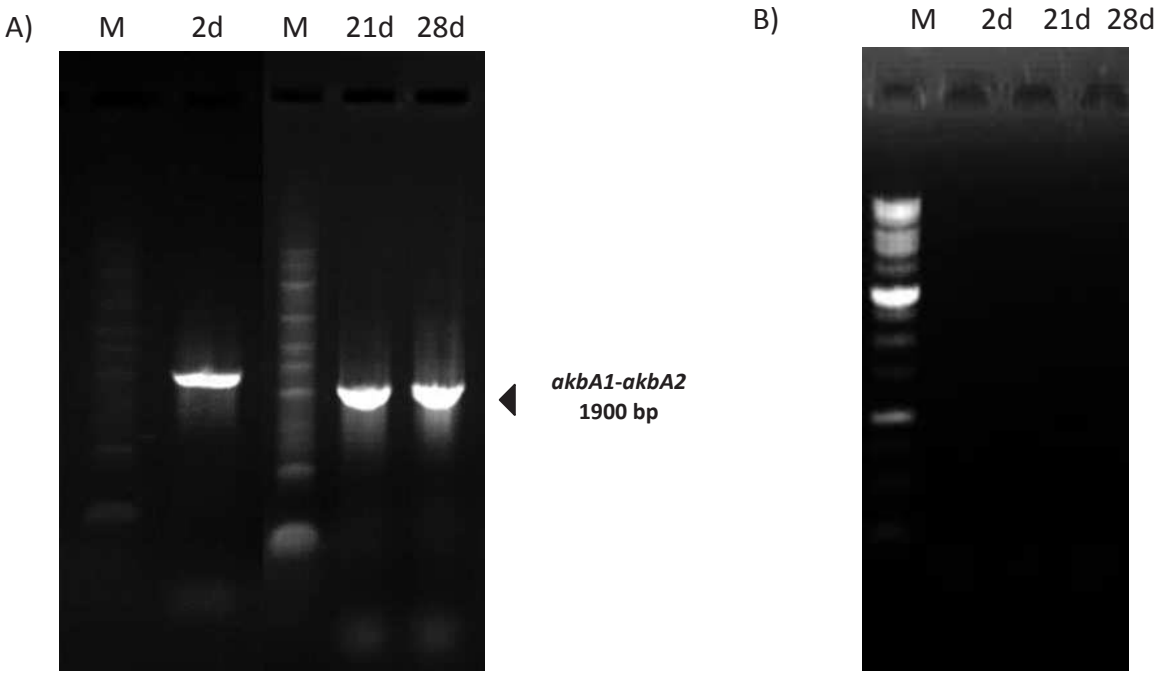

Figure 6.8 – Agarose gel of PCR-amplification of *akbA1-akbA2* genes. A) PCR-amplification using as template the total DNA extracted from soil samples from SNS-C-N-I microcosms at 2, 21, 28 days (2d, 21d, 28d); B) PCR-amplification using as template the total DNA extracted from soil samples from SNS-C-N microcosms at 2, 21, 28 days (2d, 21d, 28d); lane M: DNA marker (sigma).

The figure 6.8 panel A showed the agarose gel of 1.9 kb DNA PCR-amplification of *akbA1*-A2 genes using as template the total DNA extracted from soil samples from SNS-C-I-N microcosms at 2, 21, 28 days. Instead the PCR-amplification using as template the total DNA extracted from soil samples from SNS-C-N microcosms at 2, 21, 28 days showed no amplificates (figure 6.8 panel B).

Moreover, PCR-amplification was performed to monitor the presence of the *prmA* gene that encodes for the catalytic subunit A of a propane monooxygenase. It is well known to be able to oxidize short-chain *n*-alkane (See section Aliphatic hydrocarbons and cycloalkanes degradation in 5.3 Results and Discussion paragraph), but it is also associated to BTEX degradation pathway through the oxidation of the methyl group on the aromatic ring (See paragraph 5.1.2 Aromatic hydrocarbon aerobic metabolism within Chapter 5). This sequence appeared well conserved among different bacteria genera such as *Rhodococcus*, *Mycobaterium*, *Gordonia*, *Nocardia*, *Amycolatopsis* and *Pseudomonas* as shown in Chapter 4 (see 4.3.2 Phylogenetic classification of monooxygenases of *Rhodococcus opacus* R7 paragraph). The figure 6.9 showed the agarose gel of 0.5 kb DNA PCR-amplification of *prmA* gene using as template the total DNA extracted from soil samples from SNS-

C-I-N microcosms at 2, 16, 21, 28 days and from soil samples from SNS-C-N microcosms at 2, 21 days.

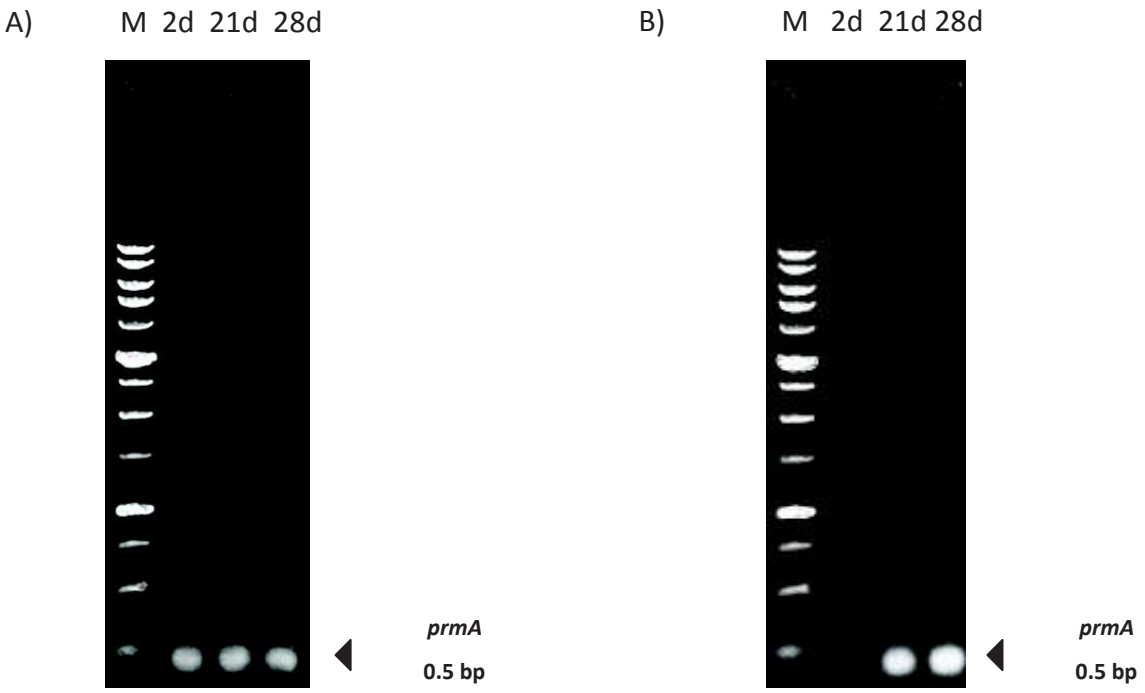

Figure 6.9 – Agarose gel of PCR-amplification of *prmA* gene. Amplification of part of the *prmA* gene sequence of around 0.5 bp. A) PCR-amplification using as template the total DNA extracted from soil samples from SNS-C-N-I microcosms at 2, 21, 28 days (2d, 21d, 28d); B) PCR-amplification using as template the total DNA extracted from soil samples from SNS-C-N microcosms at 2, 21, 28 days (2d, 21d, 28d); lane M: DNA marker (sigma).

Even if the *R. opacus* R7 strain didn't contribute much to the *o*-xylene degradation, it was possible to follow one of the most conserved sequence (*prmA* gene) encoding for a propane monooxygenase involved in this metabolism. These results show that it is possible to use this sequence in different conditions: in bioaugmented and in not bioaugmented microcosms.

A qRT-PCR method was developed to monitor not only the presence, but also the relative expression of *akb* genes in soil microcosms bon 2, 21, 28 days. Total RNA extraction from soil samples from SNS-C-I-N microcosms was carried out (figure 6.10), but the RNA samples were not clean enough to allow the evaluation of gene expression by PCR.

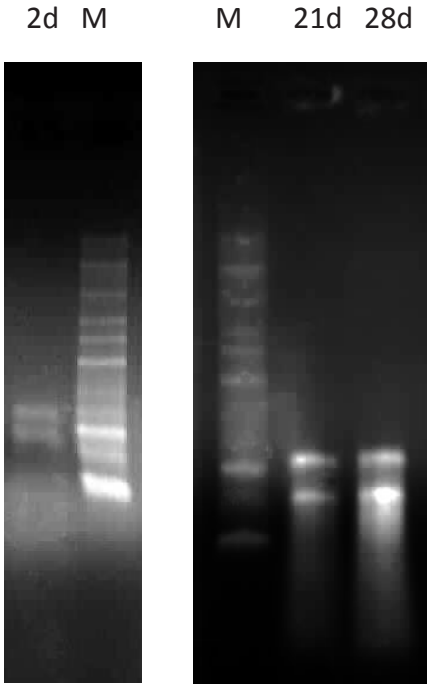

Figure 6.10– Agarose gel of total RNA extraction. Total RNA extraction from soil samples from SNS-C-N-I microcosms at 2, 21, 28 days (2d, 21d, 28d); lane M: DNA marker (sigma).

6.4 Conclusions

The aim of the present chapter was to assess the depuration potential of an hydrocarbon contaminated soil using a culture-independent approach in solid-phase microcosm experiments.

A soil deriving from an agricultural area near Lodi was sampled and characterized considering agronomic, chemicals and microbiological parameters. It revealed not an uncontaminated soil, but alkanes and carboxylic acids were found. Moreover it was rich of *o*-xylene degrading bacteria, even if no *o*-xylene was found. For this reason a new spiking method was developed and preliminary microcosm kinetic experiments were performed. On the basis of the preliminary results, long-term microcosm experiments were performed for 28 days. The biodegradation results indicated that *o*-xylene mineralization was mostly performed by the autochthonous degraders, while bioaugmentation did not enhance the remediation process. The limited effect of bioaugmentation likely reflects the inability of the applied inoculum to effectively colonize the microcosm soil. That, in turn, could be attributed to the origin of the consortia and to the presence of a rich autochthonous community [15]. Qualitative analysis to evaluate the effectiveness of sequences belonging to *akb* and *prm* clusters were performed and the results showed the presence of the most conserved gene, *prmA*, along 28 days of growth experiments in bioaugmented microcosms and without the presence of *R. opacus* R7 strain. Instead, it was possible to follow the *akb* gene

sequences only in bioaugmented microcosms. It can be concluded that it was possible to monitor the hydrocarbon contaminated soil and therefore to enhance the autodepuration soil capability.

6.5 References

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

[2] 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.

[3] 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.

[4] http://www.atsdr.cdc.gov/SPL/index.html.

[5] Nagarajann K and Loh K-C (2015) Formulation of microbial cocktails for BTEX biodegradation. Biodegradation 26:51-63.

[6] Hendrickx B, Junca H, Vosahlova J, Lindner A, Rüegg I, Bucheli-Witschel M, et al. (2006) Alternative primer sets for PCR detection of genotypes involved in bacterial aerobic BTEX degradation: distribution of the genes in BTEX degrading isolates and in subsurface soils of a BTEX contaminated industrial site. J Microbiol Methods 64:250-265.

[7] Zylstra GJ (1994) Molecular analysis of aromatic hydrocarbon degradation. In Garte SJ (Ed.), Molecular Environmental Biology. Lewis Publishers, Boca Raton, FL, 83-115.

[8] Lee J-Y, Jung K-H, Choi SH, Kim H-S (1995) Combination of the *tod* and the *tol* pathways in redesigning a metabolic route of *Pseudomonas putida* for the mineralization of a benzene, toluene, and *p*-xylene mixture. Appl Environ Microbiol 61:2211-2217.

[9] Kahraman H, Geckil H (2005) Degradation of benzene, toluene and xylene by *Pseudomonas aeruginosa* engineered with the *Vitreoscilla* hemoglobin gene. Eng Life Sci 5:363-368.

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

[11] Collina E, Bestetti G, Di Gennaro P, Franzetti A, Gugliersi F, Lasagni M, Pitea D (2005) Naphthalene biodegradation kinetics in an aerobic slurryphase bioreactor. Environ Int 31:167-171.

[12] Di Gennaro P, Collina E, Franzetti A, Lasagni M, Luridiana A, Pitea D, Bestetti G (2005) Bioremediation of di-ethylhexyl-phthalatecontaminatedsoil: a feasibilitystudy in slurry and solidphase reactors. Environ Sci Technol 39:325-330.

[13] Simkins S and Alexander M (1984) Models for mineralization kinetics with the variables of substrate concentration and population density. Appl Environ Microbiol, 47:1299-1306.

[14] Bouchez M, Blanchet D, Vandecsteele J-P (1995) Substrate availability in phenanthrene biodegradation: transfer mechanism and influence on metabolism. Appl Microbiol Biotechnol, 43:952-960.

[15] Fodelianakis S, Antoniou E, Mapelli F, Magagnini M, Nikolopoulou M, Marasco R, et al. (2015) Allochthonous bioaugmentation in ex situ treatment of crude oil-polluted sediments in the presence of an effective degrading indigenous microbiome. J Hazardous Materials 287:78–86.

Scientific Contributions

Research papers

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

- A Orro, M Cappelletti, P D'Ursi, L Milanesi, A Di Canito, J Zampolli, E Collina, F Decorosi, C Viti, S Fedi, A Presentato, D Zannoni, P Di Gennaro (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.

- J Zampolli, E Collina, M Lasagni, P Di Gennaro (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.

- P Di Gennaro, J Zampolli, I Presti, M Cappelletti, P D'Ursi, A Orro, A Mezzelani, L Milanesi (2014) Genome sequence of *Rhodococcus opacus* strain R7, a biodegrader of mono- and polycyclic aromatic hydrocarbons. Genome Annunc 2:1.

Conference contributions

- A Di Canito, J Zampolli, P Di Gennaro, P D'Ursi, L Milanesi, A Orro. Molecular mechanisms of *o*-xylene degradation in *Rhodococcus opacus* R7. Congress Società Italiana di Microbiologia Generale e Biotecnologie Microbiche (SIMGBM), Ravenna 2015.

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

Congress Società Italiana di Microbiologia Generale e Biotecnologie Microbiche (SIMGBM), Ravenna 2015.

- P D'Ursi, A Di Canito, M Uggeri, J Zampolli, P Di Gennaro , L Milanesi, A Orro. Study of the molecular mechanisms of the o-xylene monooxygenase in *Rhodococcus opacus* R7. Meeting della Bioinformatics Italian Society (BITS), Milano 2015.

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

Annex

1. Di Gennaro et al., 2014;
2. Zampolli et al., 2014;
3. Orro, Cappelletti et al., 2015;
4. Cappelletti et al., 2016.

genomeA
Journals.ASM.org

Genome Sequence of *Rhodococcus opacus* Strain R7, a Biodegrader of Mono- and Polycyclic Aromatic Hydrocarbons

P. Di Gennaro,[a] J. Zampolli,[a] I. Presti,[a] M. Cappelletti,[b] P. D'Ursi,[c] A. Orro,[c] A. Mezzelani,[c] L. Milanesi[c]

Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milan, Italy[a]; Department of Pharmacy and Biotechnology (FaBiT), University of Bologna, Bologna, Italy[b]; CNR, Institute for Biomedical Technologies, Segrate, Italy[c]

***Rhodococcus opacus* strain R7 (CIP107348) degrades several mono- and polycyclic aromatic hydrocarbons. Here, we present the high-quality draft genome sequence of strain R7, consisting of 10,118,052 bp, with a G+C content of 67.0%, 9,602 protein-coding genes, and 62 RNAs genes.**

Received 24 July 2014 **Accepted** 4 August 2014 **Published** 21 August 2014

Citation 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(4):e00827-14. doi:10.1128/genomeA.00827-14.

Address correspondence to P. Di Gennaro, patrizia.digennaro@unimib.it.

Members of the genus *Rhodococcus* demonstrate a remarkable ability to degrade a wide range of natural organic and xenobiotic compounds (1). *Rhodococcus opacus* R7 (CIP107348) was isolated from a polycyclic aromatic hydrocarbon-contaminated site in Italy because of its ability to grow on naphthalene and *o*-xylene as the only carbon and energy source (2, 3). As R7 also catabolizes a wide range of aliphatic, alicyclic, and both mono- and polycyclic aromatic hydrocarbons, it represents a strain of considerable environmental and industrial interest.

The genome sequencing of *Rhodococcus opacus* R7 was performed using 454 sequencing technology (Roche GS FLX Titanium). The total numbers of sequence reads were 312,384 from one shotgun library and 380,920 from one paired-end library. All the reads were assembled using Newbler 2.6 into 223 contigs, with an N_{50} length of 184,729 bp and an average genome coverage of 17×. Based on paired-end directional information, the contigs were further ordered into 6 scaffolds, giving a total genome size of 10.1 Mb with a G+C content of 66.0%.

The annotation was performed by using the RAST (Rapid Annotation using Subsystem Technology) server (4). Five scaffolds, constituting a total of ~1.65 Mb, are likely to be plasmids, as they carry genetic signatures typical of *Rhodococcus* plasmids.

A total of 9,602 putative open reading frames (ORFs), 62 RNAs genes, 9 rRNAs and 53 tRNAs, were predicted. RAST annotation indicates that strains *Rhodococcus jostii* RHA1 (score 501), *Rhodococcus opacus* B4 (score 409) are the closest neighbors of strain R7. Subsystem (457) categories representing the metabolisms of carbohydrates, amino acids and cofactors, vitamins, prosthetic groups, or pigments are the most abundant, and they account for 1,236, 995, or 665 proteins, respectively. 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 byphenil), 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 (5).

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession numbers CP008947, CP008948, CP008949, CP008950, CP008951, and CP008952. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This research was supported by the Italian Ministry of Education and Research through the Flagship (PB05) "Interomics," HIRMA (RBAP11YS7K), and the European "MIMOMICS" projects.

REFERENCES

1. **Martínková L, Uhnáková B, Pátek M, Nesvera J, Kren V.** 2009. Biodegradation potential of the genus *Rhodococcus*. Environ. Int. **35:**162–177. http://dx.doi.org/10.1016/j.envint.2008.07.018.
2. **Di Gennaro P, Rescalli E, Galli E, Sello G, 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. Res. Microbiol. **152:**641–651. http://dx.doi.org/10.1016/S0923-2508(01)01243-8.
3. **Di Gennaro P, Terreni P, Masi G, Botti S, De Ferra F, Bestetti G.** 2011. Identification and characterization of genes involved in naphthalene degradation in *Rhodococcus opacus* R7. Appl. Microbiol. Biotechnol. **87:**297–308. http://dx.doi.org/10.1007/s00253-010-2497-3.
4. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9:**75. http://dx.doi.org/10.1186/1471-2164-9-75.
5. **Urlacher VB, Girhard M.** 2012. Cytochrome P450 monooxygenases: an update on perspectives for synthetic application. Trends Biotechnol. **30:**26–36. http://dx.doi.org/10.1016/j.tibtech.2011.06.012.

Zampolli et al. AMB Express 2014, 4:73
http://www.amb-express.com/content/4/1/73

AMB Express
a SpringerOpen Journal

ORIGINAL ARTICLE Open Access

Biodegradation of variable-chain-length *n*-alkanes in *Rhodococcus opacus* R7 and the involvement of an alkane hydroxylase system in the metabolism

Jessica Zampolli[1,2], Elena Collina[2], Marina Lasagni[2] and Patrizia Di Gennaro[1*]

Abstract

Rhodococcus opacus R7 is a Gram-positive bacterium isolated from a polycyclic aromatic hydrocarbon contaminated soil for its versatile metabolism; indeed the strain is able to grow on naphthalene, *o*-xylene, and several long- and medium-chain *n*-alkanes. In this work we determined the degradation of *n*-alkanes in *Rhodococcus opacus* R7 in presence of *n*-dodecane (C12), *n*-hexadecane (C16), *n*-eicosane (C20), *n*-tetracosane (C24) and the metabolic pathway in presence of C12. The consumption rate of C12 was 88%, of C16 was 69%, of C20 was 51% and of C24 it was 78%. The decrement of the degradation rate seems to be correlated to the length of the aliphatic chain of these hydrocarbons. On the basis of the metabolic intermediates determined by the R7 growth on C12, our data indicated that *R. opacus* R7 metabolizes medium-chain *n*-alkanes by the primary alcohol formation. This represents a difference in comparison with other *Rhodococcus* strains, in which a mixture of the two alcohols was observed. By GC-MSD analysis we also identified the monocarboxylic acid, confirming the terminal oxidation.
Moreover, the *alkB* gene cluster from *R. opacus* R7 was isolated and its involvement in the *n*-alkane degradation system was investigated by the cloning of this genomic region into a shuttle-vector *E. coli-Rhodococcus* to evaluate the alkane hydroxylase activity. Our results showed an increased biodegradation of C12 in the recombinant strain *R. erythropolis* AP (pTipQT1-*alk*R7) in comparison with the wild type strain *R. erythropolis* AP. These data supported the involvement of the *alkB* gene cluster in the *n*-alkane degradation in the R7 strain.

Keywords: *Rhodococcus*; *n*-alkanes degradation; Alkane hydroxylase; AlkB; Enzymatic expression

Introduction

The problems associated with contaminated sites are assuming rising prominence in many countries (Vidali, 2001), due to the increase of anthropogenic pollution in the environment. Anthropic activities produce a wide amount of pollutants, such as aliphatic and aromatic polycyclic hydrocarbons; soil and water microorganisms can selectively degrade these xenobiotic compounds as the only carbon and energy source. The knowledge of the metabolic pathways of these microorganisms allows the reclamation of polluted sites. There are a lot of studies on the metabolism of these compounds in Gram-negative bacteria while Gram-positive bacteria have not been investigated to the same extent. Members of the *Rhodococcus* genus, found in many environmental niches, have a marked ability to metabolize a wide variety of xenobiotic compounds, and it is well recognized that species of this genus are key participants in the recycling of complex organic compounds (Finnerty, 1992; Larkin et al., 2005; Martínková et al., 2009). The genetic rearrangement and plasticity of the *Rhodococcus* genome have appeared to play a significant role in its adaptation to a wide variety of environmental contaminants (Leahy and Colwell, 1990; Van der Meer et al., 1992; Kulakov et al., 1998; Larkin et al., 2006; McLeod et al., 2006).

Alkane degraders are bacteria that have a very versatile metabolism, so that they can use as carbon source many other compounds in addition to alkanes (Smits et al., 2002; Margesin et al., 2003; Harayama et al., 2004). The ability of these bacteria of degrading *n*-alkanes is well

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

established, but relatively little is known about the molecular characteristics of their alkane-degradative systems.

Alkanes are saturated, linear molecules whose chain length can vary from 1 (in methane) to more than 50 carbon atoms. They are the major components of petroleum fuels which can be commonly found in contaminated environments (So and Young, 1999), indeed they constitute about 20-50% of crude oil, depending on the source of the oil. In addition, alkanes (predominantly long-chain compounds) are produced throughout the biosphere by living organisms (plants, algae and bacteria) as a waste product, a structural element, a defense mechanism, or as a chemoattractant (van Beilen et al., 2003). The aerobic degradation of these molecules starts by oxidation of one of the terminal methyl groups to generate the corresponding primary alcohol by alkane hydroxylases (AHs). The oxidation can occur on various positions: terminal or subterminal, to final conversion to a fatty acid (van Beilen et al., 2003; Ji et al., 2013). So far some alkane-degradative systems of a small number of Gram-negative bacteria have been well characterized, such as those of *Alkanivorax, Pseudomonas* and *Acinetobacter* (Alonso and Roujeinikova, 2012; Ratajczak et al., 1998; van Beilen and Funhoff, 2007; Wang and Shao, 2013). The *alk* system found in *Pseudomonas putida* GPo1, which degrades *n*-alkanes from *n*-pentane to *n*-dodecane, remains the most extensively characterized alkane hydroxylase system. It is a three-component alkane hydroxylase complex consisting of a particulate nonheme integral membrane alkane monooxygenase (AlkB) and two soluble proteins, rubredoxin (AlkG) and rubredoxin reductase (AlkT) (Alonso and Roujeinikova, 2012).

Much less is known about the alkane-degradative systems of Gram-positive bacteria. Homologs of *alkB* gene were amplified from psychrotrophic *Rhodococcus* sp. Q15 (Whyte et al., 2002), *Gordonia* sp. strain SoCg (Lo Piccolo et al., 2011) and *Rhodococcus opacus* B4 (Sameshima et al., 2008).

The *Gordonia* sp. strain SoCg unique *alkB* gene was analyzed by functional heterologous expression in *E. coli* and in *S. coelicolor,* which were shown to oxidize *n*-hexadecane to the correspondent primary alcohol, 1-hexadecanol, but no expression occurred in presence of long-chain *n*-triacontane. Similarly, the *Rhodococcus opacus* B-4 *alkB1* and *alkB2* genes were expressed heterologously in two *E. coli* recombinants which were able to convert *n*-alkanes (*n*-pentane to *n*-hexadecane) to their corresponding alcohols in anhydrous organic solvents. The start codons of two genes were changed from GTG to ATG to express them in *E. coli. Rhodococcus* sp. strain Q15 was shown to be able to mineralize *n*-dodecane and *n*-hexadecane and bioconvert them into corresponding primary and secondary alcohols at 5°C. The utilization of potential metabolic intermediates indicated that Q15 oxidizes alkanes by both the terminal oxidation pathway and the subterminal oxidation pathway (Whyte et al., 1998).

Rhodococcus opacus R7 is a Gram-positive bacterium isolated from a polycyclic aromatic hydrocarbon contaminated soil for its ability to grow on naphthalene. It is characterized by a versatile metabolism, indeed the strain is able to grow on naphthalene, *o*-xylene, and several long- and medium-chain *n*-alkanes. In previous studies, the genes involved in the degradation of naphthalene and *o*-xylene were identified, and the related metabolic pathways were characterized (Di Gennaro et al., 2010). In this work we determined the degradation of *n*-alkanes in *Rhodococcus opacus* R7, performing biodegradation kinetics in presence of *n*-dodecane (C12), *n*-hexadecane (C16), *n*-eicosane (C20), *n*-tetracosane (C24), and the metabolic pathway in presence of C12. Moreover, the *alkB* gene cluster from *R. opacus* R7 was isolated and its involvement in *n*-alkane degradation system was investigated by cloning and expression of this genomic region.

Materials and methods

Bacterial strains, growth conditions and general procedures

Bacterial strains and plasmids used in this study are listed in Table 1. *Rhodococcus opacus* R7 (CIP 107348), isolated for its ability to grow on naphtalene and *o*-xlyene (Di Gennaro et al., 2001), was grown on M9 mineral medium (Maniatis et al., 1982), supplemented with naphthalene or *o*-xylene or different *n*-alkanes in an atmosphere saturated with these compounds, as the only carbon and energy source. Growth of R7 was performed in 100 ml-flasks with 20 ml of M9 mineral medium in presence of *n*-hexane (C6), *n*-octane (C8), *n*-decane (C10), *n*-dodecane (C12), *n*-hexadecane (C16), *n*-eicosane (C20), *n*-tetracosane (C24), *n*-hexatriacontane (C36) at the concentration of 1 g/l. *n*-Eicosane (C20), *n*-tetracosane (C24), *n*-hexatriacontane (C36) were added in flasks as hexane solution evaporated over night or as finely ground powder.

Rhodococcus erythropolis AP, isolated in our laboratory (CIP 110799) for its ability to grow on diesel fluel (Maffei, 2004), was maintained on M9 mineral medium in a saturated atmosphere of diesel fuel at 30°C (Table 1). *Escherichia coli* DH5α was grown on Luria-Bertani medium at 37°C. To select *E. coli* and *Rhodococcus* trasformants antibiotics (ampicillin at 100 μg/ml for *E. coli* and tetracycline at 25 μg/ml for *Rhodococcus*) were added in the cultural medium.

Genomic DNA of *R. opacus* R7 was extracted from R7 cells as reported by Di Gennaro et al., 2010. DNA manipulation, enzymatic digests, ligation and PCR reactions were performed using standard molecular techniques (Sambrook and Russell, 1989). For recombinant plasmids extraction NucleoSpin Plasmid Kit by Machery and Nagel was used according to the manufacturer's instruction.

Table 1 Bacterial strains and plasmids used in this study

Strain or plasmid	Description	Reference or source
Strains		
Rhodococcus opacus R7	Long-medium-chain *n*-alkane degrader, *alkB*$^+$, *nar*$^+$, *gen*$^+$, *o-xyl*$^+$	Di Gennaro et al. 2001; Di Gennaro et al. 2010
Rhodococcus erythropolis AP	Diesel Fuel degrader	Maffei 2004
Escherichia coli DH5α	d*lacZ*ΔM15, *rec*A1, *end*A1, *gyr*A96, *thi*-1, *hsd*R17(rK$^-$, mK$^+$), *sup*E44, *rel*A1, *deo*R, Δ(*lac*ZYA-*arg*F)U169	Promega
Escherichia coli DH5α (pDrive-*alk*R7)	*E. coli* DH5α containing the cloning vector pDrive, *alkB* fragment, Amp^r	This study
Escherichia coli DH5α (pTipQT1-*alk*R7)	*E. coli* DH5α containing the recombinant expression vector pTipQT1, *alkB* fragment, Thio^r, Amp^r	This study
Rhodococcus erythropolis AP (pTipQT1-*alk*R7)	*R. erythropolis* AP containing the recombinant expression vector pTipQT1, *alkB* fragment, Thio^r, Tc^r	This study
plasmids		
pDrive	*E. coli* cloning vector, Amp^r	Qiagen
pDrive-*alk*R7	pDrive containing *alkB* fragment from *R. opacus* R7	This study
pTipQT1	Shuttle-vector Amp^r *E. coli* – Tc^r *Rhodococcus* spp.	Nakashima and Tamura 2004b
pTipQT1-*alk*R7	pTipQT1 containing *alkB* fragment from *R. opacus* R7	This study

Growth curves and *n*-alkanes biodegradation

Growth experiments on R7 strain were performed in presence of *n*-alkanes (1 g/l) ranging from C6 to C36 in 20 ml of M9 with OD_{600} 0.1.

To determine *R. opacus* R7 *n*-alkanes degradation, the strain was inoculated in 20 ml of M9 with OD_{600} 0.1 and supplemented with 1 g/l of *n*-alkanes (C12, C16, C20, C24). To determine abiotic loss, uninoculated flasks were also prepared. The flasks were incubated at 30°C for a maximum of 72 h. Every 24 h a flask was sacrified to evaluate *n*-alkanes degradation by determination of the residual *n*-alkane by gas chromatography–mass spectrometry (GC-MSD) after extraction with *n*-hexane. Bacterial growth was determined by OD_{600} measurements. When the strain was inoculated on *n*-alkanes with more than 16 carbon atoms, the bacteria formed clumps which included *n*-alkanes (van Beilen et al., 2002). For optical density measurements, clumps were resuspended by strong agitation.

Chromatographic analyses

Residual *n*-alkanes were extracted from the inoculated and uninoculated flasks using *n*-hexane. It occurred in 20 min by strong manual agitation (in test tubes), after washing of the flasks to take away any residual *n*-alkanes on their wall. The suspension was centrifuged at 4000 rpm for 10 min and after 10 min of settling, 2 ml were drawn from organic phase. It was conserved at −20°C in vial with teflon-coated screw caps. The organic phase extracted was diluted 1:100 for GC injection. Residual *n*-alkanes were analysed by GC-MSD using a Technologies 6890 N Network GC System, interfaced with 5973 Network Mass Selective Detector (MSD) (Agilent Technologies). A ZB-5MS capillary column was used (5% diphenyl-95% dimethylpolysiloxane 60 m x 0.25 mm, 0.25 μm; Alltech).

Analyses were carried out in split-less injection mode using helium as carrier gas at 99.99%. The injector port was set at 250°C. The oven temperature was programmed from 60°C for 3 min, then 15°C min^{-1} to 280°C, holding this temperature for 10 min. Electron impact ionization spectra were obtained at 70 eV, with recording of mass spectra from *m*/*z* 42 to 550 amu, which allows 3.5 scans s^{-1}. All analyses were carried out with three replicates, and the mean values obtained are reported.

Analysis of the metabolic intermediates from the *n*-alkane pathway

The metabolic intermediates resulting from incubation of *R. opacus* R7 on C12 were analyzed by GC-MSD after extraction with ethyl acetate. It occurred in 20 min by strong manual agitation (in test tubes) in presence of 1.5 ml HCl 1 M. The suspension was centrifuged at 4000 rpm for 10 min. After centrifugation, 1.5 ml of samples were taken from organic phase and transferred into 1.5 ml vials. For derivatization the solution was stripped under a gentle stream of nitrogen, 10 μL of derivatizing agent TMSI and Pyridine (Supelco) and hexane up to 1 ml were added. The solutions were heated at 70°C for 30 min. The derivatized samples were analyzed by GC-MSD and the extracted samples were conserved at − 20°C in vial with teflon-coated screw caps. The registered mass spectra were compared with those of the spectra library (NIST) of the instrument, and the

identification of the metabolites was confirmed by injection and analysis of the corresponding standard compound.

Kinetic modeling

The residual *n*-alkane concentrations as a function of time were modeled according to the Monod-type kinetic models as reported by Simkins and Alexander (1984). In particular, considering the substrate scarce water solubility and the biomass initial concentration (growth *viz*. exposition runs), the logistic and first-order kinetic models were applied. For logistic model, the maximum specific growth rate, μ_{max} (h^{-1}), was estimated from the microbial growth curve, plotting the natural logarithm of OD_{600} vs time and determining the slope of the linear part of the graph (Zhukov et al., 2007). The inverse yield parameter, q (ppm OD_{600}^{-1}) (Simkins and Alexander, 1984), was estimated as the ratio of the degraded *n*-alkane to OD_{600} increase. A non-linear regression was used for the fitting of the logistic model to the experimental data. The goodness of fit was evaluated on the basis of the determination coefficient, the standard error of the estimated parameter and the p-value.

For first-order kinetic model, the natural logarithm of substrate concentration *vs* time was plotted and the slope of the least-squares-regression line was determined.

Identification and cloning of the *alkB* gene cluster

The isolation of the complete coding region of the *alkB* gene cluster from *R. opacus* R7 was obtained from purified genomic DNA of the strain by PCR-amplification.

To identify the genomic region, two primers (F1, 5′-AA GGCCATGGGGCGTTAGAGCACCGCAGCTAAT-3′ and R1, 5′-ACAGCATATGACCTAGCGGGCGGCCGCGACC CG-3′) were designed on the basis of the more conserved sequences between the *alkB* gene cluster of different bacteria belonging to the *Rhodococcus* genus deposited in Genbank DataBase. It was carried out using the following program: 95°C for 3 min; 95°C for 30 sec, 68°C for 45 sec, 72°C for 5 min, for 35 cycles; and 72°C for 3 min.

DNA fragments were purified from agarose gel by the NucleoSpin Extraction II Kits by Machery and Nagel. The eluted fragment was sequenced by automated sequencing (Eurofins MWG). The nucleotide sequence of the *alkB* gene cluster of *R. opacus* R7 was deposited in the GenBank DataBase (KJ573524).

The 3.0 Kb fragment, containing the *alkB* gene cluster, was then amplified by PCR using the primers *Nco*I-*alk*R7 5′-AAGGCCATGGACGTGACGACGTCGGATA TC-3′ and *Nde*I-*alk*R7 5′-ACAGCATATGACCTAGCG GGCGGCCGCGAC-3′ in order to generate the *Nco*I-*Nde*I ends. The 3.0 Kb fragment was cloned as PCR product into the pDrive vector (Qiagen). The ligation mixture was used to transform *E. coli* DH5α by electroporation and the recombinant clones were selected on LB agar supplemented with ampicillin 100 μg/ml and IPTG 1 mM and 5-bromo-4-chloro-3-indolyl-β-D galactopyranoside (X-Gal) 40 μg/ml. White colonies were isolated and plasmid (pDrive-*alk*R7) of the recombinant clones was extracted and verified with digestion *Nco*I/*Nde*I.

Construction of the recombinant strain *R. erythropolis* AP (pTipQT1-*alk*R7) and activity of the *alkB* system

The *alkB* insert was ligated as *Nco*I-*Nde*I fragment into a shuttle-vector *E. coli-Rhodococcus*, pTipQT1 (Nakashima and Tamura 2004a). The ligation mixture was used to transform *E. coli* DH5α by electroporation with standard procedures (Sambrook and Russell, 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 (pTipQT1-*alk*R7) was isolated. The same recombinant plasmid was used as shuttle-vector to transform others *Rhodococcus* spp. strains by electroporation. We chose as host *Rhodococcus erythropolis* AP selected in our laboratory (Maffei, 2004) because the expression of the *Ptip* system efficiency was more performant in *erythropolis* species (Nakashima and Tamura, 2004b). Plasmids were introduced into strains of *Rhodococcus* spp. by electroporation using a Gene Pulser II (Biorad, Italia) set at 2.50 kV, 600 Ω, 25 μF (Treadway et al., 1999) in presence of about 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 grown at 30°C for 3–4 days. The stability of the recombinant plasmid in the *R. erythropolis* AP was verified after extraction and restriction analysis of the plasmid from all *R. erythropolis* AP (pTipQT1-*alk*R7) clones. All the clones were able to maintain the recombinant plasmid.

Moreover, before biodegradation experiments, a representative amount of colonies of the recombinant strain underwent the same procedure and we verified the maintaining of the plasmid in the following generations.

Recombinant strain *R. erythropolis* AP (pTipQT1-*alk*R7) was used for biodegradation experiments in presence of *n*-alkane C12 to evaluate the activity of the *alkB* system.

Chemicals

All chemicals used were of analytical grade. All organic solvents used were high-performance liquid chromatographic (HPLC) grade supplied from Fluka. Naphtalene, *o*-xylene and *n*–alkanes, 1-dodecanol, 2-dodecanol, dodecanoic acid were supplied from Sigma.

Results

Growth of *R. opacus* R7 on medium- and long-chain *n*-alkanes

Rhodococcus opacus R7 was isolated from a soil contaminated by polycyclic aromatic hydrocarbons, and the strain was able to grow on naphthalene, *o*-xylene (Di

Gennaro et al., 2001; Di Gennaro et al., 2010) as well as on various medium- and long-chain *n*-alkanes.

The metabolism on *n*-alkanes was investigated in presence of medium- and long-chain *n*-alkanes. Growth of R7 in M9 mineral broth in presence of *n*-hexane (C6), *n*-octane (C8), *n*-decane (C10), *n*-dodecane (C12), *n*-hexadecane (C16), *n*-eicosane (C20), *n*-tetracosane (C24), *n*-hexatriacontane (C36) (1 g/l) supplied as the only carbon and energy source was tested in 96 h. An increasing biomass accumulation was observed by spectrophotometric analysis at OD_{600} every 24 h. *R. opacus* R7 grew with *n*-alkanes ranging in length from C10 to C36, but no growth was observed in the range of C6-C8 *n*-alkanes (Table 2).

Although R7 growth on *n*-alkanes ranging in length from C20 to C36 was difficult to determine, due to the low solubility of solid *n*-alkanes in water, a certain growth was observed around the solid *n*-alkane agglomerates formed and on the wall of the flasks. Before the growth measurement, a strong agitation of the flasks was necessary to observe a homogeneous cells suspension (van Beilen et al., 2002).

The capacity of *R. opacus* R7 to degrade *n*-alkanes ranging in length from C12 to C24 was investigated as a function of time.

Biodegradation kinetics of medium- and long-chain *n*-alkanes C12-C24

To investigate the capacity of *R. opacus* R7 to degrade *n*-alkanes, kinetic runs were performed in M9 mineral medium supplemented with C12, C16, C20, C24 as the sole carbon and energy source; growth was followed at different times up to 72 h (Figure 1, a, b, c, d). The increase in biomass was quite the same for the four substrates, indeed OD_{600} reached about 0.5-0.7. GC-MSD analyses of *n*-alkane residues were performed after extraction from broths; the results showed that substrate consumption occurred parallel to biomass increase.

The consumption rate of C12 was 88%, of C16 was 69%, of C20 was 51% and of C24 it was 78%. The consumption of C12 was higher than longer *n*-alkanes: the decrement of degradation rate of these hydrocarbons seems to be correlated to the length of the aliphatic chain.

Table 2 Degradation of *n*-alkanes by *Rhodococcus opacus* R7

n-Alkane	Carbon atoms	*Rhodococcus opacus* R7
Hexane	6	-
Octane	8	-
Decane	10	+
Dodecane	12	+
Hexadecane	16	+
Eicosane	20	+
Tetracosane	24	+
Hexatriacontane	36	+

-, no growth; +, growth.

Abiotic controls were performed in presence of each hydrocarbon. The amount of abiotic loss was about 20% for C12, about 24% for C16, about 18% for C20, and about 26% for C24, respectively.

n-Alkane residual concentrations as a function of time followed a logistic curve, typical of a biodegradation process where a scarcely soluble organic compound is degraded by a growing microbial culture. From the experimental data treatment, values of the maximum specific growth rate, μ_{max}, and of the semisaturation constant, K_S, were estimated (Table 3). As expected, K_S increased from C12 to C24, indicating a decrease in the biodisponibility of the *n*-alkane for the biomass. On the contrary, μ_{max} was almost constant for C12, C16 and C20 and of the order of 0.015 h^{-1}, while it approached 0.032 h^{-1} for C24. This last value could indicate that more enzymatic systems were involved in the degradation of C24 or that a different up-take mechanism could be hypothesized.

Metabolic intermediates of *n*-alkane C12 degradation

Metabolic intermediates of *n*-alkanes in *R. opacus* R7 were determined from cultural broths during the growth on *n*-dodecane as reference substrate, chosen on the basis of the kinetic parameters determined for biodegradation runs. R7 growth on C12 was followed during the exposition to the hydrocarbon as a function of time. At time 0 h, 3 h, 5 h, 8 h and 24 h, culture samples were acidified, extracted, and analyzed by GC-MSD as reported in Material and Methods. Cultural broth analysis showed the presence of 1-dodecanol and the corresponding carboxylic acid (Figure 2) which were identified by comparison with the relative standards. We identified the 1-dodecanol only and not the mixture of 1-dodecanol and 2-dodecanol, as reported in literature for many cases of *Rhodococcus* strains (Whyte et al., 1998; Whyte et al., 2002). To confirm these data, both 1-dodecanol and 2-dodecanol were supplied separately as carbon and energy source and growth was observed only in presence of 1-dodecanol and no growth was observed in presence of 2-dodecanol. Moreover, we observed growth when the dodecanoic acid was supplied as the only carbon and energy source.

Identification and sequencing of the *R. opacus* R7 *alkB* gene cluster involved in *n*-alkanes degradation

The *alkB* gene cluster from *R. opacus* R7 genomic DNA was identified (Figure 3) and sequenced. This region was isolated by PCR with primers designed on the basis of the alignment of conserved sequences from the *alkB* gene region from different *n*-alkane degrading *Rhodococcus*

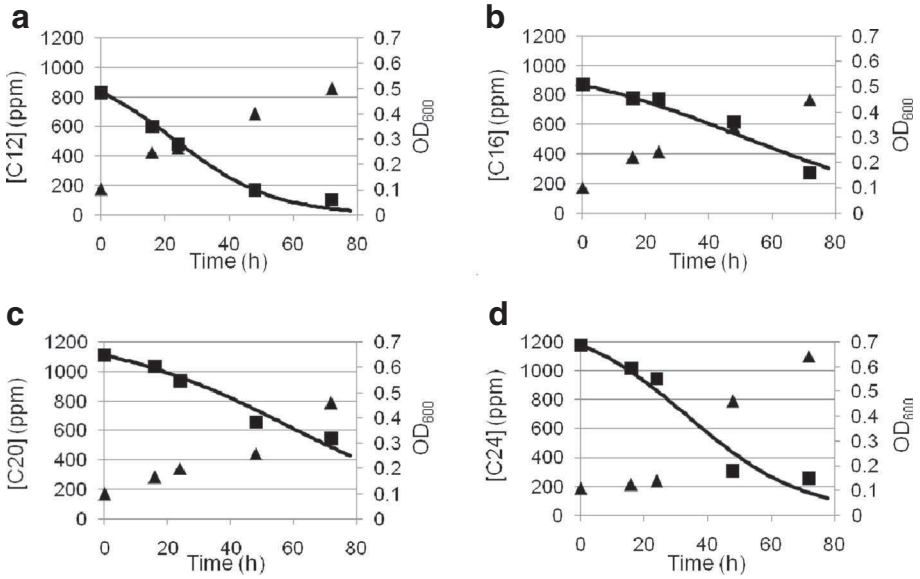

Figure 1 Kinetic analyses of *n*-dodecane (C12) (a), *n*-hexadecane (C16) (b), *n*-eicosane (C20) (c), *n*-tetracosane (C24) (d) degradation in *R. opacus* R7. Cells of *R. opacus* R7 were exposed to the *n*-alkane after growth on naphthalene and a flask each day was sacrified for the extraction and the GC-MSD determination of the residual hydrocarbon.

bacteria. The genomic 3.0 Kb fragment was isolated and sequenced. From the sequence analysis, the covering region contained four consecutive ORFs homologues to the *alkB* gene cluster components: *alkB* encoding for an alkane monooxygenase, *rubA* encoding for a rubredoxin, *rubB* encoding for a second rubredoxin, and *rubred* encoding for a rubredoxin reductase.

These sequences isolated from R7 strain were compared with the ones of other bacteria belonging to the *Rhodococcus* genus. In comparison with *Rhodococcus opacus* B4 (Sameshima et al., 2008), benzene-tolerant capable of growing on *n*-alkanes, *alkB* gene showed an identity of 94% (the corresponding protein AlkB 91%), *rubA, rubB*, and *rubred* an identity of 90%, 95% and 86% (RubA 91%, RubB 97%, RubRed 85%), respectively. In comparison with *Rhodococcus* sp. BCP1 (Cappelletti et al., 2011), capable of growing on volatile and medium-chain *n*-alkanes, *alkB* gene presented an identity of 83% (AlkB 80%), *rubA, rubB* and *rubred* an identity of 86%, 84% and 69% (RubA 79%, RubB 84%, RubRed 60%), respectively. Compared to *Rhodococcus jostii* RAH1 (McLeod et al., 2006), able to degrade polychlorinated biphenyls (PCBs), *alkB* gene presented an identity of 94% (AlkB 91%), *rubA, rubB* and *rubred* an identity of 99%, 94% and 91% (RubA 98%, RubB 97%, RubRed 91%), respectively.

Table 3 Biodegradation kinetic parameters for *n*-alkanes metabolism in *R. opacus* R7

Substrate consumption			Cellular growth		
Substrate	μ_{max} /K_S (1/h ppm)	R^2	μ_{max} (h^{-1})	R^2	K_S (ppm)
C12	(6.3 ± 0.3) 10^{-5}	0.988	0.013 ± 0.002	0.972	210
C16	(3.1 ± 0.2) 10^{-5}	0.938	0.0129 ± 0.0007	0.997	420
C20	(2.7 ± 0.2) 10^{-5}	0.952	0.017 ± 0.004	0.956	650
C24	(3.9 ± 0.4) 10^{-5}	0.952	0.032 ± 0.009	0.905	810

Construction of the recombinant strain *R. erythropolis* AP (pTipQT1-*alk*R7) and activity of the *alkB* system

The *alkB* gene cluster was isolated from R7 genomic DNA as PCR product and cloned into the pDrive vector giving the plasmid pDrive-*alk*R7. The region was isolated as *Nco*I/*Nde*I fragment and cloned into the shuttle-vector *E. coli-Rhodococcus* pTipQT1. The recombinant plasmid pTipQT1-*alk*R7 was isolated from *E. coli* DH5a and transferred into *Rhodococcus erythropolis* AP, because the *Ptip/regulator* system of the shuttle-vector is more efficient in *Rhodococcus erythropolis* species (Nakashima and Tamura 2004b). In order to verify the expression of the *alkB* gene cluster under the control of the *Ptip/regulator* system, experiments with resting cells of *Rhodococcus erytropolis* AP(pTipQT1-*alk*R7) exposed to C12 were performed. The expression of the *alk* region was determined comparing the biodegradation kinetics of the recombinant strain *R. erythropolis* AP (pTipQT1-*alk*R7) with the wild type strain *R. erythropolis* AP. Results are reported in Figure 4. The percentage of biodegradation in 6 h was near 80% in the recombinant strain and 37% in the wild type strain. From kinetics analysis we can observe that the initial

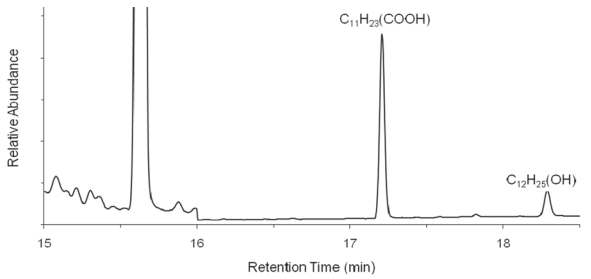

Figure 2 Cultural broth analysis by GC-MSD. Cells of *R. opacus* R7 were exposed to *n*-dodecane and at different times, the cultural broth were analysed in GC-MSD. The intermediate metabolites identified are reported in the graph.

degradation rate was higher in the recombinant strain with respect to the wild type strain indicating a difference in the activity levels of the *alkB* system. This difference was confirmed by a statistical test for the comparison of the slopes of the regression lines at 95% significance level.

Discussion

In this paper *Rhodococcus opacus* R7, a strain with a versatile metabolism, was characterized for the ability to degrade variable-chain-length *n*-alkanes.

Many alkane-degrading bacteria have been isolated and the enzyme systems that oxidize *n*-alkanes up to C16 have been characterized (Rojo, 2009; van Beilen and Funhoff, 2007, Wentzel et al., 2007). Long-chain *n*-alkanes are more persistent in the environment than the shorter but few data are available in particular concerning the metabolism of these compounds in Gram-positive bacteria (Whyte et al., 2002; Smits et al., 2002; Lo Piccolo et al., 2011).

R. opacus R7 showed the ability to grow on medium and long chain *n*-alkanes ranging from C10 to C36. Biodegradation kinetics suggested that R7 strain degraded the *n*-alkanes with a lower number of carbon atoms to a greater extent (88% C12, 69% C16, 51% C20, 78% C24). Moreover, μ_{max} was almost constant for C12, C16 and C20 and about 0.015 h^{-1}, while it approached 0.032 h^{-1} for C24. To our knowledge, no literature data are available concerning specific growth rates on single *n*-alkanes, but only Zhukov et al. reported data in presence of a mixture of *n*-alkanes (Zhukov et al., 2007). The decrement of degradation percentage of these hydrocarbons seems to be inversely correlated to the length of the aliphatic chain as already reported (van Beilen et al., 2003; Lo Piccolo et al., 2011). An exception was C24; in fact, when growing on this hydrocarbon, R7 strain showed a higher increase in OD_{600}, although the degradation percentage was similar to that observed for C12, and a higher μ_{max}. These data suggested that other enzymatic systems could be involved in C24 degradation. Another hypothesis could be a different up-take mechanism, supported by the experimental observation of a massive adhesion of R7 cells to C24 powder (van Beilen et al., 2003; Rapp et al., 2003; Lo Piccolo et al., 2011).

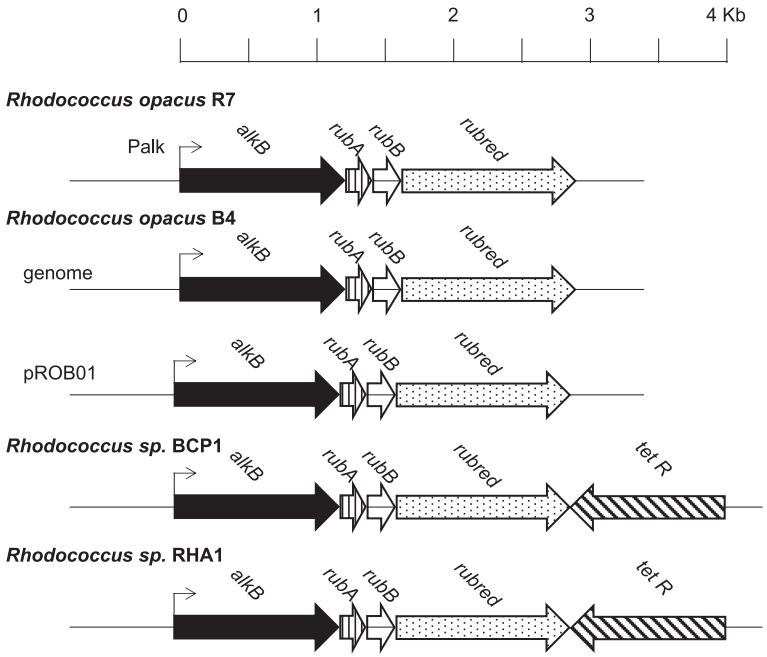

Figure 3 Gene organization of the *alkB* gene cluster of *R. opacus* R7 and comparison with equivalent clusters from other alkane-degrading *Rhodococcus* bacteria.

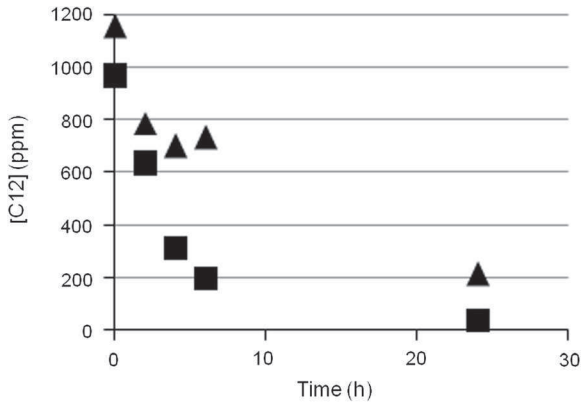

Figure 4 Kinetic analyses of *n*-dodecane degradation in *R. erytropolis* AP (pTipQT1-*alk*R7) recombinant strain (square in the graph). Cells of the recombinant strain were grown on rich medium and then, after induction with thiostrepton, were collected and exposed to *n*-dodecane. For comparison, kinetics of *n*-dodecane degradation in *R. erytropolis* AP (triangle in the graph) without the cloned fragment expressing the *alk*B gene, is also showed.

Many strains from different microbial genera are able to grow on *n*-alkanes. Alkanes are usually activated by a terminal oxidation to the corresponding primary alcohol, which is further oxidized by the alcohol and the aldehyde dehydrogenases to the corresponding acid. The resulting fatty acid enters into the beta-oxidation cycle. In some cases *n*-alkanes are metabolized via terminal as well as subterminal oxidation to the corresponding secondary alcohol. The secondary alcohol is converted to the corresponding ketone which is oxidized by a monooxygenase to an ester that is successively hydrolyzed by an esterase to an alcohol and a fatty acid (van Beilen et al., 2003; Ji et al., 2013). On the basis of the metabolic intermediates determined by the R7 growth on C12, our data indicated that *R. opacus* R7 metabolizes medium-chain *n*-alkanes by the primary alcohol formation because no trace of the secondary alcohol was found and because the strain is not able to grow on this compound. This result highlights a difference in comparison with many *Rhodococcus* strains in which a mixture of the two alcohols was observed (Whyte et al., 1998 and references therein). In the GC-MSD analysis we also identified the monocarboxylic acid, confirming therefore the activation by a terminal oxidation of the *n*-alkane as reported for *Rhodococcus* sp. MS11 by Rapp et al., 2003.

Moreover, the *alkB* gene cluster from *R. opacus* R7 was isolated and its involvement in the *n*-alkane degradation was studied. In this paper we identified the *alkB* gene cluster including an *alkB* gene, encoding for an alkane hydroxylase, two rubredoxins A and B, and a gene encoding for a rubredoxin reductase that are all required in the catalytic process as electron transfer proteins. Phylogenetic analysis of R7 *alkB* gene cluster showed a significant similarity with the homologous system of *R. opacus* B4 (Sameshima et al., 2008) and a lower identity with *Rhodococcus* sp. BCP1 (Cappelletti et al., 2011).

Although four types of *n*-alkane aerobic degradation pathways have been identified to date, the number of alkane hydroxylases that have been isolated, characterized and analyzed by structural biology techniques remains limited. The most widely characterized alkane degradation system is the AlkB of *Pseudomonas putida* GPo1 (Smits et al., 2002; van Beilen et al., 2003). Researchers have also isolated and cloned novel genes encoding AlkB from bacteria belonging to *Rhodococcus* genus, such as *Rhodococcus opacus* B4 (Sameshima et al., 2008) or *Rhodococcus* sp. Q15 (Whyte et al., 2002), but more information in these bacteria needs to be addressed. Literature data report the involvement of the alkane monoxygenase in the *n*-alkane biodegradation demonstrated by the changing of the GTG coding sequence for the expression in *E. coli* of this enzyme (Sameshima et al., 2008). Other authors were unable to show function heterologous expression of *alkB* genes in *Pseudomonas* or *E. coli* expression system, principally because the functional expression requires proper synthesis, correct folding, and proper assembly, and these are not always ensured for rhodococcal proteins (Whyte et al., 2002). On these bases, the *alkB* region identified from *R. opacus* R7 was cloned for the first time into a shuttle-vector *E. coli-Rhodococcus* and the alkane hydroxylase activity was evaluated in *R. erythropolis* AP. This strain was chosen not only because it belongs to the *Rhodococcus* genus but also because it belongs to the *erythropolis* species. In fact, the *Ptip/regulator* system of the shuttle-vector showed the best performance when expressed in *erythropolis* species as reported by Nakashima and Tamura, 2004b. In this way, we were able to overcome difficulties concerning gene expression in heterologous bacteria. Our results showed an increased biodegradation of C12 in the recombinant strain *R. erythropolis* AP (pTipQT1-*alk*R7) in comparison with the wild type strain *R. erythropolis* AP. These data supported the involvement of the *alkB* gene cluster in the *n*-alkane degradation in R7 strain. Considering the biodegradation kinetics for the C12-C24 substrate range, we can hypothesize that for the medium-chain-length *n*-alkanes an *alkB* like system plays the main role, while for *n*-alkanes longer than C20, other alkane hydroxylases could be involved.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JZ carried out the molecular biology studies and chemical analysis, participated in the sequence alignment and drafted the manuscript. EC participated in the design of the study and performed kinetic analysis. ML

performed statistical data treatments and interpretation of the data. PDG participated in the design of the study and the coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgment
We thank Prof. T. Tamura from Okkaido University of Japan to for supplying the shuttle-vector pTipQT1.

Author details
[1]Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, 20126 Milano, Italy. [2]Department of Earth and Environmental Sciences, University of Milano-Bicocca, Milano, Italy.

Received: 1 September 2014 Accepted: 5 September 2014
Published online: 30 September 2014

References

Alonso H, Roujeinikova A (2012) Characterization and two–dimensional crystallization of membrane component alkB of medium-chain alkane hydroxylase system from *Pseudomonas putida* GPo1. Appl Environ Microbiol 78:7946–7953

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

Di Gennaro P, Rescalli E, Galli E, Sello G, 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. 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

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

Harayama S, Kasai Y, Hara A (2004) Microbial communities in oil-contaminated seawater. Curr Opin Biotechnol 15:205–241

Ji Y, Mao G, Wang Y, Bartlam M (2013) Structural insights into diversity and n-alkane biodegradation mechanisms of alkane hydroxylases. Front Microbiol 4:1–13

Kulakov LA, Delacroix VA, Larkin MJ, Ksenzenko VN, Kulakova AN (1998) Cloning of new *Rhodococcus* extradiol dioxygenase genes and study of their distribution in different *Rhodococcus* strains. Microbiology 144:955–963

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

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

Leahy J, Colwell R (1990) Microbial biodegradation of hydrocarbons in environment. Microbiol Rev 54:305–315

Lo Piccolo L, De Pasquale C, Fodale R, Puglia AM, Quatrini P (2011) Involvement of an alkane hydroxylase system of *Gordonia* sp. strain SoCg in degradation of solid *n*-alkanes. Appl Environ Microbiol 77:1204–1213

Maffei F (2004) Individuazione dei fattori che stimolano la biodegradazione di idrocarburi presenti in suoli contaminati. University of Milano-Bicocca, Thesis

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

Margesin R, Labbé D, Schinner F, Greer CW, Whyte LG (2003) Characterization of hydrocarbon-degrading microbial populations in contaminated and pristine alpine soils. Appl Environ Microbiol 69:3085–3092

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

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 103:15582–15587

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

Nakashima N, Tamura T (2004b) Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-recombinant-protein expression. Appl Environ Microbiol 70:5557–5568

Rapp P, Lotte H, Gabriel-Jurgens E (2003) Degradation of alkanes and highly chlorinated benzenes, and production of biosurfactants, by a psychrophilic *Rhodococcus* sp. and genetic characterization of its chlorobenzene dioxygenase. Microbiology 149:2879–2890

Ratajczak A, Geiβdörfer W, Hillen W (1998) Expression of alkane hydroxylase from *Acinetobacter* sp. strain ADP1 is induced by a broad range of *n*-alkanes and requires the transcriptional activator AlkR. J Bacteriol 180:5822–5827

Rojo F (2009) Degradation of alkanes by bacteria. Environ Microbiol 11:2477–2490

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

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

Simkins S, Alexander M (1984) Model for mineralization kinetics with the variable of substrate concentration and population density. Appl Environ Microbiol 47:1299–1306

Smits THM, Balada SB, Witholt B, van Beilen JB (2002) Functional analysis of alkane hydroxylases from Gram-negative and Gram-positive bacteria. J Bacteriol 184:1733–1742

So CM, Young LY (1999) Isolation and characterization of a sulfate-reducing bacterium that anaerobically degrades alkanes. Appl Environ Microbiol 65:2969–2976

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

van Beilen JB, Funhoff EG (2007) Alkane hydroxylases involved in microbial alkane degradation. Appl Microbiol Biotechnol 74:13–21

van Beilen JB, Smits THM, Whyte LG, Schorcht S, Röthlisberger M, Plaggemeier T, Engesser KH, Witholt B (2002) Alkane hydroxylase homologues in Gram-positive strains. Environ Microbiol 4:676–682

van Beilen JB, Li Z, Duetz WA, Smits THM, Witholt B (2003) Diversity of alkane hydroxylase system in the environment. Oil Gas Sc Technol 58:427–440

Van der Meer JR, de Vos WM, Harayama S, Zehnder AJB (1992) Molecular mechanisms of genetic adaptation to xenobiotic compounds. Microbiol Rev 56:677–694

Vidali M (2001) Bioremediation. An overview. Pure Appl Chem 73:1163–1172

Wang W, Shao Z (2013) Enzymes and genes involved in aerobic alkane degradation. Front Microbiol 4:1–7

Wentzel A, Ellingsen TE, Kotlar H-K, Zotchev SB, Throne-Holst M (2007) Bacterial metabolism of long-chain *n*-alkanes. Appl Microbiol Biotechnol 76:1209–1221

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

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

Zhukov DV, Murygina VP, Kalyuzhnyi SV (2007) Kinetics of degradation of aliphatic hydrocarbons by bacteria *Rhodococcus ruber* and *Rhodococcus erythropolis*. Appl Biochem Microbiol 43:587–592

doi:10.1186/s13568-014-0073-4
Cite this article as: Zampolli *et al.*: **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.

PLOS ONE

RESEARCH ARTICLE

Genome and Phenotype Microarray Analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7: Genetic Determinants and Metabolic Abilities with Environmental Relevance

Alessandro Orro[1☯], Martina Cappelletti[2☯], Pasqualina D'Ursi[1], Luciano Milanesi[1], Alessandra Di Canito[3], Jessica Zampolli[3,4], Elena Collina[4], Francesca Decorosi[5], Carlo Viti[5], Stefano Fedi[2], Alessandro Presentato[2], Davide Zannoni[2], Patrizia Di Gennaro[3]*

1 Institute of Biomedical Technology, CNR, Segrate, Milano, Italy, **2** Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy, **3** Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milano, Italy, **4** Department of Earth and Environmental Sciences, University of Milano-Bicocca, Milano, Italy, **5** Department of Agrifood Production and Environmental Sciences, University of Firenze, Firenze, Italy

☯ These authors contributed equally to this work.
* patrizia.digennaro@unimib.it

OPEN ACCESS

Citation: 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(10): e0139467. doi:10.1371/journal.pone.0139467

Editor: Marie-Joelle Virolle, University Paris South, FRANCE

Received: June 29, 2015

Accepted: September 14, 2015

Published: October 1, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work has been supported by the Italian Ministry of Education and Research through the Flagship "InterOmics" (PB05) and HIRMA (RBAP11YS7K) Funding.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

In this paper comparative genome and phenotype microarray analyses of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7 were performed. *Rhodococcus* sp. BCP1 was selected for its ability to grow on short-chain *n*-alkanes and *R*. *opacus* R7 was isolated for its ability to grow on naphthalene and on *o*-xylene. Results of genome comparison, including BCP1, R7, along with other *Rhodococcus* reference strains, showed that at least 30% of the genome of each strain presented unique sequences and only 50% of the predicted proteome was shared. To associate genomic features with metabolic capabilities of BCP1 and R7 strains, hundreds of different growth conditions were tested through Phenotype Microarray, by using Biolog plates and plates manually prepared with additional xenobiotic compounds. Around one-third of the surveyed carbon sources was utilized by both strains although R7 generally showed higher metabolic activity values compared to BCP1. Moreover, R7 showed broader range of nitrogen and sulphur sources. Phenotype Microarray data were combined with genomic analysis to genetically support the metabolic features of the two strains. The genome analysis allowed to identify some gene clusters involved in the metabolism of the main tested xenobiotic compounds. Results show that R7 contains multiple genes for the degradation of a large set of aromatic and PAHs compounds, while a lower variability in terms of genes predicted to be involved in aromatic degradation was found in BCP1. This genetic feature can be related to the strong genetic pressure exerted by the two different environment from which the two strains were isolated. According to this, in the BCP1 genome the *smo* gene cluster involved in the short-chain *n*-alkanes

degradation, is included in one of the unique regions and it is not conserved in the *Rhodococcus* strains compared in this work. Data obtained underline the great potential of these two *Rhodococcus* spp. strains for biodegradation and environmental decontamination processes.

Introduction

The *Rhodococcus* genus comprises of Gram-positive, non-motile, non-sporulating, aerobic bacteria, with a high G+C content and mycolic acid-containing cell wall. This genus was firstly proposed by Zopf and revised by Tsukamura [1] and Goodfellow and Alderson [2]; it nowadays contains nearly 50 recognized species (http://www.bacterio.net/rhodococcus.html) [3, 4]. Members of *Rhodoccocus* genus are widely distributed in soil, water and marine sediments [5]. Some of them have also evolved for pathogenicity in humans, animals (i.e. *R. equi*) [6] and plants (i.e. *R. fascians*) [7]. 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 [3]. Compounds that are metabolically transformed or degraded by these bacteria are aliphatic and aromatic hydrocarbons, oxygenates, halogenated compounds, including polychlorinated biphenyls, nitro-aromatics, heterocyclic compounds, nitriles and various herbicides [5]. *Rhodococcus* spp. are able to perform steroid modifications, enantio-selective synthesis, production of amides from nitriles, and to convert plant secondary metabolites found in soil and rhizosphere, such as alkaloids, terpenes, and sterols [8]. Various *Rhodococcus* strains are also efficient at removing sulphur from coal and petroleum products [9, 10].

In line with the immense catabolic diversity shown by the members of this genus, *Rhodococcus* spp. are characterized to possess large and complex genomes, which contain a multiplicity of catabolic genes, a high genetic redundancy of biosynthetic pathways and a sophisticated regulatory network [11]. 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 [3]. Up to date, some *Rhodococcus* genomes have been sequenced with different level of completeness (http://www.ncbi.nlm.nih.gov/genome/Rhodococcus) [12]. The first genome completely sequenced belongs to *Rhodococcus jostii* RHA1 strain. This strain was isolated from lindane- contaminated soil for its exceptional ability to aerobically degrade polychlorinated biphenyls (PCBs) [13]. It was further described to utilize a wide range of aromatic compounds, carbohydrates, nitriles and steroids as sole carbon and energy sources. Analyses of the 9.7 Mb large genome of RHA1 provided the evidence of catabolic pathway redundancy and horizontal gene transfer events [14]. Recently, *Rhodococcus opacus* PD630 and B4 strains have been completely sequenced. Regarding these two strains, studies focused on the production and accumulation of energy-rich triacylglycerols in PD630 and on the tolerance to organic solvents in B4 [15, 16]. Compared to the increasing number of *Rhodococcus* spp. genomes that have been sequenced, very limited studies are available on their wide metabolic abilities and on their potentials for biodegradation activities related to genomic features. A high-throughput assessment of metabolic abilities is given by Phenotype Microarray system, which consists of 96-well plates provided with different conditions of growth (substrates, metabolites, pH conditions, toxic chemicals) for bacteria in order to generate phenotypic data [17]. This technology has the potential to accelerate the functional characterization of genes also deriving from whole-genome sequencing. This approach has been recently used to describe the respiration, viability and growth of several environmentally and clinically relevant bacteria species like

Mycobacterium, *Pseudomonas*, *Cronobacter* and *Sinorhizobium* [18–21]. *R. jostii* RHA1 and *R. opacus* PD630 were also tested for the growth on a limited number of carbon sources using Phenotype Microarray system [16]. This work focuses on the analysis of metabolic capabilities and genetic features of *Rhodococcus* sp. BCP1 and *Rhodococcus opacus* R7 whose genomes have been recently sequenced [22, 23]. *R. opacus* R7 is a Gram-positive bacterium isolated from a polycyclic aromatic hydrocarbon contaminated soil for its ability to grow on naphthalene and *o*-xylene [24]. It was further described for its ability to utilize several long- and medium-chain *n*-alkanes [25]. R7 has one of the largest bacterial genomes (around 10.1 Mb) with a G+C content of 66.0% [23]. *Rhodococcus* sp. BCP1 was selected from an aerobic butane-utilizing consortium as the prevailing isolate able to co-metabolize chloroform, vinyl chloride, and trichloroethylene [26–28]. As BCP1 also catabolizes a wide range of aliphatic, alicyclic, and carboxylated alkanes, it represents a strain of considerable environmental and industrial interest [29, 22].

In the present paper, a comparative genome analysis was performed to investigate the genomic differences between *R. opacus* R7 and *Rhodococcus* sp. BCP1 and between these two strains and those completely sequenced (*R. jostii* RHA1, *R. opacus* PD630, *R. opacus* B4) and/or taxonomically correlated (*R. pyridinivorans* SB3094). Additionally, a phenotypic screening of R7 and BCP1 strains was conducted using Phenotype Microarray with commercially available microtiter plates (pre-loaded substrates) and with microtiter plates manually prepared with additional organic/xenobiotic compounds. Phenotype microarray data were combined with genomic analysis to genetically support the metabolic features of these two bacterial strains with environmental and industrial relevance.

Materials and Methods

Genome sequencing and annotation

The genome sequencing of the two *Rhodococcus opacus* R7 (CIP107348) and *Rhodococcus* sp. BCP1 (DSM44980) strains was performed using 454 sequencing technology (Roche GS FLX Titanium). *R. opacus* R7 genome sequencing resulted in one shotgun library made of 312,384 sequence reads and one paired-end library of 380,920 sequence reads. All the reads were assembled into 223 contigs by using Newbler 2.6, with an N_{50} length of 184,729 bp and an average genome coverage of 17X. Considering *Rhodococcus* sp. BCP1, the total numbers of sequence reads were 668,686 from one shotgun library and 353,744 from one paired-end library (8-kb inserts). All the reads were assembled into 123 contigs by using Newbler 2.6, with an N_{50} length of 237,787 bp and an average genome coverage of 65X. The preliminary annotation was performed by using the RAST (Rapid Annotation using Subsystem Technology) server [30]. NCBI pipeline was used to annotate genome sequences and to make manual curation. The whole-genome shotgun sequencing projects have been deposited at DDBJ/EMBL/GenBank under the accession numbers CP008947, CP008948, CP008949, CP008950, CP008951, and CP008952 for *R. opacus* R7 and AVAE01000000 for *Rhodococcus* sp. BCP1; the Accession Numbers of the main genes discussed in this paper are reported in Table 1.

Whole-genome alignments

The whole genomes of *Rhodococcus* sp. BCP1 and *R. opacus* R7 have been compared with the genomes of *R. opacus* PD630, *R. opacus* B4 and *R. jostii* RHA1 and *R. pyridinivorans* SB3094, to find out similar and conserved regions between them and to highlight genes and potential unique functions for each genome under analysis.

Table 1. Accession Number of the main genes of *R. opacus* R7 and *Rhodococcus* sp. BCP1 strains discussed in this paper.

Gene	Homologous protein	Function	Accession Number *R. opacus* R7	Accession Number *Rhodococcus* sp. BCP1
alkB-alkB1	AlkB	Alkane monooxygenase	AIA09965.1	ADR72654.1
alkB2	AlkB2	Alkane monooxygenase	-	KDE11615.1
rubA	RubA	Rubredoxin	AIA09966.1	ADR72655.1
rubB	RubB	Rubredoxin	AIA09967.1	ADR72656.1
rubred	RubRed	Rubredoxin reductase	AIA09968.1	ADR72657.1
prmA	PrmA	Methane monooxygenase component A alpha chain	AII03499.1	KDE11344.1
prmC	PrmC	Methane monooxygenase component C	AII03498.1	KDE11343.1
prmB	PrmB	Methane monooxygenase component A beta chain	AII03497.1	KDE11342.1
prmD	PrmD	Methane monooxygenase regulatory protein	AII03496.1	KDE11341.1
akbA1a	AkbA1a	Ethylbenzene dioxygenase large subunit	AII11493.1	KDE09919.1
akbA2a	AkbA2a	Ethylbenzene dioxygenase small subunit	AII11492.1	KDE09920.1
akbA3	AkbA3	Ethylbenzene dioxygenase ferredoxin	CP008952.1	-
akbA4	AkbA4	Ferredoxin reductase	AII11490.1	KDE12339.1
akbB	AkbB	Dihydrodiol dehydrogenase	AII11489.1	KDE09922.1
akbC	AkbC	2,3-Dihydroxybiphenyl 1,2-dioxygenase	AII11058.1	KDE14642.1
akbD	AkbD	2-Hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	AII11051.1	KDE14641.1
akbE	AkbE	2-Hydroxypenta-2,4-dienoate hydratase	AII11050.1	KDE14625.1
akbF	AkbF	4-Hydroxy-2-oxovalerate aldolase	AII11049.1	-
dszA1	DszA1	Dibenzothiophene desulfurization enzyme	AII08556.1	KDE15059.1
dszA2	DszA2	Dibenzothiophene desulfurization enzyme	AII03608.1	KDE15059.1
dszB	DszB	Possible ABC sulfonate transporter	AII06125.1	KDE15056.1
dszC1	DszC1	Probable dibenzothiophene desulfurization enzyme	AII08748.1	KDE11236.1
dszC2	DszC2	Probable dibenzothiophene desulfurization enzyme	AII08273.1	KDE11237.1
rub1bis	Rub1bis	Rubredoxin	DQ846881	-
narR1	NarR1	Regulator of GntR family	ABH01023.1	KDE09916.1
narR2	NarR2	XylR-like regulator protein	ABH01024.1	KDE09917.1
rub1-rub	Rub1-Rub	Rubredoxin	ABH01026.1	KDE09915.1
rub2	Rub2	Rubredoxin	ABH01027.1	-
orf7	Orf7	Sterol-binding domain protein- unknown	ABH01028.1	KDE09918.1
narAa	NarAa	Naphthalene dioxygenase large subunit	ABH01029.1	KDE09919.1
narAb	NarAb	Naphthalene dioxygenase small subunit	ABH01030.1	KDE09920.1
narB	NarB	Cis-naphthalene dihydrodiol dehydrogenase	ABH01031.1	KDE09922.1
genC	GenC	Salicylate hydroxylase	AII11448.1; AII10777.1	-
genB	GenB	Salicylate CoA syntethase	AII11449.1; AII10778.1	-
genA	GenA	Salicylate CoA ligase	AII11450.1; AII10779.1	-
genH	GenH	Gentisate dioxygenase	AII11451.1; AII10780.1	KDE14391.1
genI	GenI	3-Maleylpyruvate Isomerase	AII11452.1; AII10781.1	KDE14392.1
genL	GenL	Unknown function	AII11453.1	KDE14393.1
bphAa	BphAa	Biphenyl-2,3-dioxygenase α subunit	AII11493.1	KDE09919.1
bphAb	BphAb	Biphenyl-2,3-dioxygenase β subunit	AII11492.1	KDE09920.1
bphAc	BphAc	Biphenyl-2,3-dioxygenase, ferredoxin component	AII08472.1	KDE10172.1
bphAd	BphAd	Biphenyl-2,3-dioxygenase, reductase	AII11490.1	KDE10578.1

(*Continued*)

Table 1. (*Continued*)

Gene	Homologous protein	Function	Accession Number *R. opacus* R7	Accession Number *Rhodococcus* sp. BCP1
bphB	BphB	Cis-2,3-dihydrobiphenyl-2,3- diol dehydrogenase	AII11489.1	KDE09922.1
bphC	BphC	2,3-Dihydroxybiphenyl- 1,2-dioxygenase	AII11058.1	KDE14642.1
akbD	AkbD	2-Hydroxy-6-oxohepta-2,4- dienoate hydrolase	AII11051.1	KDE11753.1
bphE	BphE	2-Oxopent-4-enoate hydratase	AII03622.1	KDE14453.1
bphF	BphF	4-Hydroxy-2-oxovalerate aldolase	AII03620.1	KDE14451.1
bphG	BphG	Acetaldehyde dehydrogenase	AII03621.1	KDE14452.1
badI	BadI	Naphthoate synthase	AII08541.1	KDE15145.1
badH1	BadH1	2-Hydroxycyclohexanecarboxyl-CoA dehydrogenase	AII08542.1	KDE15144.1
badH2	BadH2	2-Hydroxycyclohexanecarboxyl-CoA dehydrogenase	-	KDE12227.1
aliA	AliA	Long-chain-fatty-acid-CoA ligase	AII08543.1	KDE15143.1
badJ	BadJ	Acyl-CoA dehydrogenase	AII08544.1	KDE15141.1
pobA	PobA	*p*-Hydroxybezoate hydroxylase	AII08627.1	KDE11135.1
catA1	CatA1	Catechol 1,2 dioxygenase	AII08813.1	KDE10959.1
catB1	CatB1	Muconate cycloisomerase	AII08814.1	KDE10958.1
catC	CatC	Muconolactone isomerase	AII08815.1	KDE10957.1
catA2	CatA2	Catechol 1,2 dioxygenase	CP008947.1	-
catB2	CatB2	Muconate cycloisomerase	AII05696.1	-
pcaI	PcaI	Succinyl-CoA 3-ketoacid-coenzyme A transferase subunit B	AII09804.1	KDE10925.1
pcaJ	PcaJ	Succinyl-CoA 3-ketoacid-coenzyme A transferase subunit A	AII09803.1	KDE10926.1
pcaH	PcaH	Protocatechuate 3,4-dioxygenase beta chain	AII09802.1	KDE10927.1
pcaG	PcaG	Protocatechuate 3,4-dioxygenase alpha chain	AII09801.1	KDE10928.1
pcaB	PcaB	3-Carboxy-cis,cis-muconate cycloisomerase	AII09800.1	KDE10929.1
pcaL	PcaL	3-Oxoadipate enol-lactone hydrolase	AII09799.1	KDE10930.1
pcaF	PcaF	Acetyl-CoA acetyltransferase	AII09797.1	KDE10932.1
paaG	PaaG	Phenylacetate-CoA oxygenase, PaaG subunit	AII08339.1	-
paaH	PaaH	Phenylacetate-CoA oxygenase, PaaH subunit	AII08340.1	-
paaI	PaaI	Phenylacetate-CoA oxygenase, PaaI subunit	AII08341.1	-
paaJ	PaaJ	Phenylacetate-CoA oxygenase, PaaJ subunit	AII08342.1	-
paaK	PaaK	Phenylacetate-CoA oxygenase-reductase, PaaK subunit	AII08343.1	KDE12334.1
paaF	PaaF	Phenylacetate-coenzyme A ligase PaaF	AII08344.1	-
paaE	PaaE	Acetyl-CoA acetyltransferase	AII08335.1	KDE15146.1
paaA	PaaA	Enoyl-CoA hydratase	AII08336.1	-
paaC	PaaC	3-Hydroxyacyl-CoA dehydrogenase	AII08337.1	KDE15147.1
paaB	PaaB	Enoyl-CoA hydratase	AII08338.1	-
paaZ	PaaZ	Aldehyde dehydrogenase	AII08334.1	-
paaD	PaaD	Phenylacetic acid degradation protein, thioesterase	AII08332.1	-
paaL	PaaL	Acetate permease ActP (cation-acetate symporter)	AII08348.1	KDE10910.1
hmgA	HmgA	Homogentisate 1,2-dioxygenase	AII08874.1	KDE10817.1
fahA	FahA	Fumarylacetoacetase	AII08876.1	KDE10815.1
mai	Mai	Enoyl-CoA hydratase-isomerase	AII08877.1	KDE11332.1
orf1	Orf1	Long-chain-fatty-acid-CoA ligase	AII08879.1	KDE10810.1

doi:10.1371/journal.pone.0139467.t001

For a global alignment and visualization of the comparison of all six genomes, the Mauve tool (2.3 Version) has been used [31] that interactively shows the relative positions of sequence regions (colored blocks) that are found in more than one genome.

A more accurate analysis has been performed using the Last program [32] to obtain local pairwise all-*vs*-all alignments of all chromosome and plasmid sequences of the six genomes. With a post-process overlapping all the regions, the common regions are subtracted from the dataset to obtain a final list of regions that can be considered unique for each genome (present in one genome and not aligned in none of the other five). Only unique regions longer than 300 bp are considered. The genomic diversity in this group of bacteria is also represented in the diagram of Venn. This type of diagram represents the number of predicted protein-coding genes unique or shared amongst the genomes under investigation.

Metabolic map reconstruction

The preliminary approach to reconstruct metabolic capabilities of *R. opacus* R7 and *Rhodococcus* sp. BCP1 was to build a metabolic map using Pathway Tools software version 18.5. It is a comprehensive systems biology software environment for management, analysis, and visualization of integrated collections of genome, pathway, and regulatory data. It supports creation, curation, dissemination and Web-publishing of organism specific databases, called Pathway/Genome Databases (PGDBs), that integrates many types of data. It performs computational inferences, including prediction of metabolic pathways [33]. This software takes as input an annotated genome and compares the encoded proteins with the proteins in the MetaCyc and EcoCyc databases and maps them in metabolic pathways from BioCyc database. MetaCyc consists of a set of 2260 pathways from 2600 different organisms and their constituent enzymes, EcoCyc describes the genome and the biochemical machinery of *Escherichia coli* K-12 MG1655 and BioCyc consists of a collection of 5500 Pathway/Genome Databases (PGDBs). If the enzymes are found in the genome, the pathway is scored as present. The deduced metabolic network tends to have more pathways than actually present in the bacterium, but it supports well as a foundation element for inference [34].

An accurate comparative analysis of the molecular functions has been performed starting from the annotations computed with RAST in which each ORF function was annotated in a manually curated hierarchical taxonomy based on categories, subcategories, subsystems and roles. Categories contain more specific subcategories and subsystems which represent the actual function (a particular cellular basal metabolism or regulation mechanism) of the gene in which it plays a specifc role.

The enzymes of the subsystems represented in the RAST annotation have been mapped to the pathways of the Pathway Tools database. These pathways and the corresponding set of reactions present in the two genomes are compared in order to find differences between them and to quantify, for each pathway, how many components (enzymes) can be found in the genome annotations.

Pathway Tools and RAST have a different annotation system that does not permit to have a complete description of the pathway in terms of presence of their components (reactions or enzymes) in the genome. For this reason the results of the two annotation systems have been integrated together.

Phylogenomic analyses

In order to compare the two R7 and BCP1 genomes with other *Rhodococcus* genus members, a phylogenetic analysis of 28 different strains was performed. The dataset contains bacteria belonging to the following species: *erythropolis* (2), *rhodochrous* (2), *opacus* (4) (including R7),

wratislaviensis (2), *jostii* (2), *rhodnii* (1), *pyridinivorans* (2), *ruber* (2) and other less representative species (see results for more details).

We built the concatemer sequences of four marker genes of the 28 strains, that were previously identified as conserved and informative for the bacteria classification: *16S rRNA* gene, *secY* gene, *rpoC* gene and *rpsA* gene [35]. The sequences were aligned with the Muscle sequence alignment program [36] using default parameters and the phylogenetic tree was built with PhyML program [37] that iteratively search for the optimal tree with a Nearest Neighbor Interchange algorithm evaluating Maximum Likelihood probability at each step starting from 5 random generated trees.

Phenotype Microarray Assays on 811 Compounds

Phenotype Microarray was conducted using Biolog microplates, a tetrazolium-based growth assay developed by Biolog Incorporated (Biolog, Inc., Hayward, CA). R7 and BCP1 strains were assayed on microplates PM1 to PM4, PM9, PM10, PM11 to PM20 testing 811 different substrates such as several carbon sources, nitrogen sources, sulphur and phosphorous sources, different concentrations of ions and osmolites, a wide variety of antibiotics, antiseptics, heavy metals, other inhibitors, and pH stress. Phenotype Microarray technology uses the irreversible reduction of tetrazolium violet to formazan as a reporter of active metabolism [38]. The reduction of the dye causes the formation of a purple colour that is recorded by a charge-coupled-device camera every 15 min and provides quantitative and kinetic information about the response of the microbial cells in PM plates. All the procedures were performed as indicated by the manufacturer. Strains were grown at 30°C on BUG agar (Biolog), and then, each strain was picked with a sterile cotton swab from the agar surface and suspended in 1 mL of sterile mineral medium. Each strain suspension was added to 15 mL of physiological solution until cell density of 79% transmittance (T) was reached on a Biolog turbidimeter [20]. The culture medium was prepared for each type of PM plate: a mineral medium without carbon source was prepared for PM 1, 2 and for PM from 11 to 20; for PM 3 a mineral medium without N source; two kinds of media were prepared for PM 4, one without sulphur source and another without phosphorus source. Inocula for PM9 and PM10 were prepared as described by Viti et al., 2007. After the addition of 1% dye G (vol/vol) to the suspension, 100 μl of the mixture was inoculated into each well of the microplates. All PM microplates were incubated at 30°C in an OmniLog reader, and they were monitored automatically every 15 min. Reads were recorded for 72 h and data were analyzed using OminoLog PM software (release OM_PM_109M), which generated a time course curve for tetrazolium color developed. Each strain was analyzed at least in duplicate and the results were checked for consistency [19]. The growth was also evaluated in each well using optical density measurements at 590 nm (OD_{590}) after 72 hours.

Phenotype Microarray on organic/xenobiotic compounds

The metabolic abilities of R7 and BCP1 strains were also assayed on 41 organic/xenobiotic compounds added as the only carbon and energy source, such as aliphatic hydrocarbons, polycyclic aromatic hydrocarbons, aromatic and compounds belonging to the BTEX group, and naphthenic acids. Strains were grown at 30°C on BUG agar (Biolog), and then, each strain was picked with a sterile cotton swab from the agar surface and suspended in 1 mL of sterile mineral medium. Each strain suspension was added to 15 mL of mineral medium without carbon source until cell density of 79% transmittance was reached on a Biolog turbidimeter. 1% dye G (vol/vol) was added to each suspension before inoculation, while the substrate was supplied to the inoculated wells with a minimum of two wells of distance. The plates were incubated at 30°C in an OmniLog reader and were monitored automatically every 15 min. Readings were

recorded for 72 h and the data were analyzed using OminoLog PM software (release OM_PM_109M). Each strain was analyzed at least in duplicate and the results were checked for consistency. The growth was also evaluated using OD_{590} after 72 hours.

Statistical analysis on Phenotype Microarrays

Activity curves from Phenotype Microarray experiments for both *Rhodococcus* sp. BCP1 and *R. opacus* R7 have been analyzed. A clustering analysis has been performed with the DuctApe program [39] to group together the curves using nine characteristic curve parameters and three sigmoid-like models. Errors in the fitting confirmed that a direct calculation of activity area in this case is more accurate that a model-estimated activity based on not completely developed growth curves.

Activity values are reported for all the substrates into the standard Biolog Plate PM (1, 2, 3, 4, 9,10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20) and for the custom set of xenobiotics as described before. In this analysis the activities are scattered in a 2D plot; in the same plot three diagonal lines representing the mean (μ_A), standard deviation (σ_A) and double standard deviation ($2\sigma_A$) of the difference activity (ΔA) of the two genomes are overlapped. In this way it is easy to highlight regions of specificity for the two genomes on the tails of the gaussian distribution: *R. opacus* R7 specific compounds are in regions where $\Delta A < -\sigma_A$ and *Rhodococcus* sp. BCP1 specific compounds are on the other extreme points: $\Delta A > +\sigma_A$.

For both *R. opacus* R7 and *Rhodococcus* sp. BCP1 strains the PM activities obtained with different compounds were associated with enzymes (EC code) gathered from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [40, 41]. Enzyme not available are labeled as unknown.

The activities reported in the heat maps are normalized in a scale from low activity (red color) to high activity (green color).

Identification of genetic aspects for xenobiotic degradation pathways

The degradation pathways of some xenobiotic compounds and putative gene clusters were predicted in *R. opacus* R7 and *Rhodococcus* sp. 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 partially described and compared with those from other *Rhodococcus* strains reported in literature. The NCBI pipeline and RAST server were used for genome predictions and comparisons.

Results

Genome features of *Rhodococcus opacus* R7 and *Rhodococcus* sp. BCP1 strains

The genomes of *Rhodococcus opacus* R7 and *Rhodococcus* sp. BCP1 strains were completely sequenced [22, 23]. R7 genome was arranged in a chromosome of 8,466,345 bp and in 5 plasmids (pDG1 of 656,443 bp, pDG2 of 426,388 bp, pDG3 of 352,342 bp, pDG4 of 191,359 bp, and pDG5 of 25,175 bp), giving a total genome size of 10.1 Mb with a G-C content of 66.0%. RAST analysis of R7 genome sequence identified a total of 9,602 open reading frames (ORFs) and 62 RNAs genes (9 rRNAs and 53 tRNAs). *Rhodococcus* sp. BCP1 genome consisted of one chromosome of 6,008,321 and two plasmids (pBMC1 and pBMC2 of 120,373 bp and 103,29 bp, respectively), giving a total genome size of 6.2 Mb with a G-C content of 70.4%. A total of 6,206 open reading frames (ORFs) and 58 RNAs genes (8 rRNAs and 50 tRNAs) were annotated.

By RAST analysis, the annotated ORFs of *R. opacus* R7 and *Rhodococcus* sp. BCP1 were classified into categories as reported in Table 2. As a result of this classification, the number of R7 ORFs was double than BCP1 for each category. Moreover, 457 and 425 subsystems were found in R7 and in BCP1 strains, respectively. In R7, subsystem categories representing the metabolism of carbohydrates, amino acids and cofactors, vitamins, prosthetic groups, or pigments, are the most abundant, and they account for 1,236, 995, or 665 proteins, respectively. 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 amongst the 134 annotated are predicted to catalyze the oxidation of organic compounds with industrial and environmental relevance, such as linear alkanes, cyclic ketones, aromatic compounds (e.g., benzoate, catechol, gentisate, salycilate, and byphenil), 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. In BCP1 subsystem categories representing the metabolism of carbohydrates, amino acids and cofactors, vitamins, prosthetic groups, or pigments are the most abundant, and they account for 623, 565, or 429 proteins, respectively. In BCP1, 415 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 amongst the 134 annotated genes are

Table 2. Number of functions in each of the RAST category.

Category	*R. opacus* R7	*Rhodococcus* sp. BCP1
Cofactors, Vitamins, Prosthetic Groups, Pigments	665	429
Cell Wall and Capsule	111	80
Virulence, Disease and Defense	124	104
Potassium metabolism	17	24
Miscellaneous	115	81
Phages, Prophages,Transposable elements, Plasmids	0	7
Membrane Transport	108	92
Iron acquisition and metabolism	16	16
RNA metabolism	106	87
Nucleosides and Nucleotides	156	102
Protein Metabolism	257	230
Cell Division and Cell Cycle	35	27
Motility and Chemotaxis	5	4
Regulation and Cell signaling	75	51
Secondary Metabolism	13	7
DNA metabolism	103	124
Fatty Acids, Lipids and Isoprenoids	745	415
Nitrogen Metabolism	71	38
Dormancy and Sporulation	3	3
Respiration	221	158
Stress Response	162	138
Metabolism of Aromatic Compounds	267	128
Amino Acids and Derivatives	995	565
Sulphur Metabolism	126	58
Phosphorus Metabolism	35	32
Carbohydrates	1236	623
Total	**5767**	**3623**

doi:10.1371/journal.pone.0139467.t002

predicted to catalyze the oxidation of organic compounds with industrial and environmental relevance, 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 stereo-specific oxidation of a large number of substrates.

Metabolic network reconstruction

The metabolic network reconstruction of *R. opacus* R7 and *Rhodococcus* sp. BCP1 genomes was performed using pathway tools software 18.5. The metabolic reconstruction resulted in a model containing 1735 metabolic reactions for both R7 and BCP1 strains. As a result, 322 (R7) and 274 (BCP1) pathways, 1836 (R7) and 1660 (BCP1) enzymatic reactions, 10 (R7) and 8 (BCP1) transport reactions, 2023 (R7) and 1436 (BCP1) enzymes, 13 (R7) and 9 (BCP1) transporters were assigned by this software analysis. The identified pathways were divided in classes, as described in Table 3 and categorized into the RAST categories (Table 2). RAST annotation assigned 26 functional categories for both genomes, 457 subsystems and 6089 functional roles for *R. opacus* R7 and 425 subsystems and 3703 functional roles for *Rhodococcus* sp. BCP1. The number of functions and their distribution reflected the differences in size of the two genomes (10 Mb for *R. opacus* R7 and 6 Mb for *Rhodococcus* sp. BCP1).

Although the integration of the information of pathway tools and RAST allowed to obtain an improved pathway reconstruction for both the genomes, only around 50% of the pathways have all their enzymes represented in the annotation. This percentage increases up to 70% when pathways, represented by all the enzymes except one, are considered.

Whole-genome comparison

Whole genome sequence comparison of *R. opacus* R7 and *Rhodococcus* sp. BCP1 with a set of four other reference genomes (*R. jostii* RHA1, *R. opacus* PD630, *R. opacus* B4, *R. pyridinivorans* SB3094) was performed using the Mauve program, in the Mauve 2.3 Version and using the Last Software (S1 Fig). Aligned segments of pair of genomes have been collected in a dataset, filtered to remove regions <300 bp in length and grouped by genome. Similarity score between one reference genome and each other genome has been calculated considering the total amount of bases shared between the reference genome and each other genome under analysis.

Results show that *Rhodococcus* sp. 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 4). Considering the genome comparison amongst the *R. opacus* members, they share a

Table 3. Numbers of metabolic pathways deriving from metabolic reconstruction of *R. opacus* R7 and *Rhodococcus* sp. BCP1 genomes.

Pathway Class	Number of Pathways (R7)	Number of Pathways (BCP1)
activation/inactivation/interconversion	3	2
biosynthesis	220	217
degradation/utilization/assimilation	171	115
detoxification	5	4
generation of precursor metabolites and energy	25	21
metabolic clusters	8	8
superpathways	82	72
TOT	**322**	**274**

doi:10.1371/journal.pone.0139467.t003

Table 4. Similarity scores for *Rhodococcus* genomes under analysis. Values in the matrix represent the percent of bases shared in regions longer than 300 bp.

	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%	-

doi:10.1371/journal.pone.0139467.t004

range of similarity of 70%-90% percentage. A lower percentage of similarity resulted from the comparison of either *R. pyridinivorans* SB3094 or *Rhodococcus* sp. BCP1 with each *R. opacus* strain (around 50% and 40%, respectively). The analysis of relative positions of the sequences shared amongst the genomes (chromosome and plasmids) shows that they mainly do not maintain the same order (Table 5).

The percentage of the regions of each *Rhodococcus* strain shared amongst all the genomes under analysis, ranged between 37% and 50% values, indicating a high level of genetic variability amongst *Rhodococcus* genus strains. The genomic diversity in this group of bacteria is represented in the diagram of Venn shown in Fig 1. 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 *Rhodococcus* sp. BCP1 and *R. opacus* R7 are located on the chromosomes; this result suggests that these regions could be characteristic of the *Rhodococcus* genus.

Unique and non-unique genomes features

Genomic comparison of *R. opacus* R7 and *Rhodococcus* sp. BCP1 with four strains belonging to *Rhodococcus* genus allowed to identify unique genomic regions in their genomes (Table 6) (S1 Table). The genomic divergence was investigated in the two genomes and in the complete and well characterized sequenced genomes of *R. jostii* RHA1, *R. opacus* PD630, *R. opacus* B4 and *R. pyridinivorans* SB3094. The comparison showed that they are highly conserved except for the unique regions (longer than 300 bp); the majority of them including annotated genes.

The six strains have different amount of unique regions in the entire genomes as reported in Table 7. In particular, *Rhodococcus* sp. BCP1 showed the highest content of unique regions

Table 5. Numbers of similar regions that maintain the same order in each pair of genomes respect the total number of similar regions in R7 and BCP1 strains.

	R. sp. BCP1	*R. opacus* R7	*R. jostii* RHA1	*R. opacus* PD630	*R. opacus* B4	*R. pyridinivorans* SB3094
***R.* sp. BCP1**	-	286/25919 (1.10%)	475/23532 (2.02%)	400/17911 (2.23%)	481/22904 (2.10%)	642/12397 (5.18%)
***R. opacus* R7**	-	-	123/40390 (0.30%)	121/30862 (0.39%)	129/39122 (0.33%)	261/21603 (1.21%)
***R. jostii* RHA1**	-	-	-	220/26720 (0.82%)	233/33364 (0.70%)	553/20422 (2.71%)
***R. opacus* PD630**	-	-	-	-	195/25465 (0.77%)	462/15651 (2.95%)
***R. opacus* B4**	-	-	-	-	-	546/19708 (2.77%)
***R. pyridinivorans* SB3094**	-	-	-	-	-	-

doi:10.1371/journal.pone.0139467.t005

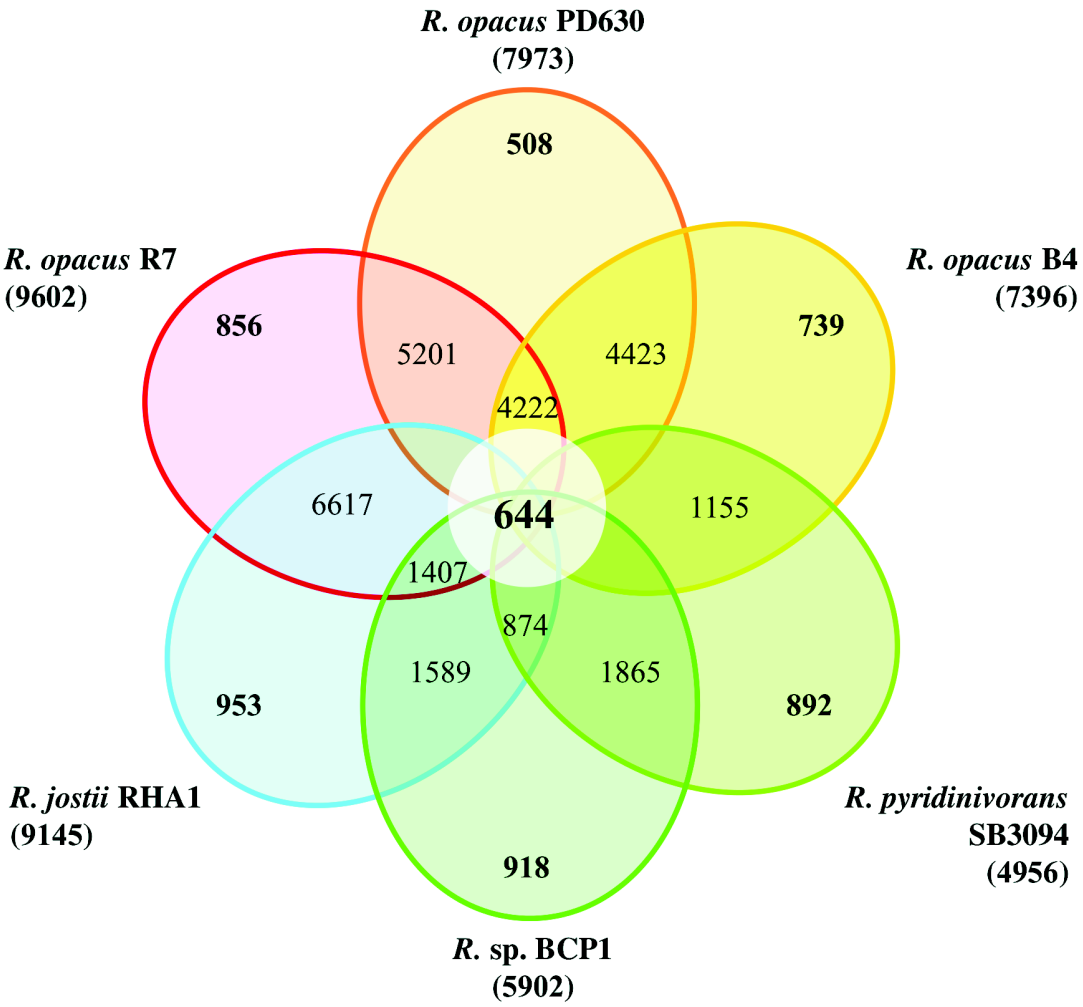

Fig 1. Venn Diagram. Genomic comparison of *R. opacus* R7 and *Rhodococcus* sp. BCP1 with other *Rhodococcus* strains, including *R. jostii* RHA1, *R. opacus* PD630, *R. opacus* B4, *R. pyridinivorans* SB4094. Each strain is represented by a colored oval. Number of predicted protein coding genes (CDSs) 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.

doi:10.1371/journal.pone.0139467.g001

(16.3%) followed by *R. jostii* RHA1 (11.41%) and *R. opacus* R7 (11.32%). Within the unique regions, a percentage of 30% and 40% of the annotated genes were predicted to code for hypothetical proteins in BCP1 and R7, respectively. Moreover, around 30% of these regions was found to be non-coding sequence. Genes encoding several enzymatic classes predicted to be involved in the organic compounds metabolism were identified in unique regions in *R. opacus* R7 and *Rhodococcus* sp. BCP1, including oxygenases, hydroxylases, dehydrogenases, hydrolases, oxidoreductases, ligases, isomerases, aldolases, ion-transporting ATPases and cytochromes P450 (S2 Table). The presence of these genes suggests different or specific metabolic capabilities of these *Rhodococcus* strains. In particular, considering the oxygenase/hydroxylase class, BCP1 unique regions contain a whole gene cluster encoding the soluble di-iron monooxygenase (*smo*) involved in the short-chain *n*-alkanes metabolism [28], two cyclohexanone monooxygenases, one cytochrome P450, and several oxygenases involved in the aromatic compound metabolism (e.g. 2,3-dihydroxyphenyl 1,2-dioxygenase, 4-hydroxyphenyl acetate

Table 6. Unique regions larger than 4 kb identified in R7 and BCP1 strains.

Region n.	*R. opacus* R7 Predicted key functions	Size (bp)	Unique gene number	*Rhodococcus* sp. BCP1 Predicted key functions	Size (bp)	Unique gene number
1	Amino acids biosynthesis	4373	3	Carotenoid biosynthesis	4962	5
2	Aromatic compounds degradation	5323	4	Unknown function	7992	7
3	Transcriptional regulators	4004	4	Amino acids metabolism	7102	5
4	Phage region	5208	2	Cell division	6351	6
5	Unknown function	8845	2	Cell division	12430	10
6	Unknown function	7095	6	Metal transport	4332	4
7	Phage region	6060	6	Unknown function	4523	4
8	Polysaccharides biosynthesis	7648	3	Cell wall biogenesis	6427	3
9	Amino acids biosynthesis	4256	2	Carbon and energy storage	4593	3
10	Unknown function	4289	2	Fatty acid metabolism	4596	4
11	Unknown function	5176	4	Unknown function	5575	5
12	Amino Acids biosynthesis	4051	3	Mobile element	12922	11
13	Phosphonate metabolism	4131	4	Aromatic compound degradation; Fatty acid degradation	4047	4
14	Unknown function	5775	4	Cell wall biogenesis; O-antigen	18068	17
15	,Monosaccharides metabolism	5707	3	Phospholipid biosynthesis	6067	7
16	Fatty acids metabolism	4195	2	Lactate utilization	4416	2
17	Antibiotic resistance; Folate Biosynthesis	4702	6	Amino acids metabolism; Fatty Acid metabolism; Phage region	30497	44
18	Monosaccharides metabolism	7569	6	Phage region	6871	12
19	Aromatic compounds degradation	4104	7	Organic acids metabolism	4815	2
20	Amino acids biosynthesis	4662	4	Fatty acids biosynthesis; DNA replication	6586	4
21	Membrane transport systems	7072	5	Mobile element	5229	6
22	Amino acids biosynthesis	4935	4	Peptidoglycan Biosynthesis	9655	4
23	Unknown function	6723	6	Mobile element	16590	7
24	Unknown function	6764	4	DNA/RNA topology; ATP/GTP binding protein	7210	3
25	Phage region	20558	16	Lipoprotein biosynthesis	4899	3
26	DNA recombination/repair	5651	2	Fatty acids biosyntehsis	4437	5
27	Unknown function	7525	1	Phage region	9451	7
28	Unknown function	6665	9	Mobile element	5618	2
29	Carbohydrates Metabolism	4823	5	PAH Hydrocarbons degradation	4556	3
30	Phage region	14443	16	Unknown function	5551	6
31	Antibiotic resistance	4233	4			
32	ABC Transporters	6368	5			
33	Amino acids biosynthesis	4221	4			
34	Unknown function	4522	4			
35	Amino acids metabolism	5411	4			
36	Aliphatic hydrocarbon degradation	5491	3			
37	Flavonoid metabolism	4180	4			
38	Carbohydrates metabolism; Aromatic compounds degradation	5293	7			
39	Unknown function	6728	9			
40	Aliphatic hydrocarbon degradation	4132	1			
41	Unknown function	4476	6			
42	Stress response	4006	4			

(*Continued*)

Table 6. (*Continued*)

Region n.	*R. opacus* R7 Predicted key functions	Size (bp)	Unique gene number	*Rhodococcus* sp. BCP1 Predicted key functions	Size (bp)	Unique gene number
43	Unknown function	9977	13			
44	Unknown function	4324	1			
45	Oxidoreductase	4899	6			
46	Mobile element	4390	1			
47	Serine/threonine protein kinase	7470	10			
48	Unknown function	7688	4			
49	Unknown function; Phage region	8499	9			

doi:10.1371/journal.pone.0139467.t006

3-monooxygenase, catechol 1,2-dioxygenase and biphenyl dioxygenase beta subunit). Conversely, in *R. opacus* R7 unique regions, two alkanal monooxygenases and one alkane-1-monooxygenase (annotated in RAST as alkanal monooxygenase), one cyclohexanone monooxygenase, and one pentachlorophenol monooxygenase were found. Other genes encoding several enzymatic classes are also located in unique regions such as kinases, peptidases, reductases, permeases, synthetases, esterases, cyclases, transcriptional regulators and membrane proteins. Interestingly, the presence of mobile elements (transposases, phage and mobile element proteins) was found, suggesting the occurrence of recombination events. Other important features of aerobic microorganisms able to degrade a wide range of organic compounds are the abilities to counteract stress conditions. Concerning this aspect, several genes coding for proteins involved in resistance to oxidative stress, metals and antibiotics were found in these regions (Table 6). The unique regions contained some heavy metal-resistance genes (i.e. lead-cadmium-zinc-mercury transporting ATPases, cobalt-zinc-cadmium resistance protein, metal-binding enzymes), stress-resistence genes (i.e.catalases, universal stress proteins, chaperons), and antibiotic-resistance genes (i.e.drug/metabolites transporters, beta-lactamase related proteins, penicillin-binding proteins).

Taxonomic classification of *Rhodococcus opacus* R7 and *Rhodoccocus* sp. BCP1

To develop a framework for resolving the phylogeny of the two *Rhodococcus* strains, we constructed a multi-locus sequence analysis (MLSA) maximum likelihood (ML) tree based on four marker genes (*16SrRNA*, *secY*, *rpoC* and *rpsA*) that were previously identified as conserved and informative for the bacteria classification [35]. The phylogenetic tree based on sequence alignments with reference strains of *Rhodococcus* genus show the distinctly defined phylogenetic positions of *Rhodococcus* sp. BCP1 and *R. opacus* R7 (Fig 2). This is the first taxonomical study

Table 7. Uniqueness characteristics of the six compared *Rhodococcus* spp. strains.

Uniqueness characteristics	*R. opacus* R7	*R.* sp. BCP1	*R. jostii* RHA1	*R. opacus* PD630	*R. opacus* B4	*R. pyridinivorans* SB3094
Lenght of unique regions (bp)	1,145,011	1,015,869	1,106,630	650,025	912,915	545,841
Unique regions (%)	11.32	16.3	11.41	7.1	10.33	10.44
ORFs in unique regions	1202	1266	1247	575	1058	1197
ORFs of hypothetical proteins in unique regions (%)	39.77	30.88	48.44	42.78	41.30	46.45
Not Annotated (%)	29.03	29.46	24.78	32.52	27.32	27.57

doi:10.1371/journal.pone.0139467.t007

of *R. opacus* R7 that correlates R7 strain to *Rhodococcus opacus* and *Rhodococcus wratislaviensis* species in a clade that also includes *R. jostii* RHA1. *Rhodococcus* sp. BCP1 clusters with *R. aetherivorans* strains in a clade including also *R. ruber* species. These results on BCP1 are in line with previous phylogenetic analyses based on *Rhodococcus* 16S rDNA genes and 1-alkane monooxygenase (AlkB) protein sequences reported by Táncsics et al., 2015 [42]. In conclusion, *Rhodococcus* sp. BCP1 can be taxonomically related to *R. aetherivorans* species. The importance of this taxonomic definition is related to the fact that BCP1 strain genome is the first complete genome available in database for *R. aetherivorans* species.

Phenotype Microarray analysis of nutritional metabolism by *Rhodococcus opacus* R7 and *Rhodococcus* sp. BCP1

In order to acquire a global understanding of the metabolic capabilities of *Rhodococcus* sp. BCP1 and *R. opacus* R7 cellular processes, we used phenotype microarrays (PM) analysis. Using PM plates 1 to 4, 379 different compounds were tested, including 190 different carbon (C) sources, 95 nitrogen (N) sources, 59 phosphorus (P) sources, 35 sulphur (S) sources. By means of the PM plates 9 to 20, the tolerance/sensitivity to different osmolytes, pH conditions and other chemicals (i.e. antibiotics, metals, preservatives etc.) was assayed. In this work, we shed light on the metabolic features of the two strains. In some cases, these metabolic properties were related to the genetic traits present in their genomes [22, 23].

The 190 carbon sources were divided in different chemical categories, including carbohydrates, carboxylic acids, nitrogen containing compounds, alcohols, amides, amines, esters, fatty

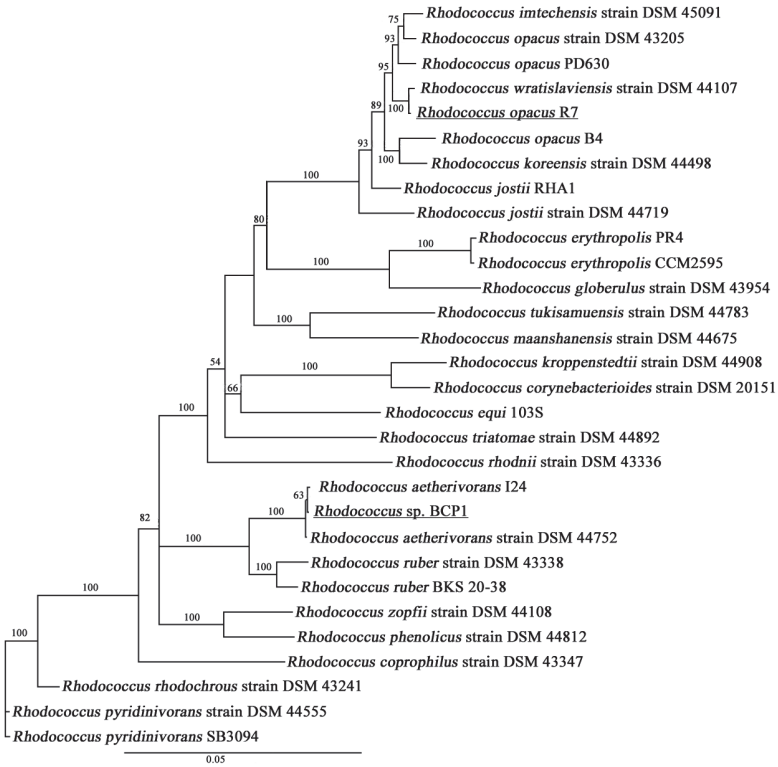

Fig 2. PhylogeneticTree. Phylogenetic analysis of *R. opacus* R7 and *Rhodococcus* sp. BCP1 based on sequence alignments with reference strains of *Rhodococcus* genus. The tree was constructed based on concatemer sequences of four marker genes of the 28 strains: *16S rRNA* gene, *secY* gene, *rpoC* gene and *rpsA* gene.

doi:10.1371/journal.pone.0139467.g002

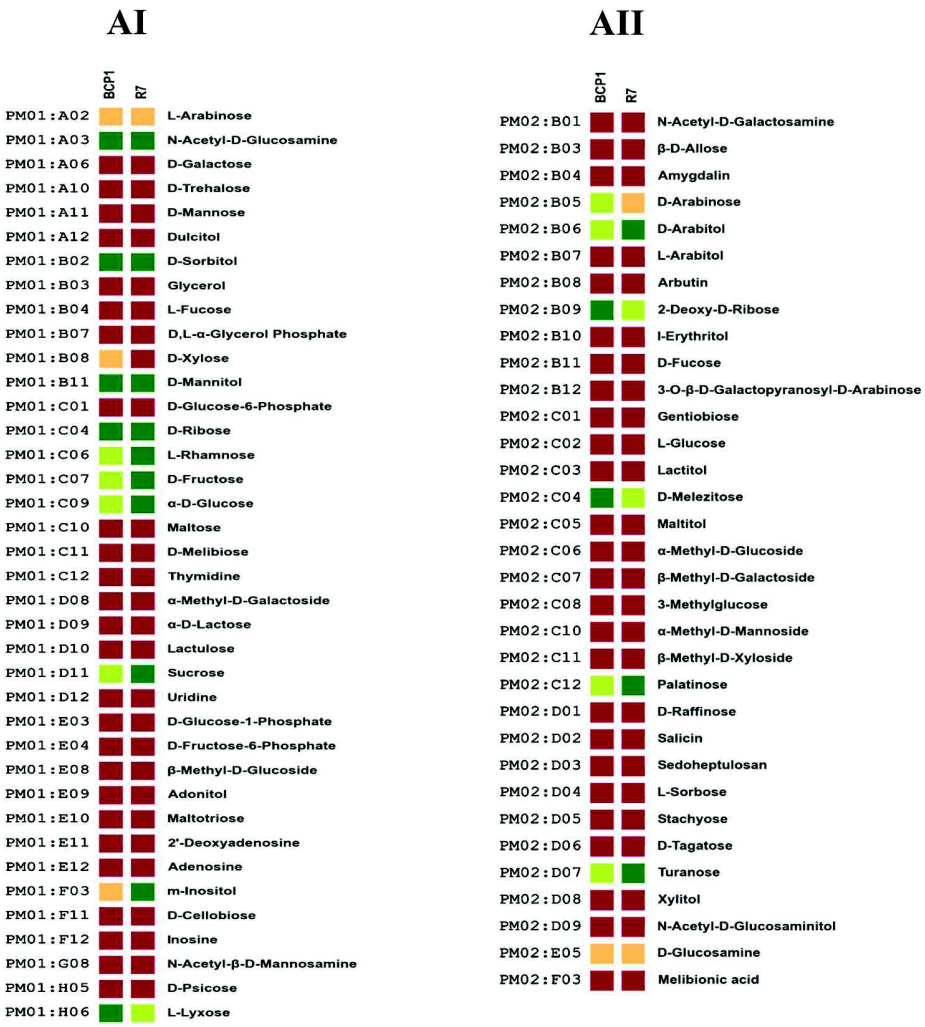

Fig 3. Phenotype Microarray PM with different carbohydrates as carbon sources. Metabolic differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of carbohydrates (AI, AII). Based on activity values of phenotype microarray analysis, threshold values were established for every plates. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.

doi:10.1371/journal.pone.0139467.g003

acids, polymers (S3 Table). R7 and BCP1 strains were metabolically active on the following monosaccharides: α-D-glucose, D-fructose, D-ribose, L-rhamnose, L-lyxose and 2-deoxy-D-ribose. Both the strains could oxidize the three disaccharides composed by glucose and fructose monomers (sucrose, palatinose and turanose) as well as the trisaccharide D-melezitose (formed by D-turanose and D-glucose). They showed high metabolic activity on the sugar alcohols D-sorbitol, D-mannitol and D-arabitol. Amongst the carbon sources utilized, R7 generally showed higher activities compared to BCP1 (69% of the carbon sources utilized by R7 gave high activity level *vs* 41% for BCP1) (Fig 3 Panel AI, AII). Both the strains could utilize a wide range of carboxylic acids (Fig 4 Panel BI, BII). Several short-chain carboxylate compounds (C2-C6) were oxidized by R7 and BCP1 strains including acetic and propionic acids as well as other common intermediates in central metabolism such as pyruvic, succinic, and citric acids. They both oxidized and grew on gluconic acid that is degraded by the pentose phosphate

pathway. Carboxylate metabolism also included the C10 dicarboxylate sebacic acid whose metabolism involves the β-oxidation pathway. β-oxidation could also explain the activity on Tween 40 and Tween 80 compounds by BCP1 and R7 (Fig 4 Panel C). Tween 20 did not support growth of the two *Rhodococcus* spp. probably because of the toxicity resulting from the hydrolytic process of this compound [16]. Interestingly, the utilization of Tween compounds is considered diagnostic of the substrate specificity of bacteria toward the use of hydrocarbons as a source of carbon and energy [43]. The complexity of fatty acids β-oxidation pathway present in BCP1 and R7 genomes is consistent with the number of genes coding for each of the oxidation steps (S4 Table).

By RAST annotation analysis, the genes predicted to be involved in central metabolic pathways for carbohydrates and organic acids utilization were more than double for R7 compared to BCP1 (Table 2) (S5 Table). Some of the metabolic differences between the two strains in terms of activity level can be therefore related to the presence of genes in R7 coding for specific functions that lack in BCP1 genome. In R7 genome, genes coding for L-rhamnose ABC-transporter components (EC 3.6.3.25), a L-rhamnose mutarotase (EC 5.1.3.322), a L-rhamnose isomerase (EC 5.3.1.14), a L-arabinose isomerase (EC 5.3.1.4), a L-rhamnose 1-phosphate aldolase (EC 4.1.2.-), a rhamnulokinase (EC 2.7.1.5) and a LacI family transcriptional regulator (EC 3.6.3.17) are clustered in one genomic region and can internalize, metabolize and regulate the rhamnose utilization in this strain. Three 2-dehydro-3-deoxygluconate kinases (EC 2.7.1.45) can phosphorylate

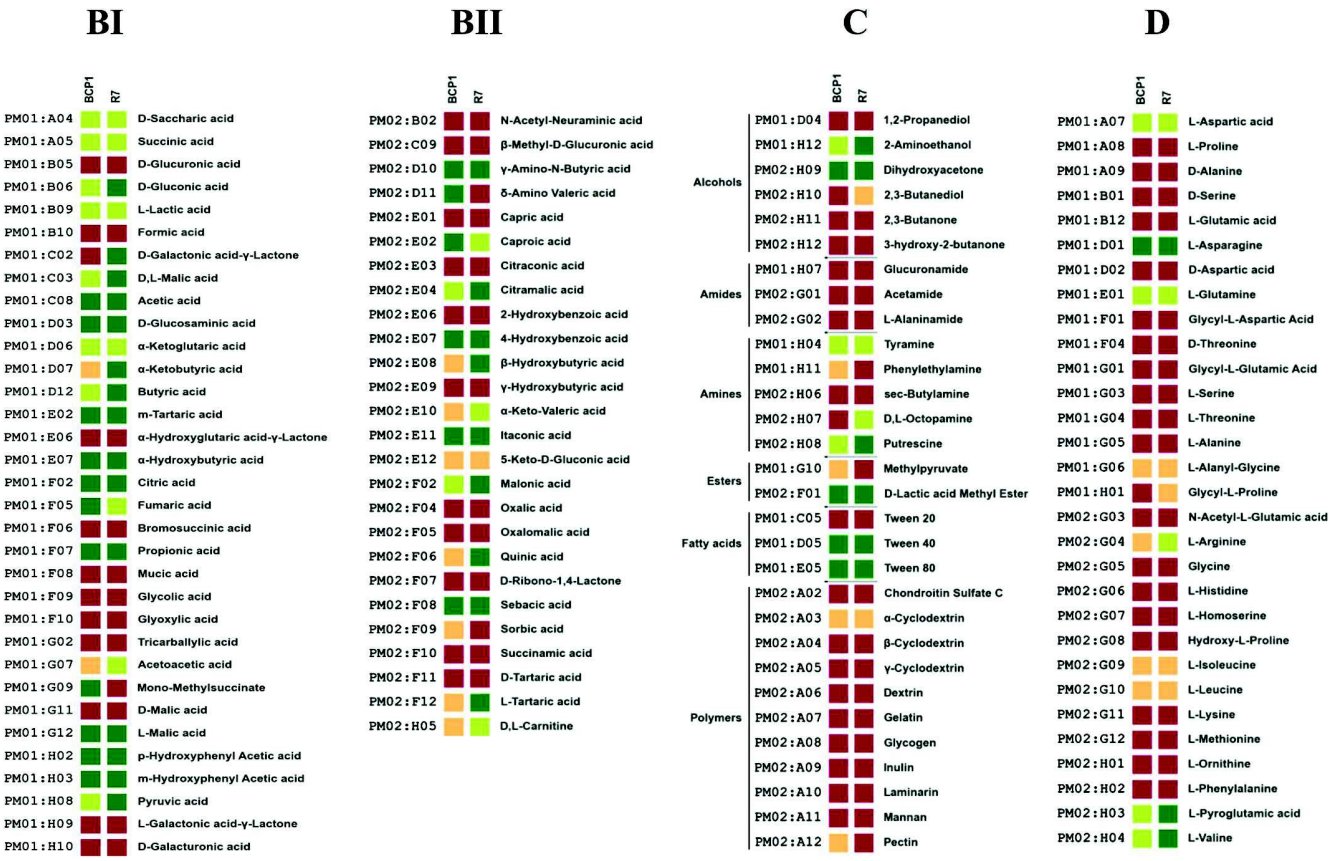

Fig 4. Phenotype Microarray PM with different carboxylic acids-alcohols, amines, amides, esters, fatty acids and polymers-amino acids as carbon sources. Metabolic differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of carboxylic acids (BI, BII), of alcohols, amines, amides, esters, fatty acids, polymers (C), and amino acids (D). Based on activity values of phenotype microarray analysis, threshold values were established for every plates. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.

doi:10.1371/journal.pone.0139467.g004

and activate the 2-dehydro-3-deoxy-D-gluconate that can be converted to D-glyceraldehyde-3-phosphate to be introduced in glycolysis pathway. The activity of a 2-dehydro-3-deoxygluconate kinase can also be associated to a gluconate dehydratase (EC 4.2.1.39) for the conversion of gluconic acid and its derivatives (such as D-glucono-1,5-lactone and 2-2-dehydro-3-deoxy-D-gluconate) into intermediates of the Entner–Doudoroff pathway. A sucrose-6-phosphate hydrolase (EC 3.2.1.26) that is able to break the sucrose 6-phosphate in α-D-glucose-6-phosphate and β-D-fructose is clustered with genes coding for several sugar-transport ABC transporters and one gene coding for an α-glucosidase (EC 3.2.1.20) that hydrolyses (1–4)-α-glucosidic linkages releasing D-glucose. Two ribokinases in R7 (EC 2.7.1.15) are predicited to catalyze the transfer of a phosphate to ribose or to 2-deoxy-D-ribose. Finally, a N-acetylglucosamine phosphotransferase (PTS) system (EC 2.7.1.69) and a N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25) can be related to the higher metabolic acitivity shown by R7 on N-acetylglucosamine.

Both R7 and BCP1 strains show metabolic activity on aromatic and hydroxylated aromatic compounds such as *p*-hydroxy benzoic acid, *m*- *p*- hydroxyphenyl acetic acid. R7 genome analysis evidenced a gene coding for a *p*-hydroxybenzoate hydroxylase (EC 1.14.13.2) located in a chromosomal region showing the same genetic organization of a chromosomal region of *R. jostii* RHA1. This gene is flanked by genes coding for a benzoate transporter, two transcriptional regulators belonging to IclR family and TetR family, and 1-alkane monooxygenase (AlkB) gene cluster. The *p*-hydroxybenzoate hydroxylase oxidizes the aromatic ring to form protocatechuate. All the genes coding for enzymes involved in protocatechuate metabolism are present in R7. In BCP1 the gene coding for the *p*-hydroxybenzoate hydroxylase is present but is located in a region not conserved amongst bacteria and not including a benzoate transporter gene. An IclR transcriptional regulator is also associated to this gene in BCP1. R7 and BCP1 genomes present six and four genes respectively coding for *p*-hydroxyphenylacetate 3-monooxygenase (EC 1.14.14.9); only R7 possesses also a gene coding for the *p*-hydroxyphenylacetate 3-monooxygenase reductase component associated with genes coding for an aromatic ring oxygenase and a phenylacetate-CoA ligase. Amongst the amino acids, L-aspartic acid, L-asparagine, L-glutamine, L-pyroglutamic acid and L-valine could be utilized as carbon sources by both strains; while only R7 showed high metabolic activity on L-arginine (Fig 4 Panel D). The amines putrescine and tyramine could be utilized by R7 and BCP1 for growth. Both the R7 and BCP1 strains showed a slight activity on the polymer α-ciclodextrin (Fig 4 Panel C). This feature can be correlated with the presence of five and two glucoamylases (EC 3.2.1.3) in BCP1 and R7 genomes, respectively. Four out of the five glucoamylases genes of BCP1 did not show any similarity with genes in *R jostii* RHA1. In R7, both the glucoamylase genes presented flanking regions homologous to those found in RHA1. Both R7 and BCP1 strains could not utilize any amide (glucuronamide, acetamide, L-alaninamide) as carbon source.

Using the PM3 plate, the two *Rhodococcus* strains were tested for their ability to grow on 95 different nitrogen sources. R7 was able to utilize 74 out of 95 nitrogen sources, while BCP1 could grow only on 52 of the tested nitrogen compounds (Fig 5 Panel AI, AII, AIII). In particular, amongst the utilized N sources, R7 and BCP1 showed high metabolic activity on 50 and 24 compounds, respectively. Twenty-six N sources were specific for R7 strain growth, while 4 nitrogen sources were used by BCP1 and did not support the growth of R7. The nitrogen sources selective for BCP1 growth were D-lysine, DL-α-amino-caprylic acid, δ-amino-N-valeric acid, ε-amino-N-caproic acid. Interestingly, δ-amino-N-valeric acid was also a carbon source utilized by BCP1 but not by R7, suggesting the presence in BCP1 of transport and/or metabolic system specific for the utilization of this amino fatty acid compound. RAST annotation reported a number of genes predicted to be involved in amino acid metabolism that was double in R7 compared to BCP1 (Table 2). However, a similar number of genes predicted for nitrate and nitrite utilization and ammonia assimilation was annotated in BCP1 and R7

genomes (29 and 33 genes, respectively) (S5 Table). The higher activity values shown by BCP1 compared to R7 on ε-amino-N-caproic acid, D,L-α-amino-caprylic acid and δ-amino-N-valeric acid cannot be related to any genetic trait. However, some of the metabolic features of R7, not observed with BCP1, could be linked to the following genetic aspects. i) Considering the asparagine/aspartate metabolism, R7 has an asparaginase (EC 3.5.1.1) that converts asparagine in aspartate and an aspartate-ammonia lyase (EC 4.3.1.1) that transforms the aspartate in fumarate and ammonia that can be used as nitrogen source; the genes coding for the asparaginase and the aspartate-ammonia lyase are clustered together with a L-asparagine permease that is missing in BCP1 genome; both BCP1 and R7 have the asparagine synthetase (EC 6.3.5.4) that produces asparagine by ligating a NH_3 group to aspartate using ATP; BCP1 genome lacks of an aspartic acid transporter. ii) Considering the arginine metabolism, R7 has genes coding for an arginase (EC 3.5.3.1) and an urease (EC 3.5.1.5) that convert arginine in urea and this in carbon dioxide and ammonia, respectively; R7 has also proton/glutamate aspartate symport protein that can mediate the import of glutamate and aspartate. iii) Considering the glutamine

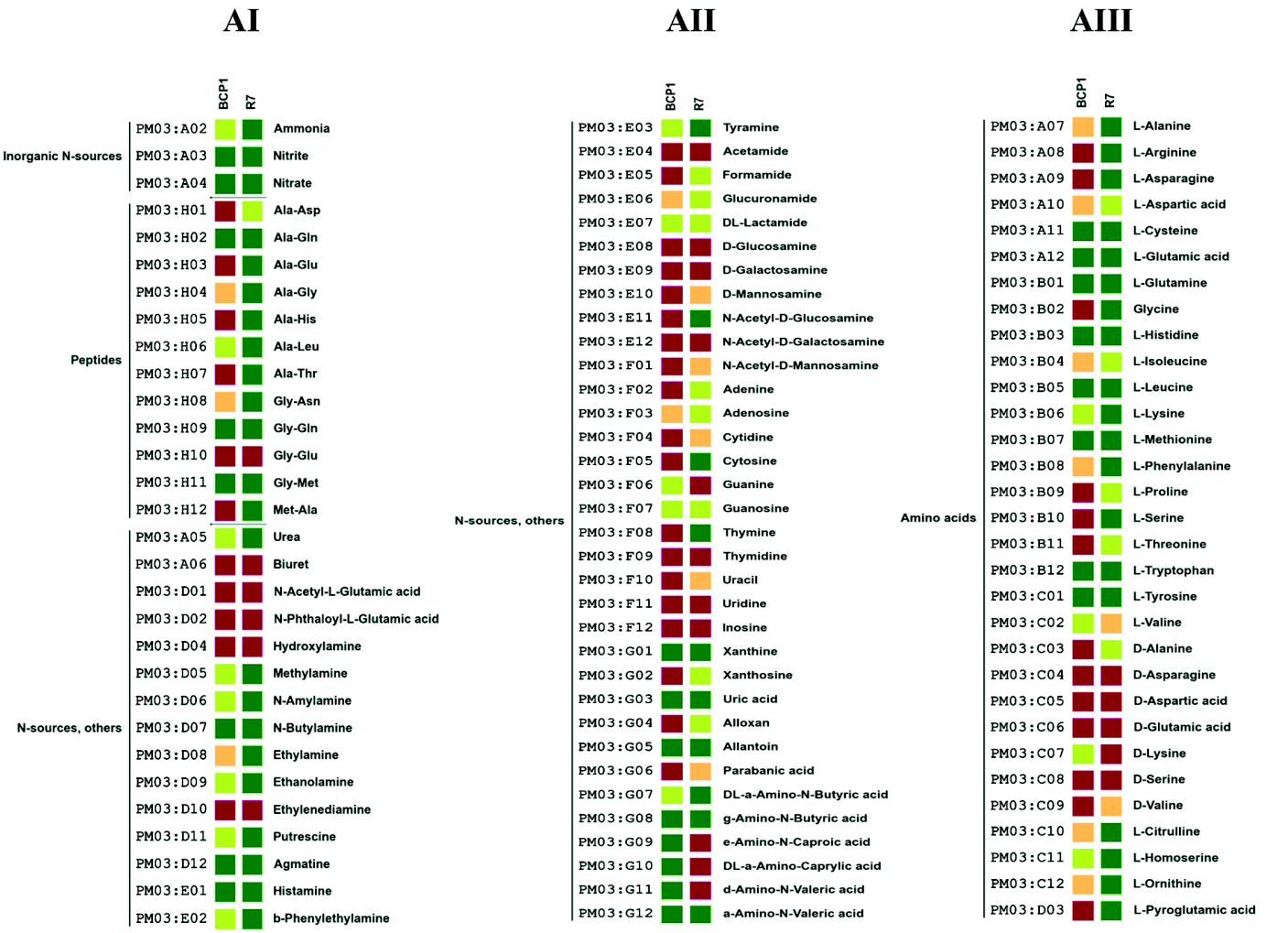

Fig 5. Phenotype Microarray PM with different Nitrogen sources. Metabolic differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of different Nitrogen sources (AI, AII, AIII). Based on activity values of phenotype microarray analysis, threshold values were established for every plates. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.

doi:10.1371/journal.pone.0139467.g005

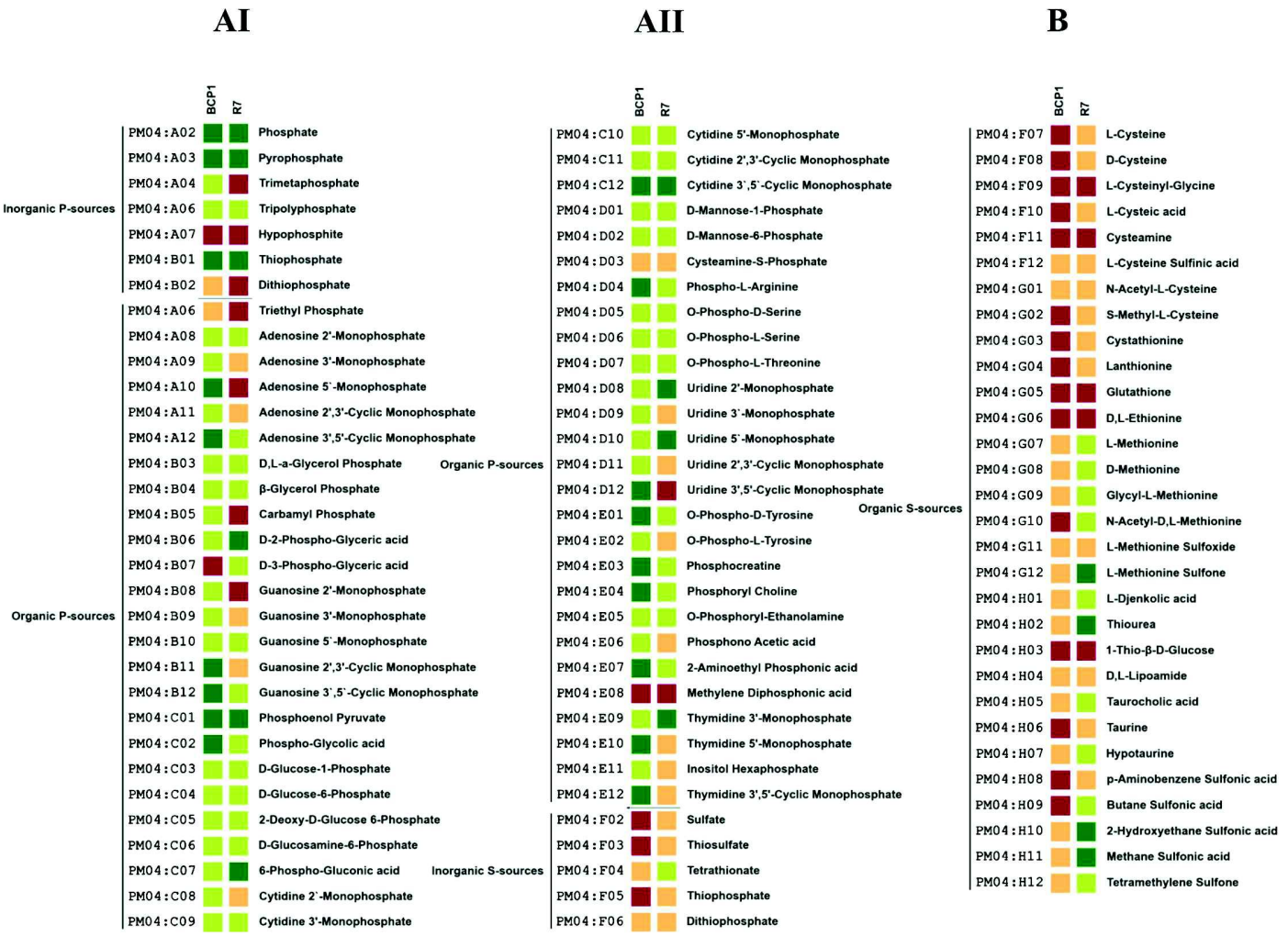

Fig 6. Phenotype Microarray PM with different Phosphorous and Sulphur Sources. Metabolic differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of different Phosphorous Sources (AI, AII) and Sulphur sources (B). Based on activity values of phenotype microarray analysis, threshold values were established for every plates. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.

doi:10.1371/journal.pone.0139467.g006

metabolism, the significantly higher activity of R7 on L-glutamine compared to BCP1 (at least double of BCP1 value) can be correlated with the presence of a L-glutaminase (EC 3.5.1.2) that hydrolyses L-glutamine in glutamate and ammonia. iv) considering the tyrosine metabolism, R7 has a tyramine:oxygen oxidoreductase (EC 1.4.-) which converts the tyramine that is included in tyrosine metabolism into 4-hydroxyphenylacetaldehyde and ammonia. This enzyme could also be related to the higher ability of R7 to utilize tyramine as carbon source.

R7 and BCP1 showed metabolic activity on most of the compounds tested in PM4 (80 and 90%, respectively) indicating a high versatility in phosphorous sources utilization (Fig 6 Panel AI, AII). A great difference was observed in sulphur sources assimilation capacity between BCP1 and R7. Indeed, BCP1 could utilize only 6 sulphur sources out of the 35 tested compounds in PM4 while R7 could assimilate 32 of these compounds (Fig 6 Panel B). This difference can be associated with the great difference between the two strains in terms of genes predicted to be involved in sulphur metabolism (S5 Table).

R7 showed more resistance to low pH compared to BCP1, particularly in presence of several amino acids. Both BCP1 and R7 showed wide resistance to osmolytes. The differences between the two strains regard: i) the higher resistance of R7 to increasing concentration of NaCl, sodium formate and sodium lactate; ii) the higher resistance of BCP1 to increasing concentration of urea (S2 and S3 Figs).

Using the PM11-PM20 plates, the two *Rhodococcus* strains were assayed for their sensitivity to 240 chemical agents (S4, S5 and S6 Figs) tested at four different concentrations. Both the *Rhodococcus* strains were resistant to the presence of several antibiotics (S4 Fig Panel AI, AII, AIII). In particular, they could grow in the presence of all the tested concentrations of the antibiotics belonging to quinolone, fluoroquinolone, cephalosporin, sulfonamide, aminocoumarin (novobiocin) classes. Amongst the tested β-lactam and glycopeptide antibiotics, only oxacillin and vancomycin, significantly inhibited the growth of both the two strains. Both R7 and BCP1 presented seven β-lactamase (EC 3.5.2.6) resistance genes in their genomes. Conversely, BCP1 and R7 showed general sensitivity to macrolide antibiotic class (i.e. erythromycin, josamycin and spiramycin) and to rifamycin group (i.e. rifampicin and rifamycin). For the tetracycline group, only minocycline was toxic for BCP1 and R7 strains. Only BCP1 genome presented one gene coding for a tetracycline resistance protein. Amongst the aminoglycoside antibiotic, both R7 and BCP1 were resistant to all the concentrations of kanamycin, spectinomycin, geneticin and dihydrostreptomycin. The two *Rhodococcus* spp. strains also showed high resistance to amikacin, gentamycin, sisomicin, tobramycin and hygromycin (S4 Fig Panel AI). R7 did not grow on any concentration of paromomycin while BCP1 was tolerant to the lowest concentration of this aminoglycoside antibiotic. Amongst the other tested chemicals (preservatives, drugs etc.) (S5 and S6 Figs Panel AI, AII and Panel AIII, AIV), R7 and BCP1 were in general resistant to the amino acid hydroxamates (protein synthesis inhibitors) and to the nucleic acid analogs. Amongst the quaternary sodium salts, R7 and BCP1 were resistant only to dequalinium. Other compounds such as ionophores (CCCP and FCCP) and the respiration inhibitor iodonitrotetrazolium chloride, inhibited the growth of the two strains even at the lowest concentration. Some of these phenotypic features can be attributed to the presence in their genomes of a high number of genes coding for multidrug resistance proteins and transport systems (20 in R7 and 16 in BCP1). Discrimination between BCP1 and R7 was shown in few cases. In particular, a higher tolerance was shown by R7 compared to BCP1 in the presence of the phenol derivative pentachlorophenol, the antifungal compound patulin and the β-lactam penicillin G. Conversely, BCP1 was more resistant than R7 to hexachlorophene (membrane electron transport inhibitor). Both BCP1 and R7 showed a general tolerance to metal salts even at the highest concentrations. Exceptions include: i) the low/medium tolerance shown by both strains to sodium *m*-periodate, antimony (III) chloride and sodium caprylate (growth observed at the lowest concentration); ii) the sensitivity to vanadate sodium salts (sodium metavanadate and sodium orthovanadate) that inhibited the growth of the two *Rhodococcus* spp. at all the tested concentrations. Thallium acetate was the only metal that had a different effect on the two strains as it inhibited R7 growth at all the concentrations while BCP1 showed tolerance to two out of the four concentrations of this metal salt (S6 Fig Panel B).

Phenotype Microarray on organic/xenobiotic compounds by *Rhodococcus opacus* R7 and *Rhodococcus* sp. BCP1

R. opacus R7 and *Rhodococcus* sp. BCP1 metabolic activity was tested on 41 organic/xenobiotic compounds supplied as sole carbon and energy source. The tested substrates belong to four chemical categories including: 1) aliphatic hydrocarbons and cycloalkanes, 2) BTEX and other aromatic compounds, 3) polycyclic aromatics (PAHs), 4) naphthenic acids and other

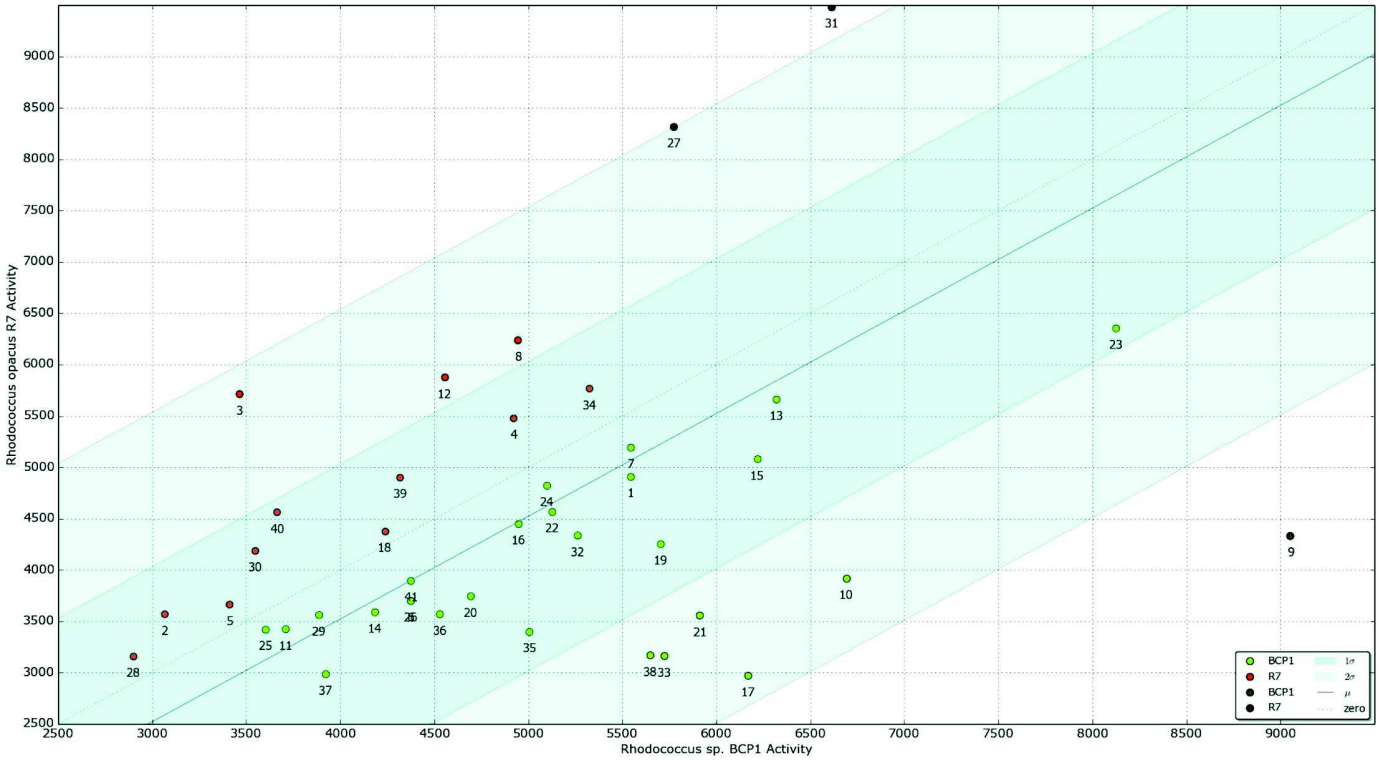

Fig 7. Phenotype Microarray PM with xenobiotic compounds (Scatter Plot). A three diagonal lines represented the mean (μ_A), standard deviation (σ_A) and double standard deviation ($2\sigma_A$) of the difference activity (ΔA) of the two genomes overlapped. It highlighted regions of specificity for the genomes of the two strains on the tails of the gaussian distribution: *R. opacus* R7 specific compounds are in regions where $\Delta A < -\sigma_A$ (red points) and *Rhodococcus* sp. BCP1 specific compounds are on the other extreme points: $\Delta A > +\sigma_A$ (green points). Numbers in the plot represent the tested chemicals reported in the Table 8.

doi:10.1371/journal.pone.0139467.g007

carboxylic acids. Phenotype Microarray analysis was conducted following both the tetrazolium-based metabolic activity and the growth measured as OD_{590} increment over 72 h. Both strains were able to grow on several hydrocarbons and their putative metabolic intermediates (Fig 7) (S6 Table). Based on the growth results in presence of several linear- and cycloalkanes, *Rhodococcus* sp. BCP1 showed a general high activity on all *n*-alkanes tested ranging from *n*-hexane to *n*-hexatriacontane; while *R. opacus* R7 had a preferential activity on *n*-dodecane, *n*-tetradecane, *n*-hexadecane and *n*-eicosane. These data confirmed previous works on BCP1 and R7 strains reported by Cappelletti et al., 2011 and by Zampolli et al., 2014 [29, 25]. Moreover, results showed that R7 strain was not able to grow in presence of odd-alkanes, except for *n*-heptadecane. Both the two strains evidenced activity on cyclohexane and cyclohexanone; however these substrates were preferentially utilized by BCP1 strain. In addition, high level of activity was revealed for both R7 and BCP1 strains on fuel oil, a mixture of different alkanes, including prystane and phytane.

Growth analyses on aromatic hydrocarbons showed that both *Rhodococcus* strains had the same capacity to degrade common aromatic compounds as toluene and ethylbenzene. In accordance with results reported by Di Gennaro et al., 2010 [24], R7 has the ability to grow only on the *ortho* xylene isomer, while BCP1 showed the capability to utilize preferentially the *meta*- and *para*- isomers. In presence of dibenzothiophene (DBT) and of 2-hydroxybiphenyl, R7 showed the highest activity.

Another class of substrates tested on *Rhodococcus* spp. strains are polycyclic aromatic hydrocarbons (PAHs) and their metabolic intermediates. Generally PAHs are metabolized by

two main reactions: the peripheral one transforming a large number of polycyclic aromatic derivatives into central intermediates; and the ring cleavage reaction of intermediates to products that are subjected to the TCA cycle. Phenotype Microarray results showed that R7 preferentially degraded PAHs compounds compared to BCP1, as for example in the case of anthracene. It was also investigated the ability of the two *Rhodococcus* strains to grow on putative PAH degradation intermediates. Results showed a good level of activity on gentisic acid, whereas a lower level was reported on salicylic acid. These experiments confirmed data from a previous work on R7 strain [24], and they suggest the identification of hypothetical intermediates of naphthalene pathway in BCP1 strain.

In presence of different naphthenic acids both strains reported a significant response with a higher activity for *Rhodococcus* sp. BCP1 on the cyclopentanecarboxylic acid (CPCA) and cyclohexanecarboxylic acid (CHCA). On the contrary, both strains responded in a similar manner to different acids that could be possible intermediates of different metabolic pathways, such as 4-phenylbutyric acid, hexanoic acids and 1,4-cyclohexane dicarboxylic acid.

The correlation between the phenotype showed by each strain and the hypothetical genotype was analysed using an adjustment of Ductape software. The enzyme functions involved in the metabolism of 19 substrates out of the 41 tested were predicted for BCP1 and for R7 strains by the software (Table 8). The majority of unidentified enzymes were related to the metabolism of *n*-alkanes and naphthenic acids. Indeed, *n*-alkanes and naphthenic acids were not associated to any known enzymes because of lacking in the KEGG's annotation system. Conversely, the metabolism of cycloalkanes and their possible metabolic intermediates was associated to the activity of monooxygenases and deyhdrogenases. For instance, the ability to utilize cyclohexanone was correlated to the activity of a cyclohexanone dehydrogenase (EC 1.3.99.14) and a cyclohexanone monooxygenase (EC 1.14.13.22).

The enzyme families predicted to be implicated in BTEX metabolism by R7 and BCP1 strains were: monooxygenases (EC 1.14.13- and 1.14.15-) involved in different hydrocarbon degradations; benzene/toluene dioxygenases (EC 1.14.12.3 and EC 1.14.12.11) and several putative enzymes involved in ethylbenzene metabolism, like the naphthalene dioxygenase (EC 1.14.12.12), the ethylbenzene hydroxylase (EC 1.17.99.2) and a putative dehydrogenase (EC 1.17-).

Concerning the activity on PAHs, amongst the 129 oxygenases/hydroxylases annotated for *R. opacus* R7 by RAST, 61 enzymes were predicted to be involved in mono- and polycyclic aromatic hydrocarbon metabolism. In BCP1, 73 oxygenases/hydroxylases amongst the 134 annotated genes, are predicted to be related to the oxidation of these organic compounds. Within PAHs, the analysis by Ductape of Phenotype Microarray showed anthracene and naphthalene to be metabolized by the same enzyme, a naphthalene dioxygenase (EC 1.14.12.12). Conversely the metabolism of phenanthrene was correlated to a PAH dioxygenase (EC 1.13.11.-). The use of gentisic acid was correlated to the gentisate 1,2-dioxygenase (EC 1.13.11.4); while the use of salicylic acid was related to salicylate 5-hydroxylase (EC 1.14.13.172). Amongst the naphthenic acids, the only substrate associated to an enzymatic class was the cyclohexanecarboxylic acid (CHCA), related to the large group of CoA synthetases (EC 6.2.1.1). Regarding this enzyme class, acetate and longer chain carboxylic acids can be converted to-CoA derivatives by CoA synthetases (EC 6.2.1.1) [16].

Genetic aspects related to the great potential for organic/xenobiotic degradation in *Rhodococcus opacus* R7 and *Rhodococcus* sp. BCP1

Based on the genome sequences of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and their annotations, functional gene clusters involved in the degradation of different organic/xenobiotic compounds were investigated. Four chemical categories of compounds were considered including:

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

Table 8. List of the xenobiotic compounds tested in phenotype microarray analysis and EC the numbers of the enzymes identified in R7 and BCP1 genomes and predicted to be involved in their metabolisms by DuctApe software.

Index	Chemical	EC numbers for *R. opacus* R7 and *Rhodococcus* sp. BCP1
1	1,4-Cyclohexane dicarboxylic acid	unknown
2	1-Adamantanecarboxylic acid	unknown
3	2-Hydroxybiphenyl	1.14.13.44; 3.13.1.3
4	4-Phenylbutyric acid	unknown
5	5,6,7,8-Tetrahydro-2-naphthoic acid	unknown
6	Benzene	1.14.12.3; 1.14.13.-
7	Cyclohexane	1.14.15.-
8	Cyclohexanecarboxylic acid	6.2.1.-
9	Cyclohexanone	1.1.1.90; 1.1.1.245; 1.3.99.14; 1.4.3.12; 1.14.13.22; 4.1.3.35
10	Cyclopentanecarboxylic acid	unknown
11	Decane	unknown
12	Dodecane	unknown
13	Eicosane	unknown
14	Ethylbenzene	1.14.12.12; 1.14.12.-; 1.17.99.2; 1.17.-.-
15	FuelOil	unknown
16	Heptadecane	4.1.99.5
17	Heptane	unknown
18	Hexadecane	unknown
19	Hexatriacontane	unknown
20	Naphthalene	1.14.12.12; 1.14.13.-; 1.14.14.1
21	Nonane	unknown
22	Octacosane	unknown
23	Tetracosane	unknown
24	Tetradecane	unknown
25	Toluene	1.14.12.11; 1.14.13.-; 1.14.15.-; 4.1.99.11
26	Tridecane	unknown
27	Anthracene	1.14.12.12; 1.14.-.-
28	Cyclohexane butyric acid	unknown
29	Cyclohexaneacetic acid	unknown
30	Decanoic acid	3.1.2.21
31	Dibenzothiophene	unknown
32	Gentisic acid	1.2.1.29; 1.2.3.1; 1.13.11.4; 1.14.13.24; 1.14.13.172; 4.1.1.62
33	Hexane	unknown
34	Hexanoic acid	3.5.1.39
35	*m*-Xylene	1.14.13.-; 1.14.15.-
36	Methyl-cyclohexanecarboxylic acid	6.2.1.-
37	*o*-Xylene	1.14.13.-; 1.14.15.-
38	*p*-Xylene	1.14.13.-; 1.14.15.-
39	Phenanthrene	1.13.11.-; 1.14.13.-
40	Salicylic acid	1.2.1.65; 1.2.1.-; 1.14.13.1; 1.14.13.172; 1.14.13.-; 2.1.1.274; 3.1.1.55; 3.7.1.8; 4.1.1.91; 4.1.1.-; 4.2.99.21
41	Trans-1,2-cyclohexane dicarboxylic acid	unknown

doi:10.1371/journal.pone.0139467.t008

database, the Rapid Annotation using Subsystem Technology (RAST server), the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database and manual curation.

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 *R. opacus* R7 and *Rhodococcus* sp. BCP1 strains [25, 28, 29]. Compared with these studies, the presence of only one copy of the *alkB* gene was confirmed in *R. opacus* R7, while in *Rhodococcus* sp. BCP1, two copies of this gene (*alkB1* and *alkB2)* were found within the chromosome (Fig 8 Panel A). 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, in addition, the *tetR* gene coding for a regulator. In BCP1 genome, an additional 1161-bp *alkB* gene (*alkB2*) was identified. This gene 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 (EC 6.2.1.3), 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 (S7 Table) with those coded by *alk*B cluster of R7 and with the ones of the most known bacteria belonging to the *Rhodococcus* genus such as *R. jostii* RAH1 or *R. opacus* PD630 [14, 16].

Two *prm* gene clusters were identified and analyzed in BCP1 (Fig 8 Panel B). The first *prm* gene cluster, constituted by the *prmA,C,B,D* 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) (*smoA,B,D,C* 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 [28]. 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% (S8 Table). Despite of the high sequence similarity between the two *prm* clusters, no growth of R7 in presence of gaseous *n*-alkanes was observed [25].

R7 and BCP1 strains showed also a considerable number of genes coding for P450 monooxygenases, 23 and 18 predicted coding sequences, respectively. We can not exclude that these putative enzymes are also involved in the aliphatic hydrocarbon degradation.

Aromatic hydrocarbons degradation. Both R7 and BCP1 strains are able to grow on different aromatic hydrocarbons, including BTEX compounds. The majority of the *Rhodococcus* genomes sequenced, as in the case of *R. jostii* RHA1, were characterized for the ability to grow on BTEX and in some cases also for their genetic organization [14, 44]. In *Rhodococcus jostii* RHA1, the *akb* gene cluster was identified as the cluster involved in the metabolism of alkylbenzenes. It contains two genes (designated *akbA1a* and *akbA1b*, approximately 6 kb apart from each other) encoding an oxygenase component large subunit, each one followed by genes (designated *akbA2a* and *akbA2b*) encoding 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. Comparing the *akb* gene cluster of RHA1 with R7, we found the same gene cluster divided in two parts and each part was allocated on two plasmids (Fig 8 Panel C): *akbA1a*, *akbA2a*, *akbA3*, *akbA4*, *akbB* genes were found on the pPDG5 plasmid; while *akbCDEF* genes, putatively coding for a complete *meta*-cleavage pathway, constituted by a *meta*-cleavage dioxygenase (AkbC), a *meta*-cleavage hydrolase product (AkbD), a hydratase (AkbE), and an aldolase (AkbF), were identified on the pPDG2 plasmid with high amino acid identity (S9 Table). An homologous sequence of *akb* cluster was also found on the chromosome and on the pPDG4 plasmid, but they showed a protein identity around 35%.

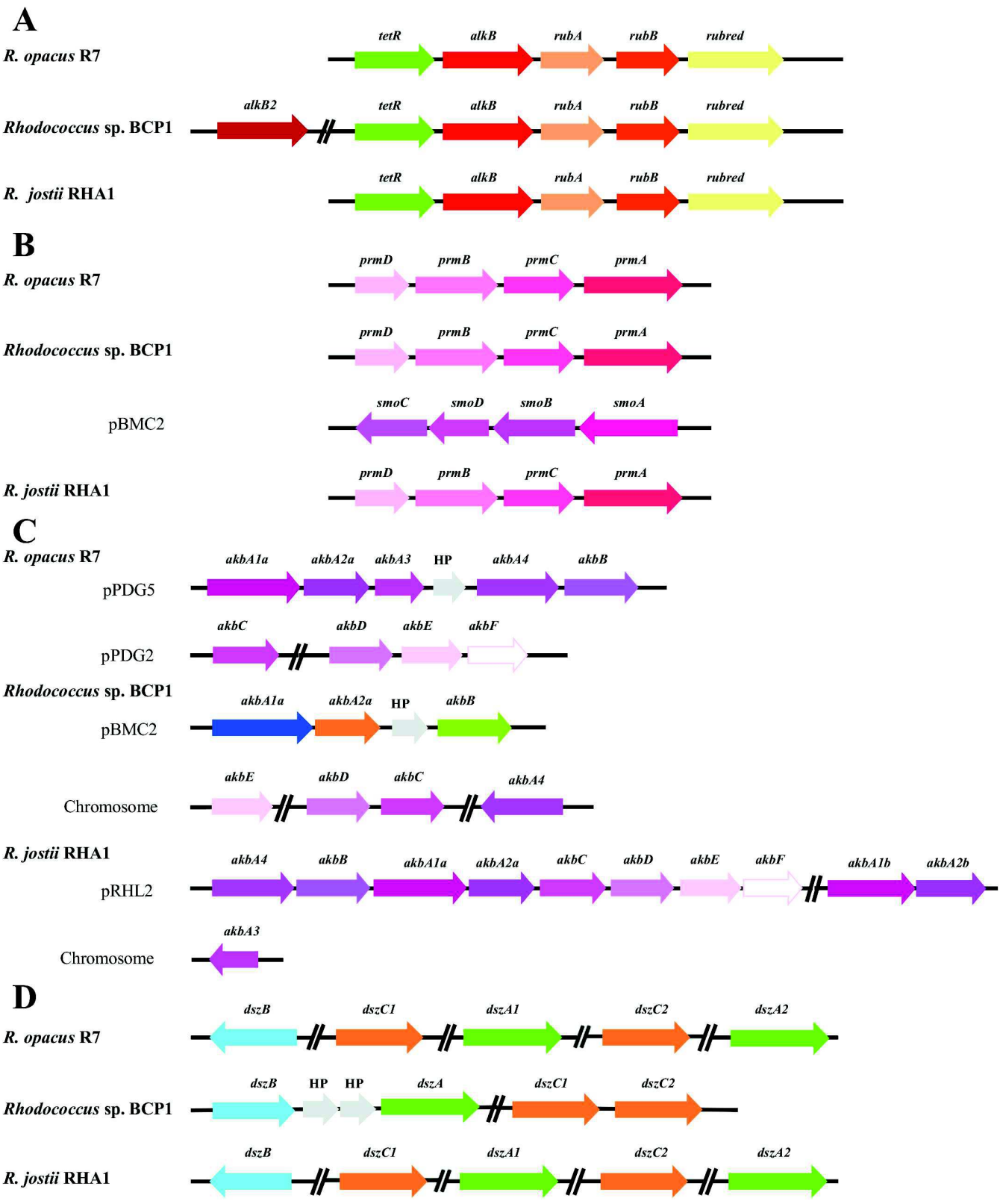
A
R. opacus R7
tetR alkB rubA rubB rubred
Rhodococcus sp. BCP1
alkB2 tetR alkB rubA rubB rubred
R. jostii RHA1
tetR alkB rubA rubB rubred
B
R. opacus R7
prmD prmB prmC prmA
Rhodococcus sp. BCP1
prmD prmB prmC prmA
pBMC2
smoC smoD smoB smoA
R. jostii RHA1
prmD prmB prmC prmA
C
R. opacus R7
pPDG5
akbA1a akbA2a akbA3 HP akbA4 akbB
pPDG2
akbC akbD akbE akbF
Rhodococcus sp. BCP1
pBMC2
akbA1a akbA2a HP akbB
Chromosome
akbE akbD akbC akbA4
R. jostii RHA1
pRHL2
akbA4 akbB akbA1a akbA2a akbC akbD akbE akbF akbA1b akbA2b
Chromosome
akbA3
D
R. opacus R7
dszB dszC1 dszA1 dszC2 dszA2
Rhodococcus sp. BCP1
dszB HP HP dszA dszC1 dszC2
R. jostii RHA1
dszB dszC1 dszA1 dszC2 dszA2

Fig 8. Comparison of *gene clusters* from R7 and BCP1 genomes correlated to aliphatic and aromatic hydrocarbon degradations. Comparative organization of genetic determinants for the tested aliphatic and aromatic hydrocarbons in *R. opacus* R7 and *Rhodococcus* sp. BCP1 with *R. jostii* RHA1 as reference strain. Predicted genes (listed in **S7**, **S8**, **S9**, **S10 Tables**) and their orientation are shown by arrow. (A) *alk* gene cluster; (B) *prm* and *smo* gene cluster (*smo* gene cluster was organized in *smoA* encoding a methane monooxygenase component A, *smoB* encoding a methane monooxygenase component B, *smoD* and *smoC* encoding a methane monooxygenase regulatory protein and a methane monooxygenase component C, respectively); (C) *akb* gene cluster; (D) *dsz* gene cluster. When not specified, it means that genes were located on chromosome. Genes with unknown or hypothetical functions were reported as HP. Double slash indicates a distances between two genes more than 1 kb within the same plasmid or chromosome.

doi:10.1371/journal.pone.0139467.g008

RAST analysis of BCP1 genome showed that coding sequences homologous to the *akb* genes were found in pBMC2 plasmid: *akbA1a* was found similar to a gene annotate as a large subunit *nah*/*bph* dioxygenase with an amino acid identity of 36%, *akbA2a* similar to a gene annotated as a biphenyl dioxygenase beta subunit (EC 1.14.12.18) with an amino acid identity of 36% and *akbB* similar to a gene annotated as a 2,3-dihydroxy-2,3-dihydro-phenylpropionate dehydrogenase (EC 1.3.1.-) with 47% protein identity.

These data indicated a redundancy of genes involved in the metabolism of the aromatic compounds in R7 strain, probably due to its isolation from a contaminated polycyclic aromatic hydrocarbons soil. The presence of multiple copies in several plasmids can derive from trans-position 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. In BCP1 only one copy of a putative cluster involved in the degradation of aromatic hydrocarbons was found.

In presence of dibenzothiophene (DBT) R7 strain showed also the highest activity 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 catalyses the stepwise *S*-oxidation of DBT, first to dibenzothiophene 5-oxide (DBTO) and then to dibenzothiophene 5,5-dioxide ($DBTO_2$); DszA catalyses the conversion of DBTO to 2-(2'-hydroxyphenyl) benzene sulphinate ($HBPSi^-$) and DszB catalyses the desulphation of $HBPSi^-$ to give HBP and sulphonate [45]. A *dsz* gene cluster was found in R7 and BCP1 genomes (Fig 8 Panel D). 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% (S10 Table). Sequence analysis revealed also two *dszC* genes (*dszC1 and dszC2*) predicted to code for dibenzothiophene desulphurization enzymes, and not so far we found only one copy of the *dszB* gene encoding an ABC sulfonate transporter protein.

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 with a percentage reported in S10 Table; 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.

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 [24]. The whole genome sequence analysis pointed out that *nar* gene cluster is located in pPDG4 plasmid in R7 strain and that the same cluster was found in pBMC2 in *Rhodococcus* sp. BCP1 (Fig 9 Panel A); this genomic region identified in pBMC2 is the same identified as *akb* gene cluster in BCP1. In S11 Table, protein identities are reported. The whole genome sequence analysis revealed the presence of *orf7* within the cluster, as in R7 strain, while any of the other six CDSs identified in R7 were found in BCP1 strain. An homologous *nar* gene cluster was not found in the reference strain *R. jostii* RHA1. The lower naphthalene catabolic

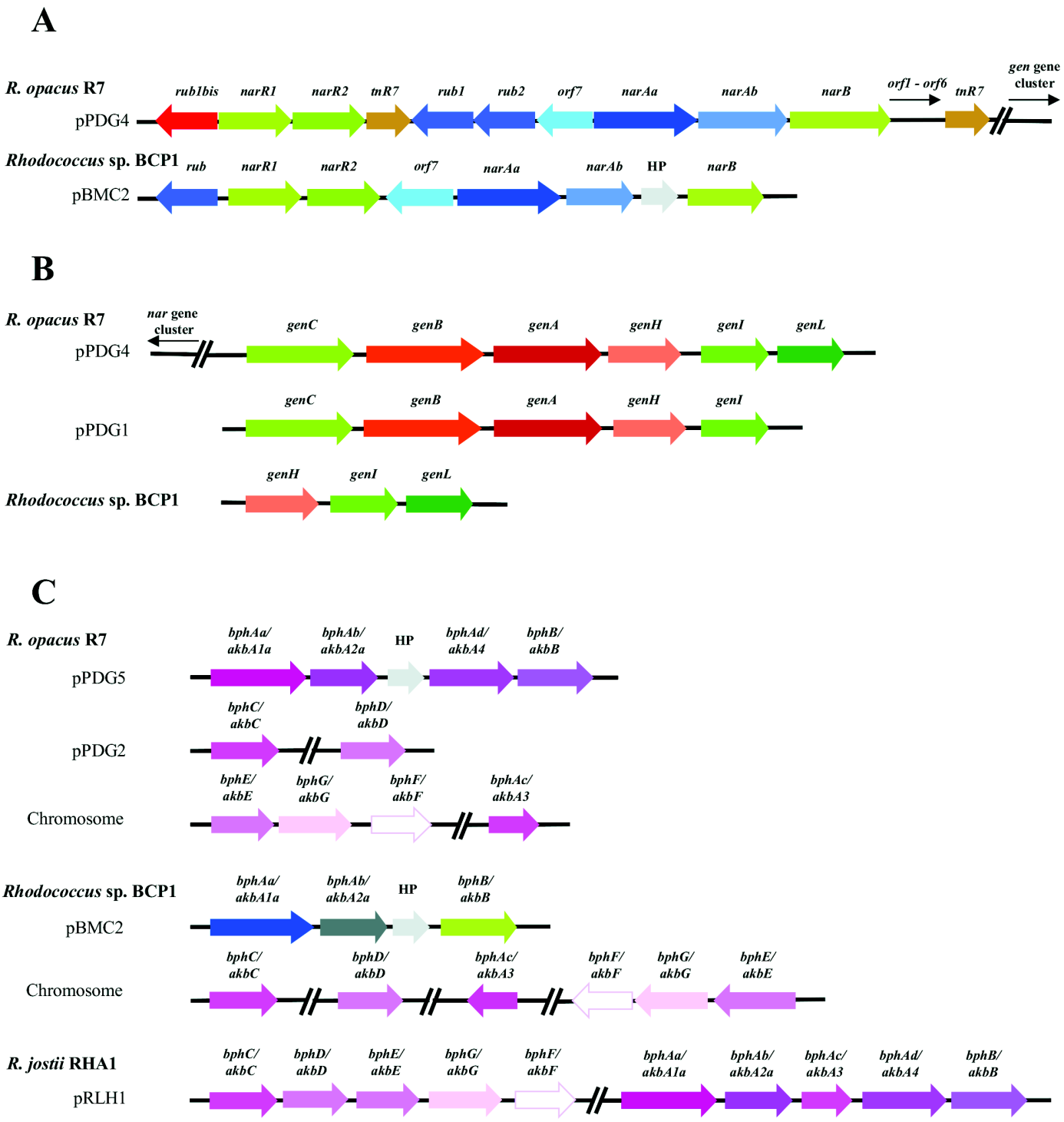

Fig 9. Comparison of *gene clusters* from R7 and BCP1 genomes correlated to polycyclic aromatic hydrocarbon degradations. Comparative organization of genetic determinants for naphthalene and biphenyl (as reference compounds of PAH and putative intermediates) in *R. opacus* R7 and *Rhodococcus* sp. BCP1 with *R. jostii* RHA1 as reference strain. Predicted genes (listed in S11, S12 **and** S13 **Tables**) and their orientation are shown by arrow. (A) *nar* gene cluster; (B) *gen* gene cluster; (C) *bph* gene cluster organization. When not specified, it means that genes were located on chromosome. Genes with unknown or hypothetical functions were reported as HP. Double slash indicates a distances between two genes more than 1 kb within the same plasmid or chromosome.

doi:10.1371/journal.pone.0139467.g009

pathway was previously investigated in R7 and it was hypothesized that the *gen* gene cluster was located far from the *nar* region as no amplification of the middle region was obtained. Genome sequence analysis confirmed these data (Fig 9 Panel B). 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 of *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 with high protein identity (around 80%), followed by *genL* gene that showed a protein identity of 48% with the homologous genes of R7. Instead, genes involved in salicylate degradation, *genA*, *B*, *C* were found in different regions of the BCP1 chromosome showing a lower amino acid identity as reported in S12 Table. This could be the reason of the low activity level shown by BCP1 during the growth on salicylic acid.

A well-known microorganism characterized for its ability to degrade biphenyls is *Rhodococcus jostii* RHA1. Genes involved in this pathway were identified in two different plasmids: pRHL1 and pRHL2 [46, 47]. In the biphenyl metabolic pathway, biphenyl 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). Thus, 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 (Fig 9 Panel C). We found different genes of this cluster in R7 chromosome and in pPDG2 and pPDG5 plasmids. 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 (S13 Table).

RAST analysis of BCP1 genome identified CDSs homologous to *bph* genes in pBMC2. In particular, only large and small subunits of a dioxygenase (*bphAa* and *bphAb* genes) and a dihydrobiphenyldiol dehydrogenase (*bphB* gene) were found with an amino acid identity around 40%; this coincides with the region identified as *akb* gene cluster and *nar* gene cluster. Also in BCP1, the genes encoding the ring-cleavage enzymes were not found near the cluster within the plasmid pBMC2; however, an homologous coding sequence was found in the chromosome with low protein identity (around 30%) (S13 Table). 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. Indeed, we can attribute the same sequence to the *akb*, *nar* and *bph* gene clusters.

Naphthenic acids degradation. Considering the ability of the two strains 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) based on cyclohexane ring i.e. cyclohexane carboxylic acid (CHCA); while none has been focused on cyclopentane ring i.e. cyclopentane carboxylic acid (CPCA). Although no genetic information was provided in these studies [48, 49], the metabolism of CHCA was described to follow two main routes: (i) aromatization of the cycloalkane ring to produce hydroxybenzoate before ring opening [48], or (ii) activation of cycloalkane ring as CoA thioester-derivative that

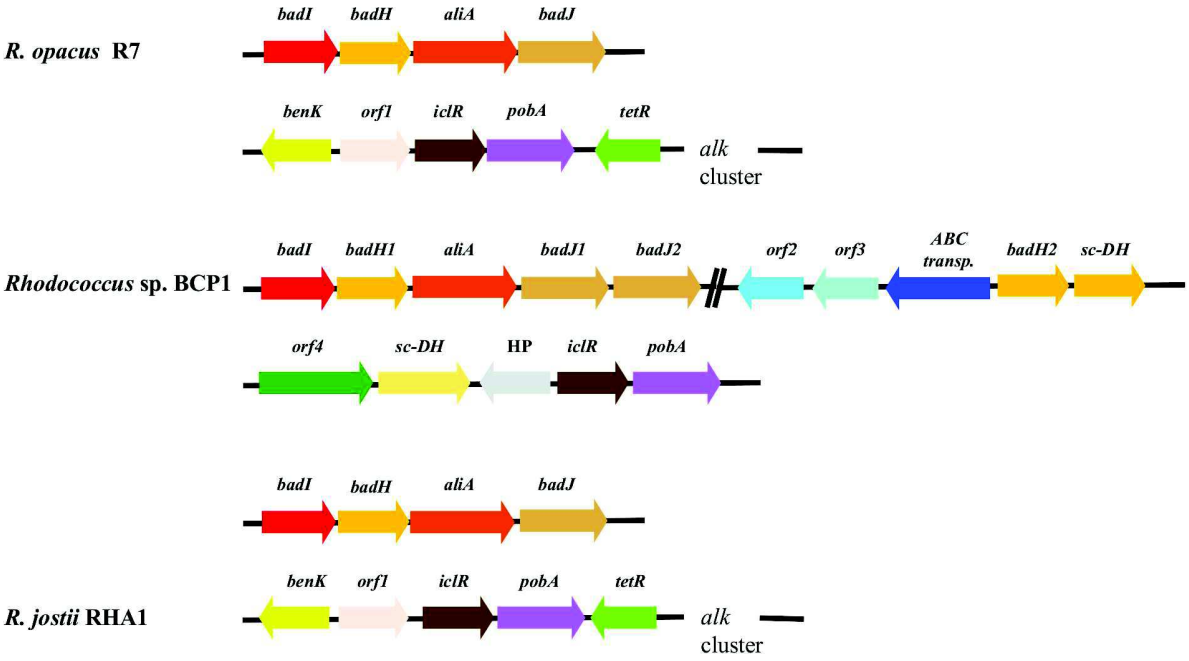

Fig 10. Comparison of *gene clusters* from R7 and BCP1 genomes correlated to carboxylated hydrocarbon degradations. Comparative organization of genetic determinants for naphthenic acids (as reference compounds of carboxylated hydrocarbons and putative intermediates) in *R. opacus* R7 and *Rhodococcus* sp. BCP1 with *R. jostii* RHA1 as reference strain. Predicted genes (listed in **S14 Table**) and their orientation are shown by arrow. The following genes encode for: *benK*, benzoate transporter; *orf1*, O-antigen acetylase; *iclR*, transcriptional regulator IclR family; *tetR*, transcriptional regulator, TetR family; *orf 2* and *orf 3*, permease; *ABC transp*, ABC transporter; *sc-DH*, probable short-chain dehydrogenase; *orf4*, permease. When not specified, it means that genes were located on chromosome. Genes with unknown or hypothetical functions were reported as HP. Double slash indicates a distances between two genes more than 1 kb within the same plasmid or chromosome.

doi:10.1371/journal.pone.0139467.g010

is further degraded through β-oxidation steps [49]. 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. Iwaki and co-workers [48] demonstrated the involvement of *pobA* product in CHCA catabolism by *Corynebacterium* strain downstream of the formation of 4-hydroxybenzoate from CHCA through several oxidation steps. One gene homologous to *pob*A 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, *pob*A gene was flanked by a gene coding for an IclR-type transcriptional regulator. The genomic region including *pob*A in R7 included 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* was different from those of R7 and RHA1 and included a permease coding gene and several oxidoreductases (Fig 10).

The activation of the cycloalkane ring as Co-A thioester-derivative was firstly described during the anaerobic degradation of benzoic acid by *Rhodopseudomonas palustris* [50]. 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 [51]. These enzymes degrade CHCA by catalyzing the following reactions: (i) activation of cyclohexanecarboxylate as cyclohexanecarboxylate-CoA by a CoA-ligase (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 catalyses the the cyclohexane ring opening with the formation of pimelyl-CoA. On the basis of the amino acid identity percentage as compared with the 2-ketocyclohexane-carboxyl-CoA dehydrogenase (BadH) of *R. palustris*, two genes *(badH1 and badH2)* were found in BCP1 genome encoding two enzymes annotated as 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase. Only one gene homologous to *badH* 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 (Fig 10) (S14 Table).

Genetic aspects related to aromatic peripheral pathways in *Rhodococcus opacus* R7 and *Rhodococcus* sp. 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. Genes responsable for such catabolic pathways have been reported in several bacteria [52], in particular in *R. jostii* RHA1. We performed a sequence comparison analysis to identify the genes predicted to be involved in these pathways in *R. opacus* R7 and in *R.* sp. BCP1. 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. These aspects might be explained on the basis of the significant difference of their genome size. Two catechol dioxygenase genes, amongst those identified on R7 chromosome, were organized in cluster (Fig 11 Panel A). 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 of *catC* gene; moreover, CatA2 (*catA2*) and CatB2 (*catB2*) were not found homologous to RHA1 genes (S15 Table). *Rhodococcus* sp. BCP1 presented only one copy of *cat* gene cluster with the same organization of RHA1 and R7; BCP1 *cat* genes showed high similarity (70–90%) with those of RHA1 and R7 strains as reported in table (S15 Table).

Both R7 and BCP1 genomes contain several genes potentially involved in protocatechuate catabolism (Fig 11 Panel B). 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 (S16 Table). 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, *pcaJI* and *pcaHGBLRF*.

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 catalized by enzymes coded by the *paa* gene cluster [53]. 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

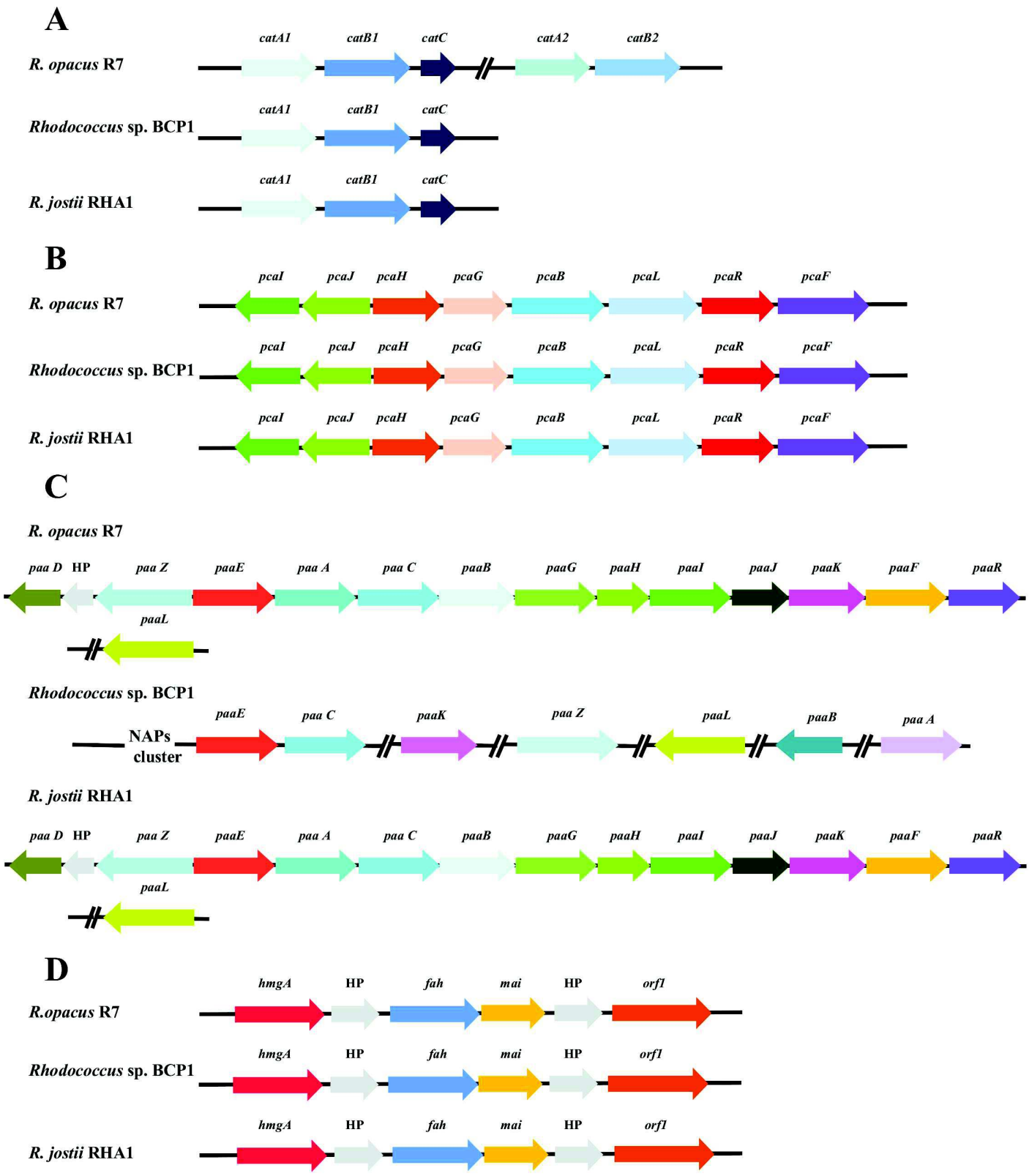
A
R. opacus R7
catA1
catB1
catC
catA2
catB2
Rhodococcus sp. BCP1
catA1
catB1
catC
R. jostii RHA1
catA1
catB1
catC
B
R. opacus R7
pcaI
pcaJ
pcaH
pcaG
pcaB
pcaL
pcaR
pcaF
Rhodococcus sp. BCP1
pcaI
pcaJ
pcaH
pcaG
pcaB
pcaL
pcaR
pcaF
R. jostii RHA1
pcaI
pcaJ
pcaH
pcaG
pcaB
pcaL
pcaR
pcaF
C
R. opacus R7
paa D
HP
paa Z
paaE
paa A
paa C
paaB
paaG
paaH
paaI
paaJ
paaK
paaF
paaR
paaL
Rhodococcus sp. BCP1
NAPs cluster
paaE
paa C
paaK
paa Z
paaL
paaB
paa A
R. jostii RHA1
paa D
HP
paa Z
paaE
paa A
paa C
paaB
paaG
paaH
paaI
paaJ
paaK
paaF
paaR
paaL
D
R.opacus R7
hmgA
HP
fah
mai
HP
orf1
Rhodococcus sp. BCP1
hmgA
HP
fah
mai
HP
orf1
R. jostii RHA1
hmgA
HP
fah
mai
HP
orf1

Fig 11. Comparison of *gene clusters* from R7 and BCP1 genomes correlated to xenobiotic peripheral pathways. Comparative organization of genetic determinants for xenobiotic peripheral pathways in *R. opacus* R7 and *Rhodococcus* sp. BCP1 with *R. jostii* RHA1 as reference strain. Predicted genes (listed in **S15**, **S16**, **S17** and **S18 Tables**) and their orientation are shown by arrow. (A) *cat* gene cluster; (B) *pca* gene cluster; (C) *paa* gene cluster; (D) *hmg* gene cluster. When not specified, it means that genes were located on chromosome. Genes with unknown or hypothetical functions were reported as HP. Double slash indicates a distances between two genes more than 1 kb within the same plasmid or chromosome.

doi:10.1371/journal.pone.0139467.g011

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 (Fig 11 Panel C). They also showed high percentage of similarity as reported in S17 Table. On the contrary, the *paa* gene cluster was not conserved in *Rhodococcus* sp. BCP1 and few genes homologous to those of RHA1, are present without a co-localization. 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 (EC 4.2.1.17), fumarylacetoacetase (EC 3.7.1.2) and a glutaryl-CoA dehydrogenase (EC 1.3.99.7). 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 (Fig 11 Panel D) (S18 Table).

Discussion

This work describes in detail the genomic, phenotypic and taxonomic features of two *Rhodococcus* spp. strains whose genomes have recently been sequenced [22, 23]. In particular, phylogenetic comparison of the two strains was performed using a multi-locus sequence analysis approach with four different taxonomic markers (*16S rRNA* gene, *secY* gene, *rpoC* gene and *rpsA* gene). *Rhodococcus* sp. BCP1 was phylogenetically related to *R. aetherivorans* species, while *R. opacus* R7 was related to both *R. opacus* and *R. wratislaviensis* species. The taxonomic correlation of BCP1 strain is in line with previous phylogenetic analysis performed using other taxonomic markers (i.e. alkane 1-monooxygenase (AlkB) [42]. These studies taken together support the belonging of BCP1 to *R. aetherivorans* species making its genome to be the first sequenced genome available of this species.

Comparative analyses have been performed amongst BCP1, R7 strains and other *Rhodococcus* strains taxonomically related. In particular, *R. opacus* PD630 and *R. opacus* B4 were considered because closely related to R7; *R. pirydinivorans* SB3094 was the most closely BCP1 related strain with a complete genome sequenced; *R. jostii* RHA1 represents the reference strain for genomic analysis of *Rhodococcus* spp. Based on core and pan-genome analysis as well as nucleotide diversity calculation, this genus displayed a considerable diversity on the genome scale. In line with the MLSA-based phylogenetic analysis, the extension of the core genome increased when we considered the genomic regions shared amongst the *R. opacus* strains; however, the highest number of shared genomic regions was in R7 and RHA1. Moreover, only a limited number of genomic regions was shared by BCP1 and SB3094, despite their phylogenetic relation. These results suggest that the extension of genomic regions shared by the two strains might be related to the adaptation to specific niches (both R7 and RHA1 were isolated from soils contaminated with aromatic compounds) more than to their taxonomic correlation. As a result of the genome comparison, the regions uniquely found in BCP1 and R7 largely included genes coding for proteins with unknown functions. Other genes found in these regions coded for transcriptional regulators, membrane transporters, oxidoreductases and mobile elements. Notably, most of the shared regions identified in *Rhodococcus* sp. BCP1 and *R. opacus* R7 were located on the chromosome, while a high percentage of unique regions was identified on the

plasmids, where a large amount of mobile elements are also present. These aspects connect the occurrence of horizontal gene transfer events to the peculiar metabolic properties that characterize each strain.

In addition to genome analysis compared to other *Rhodococcus* spp., the catabolic potentials of *R. opacus* R7 and *Rhodococcus* sp. BCP1 were assessed towards different carbon, nitrogen, sulphur and phosphorous sources and different xenobiotic compounds. This study was aimed at both investigating the peculiar abilities of these two strains and correlating the catabolic differences of BCP1 and R7 to the different habitats from which they have been isolated (butane-growing microbial consortium able to degrade chloride aliphatic hydrocarbons and polycyclic aromatic hydrocarbon contaminated soil, respectively). With this purpose, the metabolic features of *Rhodococcus* strains were systematically analyzed using a Biolog Phenotype Microarray (PM) system with both Biolog standard plates (PM1-4, PM9-PM20) and plates manually prepared by adding different organic/xenobiotic compounds to the wells (aliphatic, alicyclic and aromatic hydrocarbons, polycyclic aromatic hydrocarbons, and carboxylated hydrocarbons, i.e. naphthenic acids). The catabolic profiles of *Rhodococcus* sp. BCP1 and *R. opacus* R7 were compared to define the differences between the two strains in terms of: i) carbon, nitrogen, sulphur and phosphorous source utilization (plates PM1-4); ii) metabolic response to osmolytes and to different pH growth environments (PM9-10); iii) antibiotic resistance pattern and ability to respire in presence of toxic compounds (PM11-20); iv) ability to oxidize different organic/xenobiotic compounds (plates manually prepared). Considering results from PM1-4, we found that, despite the genomic diversity, the two strains displayed similar core carbon metabolic profiles (PM1-2 plates) (Figs 3 and 4). Both strains also showed the ability to utilize a wide range of compounds as phosphorous sources (PM4 wells A1-E12); while a great difference was shown in terms of nitrogen and sulphur compound assimilation (PM3 and PM4 wells F1-H12). In particular, R7 showed metabolic activities on a wider variety of nitrogen and sulphur sources as compared to BCP1 (Figs 5 and 6). These results are in line with the higher amount of genes predicted to be involved in sulphur and organic nitrogen metabolism in the R7 genome (S5 Table).

Previous studies described the use of Phenotype Microarray to determine the growth capacity of *R. opacus* PD630 and *R. jostii* RHA1, although the phenotypic screening was limited to the carbon sources included in PM1-2 plates and to some nitrogen sources included in PM3 [16]. Similarly to what previously reported for *R. jostii* RHA1, both *R. opacus* R7 and *Rhodococcus* sp. BCP1 strains did not show metabolic activities on galactose. By contrast, this substrate and the derived oligogalactosides (lactose, lactitol, melibionic acid, melibiose, lactulose, raffinose, and stachyose) were efficiently utilized by *R. opacus* PD630. This ability was connected with the identification in PD630 of a galactose-catabolic region that was not found in RHA1. In the present work, in line with the inability of BCP1 and R7 strains to grow on galactose and derivatives, the galactose-catabolic region detected in PD630 was not found in both the genomes under analysis. Like *R. opacus* PD630, *R. opacus* R7 was able to utilize D-galactonic acid-γ-lactone that was not degraded by BCP1 and RHA1, suggesting this catabolic capacity to be specific of strains belonging to *R. opacus* species.

The effects of different osmolites, pH ranges and toxicants on the two *Rhodococcus* spp. strains were tested by using plates PM9-PM20. Interestingly, in contrast to BCP1, R7 showed a general resistance to low pH in the presence of several amino acids (S2 and S3 Figs), probably due to the activity of decarboxylases that generate alkaline amines by the catabolism of these compounds [20]. In response to the presence of osmolytes, R7 could tolerate high concentrations of NaCl, sodium formate, and sodium lactate compared with BCP1, while BCP1 could resist to high concentrations of urea that vice versa inhibited the R7 activity. Both strains showed wide antibiotic resistance patterns including several antibiotics, metals (except for

vanadate salts) along with antiseptics, detergents and anti-microbial agents (S4, S5 and S6 Figs). The sensitivity shown by R7 and BCP1 towards macrolides, rifampin, vancomycin, novobiocin is in line with what previously described for *R. equi* [54–56]. R7 and BCP1 strains also showed sensitivity to antibiotics known to inhibit Gram-positive microorganisms like fusidic acid, pheneticillin, and to alkaloids with anti-lipase activity in *Candida rugosa* like sanguinarine and cheleritrine [57]. BCP1 differs from R7 for its ability to grow on the anti-microbial agents ethionamide and on thallium(I) acetate.

Considering the genetic aspects related to the phenotypic features observed with plates PM9-PM20, around 140 and 180 genes were annotated in the genomes of BCP1 and R7, respectively, that are involved in various stress responses (osmotic stress, detoxification, heat/cold shock, oxidative stress) (S5 Table). The resistance to stress can also be correlated to the features of *Rhodococcus* cell wall that contains high-molecular weight α-alkyl-β-hydroxy fatty acids named as mycolic acids. Indeed, mycolic acids are known to confer resistance to chemical injury, low permeability to hydrophobic antibiotics and to hydrophilic substrates, and resistance to dehydration [58]. Phenotype Microarray system was also used to test the catabolic capacity of BCP1 and R7 strains towards different categories of xenobiotic compounds belonging to: aliphatic hydrocarbons (i.e.decanoic acid, hexanoic acid), polycyclic aromatic hydrocarbons (PAHs), BTEX and other aromatic compounds (i.e. dibenzothiophene a recalcitrant component of fossil fuels), and carboxylated compounds (i.e. naphthenic acids). Both strains showed the ability to utilize a wide range of these compounds (Fig 7). Concerning *n*-alkanes, BCP1 and R7 could grow on *n*-alkanes with aliphatic chain $>C_{12}$. These metabolic capacities were described previously using only a few of the *n*-alkanes tested in the present work [25, 29]. Here we report a series of new findings on R7 and BCP1 *n*-alkanes metabolism, namely: i) R7 has lower activities than BCP1 on odd-carbon *n*-alkanes (C_{13} and C_{17}) as compared to even-carbon *n*-alkanes; ii) BCP1 has the ability to utilize tetracontane (C_{34}) and hexatriacontane (C_{36}) for growth; further, unlike R7, BCP1 could grow on *n*-alkanes ranging from *n*-hexane (C_6) to *n*-nonane (C_9). Both strains could grow on recalcitrant alicyclic hydrocarbon, cyclohexane, and on the intermediate of the cyclohexane degradation pathway [59] as cyclohexanone, although these substrates were preferentially utilized by BCP1. Amongst the naphthenic acids, R7 and BCP1 could grow and showed high activities only on cyclohexanecarboxylic acid (CHCA) and on 1,4-cyclohexanedicarboxylic acid (1,4-CHdiCA). Unlike R7, BCP1 could also grow on cyclopentanecarboxylic acid (CPCA) confirming the limited capacity of R7 to utilize odd-number chain alkanes. R7 and BCP1 strains showed a certain level of metabolic activity on trans-1,2-cyclohexanedicarboxylic acid (trans-ChdiCA), 3-methyl-1-cyclohexanecarboxylic acid (mCHCA), 4-phenylbutyric acid; while none of them could utilize the more complex molecules such as cyclohexane butyric acid (CHBA), adamantane carboxylic acid (ACA), cyclohexane acetic acid (CHAA), 1,2,3,4-Tetrahydro-2-naphthoic acid (THNA) under these growth conditions.

Notably, R7 showed its highest metabolic activity on dibenzothiophene (DBT), a fuel contaminant naturally present in crude oil, that can be used by some bacterial species mainly from the genus *Rhodococcus*, as a sulphur source [60].

A genome analysis was performed to identify gene clusters involved in the metabolism of contaminants tested in Phenotype Microarray assays. In this work, we show that *R. opacus* R7 contains multiple genes for the degradation of a large set of aromatic and polyaromatic hydrocarbons, while a lower variability in terms of genes predicted to be involved in aromatic degradation was shown by BCP1 strain. 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. Indeed, R7 was isolated from a soil contaminated with aromatic hydrocarbons, while BCP1 was isolated from a butane-utilizing consortium enriched in

short-chain aliphatic hydrocarbons. According to this, the *smo* gene cluster, previously described to code for a soluble di-iron monooxygenase involved in aliphatic chain hydrocarbon degradation [28] is included in one of the unique regions of BCP1, while it is not conserved among the *Rhodococcus* strains compared in this study. The relative dimension of the two genomes under analysis (10 Mb for R7 and 6 Mb for BCP1) can also explain the lower degree of genetic redundancy present in BCP1 compared to R7.

In conclusion, the genome analysis of *Rhodococcus* sp. BCP1 and *R. opacus* R7 along with their phenotypic characterization highlighted a 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.

Supporting Information

S1 File. Supporting Information. The S1 file contains legends.
(DOC)

S1 Fig. Mauve Diagram. Whole genome sequence comparison of *R. opacus* R7 and *Rhodococcus* sp. 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.
(TIF)

S2 Fig. Phenotype Microarray PM in presence of different osmolytes. Resistance differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of osmolytes (AI, AII, AIII). Based on activity values of phenotype microarray analysis, threshold values were established for every plates. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.
(TIFF)

S3 Fig. Phenotype Microarray PM in presence of different pH values. Resistance differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of different pH values (AIV, AV, AVI). Based on activity values of phenotype microarray analysis, threshold values were established for every plates. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.
(TIFF)

S4 Fig. Phenotype Microarray PM in presence of different antibiotics. Resistance differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of different antibiotics that were tested at four concentration (1, 2, 3, 4) according to Biolog procedure (AI, AII, AIII). Based on activity values of phenotype microarray analysis, threshold values were established for every plates. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.
(TIFF)

S5 Fig. Phenotype Microarray PM in presence of different antiseptics. Resistance differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of antiseptics (AI, AII). Based on activity values of phenotype microarray analysis, threshold values were established for every plates. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.
(TIFF)

S6 Fig. Phenotype Microarray PM in presence of other antiseptics and metals. Resistance differences among *R. opacus* R7 and *Rhodococcus* sp. BCP1 in presence of antiseptics (AIII, AIV) and metals (B). Based on activity values of phenotype microarray analysis, threshold values were established for every plates. Determined thresholds were high (green), upper middle (light green), lower middle (orange) and low (red) for high, upper middle, lower middle and low activity, respectively.
(TIFF)

S1 Table. *R. opacus* R7 unique regions deriving from the comparative genome alignments with *Rhodococcus* sp. BCP1 and *Rhodococcus* sp. BCP1 unique regions deriving from the comparative genome alignments with *R. opacus* R7.
(XLSX)

S2 Table. Enzymatic class identification for xenobiotic degradation in the unique regions of *Rhodococcus* sp. BCP1 and *R. opacus* R7 genomes.
(PDF)

S3 Table. Activity values of tested substrates in Phenotype Microarray analysis in presence of *R. opacus* R7 and *Rhodococcus* sp. BCP1. Activity values in presence of carbon sources: carbohydrates (AI, AII), carboxylic acids (BI, BII), alcohols, amides, amines, esters, fatty acids, polymers (C), amino acids (D); nitrogen sources (AI, AII, AIII); phosphorous sources (AI, AII); sulphur sources (B); osmolytes (AI, AII, AIII) and pH values (AIV, AV, AVI); antibiotics (AI, AII, AIII); antiseptics (AI, AII, AIII, AIV) and metals (B).
(XLS)

S4 Table. Enzymatic class for fatty acids β-oxidation.
(PDF)

S5 Table. RAST subsystem categories of *R. opacus* R7 and *Rhodococcus* sp. BCP1 metabolism in presence of carbon, nitrogen, phosphorous and sulphur sources.
(PDF)

S6 Table. Activity values of tested organic/xenobiotic compounds in Phenotype Microarray analysis in presence of *R. opacus* R7 and *Rhodococcus* sp. BCP1 (see Fig 7).
(XLS)

S7 Table. Comparison of predicted genes and proteins of *alk* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 8).
(PDF)

S8 Table. Comparison of predicted genes and proteins of *prm* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 8).
(PDF)

S9 Table. Comparison of predicted genes and proteins of *akb* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 8).
(PDF)

S10 Table. Comparison of predicted genes and proteins of *dsz* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 8).
(PDF)

S11 Table. Comparison of predicted genes and proteins of *nar* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 9).
(PDF)

S12 Table. Comparison of predicted genes and proteins of *gen* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 9).
(PDF)

S13 Table. Comparison of predicted genes and proteins of *bph* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 9).
(PDF)

S14 Table. Comparison of predicted genes and proteins of naphthenic acids cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 10).
(PDF)

S15 Table. Comparison of predicted genes and proteins of *cat* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 11).
(PDF)

S16 Table. Comparison of predicted genes and proteins of *pca* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 11).
(PDF)

S17 Table. Comparison of predicted genes and proteins of *paa* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 11).
(PDF)

S18 Table. Comparison of predicted genes and proteins of *hmg* cluster of *R. opacus* R7 and *Rhodococcus* sp. BCP1 and comparison with *R. jostii* RHA1 homologous proteins (see Fig 11).
(PDF)

All the files of supporting informations can be downloaded from the database at www.itb.cnr.it/rhodococcus

Acknowledgments

This work has been supported by the Italian Ministry of Education and Research through the Flagship "InterOmics" (PB05) and HIRMA (RBAP11YS7K) Projects.

Author Contributions

Conceived and designed the experiments: PDG DZ LM CV SF. Performed the experiments: AO MC PDU ADC JZ FD. Analyzed the data: AP EC ADC MC. Contributed reagents/materials/analysis tools: PDG LM DZ. Wrote the paper: PDG AO MC. Substantial contributions to the conception or design of the work: AO MC PDG LM DZ. Contributions to the experiments: ADC JZ PDU CV FD SF. Contributions to the acquisition, analysis, or interpretation of data for the work: AP EC MC AO. Contributions to the Drafting the work: AO MC PDG.

References

1. Tsukamura M. Necessity for differentiation of rhodochrous group from tubercle bacilli. Jpn J Microbiol. 1974; 18(1): 94–7. PMID: 4212191
2. Goodfellow M, Alderson G. The actinomycete-genus *Rhodococcus*: a home for the "*rhodochrous*" complex. J Gen Microbiol. 1977; 100(1): 99–122. PMID: 874450
3. Larkin MJ, Kulakov LA, Allen CCR. Genomes and Plasmids in *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus*. Springer-Verlag Berlin Heidelber. 2010. pp. 73–90.
4. http://www.bacterio.net/rhodococcus.html

5. Martínková L, Uhnáková B, Pátek M, Nesvera J, Kren V. Biodegradation potential of the genus *Rhodococcus*. Environ Int. 2009; 35(1): 162–77. doi: 10.1016/j.envint.2008.07.018 PMID: 18789530
6. Letek M, González P, MacArthur I, Rodríguez H, Freeman TC, Valero-Rello A, et al. The genome of a pathogenic *Rhodococcus*: cooptive virulence underpinned by key gene acquisitions. PLoS Genet. 2010; 6(9): e1001145. doi: 10.1371/journal.pgen.1001145 PMID: 20941392
7. Finnerty WR. The biology and genetics of the genus *Rhodococcus*. Annu Rev Microbiol. 1992; 46: 193–218. PMID: 1444254
8. Larkin MJ, Kulakov LA, Allen CCR. Biodegradation by members of the genus *Rhodococcus*: biochemistry, physiology, and genetic adaptation. Adv Appl Microbiol. 2006; 59: 1–29. PMID: 16829254
9. van der Geize R, Dijkhuizen L. Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. Curr Opin Microbiol. 2004; 7(3): 255–61. PMID: 15196492
10. Larkin MJ Kulakov LA, Allen CCR. Biodegradation and *Rhodococcus*–masters of catabolic versatility. Curr Opin Biotechnol. 2005; 16(3): 282–290. PMID: 15961029
11. Alvarez HM. Central metabolism of the species of the genus *Rhodococcus*. In: Alvarez HM, editor. Biology of *Rhodococcus*. Springer-Verlag Berlin Heidelber. 2010. pp. 91–108
12. http://www.ncbi.nlm.nih.gov/genome/Rhodococcus
13. 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(12): 4510–3. PMID: 16535201
14. McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, et al. The complete genome of *Rhodococcus sp*. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci U SA. 2006; 103(42): 15582–15587.
15. Sameshima Y, Honda K, Kato J, Omasa T, Ohtake H. Expression of *Rhodococcus opacus alkB* genes in anhydrous organic solvents. J Biosci Bioeng. 2008; 106(2): 199–203. doi: 10.1263/jbb.106.199 PMID: 18804065
16. Holder JW, Ulrich JC, DeBono AC, Godfrey PA, Desjardins CA, Zucker J, et al. Comparative and functional genomics of *Rhodococcus opacus* PD630 for biofuels development. PLoS Genet. 2011; 7(9): e1002219. doi: 10.1371/journal.pgen.1002219 PMID: 21931557
17. Bochner BR. New technologies to assess genotype-phenotype relationships. Nat Rev Genet. 2003; 4 (4): 309–14. PMID: 12671661
18. Yan Q, Power KA, Cooney S, Fox E, Gopinath GR, Grim CJ, et al. Complete genome sequence and phenotype microarray analysis of *Cronobacter sakazakii* SP291: a persistent isolate cultured from a powdered infant formula production facility. Front Microbiol. 2013; 4: 256. doi: 10.3389/fmicb.2013.00256 PMID: 24032028
19. Biondi EG, Tatti E, Comparini D, Giuntini E, Mocali S, Giovannetti L, et al. Metabolic capacity of *Sinorhizobium* (*Ensifer*) *meliloti* strains as determined by phenotype MicroArray analysis. Appl Environ Microbiol. 2009; 75(16): 5396–404. doi: 10.1128/AEM.00196-09 PMID: 19561177
20. 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(3): 553–9. PMID: 17385890
21. Khatri B, Fielder M, Jones G, Newell W, Abu-Oun M, Wheeler PR. High throughput phenotypic analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* strains' metabolism using biolog phenotype microarrays. PLoS ONE. 2013; 8(1): e52673. doi: 10.1371/journal.pone.0052673 PMID: 23326347
22. Cappelletti M, Di Gennaro P, D'Ursi P, Orro A, Mezzelani A, Landini M, et al. Genome sequence of *Rhodococcus* sp. strain BCP1, a biodegrader of alkanes and chlorinated compounds. Genome Announc. 2013; 1(5): e00657–13. doi: 10.1128/genomeA.00657-13 PMID: 24158549
23. Di Gennaro P, Zampolli J, Presti I, Cappelletti M, D'Ursi P, Orro A, et al. Genome sequence of *Rhodococcus opacus* strain R7, a biodegrader of mono- and polycyclic aromatic hydrocarbons. Genome Announc. 2014; 2(4): e00827–14. doi: 10.1128/genomeA.00827-14 PMID: 25146139
24. 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(1): 297–308. doi: 10.1007/s00253-010-2497-3 PMID: 20195856
25. 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. doi: 10.1186/s13568-014-0073-4 PMID: 25401074
26. Frascari D, Pinelli D, Nocentini M, Fedi S, Pii Y, Zannoni D. Chloroform degradation by butane-grown cells of *Rhodococcus aetherovorans* BCP1. Appl Microbiol Biotechnol. 2006; 73(2): 421–8. PMID: 17058077

27. Cappelletti M, Frascari D, Zannoni D, Fedi S. Microbial degradation of chloroform. Appl Microbiol and Biotechnol. 2012; 96(6): 1395–1409.
28. Cappelletti M, Presentato A, Milazzo G, Turner R J, Fedi S, Frascari D, et al. 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. 2015; 6: 393. doi: 10.3389/fmicb.2015.00393 PMID: 26029173
29. Cappelletti M, Fedi S, Frascari D, Ohtake H, Turner RJ, Zannoni D. 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. 2011; 77(5): 1619–27. doi: 10.1128/AEM.01987-10 PMID: 21193665
30. 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. doi: 10.1186/1471-2164-9-75 PMID: 18261238
31. Darling AE, Mau B, Perna NT. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE. 2010; 5(6): e11147. doi: 10.1371/journal.pone.0011147 PMID: 20593022
32. Kiełbasa SM, Wan R, Sato K, Horton P, Frith MC. Adaptive seeds tame genomic sequence comparison. Genome Res. 2011; 21(3): 487–493. doi: 10.1101/gr.113985.110 PMID: 21209072
33. Paley SM, Latendresse M, Karp PD. Regulatory network operations in the Pathway Tools software. BMC Bioinformatics. 2012; 13: 243. doi: 10.1186/1471-2105-13-243 PMID: 22998532
34. Wang Y, Lilburn TG. Biological Resource Centers and Systems Biology. Bioscience. 2009; 59(2): 113–125. PMID: 20157346
35. Adékambi T, Butler RW, Hanrahan F, Delcher AL, Drancourt M, Shinnick TM. Core gene set as the basis of multilocus sequence analysis of the subclass *Actinobacteridae*. PLoS ONE. 2011; 6(3): e14792. doi: 10.1371/journal.pone.0014792 PMID: 21483493
36. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl Acids Res. 2004; 32(5): 1792–1797. PMID: 15034147
37. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010; 59(3): 307–21. doi: 10.1093/sysbio/syq010 PMID: 20525638
38. Bochner BR, Gadzinski P, Panomitros E. Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. Genome Res. 2001; 11(7): 1246–1255. PMID: 11435407
39. Galardini M, Mengoni A, Biondi EG, Semeraro R, Florio A, Bazzicalupo M, et al. DuctApe: a suite for the analysis and correlation of genomic and OmniLog™ Phenotype Microarray data. Genomics. 2014; 103(1): 1–10. doi: 10.1016/j.ygeno.2013.11.005 PMID: 24316132
40. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucl Acids Res. 2000; 28 (1): 27–30. PMID: 10592173
41. Kanehisa M, Goto S, SatoY, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2014; 42(Database issue): D199–D205. doi: 10.1093/nar/gkt1076 PMID: 24214961
42. Táncsics A, Benedek T, Szoboszlay S, Veres PG, Farkas M, Máthé I, et al. The detection and phylogenetic analysis of the alkane 1-monooxygenase gene of members of the genus *Rhodococcus*. Syst Appl Microbiol. 2015; 38(1): 1–7. doi: 10.1016/j.syapm.2014.10.010 PMID: 25466921
43. Kostka JE, Prakash O, Overhol WA, Green SJ, Freyer G, Canion A, et al. Hydrocarbon-degrading bacteria and the bacterial community response in gulf of Mexico beach sands impacted by the deepwater horizon oil spill. Appl Environ Microbiol. 2011; 77(22): 7962–7974. doi: 10.1128/AEM.05402-11 PMID: 21948834
44. 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(6): 1841–7. doi: 10.1007/s00253-009-2418-5 PMID: 20082074
45. Oldfield C, Pogrebinsky O, Simmonds J, Olson ES, Kulpa CF. Elucidation of the metabolic pathway for dibenzothiophene desulphurization by *Rhodococcus* sp. strain IGTS8 (ATCC 53968). Microbiology. 1997; 143(9): 2961–73.
46. Masai E, Sugiyama, Iwashita N, Shimizu S, Hauschild JE, Hatta T, et al. The *bphDEF meta*-cleavage pathway genes involved in biphenyl/polychlorinated biphenyl degradation are located on a linear plasmid and separated from the initial *bphACB* genes in *Rhodococcus* sp. strain RHA1. Gene. 1997; 187 (1): 141–149. PMID: 9073078
47. Patrauchan MA, Florizone C, Eapen S, Gómez-Gil L, Sethuraman B, Fukuda M, et al. Roles of ring-hydroxylating dioxygenases in styrene and benzene catabolism in *Rhodococcus jostii* RHA1. J Bacteriol. 2008; 190(1): 37–47. PMID: 17965160

48. Iwaki H, Saji H, Abe K, Hasegawa Y. 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. 2005; 20(3): 144–150.

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

50. Pelletier DA, Harwood CS. 2-Hydroxycyclohexanecarboxyl coenzyme A dehydrogenase, an enzyme characteristic of the anaerobic benzoate degradation pathway used by *Rhodopseudomonas palustris*. J Bacteriol. 2000; 182(10): 2753–60. PMID: 10781543

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

52. Pérez-Pantoja D, Donoso R, Agulló L, Córdova M, Seeger M, Pieper DM, et al. Genomic analysis of the potential for aromatic compounds biodegradation in *Burkholderiales*. Environmental Microbiology. 2012; 14(5): 1091–1117. doi: 10.1111/j.1462-2920.2011.02613.x PMID: 22026719

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

54. Burton AJ, Giguère S, Sturgill TL, Berghaus LJ, Slovis NM, Whitman JM, et al. Macrolide- and rifampin-resistant *Rhodococcus equi* on a Horse Breeding Farm, Kentucky, USA. Emerg Infect Dis. 2013; 19(2): 282–285. doi: 10.3201/eid1902.121210 PMID: 23347878

55. Nordmann P, Ronco E, Nauciel C. Role of T-lymphocyte subsets in *Rhodococcus equi* infection. Infect Immun. 1992; 60(7): 2748–2752. PMID: 1351881

56. von Graevenitz A, Pünter-Streit V. Development of a new selective plating medium for *Rhodococcus equi*. Microbiol and Immunol. 1995; 39(4): 283–284.

57. Grippa E, Valla R, Battinelli A, Mazzanti G, Saso L Silvestrini B. Inhibition of *Candida rugosa* lipase by berberine and structurally related alkaloids, evaluated by high-performance liquid chromatography. Biosci Biot Bioch. 1999; 63(9): 1557–1562.

58. Barry CE, Lee RE, Mdluli K, Sampson AE, Schroeder BG, Slayden RA, et al. Mycolic acids: structure, biosynthesis and physiological functions. Prog Lipid Res. 1998; 37(2–3): 143–79. PMID: 9829124

59. Yi T, Lee EH, Ahn YG, Hwang GS, Cho KS. Novel biodegradation pathways of cyclohexane by *Rhodococcus* sp. EC1. J Hazard Mater. 2011; 191(1–3): 393–6. doi: 10.1016/j.jhazmat.2011.04.080 PMID: 21571424

60. McFarland BL. Biodesulfurization. Curr Opin Microbiol. 1999; 2(3): 257–264 PMID: 10383871

ELSEVIER

Institut Pasteur

Research in Microbiology xx (2016) 1–8

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 ■ ■ ■

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

Please cite this article in press as: Cappelletti M, et al., Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.06.008

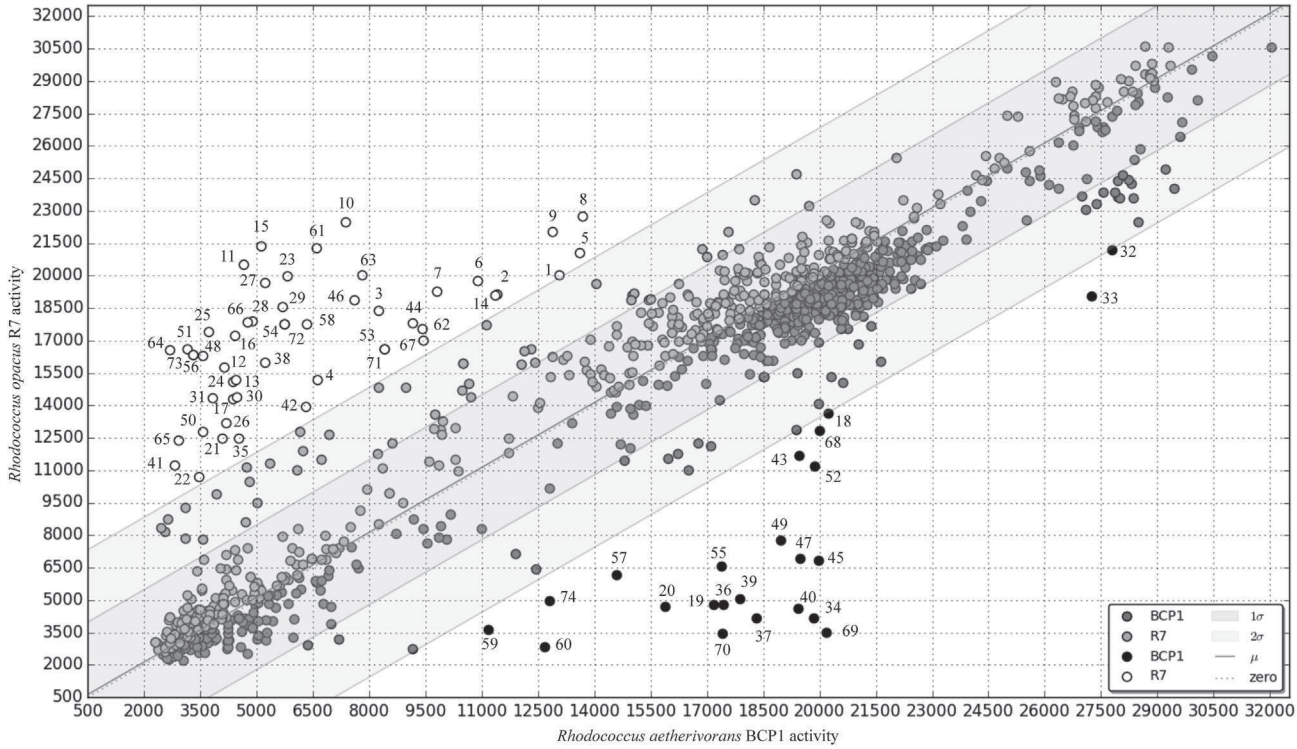

Fig. 1. Scatter plot comparing the metabolic activities of *R. opacus* R7 and *R. aetherivorans* BCP1 from PM plates 9–20. The metabolic activities of BCP1 and R7 are represented as dots on a 2D scatterplot. Plot bands represent mean (μ_A), standard deviation (σ_A) and double standard deviation ($2\sigma_A$) 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

Please cite this article in press as: Cappelletti M, et al., Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.06.008

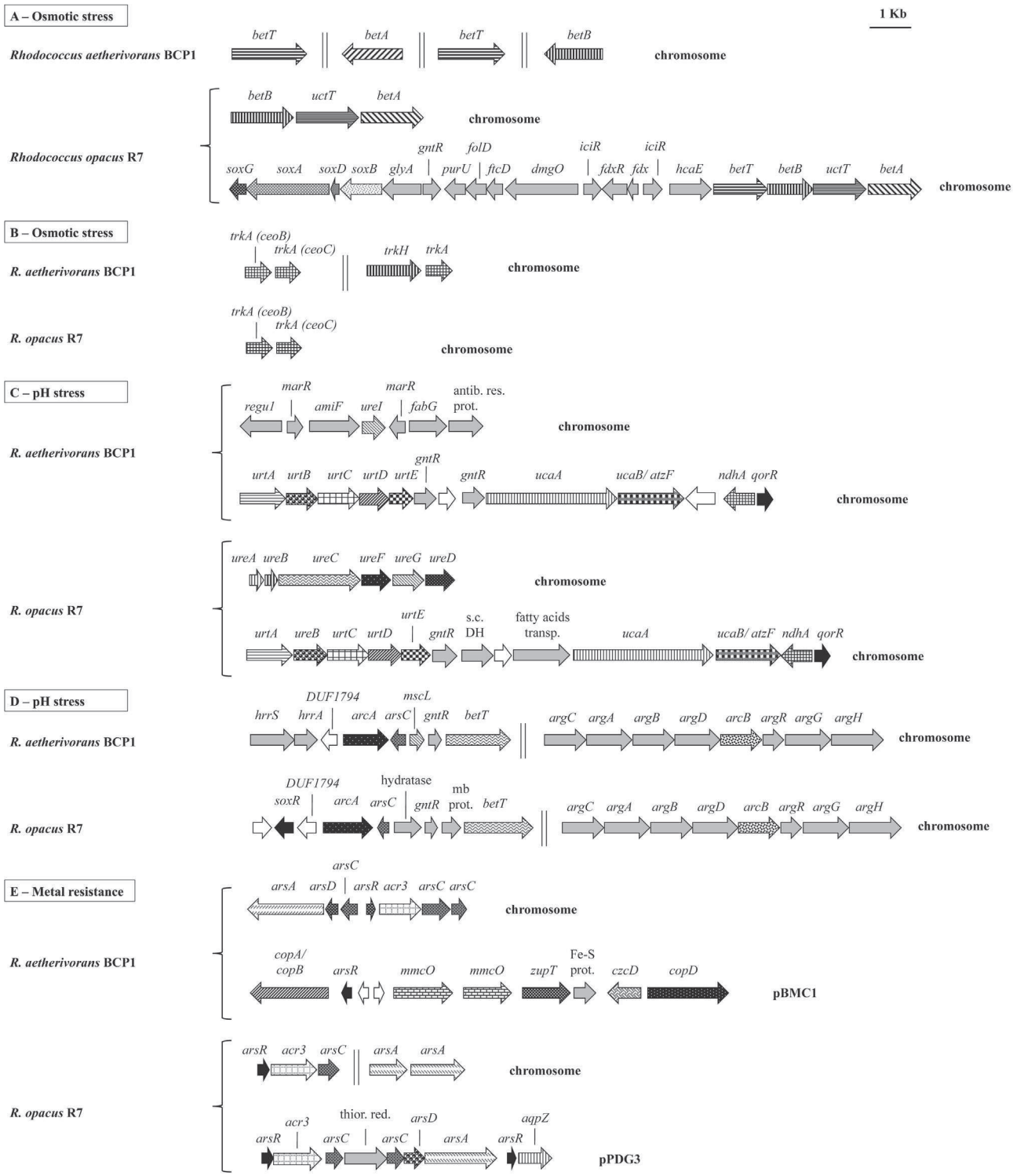

Fig. 2. Representative gene clusters involved in the *R. opacus* R7 and *R. aetherivorans* BCP1 stress responses whose presence or organization differentiates the two strains (except for *arcB* and the flanking genes). Genes are visualized as arrows. The predicted protein products of the genes are listed 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.

Please cite this article in press as: Cappelletti M, et al., Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.06.008

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

Please cite this article in press as: Cappelletti M, et al., Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.06.008

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).

Please cite this article in press as: Cappelletti M, et al., Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.06.008

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.

Please cite this article in press as: Cappelletti M, et al., Phenotype microarray analysis may unravel genetic determinants of the stress response by *Rhodococcus aetherivorans* BCP1 and *Rhodococcus opacus* R7, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.06.008