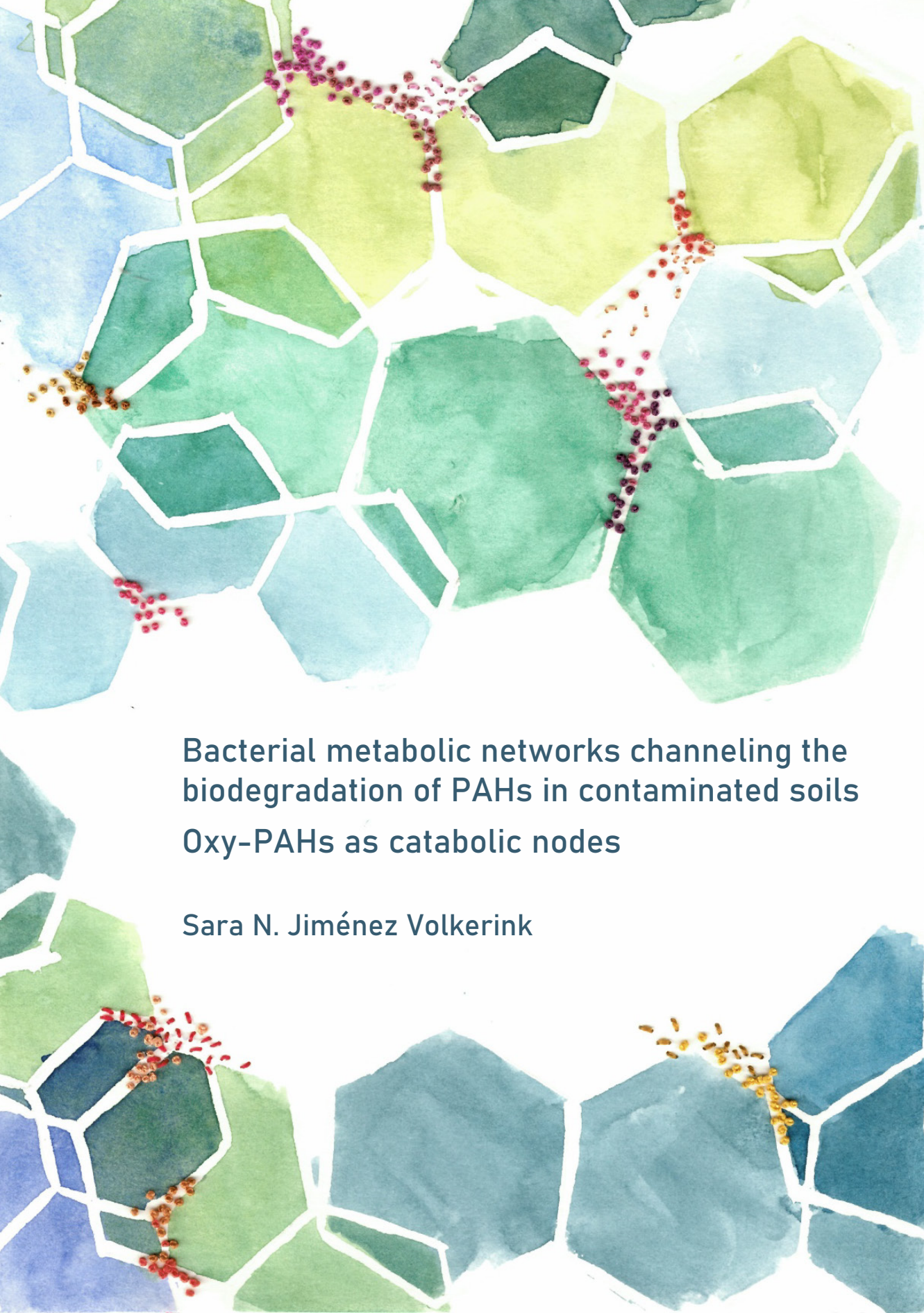# Bacterial metabolic networks channeling the biodegradation of PAHs in contaminated soils. Oxy-PAHs as catabolic nodes

Sara Nienke Jiménez Volkerink

Bacterial metabolic networks channeling the biodegradation of PAHs in contaminated soils

Oxy-PAHs as catabolic nodes

Sara N. Jiménez Volkerink

UNIVERSITAT DE BARCELONA

Departament de Genètica, Microbiologia i Estadística
Secció de Microbiologia, Virologia i Biotecnologia
Facultat de Biologia

Bacterial metabolic networks channeling the biodegradation of PAHs in contaminated soils. Oxy-PAHs as catabolic nodes

Memòria presentada per Sara Nienke Jiménez Volkerink per optar al grau de Doctora per la Universitat de Barcelona

Programa de Doctorat (EEES) de Biotecnologia

V. i p. de la directora i tutora

Magdalena Grifoll Ruiz

V. i p. del director

Joaquim Vila Grajales

Doctoranda

Sara Nienke Jiménez Volkerink

Barcelona, 2021

Thesis Cover: Artistic representation of the bacterial networks for the biodegradation of PAHs. Chemical structures of PAHs were painted in watercolors and bacteria were hand-embroidered with cotton thread. The artwork was entirely designed and created by the author.

Life did not take over the world by combat, but by networking.

— Lynn Margulis.

Acknowledgments

Estas palabras son las últimas que escribo y seguramente sean las más importantes. Una tesis no se hace sola. Y es que, igual que para las bacterias, la unión hace la fuerza. Me siento tremendamente afortunada por las personas que me han acompañado en este camino y que, sin duda, han hecho que haya sido tan feliz durante esta etapa.

En primer lloc, vull agrair als meus directors de tesi. Per obrir-me les portes al món de la ciència, per tot el que m'heu fet créixer tant a nivell professional com personal, per la vostra exigència i al mateix temps pel vostre suport, per la vostra dedicació i empenta. La passió, la motivació i l'entusiasme que teniu pel que feu són contagioses. **Magda**, tot va començar quan vas fer-me descobrir el fascinant món de la biodegradació en les teves classes de *Biotecnologia Ambiental* i vaig voler seguir aprofundint fent el TFG amb tu. I després el TFM, i el doctorat. Gràcies per donar-me la oportunitat d'aprendre tant. Quan vaig començar gairebé no sabia ni el que era una PCR, per posar un exemple. Gràcies per tots el moments mirant cromatogrames, espectres, plaques, gaudint-ho com si fossis una nena, i per la força que tens per tirar les coses endavant passi el que passi. **Quim**, tot i la distància durant els primers anys de la tesi, sempre has estat per mi en qualsevol moment. Sempre disposat a ajudar en el que calgui. Gràcies per les hores i hores de discussions científiques tan inspiradores i per la il·lusió que poses pensant en nous experiments i l'alegria compartida quan surten bé les coses.

I would like to thank **Hauke Smidt** for giving me the chance to work in his lab in the Wageningen University & Research. Despite it was only for three months, it has been an enriching experience both for the development of my Thesis and for myself. Thank you for the interest you put in my work and your enthusiasm.

També vull donar les gràcies a la **Cristina Minguillón**, per acollir-me durant unes setmanes al seu laboratori al Campus de Torribera i ajudar-nos a intentar sintetitzar el metabòlit de l'antraquinona i amb les anàlisis de RMN. Gràcies per la teva amabilitat i paciència.

A **José Julio Ortega,** gracias por tirar adelante el artículo de los girasoles y por tus disertaciones sobre biodisponibilidad. Y a **Rosa**, por haber tenido tanta paciencia con los pedidos y las facturas.

Siempre estaré agradecida a todas las personas que forman o han formado parte del laboratorio 5. **José María**, gracias por tu cordialidad y por tus pequeños chistes, fue un placer trabajar contigo en la búsqueda del *pico*. **Marga**, tu has estat un pilar

Abstract

Polycyclic aromatic hydrocarbons (PAHs) are major environmental pollutants in a number of point source contaminated sites, where they are found embedded in complex mixtures containing different polyaromatic compounds. The application of bioremediation technologies is often constrained by unpredictable end-point concentrations, enriched in recalcitrant high molecular weight (HMW)-PAHs, and by the formation of generally overlooked PAH transformation products, such as oxygenated-PAHs (oxy-PAHs). The starting hypothesis of the Thesis was that cometabolic interactions and partial oxidation processes could play an essential role in channeling PAH carbon fluxes, with oxy-PAHs acting as potential catabolic nodes. The general objective is to expand the knowledge on the metabolic networks that drive the biological removal of PAHs in contaminated soils. Benz(*a*)anthracene (B*a*A) was used as a model to unravel the microbial populations and functions involved in the biodegradation of HMW-PAHs. The combination of DNA-SIP and shotgun metagenomics of ^{13}C-labeled DNA allowed the unequivocal identification and functional analysis of the key B*a*A-degrading phylotype, a member of the novel genus *Immundisolibacter*. Analysis of the corresponding metagenome assembled genome (MAG) revealed a highly conserved genetic organization unique within this genus, including novel aromatic ring-hydroxylating dioxygenases. The influence of other HMW-PAHs on B*a*A degradation was ascertained in soil microcosms spiked with B*a*A and fluoranthene (FT), pyrene (PY) or chrysene (CHY) in binary mixtures. Co-incubation of PAHs resulted in a major delay in the removal of the less soluble PAHs and an increased formation of benz(a)anthracene-7,12-dione (B*a*AQ), the ready oxidation product from B*a*A, which was associated to relevant microbial interactions. To elucidate the mechanisms driving oxy-PAH biodegradation we isolated a 9,10-anthraquinone (ANTQ)-degrading bacterial strain, *Sphingobium* sp. AntQ-1. The metabolomic, genomic and transcriptomic characterization of the isolate served to reconstruct the ANTQ catabolic pathway, initiated by two sequential Baeyer-Villiger oxidations. Essential genes for the biodegradation of ANTQ were located in the megaplasmid *pANTQ-1*. The environmental relevance of the strain and the identified degradative mechanisms were confirmed by qPCR assessment during a previous biostimulation experiment of the creosote-contaminated soil. The metabolic networks involved in oxy-PAH biodegradation were investigated using a B*a*AQ-degrading microbial consortium obtained by enrichment in sand-in-liquid cultures with B*a*AQ as sole carbon source. The integration of data from metabolomic and metagenomic functional gene analyses revealed that the B*a*AQ metabolic pathway was probably initiated by the Baeyer-

Villiger monooxygenases encoded in *pANTQ-1*, indicating horizontal gene transfer phenomena. Our results suggest that Baeyer-Villiger oxidations, infrequent during PAH-biodegradation, could be a relevant mechanism for the processing of oxy-PAHs in contaminated sites, thus contributing to mitigate the potential risk of their accumulation. Further analysis of the B*a*AQ-degrading community MAGs also provided an insight into the potential roles and interactions within the consortium members. Several potential auxotrophies were detected, indicating that relevant interactions were taking place within the community members, not only to provide suitable carbon and energy sources, but also to supply essential nutrients and cofactors.

Contents

Chapter 2: Polyomic elucidation of the 9,10-anthraquinone biodegradation pathway by a newly isolated oxy-PAH degrading specialist from PAH-contaminated soil 65

Abbreviation list

ANI	Average Nucleotide Identity
ANT	Anthracene
ANTQ	9,10-anthraquinone
B*a*A	Benz(*a*)anthracene
B*a*AQ	Benz(*a*)anthracene-7,12-dione
BV	Baeyer-Villiger
BVMO	Baeyer-Villiger monooxygenase
CA	Correspondence Analysis
cDNA	Complementary Deoxyribonucleic Acid
CDS	Coding Sequence
CHY	Chrysene
DCM	Dichloromethane
DEG	Differentially Expressed Gene
DGGE	Denaturing Gradient Gel Electrophoresis
DNA	Deoxyribonucleic Acid
FT	Fluoranthene
GC-FID	Gas Chromatography – Flame Ionization Detector
GC-MS	Gas Chromatography – Mass Spectrometry
GN	Gram Negative
GP	Gram Positive
GTDB	Genome Taxonomy Database
HMW-PAH	High Molecular Weight Polycyclic Aromatic Hydrocarbon
HPLC	High Pressure Liquid Chromatography
HPLC-ESI-HRMS	High Pressure Liquid Chromatography – Electrospray Ionization – High Resolution Mass Spectrometry
IRNAS-CSIC	Instituto de Recursos Naturales y Agrobiología de Sevilla – Consejo Superior de Investigaciones Científicas
LMW PAH	Low Molecular Weight Polycyclic Aromatic Hydrocarbon

MAG	Metagenome Assembled Genome
MGP	Manufactured Gas Plant
MM	Mineral Medium
NAPL	Non-Aqueous Phase Liquid
NIST	National Institute of Standards and Technology
NMR	Nuclear Magnetic Resonance
N-PAC	Nitrogen containing Polycyclic Aromatic Compound
NSO-PACs	Nitrogen/Sulfur/Oxygen containing Polycyclic Aromatic Compound
OTU	Operational Taxonomic Unit
Oxy-PAH	Oxygenated Polycyclic Aromatic Hydrocarbon
PAC	Polycyclic Aromatic Compound
PAH	Polycyclic Aromatic Hydrocarbon
PCR	Polymerase Chain Reaction
PY	Pyrene
qPCR	Quantitative Polymerase Chain Reaction
RDP	Ribosomal Database Project
RHD	Ring Hydroxylating Dioxygenase
RHO	Ring Hydroxylating Oxygenase
RNA	Ribonucleic Acid
rRNA	Ribosomal Ribonucleic Acid
RT-qPCR	Reverse Transcription – Quantitative Polymerase Chain Reaction
SIP	Stable Isotope Probing
TCA	Tricarboxylic Acid
TPM	Transcript per Million
US-EPA	United States Environmental Protection Agency
WHC	Water Holding Capacity

Introduction

Introduction
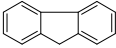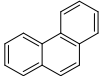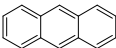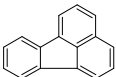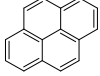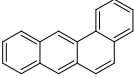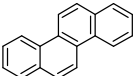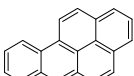
Nature, Source and Environmental Fate of PAHs

Polycyclic aromatic hydrocarbons (PAHs) are a heterogeneous group of hydrophobic organic compounds consisting of two or more fused benzene rings in linear, angular or cluster arrangements. PAHs are ubiquitously distributed in soils, water and sediments because of incomplete combustion of organic matter (e.g., forest fires, traffic exhaust, waste incineration) and subsequent atmospheric deposition (Srogi, 2007). However, they reach particularly high concentrations in industrial soils and marine environments affected by point source pollution due to accidental leakage or spillage associated with the extraction, transformation, transport or use of fossil fuels, of which they are major constituents.

The distinctive physicochemical properties of PAHs (Table 1) determine their environmental distribution, fate and (eco)-toxicological properties. The resonance energy featured by the delocalized π-electron distribution of their aromatic rings and the lack of electron-withdrawing substituents confer PAHs a high resistance to nucleophilic attack. Their low aqueous solubility together with their high solid-water and octanol-water partition coefficients constrain their availability in water, promoting their accumulation in the solid phases of soils and sediments as well as in the non-aqueous phase liquids (NAPLs) in which they are usually embedded. This recalcitrance to biochemical processing and low bioavailability result in environmental persistence, which increases with molecular weight. PAHs include acute toxicants, known carcinogens, endocrine disruptors and immune suppressors (Kim et al., 2013). They are bioaccumulated in animal tissues and, thus, biomagnified through food chains (Alexander et al., 2008). Acute toxicity is generally associated to low molecular-weight (LMW) PAHs (composed by two or three rings) (e.g., naphthalene), while the high molecular weight (HMW) PAHs (composed of four or more rings) include genotoxic compounds and known carcinogens (e.g., benzo(*a*)pyrene) (Kanaly & Harayama, 2010). It has been demonstrated that PAHs interfere on the activities of hormone metabolizing enzymes of the thyroid glands, posing potential adverse effects on reproduction and immune function (Oostingh et al., 2008). Thus, PAH-polluted soils and sediments pose a serious risk for public health and are devoid of their ecological functions, often being considered a hazardous waste (Landrigan et al., 2018).

Table 1. Structure and relevant physicochemical and toxicological properties of representative PAHs.

Name	Chemical structure	Water[a] solubility	Vapor[a] pressure	Log K_{ow}	Toxicity[b]	Carcino-genicity[c]
Naphthalene		31	11.3	3.35	1110	2B
Acenaphthene		3.90	0.287	3.92	10000	3
Fluorene		1.69	0.080	4.18	>16000	3
Phenanthrene		1.15	0.016	4.46	700	3
Anthracene		0.043	0.001	4.45	3200	3
Fluoranthene		0.26	0.0012	5.16	10-30	3
Pyrene		0.135	0.0006	4.88	2700	3
Benz[a]anthracene		0.009	0.00003	5.76	n.d.	2B
Chrysene		0.002	0.0000008	5.81	n.d.	2B
Benzo[a]pyrene		0.00162	0.0001	6.13	n.d.	1

[a] Water solubility (mg/L) and vapor pressure (Pa) are at 25ºC.

[b] Toxicity is reflected as LD50 (mg/kg) for oral ingestion by rats.

[c] International Agency for Research on Cancer (IARC) Classification (1: carcinogenic to humans, 2A: probably carcinogenic to humans, 2B: possibly carcinogenic to humans, 3: not classifiable as carcinogenic to humans)

n.d. = no data available

Source: Achten and Andersson, 2015.

Polycyclic aromatic compounds: beyond the 16 priority PAHs

PAHs are major soil contaminants especially in former wood processing facilities using creosote and manufactured gas plant (MGP) sites (Kuppusamy et al., 2016). According to the United States Environmental Protection Agency (US EPA), PAHs have been found in over 250 Superfund sites, and are the predominant pollutant class at an estimated 45,000 former MGP sites in the United States. The European

Environmental Agency has also identified PAHs as the main contaminants in more than 13% of EU-polluted sites (Panagos et al., 2013). Current analytical data and risk assessment procedures for those contaminated sites are exclusively based on the detection and quantification of the 16 PAHs included in the 42-years old US EPA list of priority pollutants (Keith & Telliard, 1979), with generic reference levels for these compounds being included in most European national regulations. However, this group does not include other polycyclic aromatic compounds (PACs), such as alkylated PAHs, oxygenated PAHs (oxy-PAHs) or heterocyclic compounds containing nitrogen, sulfur or oxygen (NSO-PACs). These compounds are always detected associated to PAHs in contaminated sites, and accumulate in soils and sediments (Larsson et al., 2018). Alkylated PAHs, which are far more abundant than the non-alkylated PAHs in petrogenic mixtures, have similar physicochemical and toxicological properties (Achten & Andersson, 2015), and are more resistant to biodegradation. Oxy-PAHs and heterocyclic compounds are also part of fossil fuels. For instance, creosote contains an 85% of PAHs, 10% phenolic compounds and 5% NSO-PACs (Mueller et al., 1989). Despite they are found at lower concentrations than PAHs, novel non-target analytical techniques are revealing a higher abundance and diversity than originally expected (Tian et al., 2017). Oxy-PAHs are also formed as a result of photo-oxidation or microbial transformation of the parent PAHs (see below). Both groups of compounds have higher water solubility (i.e., mobility) than the parent PAHs and some present higher toxicity (Achten & Andersson, 2015; McCarrick et al., 2019), having been identified in the most mutagenic organic fractions in soils, river and marine sediments (Grifoll et. al, 1990; Titaley et al., 2020). However, they lack any specific regulation, and are systematically overlooked by monitoring and risk assessment procedures. For all these reasons, some scientists claim the current list of priority pollutants should be updated and consider analytical targets representative of these classes of compounds (Andersson & Achten, 2015).

Bioremediation of PAH-contaminated soils

Of all the biogeochemical processes involved in their environmental fate, microbial biodegradation by the naturally occurring microbial communities is the only one that guarantees PAH removal by their eventual conversion to microbial biomass, CO_2 and H_2O. Bioremediation, exploiting these natural biodegradation processes, is considered the most cost-efficient technology for the clean-up of polluted sites, presenting low environmental footprints (e.g., CO_2 emissions, space) and being capable of restoring key natural soil functions (Gillespie & Philp, 2013). Indeed, this technology is now widely applied for the decontamination of PAH-contaminated soils (Kuppusamy et al., 2016). However, its use still faces some constraints: i) slow biodegradation and unpredictable endpoint concentrations, mainly enriched in

HMW-PAHs, due to both contaminant mass transfer and microbial catabolic restrictions; ii) the lack of adequate monitoring tools to accurately assess the occurrence of active biodegradation processes *in situ*; and, iii) the inadequate risk assessment policies based exclusively on the concentration levels for the 16 regulated PAHs, that neglect the co-occurrence and fate of other ecotoxicologically relevant PACs generally present in PAH-polluted sites. To overcome these limitations, it is essential to understand in depth the microbial processes that determine the fate of PAHs *in situ*, moving forward from the traditional "black box" perspective to an actual environmental biotechnology.

PAHs are found in the environment as part of complex mixtures exposed to diverse degrading microbial communities whose members harbor different, versatile and often interconnected metabolic pathways (Vila et al., 2015). However, the *in situ* biodegradation of PAHs not only relies on the microbial catabolic networks available in the impacted zone. Contaminants and microorganisms interact with the environmental matrix where they are embedded, which influences their distribution and behavior, and are affected by other environmental factors that strongly condition contaminant biodegradability.

Factors Modulating PAH Degradation in soils

The biodegradability of PAHs in terrestrial environments depends on a number of intensively interacting factors (Haritash & Kaushik, 2009). The first factor is the presence and abundance of a specialized microbiome with the necessary catabolic functions, which is directly related to the history of previous exposure to the pollutant(s). Other site-specific factors modulating the rate and extent of degradation of PAHs *in situ* act on the catabolic activity and growth of PAH-degrading microorganisms, the bioavailability of the substrates, or both (Figure 1).

Although a high PAH concentration may inhibit biodegradation due to toxicity, the concentration of PAH bioavailable for microorganisms must be higher than the threshold substrate concentration required, for example, for the transcriptional activation of microbial promoters, the induction of the expression of biodegradation genes or the associated microbial growth. In PAH-contaminated soils, the adverse effect of an excessive concentration is especially relevant for compounds presenting high water solubility and acute toxicity to microorganisms, such as naphthalene. In contrast, the relatively low proportions of HMW compounds like benzo(*a*)pyrene in PAH-mixtures, and their low water solubility, could play a role in their slow biodegradation rates. This can have serious implications for risk management, as threshold values for microbial activity can be allocated well above the stringent reference levels defined for these carcinogenic compounds in some national

regulations. Those levels vary depending on the countries and the environmental matrix. For example, the Dutch soil policy, used as a model for most European countries, stablishes a value of 1.5 mg/kg as the maximum reference level for the sum of 10 representative PAHs (including compounds from 2 to 5 rings) in soils for recreational use. Since PAHs are constituents of complex mixtures, substrate interactions during biodegradation are crucial to determine the fate of PAHs, in particular of HMW PAH compounds with five-rings or more. These HMW PAHs are considered to be exclusively degraded by cometabolic oxidation during the utilization of lower molecular weight compounds (Juhasz & Naidu, 2000). Conversely, competitive inhibition between substrates has also been observed, for example during the degradation of phenanthrene by pseudomonads in the presence of naphthalene, methylnaphthalene and fluorene (Stringfellow & Aitken, 1995). The
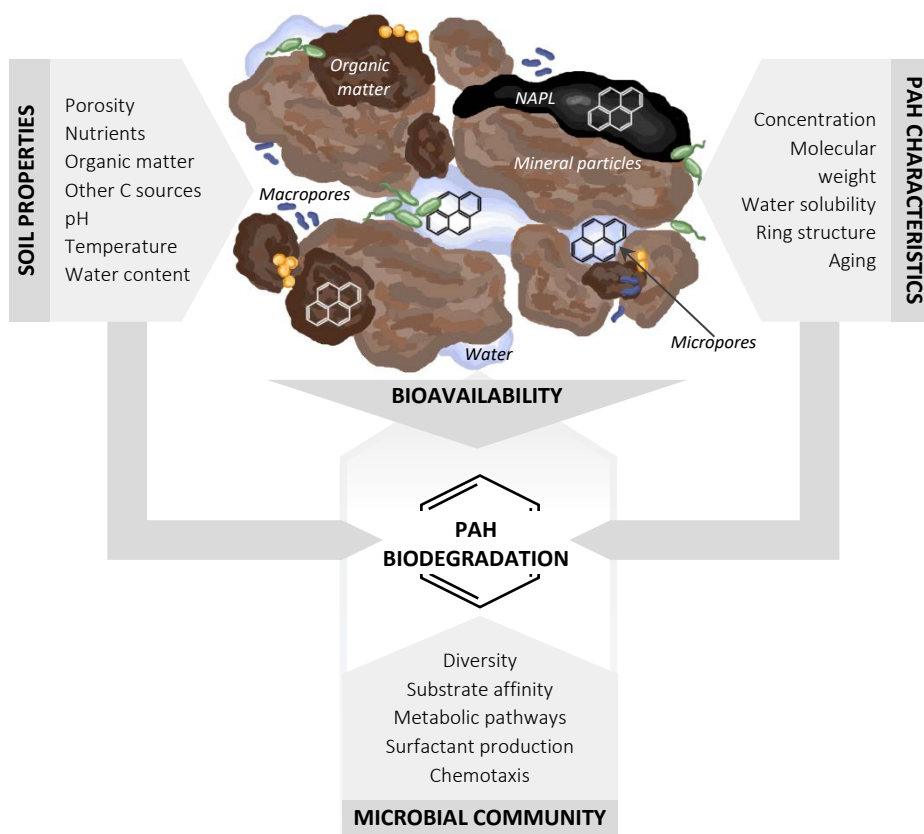
Figure 1. Schematic overview of the factors controlling PAH biodegradation and bioavailability in soils. Three major groups of parameters are considered, including those related with the soil properties, the chemical characteristics and the naturally occurring microbial communities.

structure and physicochemical properties of a given PAH also influence its environmental biodegradability. In general, it is assumed that the recalcitrance of PAHs increases with increasing molecular weight and number of substituents in the molecule. Therefore, within mixtures, LMW PAH compounds are more readily and extensively degraded than HMW PAHs; also, non-methylated PAHs are more susceptible to biodegradation than their corresponding methylated congeners. These trends have been observed for the degradation of crude oil and its derivatives, during laboratory biodegradation experiments using environmental microcosms (Wang et al. 1998) and mixed (Vila et al., 2010) or pure bacterial liquid cultures (Vila & Grifoll, 2009).

The low water solubility of PAHs is a main factor that hinders their biodegradation in soils and sediments, having a direct impact on bioavailability. Indeed, their high hydrophobicity is the source of their strong tendency to sorb onto solid surfaces, such as clays or soil organic matter (Ortega-Calvo et al., 2013). Another relevant aspect is that PAHs are normally found in complex materials, such as NAPLs, like creosote or coal tar. Considering the high partition coefficients of PAHs between organic phases and water (K_{ow}) (Table 1), PAHs remain entrapped into NAPLs or soil organic matter and their slow rates of transfer to the water phase limits their availability to microorganisms, and thus biodegradability. It is worth noting that hydrophobicity of PAHs increases almost logarithmically with increasing molecular mass, these phenomena having thus a stronger influence on the environmental fate of HMW PAHs.

Bioavailability represents the accessibility of a chemical for biotransformation and toxicity (Ortega-Calvo et al., 2015). In terms of biodegradation, the bioavailable fraction is the portion of a chemical that can be taken up or transformed by living organisms during the course of bioremediation and is considered one of the most crucial factors that limits the environmental degradation of hydrophobic organic contaminants. In addition to the intrinsic physicochemical properties of PAHs, other parameters related to the characteristics of the soil or the microorganisms may influence their mass transfer. The composition of the soil (e.g., mineral and organic matter content), its structure (e.g., porosity) and hydrogeology (e.g., contaminant mixing) can influence the dissolution, desorption, diffusion and distribution of the contaminant. At the same time, the degree of interaction of the PAHs with the soil matrix and its transformation will be highly influenced by the time of exposure, in a phenomenon known as aging and weathering of the contaminant (Boulangé et al., 2019). The longer the contact of the PAHs with the soil, the more irreversible will be their sorption and higher their penetration into the finer soil aggregates, thus decreasing the accessibility of bacteria to the contaminants. During bioremediation, one strategy to enhance the bioavailability of the slowly desorbing PAHs is the

application of surfactants to the soil, which has resulted in variable outcomes (Ortega-Calvo et al., 2013). However, a research topic of increasing interest is how bacteria can actively modulate their accessibility to the low bioavailable fraction of PAHs. Natural soil bacteria have developed mechanisms to overcome the limitations posed by low substrate bioavailablity. These mechanisms include the production of surface-active compounds to increase the mass transfer of PAHs into the water phase by pseudosolubilization (Posada-Baquero et al., 2019); their active movement towards the source of contaminant favored by chemotactic behavior (Ahmad et al., 2020) and/or their migration along fungal mycelia (Furuno et al., 2010); or the increase of PAH bioavailability by the direct contact of bacterial cells with contaminant NAPLs (Garcia-Junco et al., 2003).

Biodegradation of PAHs

Microorganisms are capable of metabolically transform or mineralize virtually all natural and anthropogenic compounds under a broad range of environments and physicochemical conditions (e.g., pH, temperature, redox potential). Ideally, PAH biodegradation will involve their progressive oxidation and break down until their complete mineralization to H_2O and CO_2 in aerobic environments or their conversion to CH_4 under strict anaerobic conditions. As explained above, the rate and extent of biodegradation is generally assumed to result from the combination of a number of variables, such as the environmental conditions, the number and versatility of the degrading microorganisms present, and the nature, chemical structure and bioavailability of the chemical compounds being degraded.

To estimate the biodegradability of organic contaminants, the most common approach is to conduct laboratory incubations of pure cultures, microbial consortia or environmental samples in the presence of the compound of interest. Biodegradation is then measured from the application of physiological and chemical tests, such as cell growth, decrease in substrate concentration, consumption of electron acceptors or the accumulation of intermediate metabolites or metabolic end-products (e.g., mineralization to CO_2). These tests give an accurate idea of the potential biodegradability of the contaminant under specific incubation conditions; however, the actual mechanisms involved, including transformation reactions, potential accumulation of partially oxidized intermediates, or the microbial degrading populations and their metabolic synergies during environmental sample incubations, are generally not investigated (Vila et al., 2015).

To understand the factors and mechanisms underlying environmental biodegradability of pollutants, scientist have addressed all these questions with increasingly complex and complementary study approaches, spanning from pure

cultures or soil microcosms to *in situ* biodegradation experiments. These approaches combine environmental and molecular microbiology, organic chemistry, biochemistry and molecular systems biology (Figure 2).
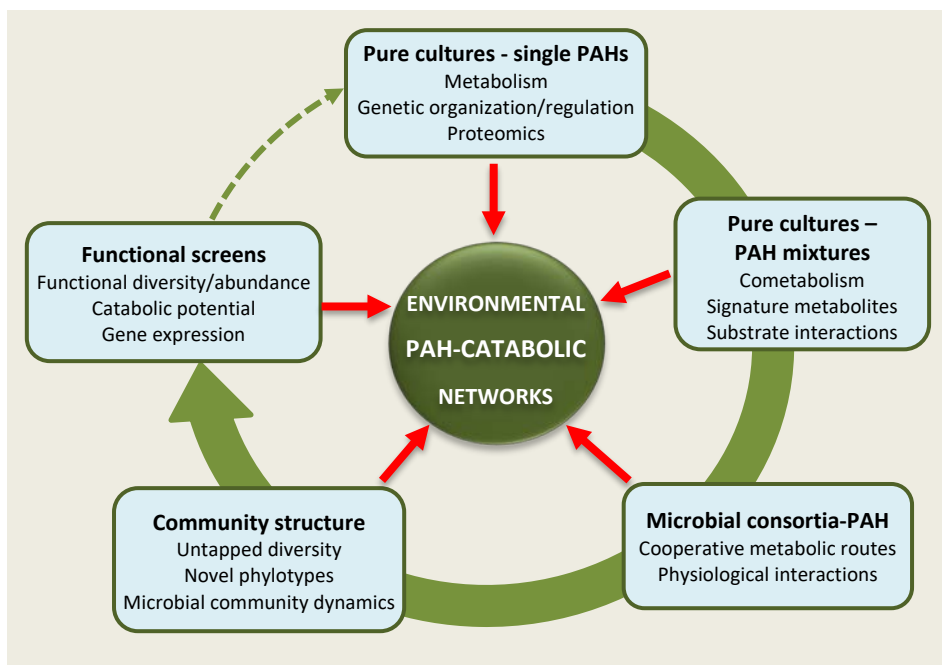
Figure 2. Information flow in the complementary scientific approaches for the progressive unraveling of the terrestrial and marine microbial networks involved in PAH degradation. The integration of taxonomical and functional data from pure cultures, microbial consortia, and environmental communities after exposure to single PAHs or PAH mixtures will enable to reconstruct the actual processes occurring in polluted sites. Insights on novel phylotypes and functions from community analysis will provide feed-back information and define targets for directed strain isolation, thus closing the circle and strengthening the information flow. Reproduced from Fig 1 in Vila *et al.* (2015). Bacterial PAH degradation in marine and terrestrial habitats. *Current Opinion in Biotechnology* 33: 95-102, with permission from Elsevier.

Fundamentals of PAH Aerobic Biodegradation

The classical approach to study microbial metabolism of PAHs has been the reconstruction of pathways after the identification of metabolites accumulated during incubations of pure cultures of bacteria or fungi in the presence of individual substrates. Those efforts have been mainly devoted to the elucidation of transformation pathways under aerobic conditions (Kanaly & Harayama, 2000; Bamforth & Singleton, 2005). As a result, the biochemical reactions for the

degradation of LMW PAHs in aerobic microorganisms are well established (Figure 3). The microbial transformation of PAHs can serve two major functions: i) the assimilative biodegradation, that yields carbon and energy for the growth of the degrading organisms, and generally results in the complete or partial mineralization of the compound; or ii) detoxification reactions, which aim to transform PAHs into more water-soluble molecules to facilitate their excretion from the cytoplasm. In general, the first mechanism has been associated to the actions of bacteria, while detoxification reactions have been linked to the activity of fungi (Cerniglia, 1992).

Cellular Metabolic Pathways for PAH Biodegradation in Aerobic Bacteria

The ability to use PAHs as sole carbon and energy source under aerobic conditions is a widespread trait within the prokaryotic world. Researchers have associated the ability to degrade PAHs to bacteria belonging to a variety of phyla, including three major classes of *Proteobacteria* (alpha, beta and gamma), *Firmicutes*, *Actinobacteria*, *Bacteroidetes*, *Cyanobacteria*, and even some halophilic *Archaea* (Ghosal et al., 2016). PAH-degrading bacteria have been found in diverse ecological niches, including those in extreme habitats, and experience a bloom in response to eventual hydrocarbon exposure. Bacteria utilize hydrocarbons as sources of carbon and energy for growth. Thus, ideally, biodegradation of PAHs involves their progressive oxidation and breakdown to furnish building blocks for cell biosynthesis, and electrons for ATP production through the aerobic respiratory chains. In addition to being the final electron acceptor in these pathways, oxygen acts as a co-substrate for the initial hydroxylation and cleavage of the aromatic rings mediated by different sets of oxygenases.

The first step in the bacterial aerobic degradation of PAHs is the hydroxylation of one of the aromatic rings via a ring hydroxylating dioxygenase (RHD) that introduces the two atoms of molecular oxygen in vicinal carbons to form a cis-dihydrodiol. RHD are multicomponent enzyme systems generally constituted by reductase, ferredoxin and terminal oxygenase subunits (Mallick et al., 2011). The *cis*-dihydrodiol then is rearomatized by the action of a dehydrogenase to yield the corresponding diol. The ring is now ready to be attacked either by an *ortho-* or a *meta*-cleavage dioxygenase that opens the aromatic ring to produce, respectively, a *cis,cis*-muconic acid or a hydroxymuconic acid semialdehyde analogue. Further molecular rearrangements and the action of hydrolases release 2-3 carbon fragments that furnish the central metabolism. Subsequent similar reactions result in the progressive mineralization of the molecule (Figure 3).

The first metabolic pathway for the degradation of PAHs was described for the bicyclic compound naphthalene utilized as a growth substrate by a strain of
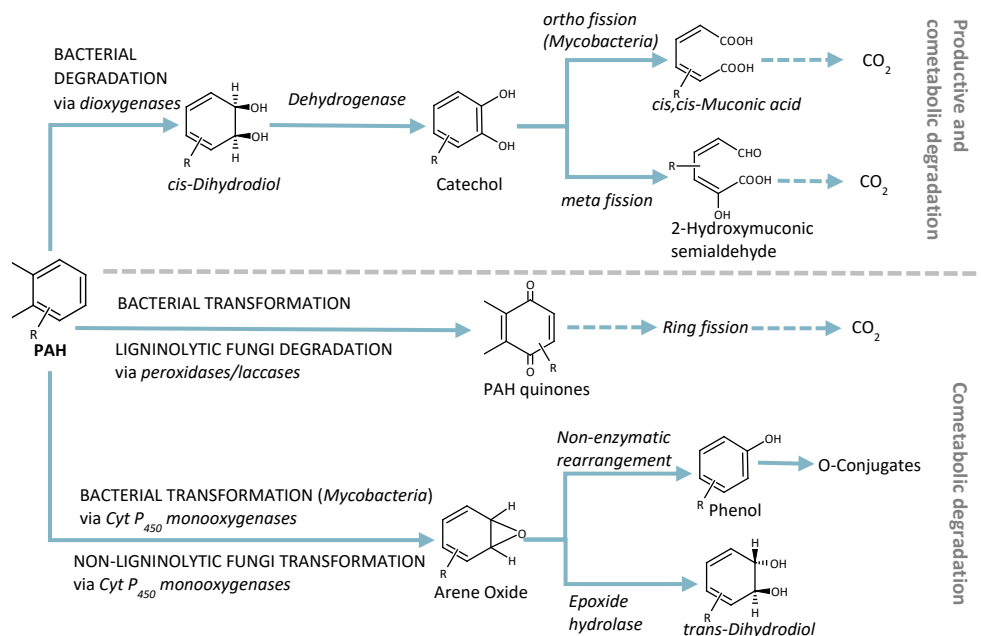
Figure 3. Main biochemical pathways proposed for the aerobic microbial degradation or transformation of polycyclic aromatic hydrocarbons. Modified from Fig 4 in Cerniglia (1992) Biodegradation of polycyclic aromatic hydrocarbons. Biodegradation 3: 351-368.

Pseudomonas putida (Davies & Evans, 1964). The catabolic genes (*nah*) were located in a large conjugative plasmid, NAH7, and were organized in two operons (Yen & Gunsalus, 1985). The first operon contained the genes encoding for the enzymes of the upper pathway, which after meta-cleavage and pyruvate release, produced salycilic acid. The second operon included the genes encoding for the lower pathway, a series of enzymes that subsequently converted this aromatic hydroxyacid to intermediates of the central metabolism. Both operons were positively regulated by a common transcriptional regulator (*NahR*) that responded positively to the presence of salicylate. Similar linear arrangements have been described for different pseudomonads, some of which are able to degrade the three ring compounds phenanthrene and anthracene by analogous routes (Sakshi & Haritash, 2020).

Similar organization in upper and lower metabolic pathways has been generalized for PAC degradation by most bacterial taxa, although genes encoding those pathways may not be necessarily arranged into linear operons. In order to accommodate a variable number of polyaromatic substrates, bacteria can possess one or more upper pathways, that may include reactions atypical in the

abovementioned canonical PAH-degradation mechanisms described from the study of pseudomonads. Those routes channel PAC degradation towards a limited number of monoaromatic intermediates (ie. cathecol, protocatechuic acid, gentisic acid or salicylic acid), which are processed following well-stablished lower pathways to furnish the central metabolism with TCA intermediates.

Another important group of PAH-degrading bacteria are the sphingomonads, especially those included in the genera *Sphingomonas*, *Sphingobium* and *Novosphingobium*. Sphingomonads are frequently isolated from soils by their capability to utilize 3- or 4-ring PAHs (mainly phenanthrene or fluoranthene), showing a great versatility for the utilization of other aromatic substrates (Stolz, 2009). Their degradative genes are also located in large catabolic plasmids, however, unlike pseudomonads, they are not organized in discrete operons but scattered in the genome and often flanked by transposons (Pinyakong et al., 2003). In fact, the catabolic versatility of these bacteria seems to rely in both, the possession of a variety of different and up-regulated ring-hydroxylating dioxygenases (RHD) and the relaxed specificity of those enzymes. Thanks to the whole genome sequencing of different PAH-degrading microorganisms and further -omics and systems biology studies (Schuler et al., 2009; Aylward et al., 2013; Khara et al., 2014; Verma et al., 2014), the complex network of catabolic reactions and its regulation in a single cell can now be decoded at the system level, changing the perspective from the traditional linear routes to highly branched pathways. For example, *Sphingobium* sp. KK22 grows on phenanthrene and transforms the HMW-PAHs fluoranthene, benzo(*a*)anthracene and benzo(*k*)fluoranthene, initiating the oxidation of these substrates at different positions (Maeda et al., 2020). *Sphingobium* sp. PNB, capable of attacking a wide range of mono- and polyaromatic substrates harbors seven sets of ring hydroxylating oxygenases (RHO) with different substrate specificities that interact with a sole set of downstream degradative genes (Roy, 2013). The marine bacterium *N. pentaromativorans* US6-1 harbors a variety of degradation pathways mediated by genes located either in the large plasmid pLA1 or in its chromosome, that are separately induced by PAHs and monoaromatic compounds (Yun et al., 2014). This plasticity in their catabolic networks and their short duplication times confer sphingomonads a pronounced selective advantage towards a wide range of aromatic pollutants, contributing to their fast response to sudden PAH-pollution and nutrient supply (r-strategists).

The bacterial metabolism of HMW PAHs with 4 aromatic rings such as pyrene was unknown until the isolation of *Mycobacterium vanbalenii* PYR-1 by the group of Cerniglia (Kim et al., 2007). In fact, Actinobacteria, with slow growth rates and their ability to adhere to hydrophobic surfaces, are highly adapted to the poorly bioavailable HMW substrates (k-strategists) (Kanaly & Harayama, 2010). Like

sphingomonads, mycobacterial strains are highly versatile degraders, being able to attack not only a wide range of PAHs (from 2- to 4-rings) but also efficiently degrade alkanes including branched and long chain compounds. Productive routes are initiated by RHDs to form *cis*-dihydrodiols, which are subsequently dehydrogenated to diols. However, these diols can be opened by both, *meta*-cleavage or *ortho*-cleavage dioxygenases. Furthermore, PAHs may undergo an initial attack by P450 monooxygenases to give *trans*-dihydrodiols and phenols which can be eventually methylated. This variety of reactions result in ramified catabolic routes where, simultaneously, part of the PAH molecules undergo productive reactions that lead to intermediates of the central metabolism and mineralization, and another part are transformed to atypical products that can accumulate, and whose environmental fate is often unknown (Figure 5). Recent studies have shed light on the metabolic organization of mycobacteria, although regulatory mechanisms remain to be described (Kweon et al., 2015). Similarly to sphingomonads, PAH-degrading
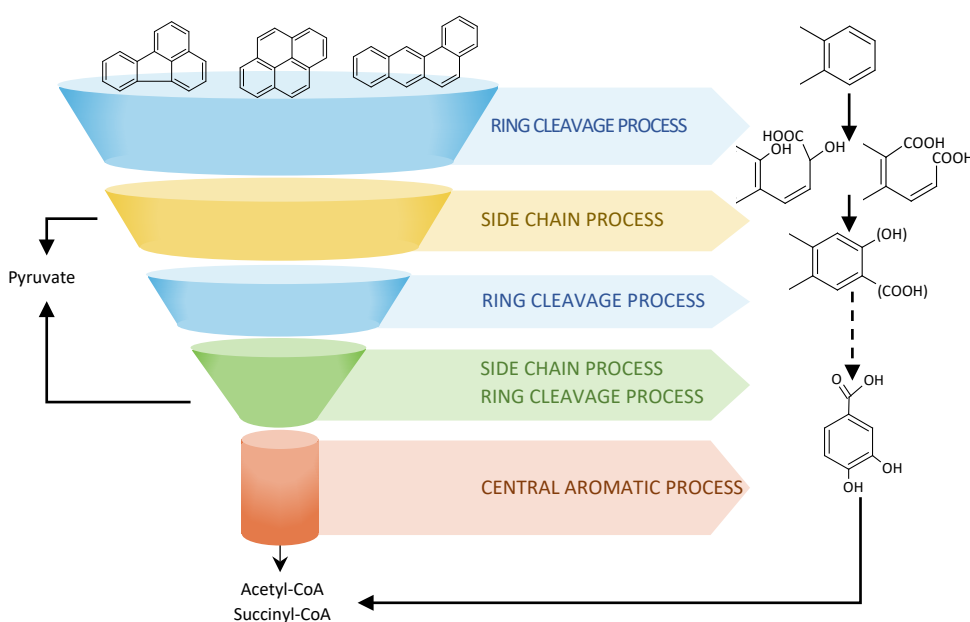
Figure 4. Funnel effect of the PAH metabolic network of *Mycobacterium vanbaalenii* PYR-1. This scheme represents how within a cell several PAHs can be processed by the different initial ring hydroxylating dioxygenases, in a process favored by the relaxed specificity of these enzymes. After ring opening and side chain processing, smaller molecules can be channeled to the central metabolism using mechanisms common for different polyaromatic substrates. Adapted from fig 9 in Kweon et al. (2011) Polycyclic aromatic hydrocarbon metabolic network in *Mycobacterium vanbaaleni* PYR-1. Journal of Bacteriology 193: 4326-4337, with permission from ASM.

mycobacterial cells harbor several RHDs with different range of substrate specificities. The pyrene-degrading *Mycobacterium vanbaalenii* PYR-1 (Kweon et al., 2011) has 21 genes encoding for ring hydroxylating oxygenases. The RHDs *NidAB* and *NidA3B3*, despite their relaxed specificity, present their highest conversion rates for pyrene and fluoranthene, respectively. These enzymes have evolved to specifically accommodate HMW-PAHs in the large substrate-binding pockets of their active sites, thus satisfying spatial requirements for their efficient and regiospecific dihydroxylation. The combination of high-throughput proteomic analyses with genomic and metabolomic data, has served to reconstruct a comprehensive and hierarchical metabolic network (183 intermediates, 224 chemical reactions) in funnel organized modules, with a diversity of ring-cleavage and side chain processes that progressively narrow down to central metabolism; thus providing a systems-wide perspective on the biodegradation of PAHs (Figure 4). The evolution of those highly flexible metabolic networks relies on horizontal gene transfer mechanisms facilitated by the fact that, as mentioned above, the corresponding genes are frequently encoded within mobile genomic material, such as plasmids or transposon-flanked genomic regions (DeBruyn et al., 2012).

Cell Cometabolic Reactions Drive Cooperative Degradation of Environmental PAH Mixtures

Knowledge gathered from the actions of pure cultures or enzymes towards single substrates is a first and fundamental step in the progressive approach for comprehending the ecosystem networks that determine the environmental fate of PAHs. However, in soils and sediments, PAHs are found components of complex organic matrices (NAPLs), and the relaxed specificity of the degrading enzymes, especially RHDs, anticipates that degrading populations will simultaneously act on a variety of components. Depending on the capability of the subsequent enzymes of the metabolic pathways to accommodate the products of the preceding reactions, the single degrading strains will produce a spectrum of products oxidized at different extents. In some cases, the partial oxidation will permit the bacterial cell to release carbon fragments to furnish the central metabolism and obtain energy; in others, the reactions will be non-productive and fall in the category of cometabolic reactions (Nzila, 2013). In this case, the strain will use some PAHs for growth while deriving part of the energy to produce these partially oxidized products without gain of energy or carbon at the single cell level. This was first demonstrated in the degradation of creosote-PAHs by a strain of *Bulkholderia* in the middle 90's (Grifoll et al., 1995).

In addition to the relaxed specificity of RHD in dihydroxylating a variety of substrates, these enzymes can also act as monooxygenases catalyzing the
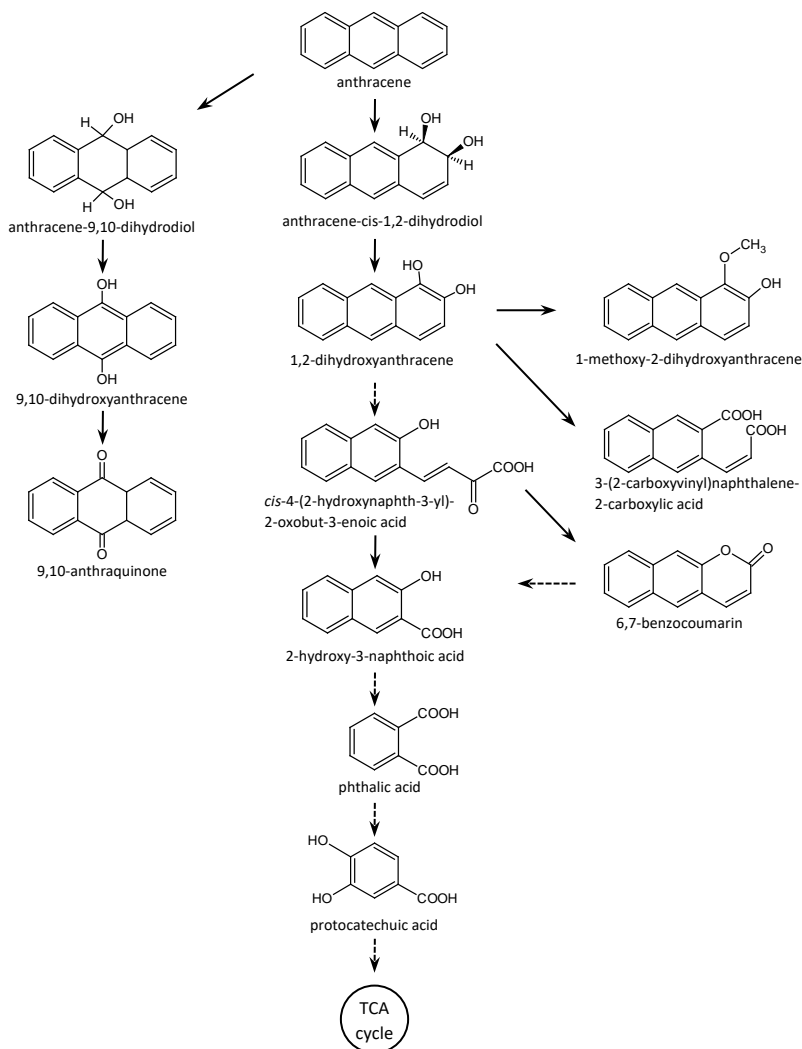
Figure 5. Proposed pathways for the biodegradation of anthracene by mycobacteria. The central route is a productive degradation pathway initiated by a ring hydroxylating dioxygenase (RHD) that leads to intermediates of the central metabolism (TCA cycle) and mineralization. Lateral branches encompass the transformation of anthracene to products whose environmental fate has not been elucidated. Dashed arrows indicate more than one successive step.

hydroxylation of both, methylenic groups in napththenoaromatic compounds (i.e., fluorene, acenaphthene) to give the corresponding ketones or quinones (Figure 5), and the methyl groups in alkylated PAHs, which results in the formation of aromatic carboxylic acids (Mallick et al., 2011). For example, the fortuitous action of naphthalene 1,2-dioxygenase on fluorene produces 9-fluorenone, while on

acenapthene produces acenaphthenone, acenaphthoquinone and 1,8-napthtalene dicarboxylic acid (Grifoll et al., 1995). Mycobacteria mineralize anthracene by dioxygenation at the C-1 and C-2 positions and further ring *meta-* or *ortho-*cleavage, however, *Mycobacterium vanbaalenii* PYR-1 can also catalyze 9,10-dioxygenation resulting in the formation of the dead-end product 9,10-anthraquinone (Moody et al. 2001) (Figure 5). As mentioned above, the action of the actinobacterial P450 monooxygenases on PAHs may also result in oxy-PAH formation. *Mycobacterium vanbaalenii* PYR-1 converts pyrene to hydroxypyrenes, dibenzothiophene to dibenzothiophene sulfoxide and 7-methylbenz(a)anthracene to the corresponding methanol derivative as a result of the activity of P450 monooxygenases (Brezna et al., 2006).

The actions of high molecular weight degrading mycobacteria during the biodegradation of complex PAH-mixtures were first evaluated using the pyrene-degrading strain *Mycobacterium gilvum* AP1 (Vila et al., 2001), isolated from a crude oil contaminated beach. When incubated in the presence of creosote-PAHs, the strain simultaneously attacked most of the predominant compounds in the mixture (including 2-, 3-, and 4-ring compounds), using some of these products as growth substrates (e.g. fluoranthene, pyrene and phenanthrene) while transforming others (e.g. naphthalene, fluorene or acenaphthene). As a result, a number of oxidized PAHs, including products from cometabolic and partial degradation reactions, accumulated in the medium (López et al., 2008). When incubated with an oil derivative, such as a heavy fuel oil, strain AP1 degraded almost completely the aliphatic fraction (C12-C40), which favored an extensive and selective degradation of 2-, 3- and 4-ring PAH families (Vila & Grifoll, 2009). Identification of signature metabolites indicated utilization of phenanthrene, pyrene, fluoranthene and alkyl derivatives by known assimilatory metabolic routes, while other components were co-oxidized.

Environmental PAH-Degrading Microbial Communities - Untapping Microbial Key Players

Culture-dependent approaches based on the isolation of bacteria with the ability to grow on PAHs and their further catabolic, genomic or proteomic characterization have provided fundamental knowledge on the mechanisms involved in PAH biodegradation. However, culture-based analyses give a biased perspective of the environmental microbial communities, underestimating their actual diversity and overlooking potential interactions between their components. During the last two decades, culture-independent methods have been increasingly applied to investigate the microbial communities present in PAH-impacted environments, or the structure of microbial consortia selected for their broad PAH-degrading

capabilities (Vila et al., 2015). The most widely used approach for the analysis of the microbial community structure is based on PCR amplification of the 16S rRNA genes, followed by fingerprinting methods (e.g., DGGE), clone library analysis, or more recently, high throuput sequencing of tag-encoded libraries.

The application of those molecular tools has revealed that, during active PAH degradation in contaminated soils, a limited number of taxonomic groups predominate (Lladó et al., 2009; Singleton et al., 2013), including members of sphingomonads (*Sphingomonas*, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*), *Bulkholderia*, *Pseudomonas*, or *Mycobacterium*, the latter generally present with a lower prevalence. Similarly, in marine environments, studies analyzing the microbial community composition in the water plume or sediment samples from harbors, or monitoring the microbial community in the aftermath of massive oil spills (e.g., Deepwater Horizon, Prestige or Nakhodka), revealed blooms of a group of highly specialized hydrocarbonoclastic bacteria after exposure to petroleum hydrocarbons (Fuentes et al., 2016; Head et al., 2006). Within this group, members of *Gammaproteobacteria*, such as *Alcanivorax*, *Colwellia*, *Cycloclasticus*, *Marinobacter*, *Methylophaga*, *Neptunomonas*, *Thalassolituus* or *Oleispira*, clearly predominate; and *Colwellia*, *Cycloclasticus*, *Marinobacter*, *Methylophaga* and *Neptunomonas* have been associated to PAH-biodegradation. This type of approach provides insights into the taxonomical composition of the microbial communities, but the role of the detected phylotypes is merely inferred from their phylogenetic affiliation and the hydrocarbon-degrading capabilities described for previously isolated relatives. Therefore, the function for an often substantial proportion of yet to be cultured members of the microbial communities is difficult to assign (Langille et al., 2013).

Functional screens are based on the detection and quantification of functional marker genes encoding for key catabolic enzymes. These analyses complement phylogenetic studies providing functional insights on the microbial communities. Based on genomic information retrieved from bacterial isolates, several sets of primers and probes have been defined to target bacterial biomarkers for PAH-biodegradation, mainly the genes encoding for ring-hydroxylating or ring-cleavage dioxygenases (Meynet et al., 2015). The most common approach is the quantitative amplification by qPCR using degenerate primers, which provides estimates on the abundance of PAH-degrading microbial communities (Cébron et al., 2008). Alternatively, the hybridization of total DNA on functional microarrays (Nostrand et al., 2012), such as the GeoChip (Glomics, Inc., Norman, OK), can provide an overview of the catabolic potential present in PAH-contaminated environments.

Most qPCR-based functional studies target the large-subunit (α-subunit) of RHDs, which contains the catalytically relevant components and determines the substrate specificity of the enzyme. PAH-RHDs are currently classified into five different clusters, defined on the basis of substrate specificity of the enzymes and the amino acid sequence of the α-subunits (Chakraborty et al., 2012). Considering this diversity, the published primer sets for qPCR have limited effectiveness, and either target for RHDs of a limited number of taxa, or have such high degree of degeneracy that compromises their specificity (Iwai et al., 2011). To address this issue and quantify RHD gene abundances in PAH-contaminated environments, researchers have used different approaches. These include the utilization a discrete number of primer sets encompassing the major clusters of RHD genes from Gram-positive (i.e., *nidA*) and Gram-negative bacteria (i.e., *nah*); or apply a relatively large number of primer sets targeting a range of RHD genes according to their substrate specificities, or their classification in different clades and subclades (Meynet et al., 2015). Microarrays constitute a high-throughput approach to analyze the functional diversity in environmental samples that avoids the need of PCR amplification. The most comprehensive functional gene array currently available is GeoChip (He et al., 2010), that in its current version contains 167,044 distinct probes that cover 395,894 coding sequences from about 1500 functional gene families involved in a broad range of biogeochemical processes (including PAH biodegradation). Different versions of GeoChip have been applied to characterize the changes in the catabolic capabilities of microbial communities occurring in the deep-sea water plume in response to oil spills, such as in the Gulf of Mexico or the Eastern Mediterranean Sea. Those works highlighted the tremendous enrichment of genes related to oil biodegradation, indicating a rapid microbial adaption to oil contamination, and the strong oil biodegradation potential of naturally occurring deep-sea microbial communities (Liu et al., 2017).

Despite their unquestionable usefulness, functional approaches are still constrained by the availability of genomic information, requiring a constant update to encompass the increasing flow of genomic and metagenomic data feeding databases. Moreover, most of the genomes currently available in repositories correspond to a limited number of isolated taxa that have traditionally attracted most of the scientific attention, which biases the results obtained from functional screens towards the detection of those groups. To overcome this limitation, the gap of genomic information needs to be filled by providing insights on still uncultured PAH-degrading microorganisms. Efforts should be directed to the isolation of bacteria detected by molecular methods using custom-made media to meet their physiological requirements; and to the application of -omic approaches (e.g., shotgun metagenomics) combined to isotopic labeling (e.g., nucleic acid-based

stable isotope probing (SIP) technique) can provide genomic insights into active PAH-degraders present in environmental samples circumventing the need of their isolation (Coyotzi et al., 2016).

SIP was developed in the 2000's, and has been increasingly applied during the last decade (Madsen, 2006). The method relies on the application of a stable isotope labelled tracer (e.g., ^{13}C-PAHs) to an environmental sample. After incubation, the microorganisms involved in the active assimilation of the tracer will have incorporated the stable isotopes in their macromolecules, and those ^{13}C-enriched macromolecules, mainly DNA or RNA, can be selectively recovered and purified from the natural ^{12}C-containing molecules by density-gradient centrifugation. Analysis of the ^{13}C-enriched fraction of nucleic acids by 16S rRNA gene-based community analysis methods or shotgun metagenomics provide valuable information on the major PAH degraders present in environmental samples, or the genomic mechanisms underlying their capabilities. Therefore, DNA-SIP has represented a decisive step further towards the linking of microbial communities to their functions. The high cost of ^{13}C-labeled PAHs has limited the use of SIP, with most studies focusing on naphthalene and phenanthrene. However, their application has provided novel insights into PAH-degrading communities both in marine environments and soils (Vila et al., 2015). For example, application of DNA-SIP revealed that previously overlooked groups of *Betaproteobacteria* (*Acidovorax*, *Rhodopherax*, *Pigmentiphaga* and *Hydrogenophaga*) can have a crucial role in the early response to PAH-exposure in contaminated industrial soils (Jones et al., 2011). Regarding marine environments, DNA-SIP identified the marine phytoplankton as a natural biotope for 2-ring and 3-ring PAH-degrading bacteria, guiding the directed isolation of novel PAH-degrading strains affiliated to the *Gammaproteobacteria* and *Bacteroidetes* (Gutierrez et al., 2011). The most comprehensive DNA-SIP work applied to PAH-biodegradation has been developed by the research group of Dr. Aitken at the University of North Carolina. Appling DNA-SIP to soils from gasworks sites in North Carolina, this group has produced results on additional substrates, such as the PAHs anthracene, fluoranthene, pyrene and benzo(*a*)anthracene, or the oxy-PAH anthraquinone. In addition to the key role of Betaproteobacteria (*Acidovorax*, *Pigmentiphaga*), these authors identified novel groups of uncultured anthracene- and pyrene-degrading *Proteobacteria* (Singleton et al., 2006; Jones et al., 2011). The abundance of these groups, and their response to biostimulation, was demonstrated by high throughput sequencing of 16S rRNA gene libraries during simulated bioremediation experiments, highlighting how culture-based studies minimize the true diversity of metabolic capabilities in complex contaminated systems. It is worth noting that, based on the information collected from this DNA-SIP approach, this work has been recently culminated with the isolation of two novel

pyrene-degrading *Proteobacteria*, *Rugosibacter aromaticivorans* (Corteselli et al., 2017a) and *Immundisolibacter cernigliae* (Corteselli et al., 2017b), which define a novel betaproteobacterial genus and a new gammaproteobacterial order, respectively. Based on the genomic information retrieved from those isolates, a major role during PAH biodegradation at contaminated sites has been attributed to phylogenetically close relatives in soils from France, the Czech Republic and Spain, suggesting the relevance and global distribution of these new groups of PAH-degrading bacteria (Duarte et al., 2017; Tauler et al., 2016).

Another interesting aspect highlighted by the application of microbial ecology methods to contaminated soils is the functional redundancy of environmental PAH-degrading communities (Louca et al., 2018). Applying DNA-SIP, Dunlevy et al. (2013) identified functional redundancy for anthracene degradation in the natural community found in aged PAH-contaminated soil, showing clearly distinctive microbial communities when comparing treated and untreated soil. A similar phenomenon had been previously observed in another soil sample for the degradation of the LMW PAH naphthalene (Guazzaroni et al., 2013). These works revealed how, within a contaminated environment, different microbial communities with the ability to degrade one particular compound can coexist. Their specific induction/repression may be thus dependent on the stimulation approach, and characterization of the communities present in the untreated sample may fail to accurately predict the most relevant organisms following a biological treatment scheme.
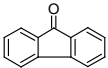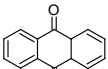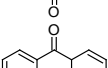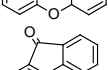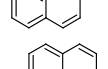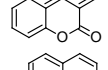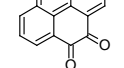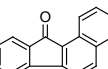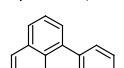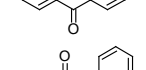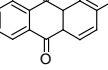
Oxygenated PAHs in PAH-contaminated soils

Sources, occurrence and environmental fate of oxy-PAHs

Oxygenated PAHs (oxy-PAHs) are partial oxidation products of PAHs with one or more carbonylic oxygen(s) attached to the aromatic ring structure, including polyaromatic ketones, quinones or lactones, that in some cases also contain alkyl and hydroxyl groups (Table 2). These compounds are constituents of fossil fuels and are emitted from the same sources as PAHs. However, they can also be readily formed as stable intermediates and dead-end products from PAHs by photooxidation, chemical oxidation or microbial transformation (Lundstedt et al., 2007). The contribution of each of these sources to the demonstrated presence of partially oxidized PAHs as co-contaminants in all the studied PAH polluted sites is not known and may be site specific. Indeed, many oxy-PAHs have been found at significant levels in PAH-contaminated soils with comparable concentrations to PAHs (Lundstedt et al., 2014; Larsson et al., 2018).

Table 2. Structure and relevant physicochemical and toxicological properties of representative oxy-PAHs.

Name	Chemical structure	Water sol.	Vapor pressure	Log K_{ow}	Tox.	Carc.[a]	Gen.
Indan-1-one		1430	3.9	2.11	n.d.		
Fluoren-9-one		25.3	0.076	3.58	n.d.		
Anthracene-9,10-dione		1.35	0.000015	3.39	>5000	2B	
Xanthenone		4.52	0.0078	3.39	>500		
Cyclopenta[*def*]phenanthrene-4-one		1.80	0.001	3.81	n.d.		
4-oxapyrene-5-one		2.32	0.00053	2.58	n.d.		+[b]
Pyrene-4,5-dione		0.663	0.00036	4.15	n.d.		
Benzo[*a*]fluorenone		0.0216	0.00005	4.73	n.d.		+[c]
*7H-*benzo[*de*]anthracene-7-one		0.240	0.00003	4.81	n.d.		+[b]
Benz[*a*]anthracene-7,12-dione		0.0289	0.000005	4.40	n.d.		+[c]
Naphthacene-5,12-dione		0.0228	0.000005	4.52	n.d.		+[c]
Benzo[*cd*]pyrene-6-one		0.005	0.000002	5.31	n.d.		+[c]

[a] International Agency for Research on Cancer (IARC) Classification (1: carcinogenic to humans, 2A: probably carcinogenic to humans, 2B: possibly carcinogenic to humans, 3: not classifiable as carcinogenic to humans)

[b] Identified in genotoxic/mutagenic soil extracts post-bioremediation in Titaley et al., 2020.

[c] Identified as potential genotoxic in Clergé et al., 2019.

n.d. = no data available

Water solubility (mg/L) and vapor pressure (Pa) are at 25ºC. Toxicity is reflected as LD50 (mg/kg) for oral ingestion by rats. Source: Achten and Andersson, 2015.

Because oxy-PAHs contain oxygen atoms, they are more polar than unsubstituted PAHs, consequently, they are more water soluble and less lipophilic than the PAHs with the same number of aromatic rings (Table 2). Thus, oxy-PAHs have higher mobility and bioavailability in the environment than PAHs. However, they have been described to be persistent in soils as some works have demonstrated their accumulation during enhanced PAH degradation (Lundstedt et al., 2007). As discussed above, oxy-PAHs can be produced as final products of branched metabolic pathways or by cometabolic reactions in PAH-mixtures and become dead-end products whose further utilization has not been described for most of the studied PAH-degrading strains (Figure 5). The type of initial attack and extent of oxidation of one substrate could be conditioned by the presence and abundance of others, thus determining if it is channeled through productive pathways and mineralized, or partially oxidized. Nevertheless, those partially oxidized products could be important metabolic nodes in environmental networks. It is to be expected that dead-end products for one strain could be substrates for others leading to cross-feeding networks and metabolic interactions.

Actually, some studies have reported the biodegradation of oxy-PAHs in soils after their transient accumulation during biodegradation of PAHs (Lundstedt et al., 2003; Wilcke et al., 2014). However, the microbial populations and mechanisms driving oxy-PAHs degradation is a key issue that has been seldom addressed. PAH-degrading mycobacteria were the only organisms that had been investigated until recently for metabolizing a PAH-quinone, 4,5-pyrenequinone, and, in this case, the quinone underwent reductive reactions to form the corresponding diol, that could be channeled again through the pyrene degradation pathway (Kim et al., 2004). Conversely, the application of DNA-SIP to a PAH-contaminated soil revealed that the bacterial populations associated to anthracene degradation before and after treatment in a slurry bioreactor differed from those responsible for anthraquinone degradation, suggesting the existence of a specialized quinone-degrading community (Rodgers-Vieira et al., 2015). This opens the door to the important role that cometabolism may play in channeling single PAHs through environmental cooperative metabolic networks involving different interacting microbial populations that use oxy-PAHs as catabolic nodes.

Genotoxic and mutagenic risks of oxy-PAHs

Although knowledge on the toxicity of oxy-PAHs is not as extensive as for PAHs, there is increasing evidence for their (geno)toxic, mutagenic and carcinogenic properties, in some cases even presenting greater toxicity than their parent compounds (Clergé et al., 2019; McCarrick et al., 2019). In fact, oxy-PAHs have been suggested to be the causative agents for the eventual increase in genotoxicity

observed in PAH-contaminated soils after biological treatment (Titaley et al., 2020). Moreover, a recent study applying non-target analysis combined with tools of metabolomics identified a genotoxic lactone resulting from the dioxygenation of pyrene on an atypical ring (13% of the initial PAH) as a major contributor to the genotoxicity observed in a PAH-contaminated soil after treatment in a bioreactor (Tian et al., 2017). This product was later identified in two other creosote-polluted soils after biostimulation, and its microbial origin was confirmed by metabolomic screens of pyrene-cultures from two widespread pyrene-degrading strains. Since oxy-PAHs may be more toxic than the parent PAHs and they have a greater bioavailability and mobility to exert their toxic effect, their potential contribution to risk needs to be taken into account. To minimize the risk associated to the accumulation of oxy-PAHs during bioremediation, further research efforts should address the microbial processes and interactions underlying their *in situ* formation and further possible reutilization, and the factors that modulate those processes.

Background of this Thesis

The investigations gathered in this PhD Thesis are a continuation of previous work conducted by the Research Group on Biodegradation and Bioremediation of the University of Barcelona on a creosote-contaminated soil (Tauler, 2015). This soil was sampled from a former wood-treatment facility in Andújar (Jaén, Spain) with a record of creosote contamination over 100 years. The soil, selected as a model PAH-contaminated environmental matrix, presented an extremely high PAH concentration, about 30,000 ppm of total 16 EPA PAHs, and the existence of a well-adapted PAH-degrading community. In order to determine the microbial processes driving PAH removal *in situ*, a biostimulation treatment was performed on this soil consisting of laboratory-scale dynamic biopiles amended with nutrients (300:10:1 C:N:P ratio) and adjusted water content (40% water holding capacity). After 150 days of incubation, extensive removal of PAHs was observed (93% degradation). However, a residual fraction (approximately 700 ppm of PAHs) especially enriched in HMW-PAHs remained in the soil. The biodegradation of PAHs was sequential as LMW-PAHs rapidly disappeared in the first place followed by the depletion of HMW-PAHs (Figure 6). Degradation kinetics of 4-ring PAHs grouped in highly similar pairs suggesting they were attacked by the same microbial populations.

Concomitant with the highest PAH-degradation rates there was a transient peak of accumulation of PAH-transformation products, which were later removed by the
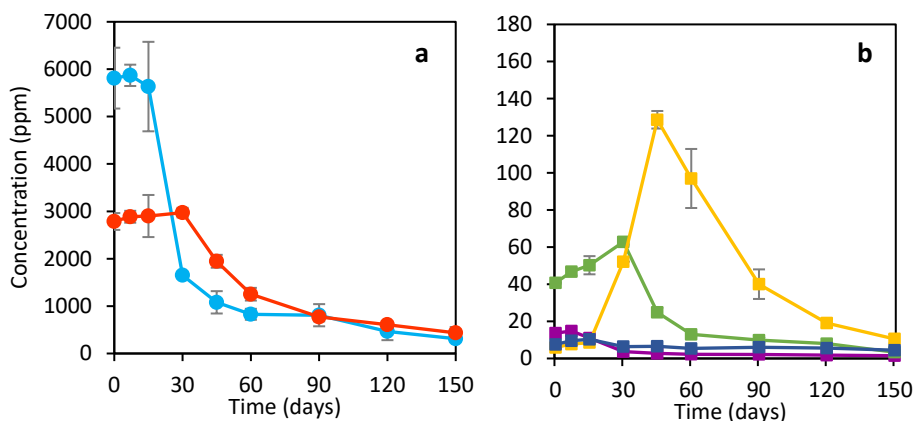
Figure 6. Concentration (mg·kg dry soil⁻¹) in the biostimulation assay of (a) LMW-, in blue, and HMW-PAHs, in red; and (b) selected oxy-PAHs, i.e. 9-fluorenone, in purple; 9,10-anthraquinone, in green; cyclopenta[*def*]phenanthrene-4-one, in yellow; benzo[*a*]fluorenone, in dark blue.

microbial populations present in the soil (Figure 6). The initial concentration of oxy-PAHs was of approximately 93 ppm. Some oxy-PAHs were removed in the course of the incubation, while others transiently accumulated and were later degraded. For example, the concentration of anthracene-9,10-dione, the most abundant oxy-PAH at the beginning of the incubation (41 ppm), increased during the first 30 days and subsequently dropped. Conversely, oxy-PAHs accumulated in the dry controls, where no microbial activity was detected, reaching a total concentration of 491 ppm. This highlights the existence of microbial populations with the ability to attack oxy-PAHs embedded in the soil networks that need the proper conditions to exert their function.

The degradation kinetics of PAHs, oxy-PAHs and N-PACs, together with the formation and/or accumulation of acidic products were correlated with key phylotypes and community shifts. A real-time insight into the community dynamics was obtained from the combined analysis of changes in global (genes) and active (transcripts) microbial communities, both at the phylogenetic (16S rRNA) and functional (RHD-coding genes) level. The bacterial key players involved in the initial removal of LMW-PAHs was attributed to members of *Pseudomonas*, *Pseudoxanthomonas*, *Achromobacter* and *Olivibacter*, while the disappearance of HMW-PAHs was associated to uncultured members of *Gammaproteobacteria*, and members of *Sphingobium* and *Mycobacterium*. However, the specific populations and functions driving the biodegradation of each particular PAC remained unclear as different populations were simultaneously attacking several compounds.

Framework of this Thesis

This PhD Thesis has been developed in the frame of three research projects focused on the biodegradation of organic contaminants in soils: "Functional synergies between microorganisms and plants in sustainable remediation of soils contaminated by PAHs" (CGL2013-44554-R), "From bioavailability science to soil remediation: Sustainable stimulation of biological networks for enhanced pollutant carbon turnover" (CGL2016-77497-R), and "Towards a risk-based modulation of carbon fluxes involved in the biological removal of organic contaminants in soils: microbial networks" (PID2019-109700RB-C22). These projects were funded by the Spanish Government and were conducted in cooperation with the research group of Dr. José Julio Ortega-Calvo from IRNAS-CSIC (Sevilla). As part of this collaboration, a joint experiment was conducted, in which the PhD candidate performed the microbial community analysis, and culminated in a manuscript published in Science of the Total Environment in 2020 (Annex).

The PhD candidate also contributed to a book chapter on the state of the art of PAH biodegradation in the recently published 3rd edition of Comprehensive Biotechnology (Annex). Part of this book chapter has been included in the general Introduction of this Thesis.

The DNA-stable isotope probing experiment in Chapter 1 was conducted in collaboration with the research group lead by Prof. Dr. Michael D. Aitken, at the Department of Environmental Science and Engineering from the University of North Carolina (Chapel Hill, USA). Metagenomic analysis of the ^{13}C-labeled DNA was performed by the PhD candidate at the University of Barcelona.

The purification of metabolites and NMR analysis in Chapter 2 were performed in collaboration with Prof. Dr. Cristina Minguillón, from the Department of Nutrition, Food Science and Gastronomy, Faculty of Pharmacy and Food Science, of the University of Barcelona.

Part of the genomic and transcriptomic analysis included in Chapter 2 and the metagenomic analysis included in Chapter 3 were initiated during a 3-month stay in Wageningen University & Research (The Netherlands). The stay was conducted under the supervision of Prof. Dr. Hauke Smidt, head of the Microbial Ecology Group at the Laboratory of Microbiology, Department of Agrotechnology and Food Sciences. The title of the project was "Omics for the understanding of the microbial biodegradation of polycyclic aromatic compounds". The stay was funded by the Spanish Ministry of Science, Innovation and Universities.

Objectives

Objectives

The general objective of this Thesis is to expand the knowledge on the metabolic networks that drive the biodegradation of polycyclic aromatic hydrocarbons in contaminated soils. The ultimate goal is to address the current limitations of bioremediation technologies, by determining the mechanisms involved in the removal of high molecular weight (HMW)-PAHs, often presenting high endpoint concentrations after bioremediation, and by ascertaining the formation and fate of generally overlooked partially oxidized products (oxy-PAHs). For this, a creosote-contaminated soil with a well-established PAH-degrading microbial community is used as a model. The integration of the information gathered from metabolomic, genomic and transcriptomic approaches on single strains, bacterial consortia and environmental communities, will shed light on the complex PAH-degrading microbial networks present in the soil. The starting hypothesis of this Thesis is that cometabolic interactions and partial oxidation processes can play an essential role in funneling PAH carbon fluxes, with oxy-PAHs acting as potential catabolic nodes. Within this global aim, the specific objectives to be developed in this Thesis are:

I. To unravel the microbial populations and functions involved in the biodegradation of the 4-ring PAH benz(*a*)anthracene. Unequivocal identification and functional analysis of B*a*A-degrading bacteria will be accomplished by combining DNA-SIP with shotgun metagenomics of ^{13}C-labeled DNA. Eventual substrate interactions with other HMW-PAHs will be ascertained by integrating degradation kinetics, metabolomic screens and 16S rRNA amplicon metagenomic sequencing in soil microcosms spiked with B*a*A and fluoranthene, pyrene or chrysene in binary mixtures.

II. To elucidate the mechanisms driving oxy-PAH biodegradation on the basis of the directed isolation of a novel 9,10-anthraquinone-degrading bacterial strain. The ANTQ metabolic pathway will be reconstructed by a polyomic approach involving the metabolomic, genomic and transcriptomic characterization of the isolate. The environmental relevance of the strain and the identified degradative mechanisms will be evaluated by qPCR analysis during biostimulation of the creosote-contaminated soil.

III. To establish if distinctive microbial populations are responsible for the degradation of B*a*A and its ready oxidation product, benz(*a*)anthracene-7,12-dione. The B*a*AQ-degrading subpopulations will be enriched to establish a stable microbial consortium. Metabolomic and metagenomic functional gene analyses will serve to determine the B*a*AQ metabolic pathway and to give insight into the potential roles and interactions within the consortium members.

Chapter 1

Bacterial benz(a)anthracene catabolic networks in soils are influenced by HMW-PAHs as co-substrates

Bacterial benz(*a*)anthracene catabolic networks in soils are influenced by HMW-PAHs as co-substrates

Introduction

Bioremediation is a widely applied technology for the decontamination of soils impacted by polycyclic aromatic hydrocarbons (PAH). However, its success is often constrained by end-point residues enriched in high molecular weight (HMW) PAHs exceeding regulatory limits. HMW PAHs, with 4 or more fused benzene rings, are of special concern due to their genotoxic, mutagenic and carcinogenic properties (Ghosal et al. 2016) and their potential bioaccumulation (van der Oost et al. 2003). These compounds present a very low water solubility and high octanol/water partition coefficients that limit their bioavailability. These physicochemical properties, along with their structural complexity and strong molecular bonds, make them very resistant to extensive microbial degradation in soils, thus contributing to their long persistence in the environment (Kanaly & Harayama, 2000). Despite a number of soil bacterial strains have the capacity to mineralize four-ring PAHs (Kanaly & Harayama, 2010), biodegradation of HMW PAHs in polluted soils is assumed to be accomplished by the cooperation of different microbial populations undertaking interconnected metabolic pathways (Vila et al., 2015). Recent studies identified previously unknown metabolites of HMW-PAHs in contaminated soils after bioremediation, highlighting that the current knowledge of PAH biodegradation pathways gathered from the study of isolated strains on single compounds can be limited (Tian et al., 2017, 2018). Furthermore, PAHs are constituents of complex mixtures in contaminated soils, such as coal and crude oil derivatives. As a result, interactions between components within these mixtures can be crucial to determine the fate of PAHs, in particular of HMW PAH. Substrate interactions have been observed during the biodegradation of PAH mixtures by pure cultures and defined consortia (Stringfellow and Aitken, 1995; Mueller et al. 1989). However, little is known about the potential interactions occurring during biodegradation of PAH mixtures by complex soil microbial communities, where microbial synergism and/or competition may occur, thus influencing the final composition of the endpoint residue enriched in HMW-PAHs.

The bacterial biodegradation of benz(*a*)anthracene (B*a*A) has been basically addressed with single strains, with a special focus on sphingomonads (Jerina et al., 1984; Sohn et al., 2004; Baboshin et al., 2008; Kunihiro et al., 2013; Jouanneau et al., 2016) and mycobacteria (Moody et al., 2005; Kim et al., 2006). Most of these isolates were not able to grow on B*a*A as sole carbon and energy source and could

only transform it due to the relaxed specificity of their aromatic ring-hydroxylating dioxygenase (RHD) and monooxygenase enzymes. As a result of this enzymatic promiscuity, the biotransformation of B*a*A has been reported at multiple positions within a single strain. For instance, the oxidation of B*a*A by *Sphingobium* sp. KK22 occurred via both the angular kata and linear kata ends of the molecule, at the 1,2-, 3,4-, 8,9- and 10,11- positions (Kunihiro et al., 2013). *Mycobacterium vanbaalenii* PYR-1 preferentially attacked B*a*A at the 10,11- positions but could also transform it at the 1,2- and 5,6- positions, giving the corresponding diols, and at the 7,12- positions, producing benz(*a*)anthracene-7,12-dione (Moody et al. 2005, *see* Annex – Figure S1.1). However, information on the microbial populations and mechanisms that modulate the transformation and eventual mineralization of B*a*A *in situ* in contaminated soils is scarce.

In this study, we aimed to unravel the microbial populations and functions involved in the biodegradation of benz(*a*)anthracene in a historically creosote-contaminated soil from the South of Spain. Unequivocal identification and functional analysis of B*a*A-degrading bacteria was accomplished by combining DNA-SIP, using U^{13}C-benz(*a*)anthracene, with shotgun metagenomics of ^{13}C-labeled DNA. On the other hand, soil microcosm incubations were conducted spiking B*a*A and three other HMW-PAHs (fluoranthene, pyrene and chrysene) individually, or B*a*A in binary mixtures with those. By integrating biodegradation kinetics, metabolomic screens and 16S rRNA amplicon metagenomic sequencing data, we ascertained eventual substrate interactions that were associated to changes in microbial populations.

Materials and methods

Chemicals

Benz(*a*)anthracene (B*a*A, 99%), fluoranthene (FT, >97%), pyrene (PY, 98%) and chrysene (CHY, 98%) were purchased from Sigma-Aldrich Chemie (Steinheim, Germany). Solvents were obtained from J.T. Baker (Deventer, The Netherlands). Inorganic salts and other chemicals were acquired from Panreac Química S.A.U (Barcelona, Spain) or Merck (Darmstad, Germany). Solvents (organic residue analysis grade) were obtained from J.T. Baker (Deventer, The Netherlands). All chemicals and solvents were of the highest purity available. Uniformly ^{13}C-labelled (U-^{13}C)-B*a*A was synthesized by methods described elsewhere (Zhang et al., 2011) and kindly provided by Z. Zhang and A. Gold.

Soil

The creosote contaminated soil used in this study resulted from a lab-scale biostimulation treatment (Tauler, 2015) of a heavily contaminated soil from a historical wood-treating facility in southern Spain. The soil was treated by aeration, adjustment of the water content to 40% of its water-holding capacity (WHC) and nutrient addition (urea and K_2HPO_4 to a C:N:P ratio of 300:10:1). After 150 days of incubation, the final concentration of Σ17 PAHs was 747 ± 95 mg/kg dry soil (Annex – Table S1.1). The remaining treated soil was kept at 4ºC until it was used here.

DNA-Stable Isotope Probing (SIP) using U^{13}C-benz(*a*)anthracene

Incubations were conducted in 40-ml amber glass EPA vials with Teflon-lined stoppers containing 1 g of treated creosote-contaminated soil previously spiked with B*a*A as described by others (Jones et al., 2008). Briefly, 20% of the total air-dried soil was added to a glass beaker and spiked with a B*a*A solution (10 mg/mL in acetone). The content was mixed manually for 1 min using a clean stainless-steel spatula. The beaker was covered loosely with aluminum foil and placed in a fume hood for 24 h for solvent evaporation. The remaining 80% of the air-dried soil was subsequently mixed into the beaker in 20% increments for 30 s per increment. The final concentration of added B*a*A in the soil was 1 mg/g. Two replicate vials were prepared for each incubation time point and the whole content was used for residual B*a*A quantification. Two replicate flasks prepared in the same manner were acidified with phosphoric acid and used as killed controls to account for abiotic B*a*A losses at the end of the incubation. Soil moisture was adjusted to 40% WHC using sterile water supplemented with urea and K_2HPO_4 to meet a 300:10:1 C:N:P molar ratio relative to B*a*A carbon. Two additional sets of replicates were prepared in the same manner with unlabeled or [U-^{13}C]-benz(*a*)anthracene for SIP incubations that were incubated in parallel for the duration of the experiment. All vials were incubated static at 25ºC in the dark.

Duplicate soil incubations were sacrificed at each incubation time to determine residual B*a*A concentration, and results were used to determine the endpoint for the SIP experiment. 0.25 g of soil were transferred to a 10 mL conical glass tube with Teflon-line screw cap and mixed with an equal amount of anhydrous sodium sulphate. Extraction was performed with 2 mL of a 1:1 DCM:acetone mixture by vortexing for 1 min at maximum speed, soil decantation, and vortexing again for and additional minute. Solvent was recovered after centrifugation at 1500 rpm for 5 min and transferred to a 10 mL volumetric flask. The extraction process was repeated 4 times. The collected extracts were brought up to a final volume of 10

mL, and stored at 4 °C until further analysis. Prior to analysis, samples were diluted in methanol (1:10) and filtered through a nylon syringe filter (0.22 μm). BaA quantification was achieved by HPLC with fluorescence detection as described elsewhere (Richardson et al., 2011).

DNA extraction, isopycnic separation and 16S rRNA clone libraries

Soil from incubations in the presence of unlabeled or ^{13}C-labeled BaA was divided into portions containing approximately 250 mg (dry weight). Each 250-mg aliquot of soil was extracted using the DNeasy PowerSoil Kit (Qiagen, USA). DNA extracts from each 250-mg aliquot of a given replicate were eluted in Tris-EDTA (TE; 10 mM Tris-HCl, 1 mM EDTA; pH = 8.0) buffer and combined to give a final volume of 500 μL. DNA was mixed with 5.2 ml of a cesium chloride solution in TE (1.81-1.82 g/mL) in a 6-ml polyallomer Ultracrimp tube (Kendro Laboratory Products, Newtown,CT). The final filling volume of the tube was 5.7 mL with a density of 1.73-1.74 g/mL. Tubes were crimp-sealed and ultracentrifuged in a Sorvall (Newtown, CT) RC70 ultracentrifuge using a TV-1665 rotor (Sorvall) at 145,200 g for 48 hours at 20ºC. After centrifugation, 24 fractions of 250 μL each were collected from the bottom of each tube. For each fraction, CsCl density was determined using a refractometer AS200 (Reichert Technologies) in relation to a six-point calibration standard, total DNA was purified by ethanol precipitation (Green & Sambrook, 2016) and quantified using a Nano-Drop 3300 fluorospectrometer (NanoDrop Products, Wilmington, DE) with the Quant-iT PicoGreen double-stranded DNA (dsDNA) assay kit (Invitrogen, Eugene, OR).

To identify fractions containing amplifiable DNA and identify shifts in the community DNA resulting from isotopic enrichment, PCR was performed with 16S rRNA gene primers 341FGC/517R. PCR was done with a Mastercycler gradient thermocycler (Eppendorf, Hauppauge, NY) using GoTaq Master Mix (Promega, Madison, WI) in a 25 μL reaction mixture. PCR products were visualized in 1.5% agarose gels and used for denaturing gradient gel electrophoresis (DGGE) analysis in 10% acrylamide gels with a linear denaturing gradient of 35 to 60% on a DCode universal mutation detection system (Bio-Rad Laboratories, Hercules, CA). Fractions that corresponded to heavy DNA in extracts from each flask containing ^{13}C-labeled substrate were identified based on total DNA concentration and DGGE analysis.

The DNA from ultracentrifuge fractions containing heavy DNA from a given replicate incubation was pooled to generate a 16S rRNA gene clone library for that replicate. DNA from the corresponding fractions in replicate incubations with unlabeled benz(a)anthracene were also pooled and used as controls to assess the

absence of amplification. PCR was performed using 1 μL of pooled DNA as a template, primers 27F and 1492R (Weisburg et al., 1991) and the GoTaq Master Mix in a 25 μL reaction. Clone libraries were produced from the amplified products using the TOPO TA cloning kit for sequencing (Invitrogen, Carlsbad, CA). 50 clones from each replicate library were picked at random and the plasmids were purified with the GenJET Plasmid Miniprep kit (Thermo Scientific). All inserts were partially sequenced using universal primer 27F. Selected clones representing clusters of highly similar sequences were also sequenced with universal primer 1492R (Genewiz, Research Triangle Park, NC).

Shotgun metagenomic sequencing of ^{13}C-labeled DNA and analysis

Metagenome sequencing of ^{13}C-labeled DNA recovered from pooled fractions containing heavy DNA of one replicate SIP incubation was performed using an Illumina PE150 NovaSeq 6000 platform at Novogene Co. (Cambridge, UK). Adapters removal and quality filtering was done on the received paired-end raw reads using Trimmomatic v.0.38 (ILLUMINACLIP:TrueSeq3-PE:2:30:10:8, LEADING:3 TRAILING:3, MINLEN:50, SLIDINGWINDOW:4:20, AVGQUAL:20). Raw read and trimmed read data quality was checked using FastQC (v0.11.8). Trimmed reads were assembled using metaSPAdes v3.9.0 (-meta option) and testing kmer sizes of 21, 33 and 55. Assembled scaffolds were clustered into bins with MaxBin2 v.2.2.7 using trimmed reads for mapping and a minimum contig length of 1000 bp. Relative abundance of each bin was inferred by mapping reads to each bin with Bowtie2 v.2.4.2. Bin quality assessment was performed with CheckM v.1.0.18 for contamination and completeness and the bins were referred to as metagenome-assembled genomes (MAGs). Taxonomic classification of the MAGs was done using the Genome Taxonomy Database toolkit GTDB-Tk v.1.1.0. Functional annotation of the most abundant and complete MAGs was performed using Prokka v.1.14.6. The GhostKOALA annotation server was used to reconstruct metabolic pathways.

B*a*A biodegradation in soil microcosms in the presence of HMW-PAHs as co-substrates

The treated creosote polluted soil was used as an inoculum (10% w/w) in soil microcosms containing sterile agricultural soil spiked either with B*a*A, fluoranthene (FT), pyrene (PY) or chrysene (CHY) individually, or in mixtures of B*a*A with the remaining PAHs as cosubstrates (B*a*A+FT, B*a*A+PY or B*a*A+CHY). The final concentration of each PAH in the microcosms was 200 ppm. For each condition, triplicate microcosms were set up consisting of hermetic glass jars containing 70 g of soil with the humidity adjusted to a 40% of the soil WHC. Microcosms (27 in total) were incubated for 60 days at 25ºC in the dark. Before

set-up, the agricultural soil was homogenized by grinding, sieving through a 2 mm mesh and mixing, and was submitted to four cycles of autoclave sterilization (45 min, 121ºC, 1 bar of overpressure) conducted in 24h intervals. Effective soil sterilization was verified by negative results in the quantification of heterotrophic microbial populations by MPN (data not shown). After 0, 15, 30, 45 and 60 days, samples were collected from each jar: 5 g were stored at -20ºC for chemical analysis; and 1 g was stored at -80ºC for molecular analysis. Before sampling, jars were weighed to check if water content remained constant and soil was homogenized to ensure the collection of a representative sample.

Analysis of PAHs and PAH-transformation products

2 grams of selected soil samples (0, 15, 30, 45 and 60 days) stored at -20ºC were mixed with 2 g of Na_2SO_4 in a glass centrifuge tube. 200 μL of an *ortho*-terphenyl solution (1 mg/mL) were spiked as internal standard to each tube, and 10 mL of a DCM:acetone (2:1) mixture were added as extraction solvent. After proper mixing, tubes were sonicated in an ultrasonic bath for 10 minutes and centrifuged for 20 minutes at 1,200 rpm. Solvent was recovered and filtered to remove residual soil particles (Whatman Grade No.1, GE Healthcare, UK). This extraction process was repeated 5 times. Soil extracts were concentrated with a rotatory evaporator to a final volume of 1 mL and stored at -20ºC for further analysis. Extracts were fractionated using BakerBond® solid phase extraction columns (JT Baker) packed with 2.5 g of silica, previously activated at 125ºC for 24 h and deactivated with water (5%). Two fractions were eluted; F1, was eluted with 13 mL of a hexane:dichloromethane (80:20) mixture and 2 mL of pure dichloromethane, containing PAHs; F2, was eluted using 1 mL of DCM followed by 6 mL of methanol, and contained polar compounds. All were concentrated by rotatory evaporation, transferred to 1 mL of DCM and stored at -20ºC. The extracted fractions were analyzed by GC-FID on a Shimadzu GC 2010 gas chromatograph (Shimadzu Corporation, Kyoto, Japan), equipped with a flame ionization detector and a DB-5 column (30 m, 0.25 mm, 0.25 μm, Agilent Technologies). Quantification was performed by using a five-point self-prepared standard calibration curve consisting of B*a*A, FT, PY, CHY, *o*-terphenyl, and B*a*AQ.

Nucleic acids extraction from soil microcosm samples and reverse-transcription

Total DNA and RNA were extracted from two separate soil aliquots from the microcosms. DNA was extracted from an aliquot of 0.25 g using the DNeasy PowerSoil Kit (Qiagen, USA). For RNA extraction, 0.5 g of soil were processed using the RNeasy PowerMicrobiome Kit (Qiagen, USA). Prior to analysis, RNA samples

were treated with DNase I RNase-free (TermoFisher, USA) in a total reaction of 10 µL including 8 µL of RNA extract, 1 µL DNase (1U) and 1 µL buffer. 1 µL of EDTA 50mM was added to inactivate DNase when the reaction was completed. The absence of contaminant DNA was confirmed by PCR using universal 16S rRNA primers (27F and 1492R). RNA was subsequently reverse-transcribed to cDNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA).

16S rRNA Amplicon Metagenomic Sequencing

Paired end reads (2x250bp) of the V4 region of the 16S rRNA from triplicate DNA samples of selected data points (0, 15 and 30 days) were obtained using Illumina NovaSeq equipment and reagents at Novogene Co. (Cambridge, UK). The received raw data was processed using mothur 1.45.3 following the MiSeq standard operating procedure (Kozich et al., 2013, last accessed June 2021). The MiSeq libraries produced a total of 3,762,443 quality sequences with an average of 62,009 reads per sample. Sequences were aligned using SILVA v138 reference files and clustered into operational taxonomic units (OTUs) using a 99% sequence identity cutoff. The OTU taxonomic affiliations were assigned to each representative sequence using the RDP (Ribosomal Database Project) database v9. Rarefaction curves, library coverages and α-diversity indexes (Simpson's inverse index) were estimated normalizing the number of reads/sample to that of the least represented sample (52,825 reads). Representative sequences for each OTU were obtained using the get.oturep command in mothur. Linear discriminant analysis Effect Size (LEfSe) was performed on the Galaxy server with default settings to determine the OTUs most likely to explain differences between incubation conditions (Segata et al., 2011).

Quantitative PCR (qPCR) amplification

The quantitative evolution of the bacterial community present in the microcosms was monitored by real-time PCR (qPCR). Total bacterial populations (DNA) and total active bacterial populations (cDNA) were quantified by amplification of 16S rRNA gene using universal primers 341F and 534R. The most relevant bacterial populations identified during the metagenomic 16S rRNA amplicon sequencing analysis (see below) were monitored by qPCR using specific primers targeting their 16S rRNA gene sequence. The selected phylotypes were *Immundisolibacter*, *Pseudoxanthomonas*, *Mycobacterium* and *Sphingobium*. qPCR was conducted on an Applied Biosystems StepOne Real-Time PCR system (Applied Biosystems, USA) and further analyzed using the StepOne Software v.2.3. Reactions with a total volume of 20 µL were prepared in 96 well plates, containing PowerUp SYBR Green

Master Mix (Applied Biosystems, CA, USA), 1 μL of template and 4 pmol of each primer. The primer pairs and conditions used for each determination are described in Table 1.1. Primer specificities were validated *in silico* by using NCBI Primer-BLAST and tested experimentally by PCR amplification on DNA and cDNA samples containing the target sequence. Quantification was achieved against a standard calibration curve of 8 points. Standards were obtained by cloning almost full length 16S rRNA gene fragments using a pGEM-T Easy Vector System (Promega, Madison, USA). For total bacteria quantifications a fragment of the 16S rRNA gene of *E. coli* CECT 101 was used. For reactions targeting specific phylotypes, we used plasmids recovered from clone library analyses performed in previous experiments. All qPCR reactions had an efficiency ranging between 85% and 110% and a slope between -3.2 and -3.8. Primer specificity was also assessed by the observation of a single peak during melt curve analysis.

Table 1.1. Characteristics of qPCR primer sets used in this study.

Primer	Sequence (5' > 3')	Annealing temperature	Reference
341F	CCTACGGGAGGCAGCAG	55ºC	Muyzer et al., 1993
534R	ATTACCGCGGCTGCTGG		
Immundi F	AATACCGCATACGCCCTACG	60ºC	This study
Immundi R	GGCTTGGTGAGCCATTACCT		
Pxanth F	AGGCCTCACGCGATTAGATG	60ºC	This study
Pxanth R	CTGGACCGTGTCTCAGTTCC		
MycoPYR F	ACTGAGATACGGCCCAGACT	60ºC	This study
MycoPYR R	TCACGAACAACGCGACAAAC		
SphFT F	CTGGGGTGCATGGCATTTCA	60ºC	This study
SphFT R	GAAAGTGTCTCCACGATCCG		

Statistical analysis

Data was subjected to analysis of variance (ANOVA) and Tukey *post-hoc* multiple comparisons test on the IBM SPSS Statistics v.23 software, with mean differences significant at $P<0.05$.

Results and discussion

Insight into the B*a*A-degrading community by DNA-SIP combined with shotgun metagenomics

Unequivocal identification of microbial communities and functions involved in benz(*a*)anthracene assimilation in the creosote contaminated soil was achieved by
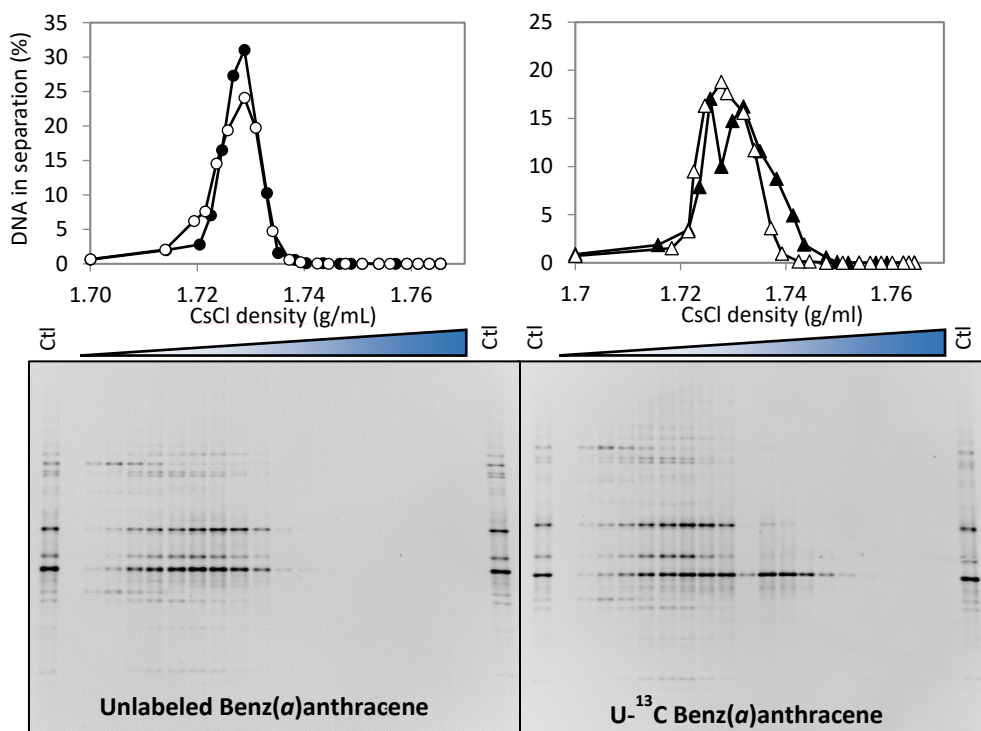
Figure 1.1. Amount of DNA and CsCl density (g/L) of the collected fractions after isopycnic centrifugation of duplicate samples from SIP incubations with unlabeled and U-13C B*a*A (top) and DGGE profiles of PCR-amplified 16S rRNA gene fragments from the collected fractions (bottom).

DNA-stable isotope probing combined with shotgun metagenomic sequencing of 13C-labeled DNA. After effective B*a*A removal (35% of the initial concentration), total DNA was extracted from parallel incubations with unlabeled- or 13C-labeled-BaA and separated into density fractions by isopycnic centrifugation in CsCl gradients. Fractions containing 13C-enriched DNA were identified by DNA quantification and community shifts in DGGE analysis. DGGE patterns revealed a shift in heavier fractions (density ranging from 1.735 to 1.745 mg·L-1), which showed a noticeable enrichment of one particular band (Figure 1.1).

Clone libraries of partial bacterial 16S rRNA gene sequences from the isotopically-enriched heavy fractions were created for each of the duplicate 13C-labeled B*a*A incubations to determine degraders of that compound. The vast majority of sequences (84%) were related to a member of *Immundisolibacter* (Table 1.2). This indicates that most of the 13C from B*a*A was assimilated by this phylotype (BA1), thus becoming evident that this member of *Immundisolibacter* is the primary

Table 1.2. Analysis of the 16S rRNA sequences recovered in the clone library of ^{13}C-labeled DNA samples from SIP incubations.

Phylotype	Freq. (%)	RDP classifier[a]	Closest relative (Accession number, %)	Closest type strain (Accession number, %)
BA1	84%	*Immundisolibacter*	Uncultured bacterium clone SPN400-90day-75 (MF314620, 100%)	*Immundisolibacter cernigliae* TR3.2 (NR_156801, 99.4%)
BA2	10%	*Sphingobium*	*Sphingobium ummariense* RL-3 (NR_044171, 100%)	*Sphingobium ummariense* RL-3 (NR_044171, 100%)
BA3	4%	*Olivibacter*	*Olivibacter* sp. JCM 28640 (LC133671, 99.7%)	*Olivibacter soli* Gsoil 034 (NR_041503, 99.4%)
BA4	2%	*Achromobacter*	Betaproteobacteria bacterium S6_AA1 (KT907033, 99.8%)	*Achromobacter marplatensis* B2 (NR_116198, 99.6%)

[a] Genus assignation with a probability of 100%.

player in the utilization of B*a*A in the creosote contaminated soil. The recently isolated *Immundisolibacter cernigliae* (Corteselli et al., 2017), which defined a new gammaproteobacterial order, had been previously related to the degradation of anthracene, fluoranthene, pyrene and benz(*a*)anthracene using DNA-SIP (Jones et al., 2011). Other less abundant phylotypes were detected in the clone libraries, corresponding to members of *Sphingobium* (10%), *Olivibacter* (4%) and *Achromobacter* (2%). *Sphingobium* BA2 shared a high 16S rRNA gene sequence identity (>99%) with *Sphingobium* strains isolated from contaminated environments, for instance, the hexachlorocyclohexane-degrading *S. ummariense* RL-3 (Accession number: NR_044171; Singh & Lal, 2009), the bisphenol A-degrading *Sphingobium* sp. SO1a (AB453305; Matsumura et al., 2009) and *S. cloacae* LH227, able to grow on phenanthrene and biotransform other PAHs (AY151393; Bastiaens et al., 2000). This *Sphingobium* phylotype could also be utilizing B*a*A as a growth substrate, but in a minor extent than *Immundisolibacter* BA1. However, we recently identified some members of *Sphingobium* as key players in the degradation of oxy-PAHs in the creosote contaminated soil, including B*a*A ready oxidation product benz(*a*)anthracene-7,12-dione (B*a*AQ) (*see* Chapter 3). Considering that a part of B*a*A could be transformed into B*a*AQ by other members of the community through cometabolic reactions, it is plausible that *Sphingobium* BA2 could also be growing on this oxy-PAH. As for *Olivibacter* and *Achromobacter,* their low sequence abundance in the clone library and the 16S rRNA sequence similarity (>99%) to rhizospheric (NR_041503, KR055002) and plant endophytic (KF844050, KR055002) bacteria from pristine environments suggest they could be utilizing intermediate metabolites of B*a*A metabolism.

Shotgun metagenomic sequencing of ^{13}C-labeled DNA generated a total of 25.9 million quality-filtered paired-end reads that were assembled into 46,337 contigs (largest contig 102,172 bp, N50 1,874 bp) and binned into 13 metagenomic bins that resulted in 10 metagenome-assembled genomes (MAG) (Table 1.3). As expected, the most abundant bin corresponded to *Immundisolibacter*, most closely related to type strain *Immundisolibacter cernigliae* (GCF_001697225.1, 94.34% ANI). Further analysis of specific CDS revealed that bins 2, 3 and 4 corresponded to *Immundisolibacter*, thus, they were considered as part of the *Immundisolibacter* sp. MAG1. Other two less abundant MAGs included members of *Pseudoxanthomonas* (MAG2) and *Sphingobium* (MAG3), along many other incomplete and low-quality genomes (MAGs 4-10). No MAGs were related to either *Olivibacter* or *Achromobacter*, as would have been expected from the 16S rRNA clone library analysis. Nevertheless, partial 16S rRNA sequences recovered from the shotgun metagenomic sequencing included members of *Immundisolibacter*, *Sphingobium*, *Pseudoxanthomonas* and *Olivibacter*, showing a 100% identity to sequences retrieved either from the amplicon metagenomic sequencing from the soil microcosms or the clone library from the DNA-SIP incubations. As for *Achromobacter*, the low sequence abundance in the clone library and the absence in shotgun metagenomic data suggests the recovered 16S rRNA sequences were not representative of organisms capable of incorporating the labeled carbon source (Singleton et al., 2005).

Functional analysis of *Immundisolibacter* sp. BA1 metagenome-assembled genome

Considering its key role in the biodegradation of B*a*A, a metagenomic functional gene analysis of the *Immundisolibacter* sp. BA1 metagenome-assemble genome (MAG) was conducted. The annotated *Immundisolibacter* sp. BA1 MAG included 3,726 CDS (2,613 in bin_1, 227 in bin_2, 303 in bin_3 and 583 in bin_4), 48 tRNA genes and 4 rRNA genes, including a partial 16S rRNA gene sequence in bin_1 with a 100% identity to BA1. Functional gene annotation revealed a large repertoire of dioxygenases and monooxygenases with potential activity on aromatic compounds (Table 1.4). Genes coding for aromatic ring-hydroxylating dioxygenases (RHD) were grouped according to Prokka annotation and phylogenetic analysis of the amino acid sequences (Annex – Figure S1.2). The *Immundisolibacter* sp. BA1 MAG presented an astonishing vast number of aromatic RHD coding genes, including genes encoding for putative naphthalene/biphenyl/benzene dioxygenases, benzoate/p-cumate dioxygenases, anthranilate/terephthalate dioxygenases and carbazole dioxygenases. Consequently, it is difficult to predict which genes encode the enzymes responsible for catalyzing the biodegradation of B*a*A.

Table 1.3. Features for all metagenomic bins recovered from shotgun metagenomic sequencing of ^{13}C-DNA from B*a*A SIP incubation.

Bin	MAG	Genome size (bp)	GC %	Compl.[a] (%)	Contam.[a] (%)	Relative[b] abundance (%)	GTDB Classification	Closest bacterial genome (ANI [%])
1	1	2,762,357	68.7	93.5	0.6	72.7	*Immundisolibacter*	*Immundisolibacter cernigliae* GCF_001697225.1 (94.34)
2	1	242,610	66.1	0.0	0.0	4.0	n.m.	n.m.
3	1	296,480	65.2	0.0	0.0	3.4	n.m.	n.m.
4	1	648,267	61.9	0.0	0.0	3.4	n.m.	n.m.
5	2	3,588,785	67.5	94.5	3.1	4.8	*Pseudoxanthomonas*	*Pseudoxanthomonas spadix* GCF_000233915.3 (98.94)
6	3	7,419,625	64.9	96.1	28.6	3.3	*Sphingobium*	*Sphingobium ummariense* GCF_000447205.1 (97.97)
7	4	3,539,026	68.7	80.6	47.0	0.6	*Sphingobium*	n.m.
8	5	3,910,887	69.0	27.6	1.7	1.9	*Pseudomonas*	n.m.
9	6	4,479,184	65.5	59.0	26.9	0.7	*Sphingobium*	n.m.
10	7	3,197,202	73.9	64.7	32.8	0.3	*Gemmobacter*	n.m.
11	8	6,382,620	61.4	34.0	9.9	0.7	*Bradyrhizobium*	n.m.
12	9	2,142,401	69.4	44.2	10.3	0.2	*Sphingomonadaceae*	n.m.
13	10	5,943,145	65.2	65.5	38.6	0.6	*Bradyrhizobium*	n.m.

[a] Genome completeness and contamination were calculated with CheckM v.1.0.18

[b] Relative abundance was inferred by mapping paired-end reads to each bin with Bowtie2.

n.m. = no match

Table 1.4. Number of genes encoding dioxygenase and monooxygenase enzymes related to the metabolism of aromatic compounds and alkanes in the *Immundisolibacter* sp. MAG.

Aromatic Ring-Hydroxylating Dioxygenases

	α-subunit	β-subunit	Fer.	F.R.
Napthalene/biphenyl 2,3-/benzene 1,2- dioxygenase	13	12	4	
Benozate/p-Cumate 2,3- dioxygenase	20	18		1
Anthranilate/Terephthalate 1,2-dioxygenase	17	11	1	
Carbazole 1,9a-dioxygenase (*CarA*)	3	1		
Phthalate 4,5-dioxygenase (*pht2*)	1			

Aromatic Ring-Cleaving Dioxygenases

1,2-dihydroxynaphthalene dioxygenase (*nahC*)	4
Biphenyl-2,3-diol 1,2-dioxygenase (*bphC*)	4
Catechol 2,3-dioxygenase (*xylE*)	7
Protocatechuate 4,5-dioxygenase (*ligB*)	3

Monooxygenases

Baeyer-Villiger monooxygenase	5
Alkane 1-monooxygenase (*alkB*)	5
Cyt P450	1

Fer. = ferredoxin, F.R. = ferredoxin reductase

Gene names in brackets.

A recent study applied stable isotope-assisted metabolomics (SIAM) to the same creosote contaminated soil used here and identified B*a*A biotransformation products (Tian et al., 2018). In the soil, B*a*A biodegradation mainly followed the preferential metabolic pathway previously described for *Sphingobium yanoikuyae* strains B8/83 and B1 (Jerina et al., 1984; Mahaffey et al., 1988), initiated by dioxygenation at positions 1 and 2, followed by *meta* cleavage leading to the formation of 1-hydroxy-2-anthracenoic acid, an intermediate also observed in *Sphingobium* sp. KK22 (Kunihiro et al., 2013). The route would continue with a further oxidation to a dihydrodiol, reported in this study for the first time, followed by a dehydrogenation and ring cleavage to produce two different isomers of dihydroxynaphthalene dicarboxylic acid. Although this pathway was proposed during an active bioremediation process of the creosote contaminated soil, with the action of the entire soil microbial community, we could link it to *Immundisolibacter* sp. BA1 as it has been proven to be the primary B*a*A-degrading bacterium. Interestingly, the MAG possessed all of the genes necessary for downstream biotransformation of 1,2-dihydroxynaphthalene (*nahC*, *nahD* and *nahD*) and *cis*-2,3-dihydrobiphenyl-2,3-diol (*bphB*, *bphC* and *bphD*). The GhostKOALA annotation also showed that *Immundisolibacter* sp. holds the complete catechol *meta*-cleavage pathway (*xylEFGHIJ*), harboring 7 different genes

Immundisolibacter sp. BA1 MAG (bin 1)
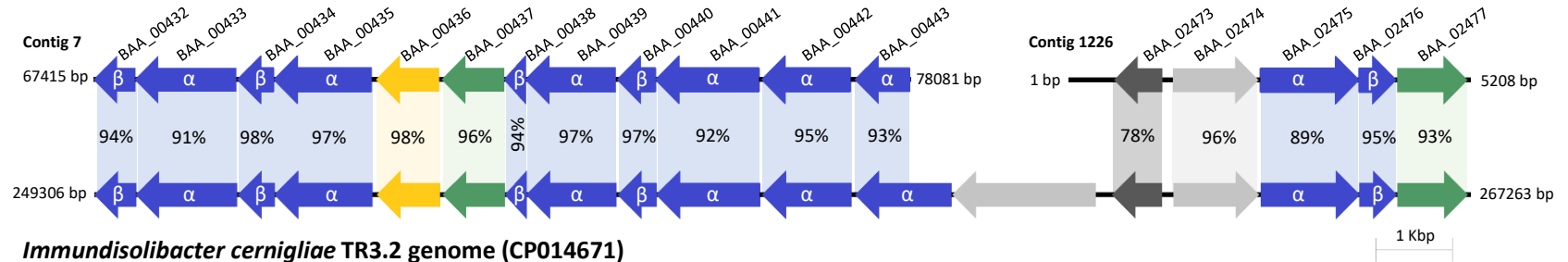
Immundisolibacter cernigliae TR3.2 genome (CP014671)

Locus tag	Contig	Start	End	Strand	Length (bp)	Annotation
BAA_00432	7	67415	67957	-	543	Biphenyl 2,3-dioxygenase subunit beta
BAA_00433	7	67954	69279	-	1326	Benzene 1,2-dioxygenase subunit alpha
BAA_00434	7	69291	69770	-	480	p-cumate 2,3-dioxygenase system, small oxygenase component
BAA_00435	7	69770	71047	-	1278	Anthranilate 1,2-dioxygenase large subunit
BAA_00436	7	71106	71930	-	825	2-hydroxy-6-oxo-6-(2'-aminophenyl)hexa-2,4-dienoic acid hydrolase
BAA_00437	7	71964	72779	-	816	Protocatechuate 4,5-dioxygenase beta chain
BAA_00438	7	72781	73062	-	282	Small subunit meta-cleavage dioxygense
BAA_00439	7	73059	74237	-	1179	Carbazole 1,9a-dioxygenase, terminal oxygenase component CarAa
BAA_00440	7	74256	74759	-	504	2-halobenzoate 1,2-dioxygenase small subunit
BAA_00441	7	74756	76111	-	1356	2-halobenzoate 1,2-dioxygenase large subunit
BAA_00442	7	76144	77307	-	1164	Carbazole 1,9a-dioxygenase, terminal oxygenase component CarAa
BAA_00443	7	77362	78081	-	720	Anthranilate 1,2-dioxygenase large subunit
BAA_02473	1226	595	1221	-	627	TetR/AcrR family transcriptional regulator
BAA_02474	1226	1366	2481	+	1116	hypothetical protein
BAA_02475	1226	2505	3794	+	1290	p-cumate 2,3-dioxygenase system, large oxygenase component
BAA_02476	1226	3791	4279	+	489	p-cumate 2,3-dioxygenase system, small oxygenase component
BAA_02477	1226	4291	5208	+	918	Biphenyl-2,3-diol 1,2-dioxygenase

Figure 1.2. Genetic organization of the PAH-degrading gene cluster of *Immundisolibacter* sp. BA1 MAG compared to the gene cluster in the *Immundisolibacter cernigliae* TR3.2 genome. Percentages represent nucleotide sequence identity. In blue, aromatic ring-hydroxylating dioxygenases; in green, aromatic ring-cleaving dioxygenases; in yellow, enzymes of the lower aromatic pathway; in dark grey, transcriptional regulators; in light grey, hypothetical proteins. Annotation of the genes conforming the cluster are listed in the table (Prokka v.1.14.6).

coding for catechol 2,3-dioxygenases, suggesting that the B*a*A metabolic pathway might convey to catechol. Moreover, the lack of genes encoding gentisate 1,2-dioxygenase supports the hypothesis that salicylate would be channeled via catechol, probably by the action of proteins annotated as anthranilate 1,2-dioxygenases (Jouanneau et al., 2007).

Further deepening into the functional annotation data of the *Immundisolibacter* sp. BA1 MAG revealed the presence of a catabolic gene cluster composed of genes coding for 7 different α-subunits of RHD, 5 RHD β-subunits, 2 aromatic ring-cleaving dioxygenases (one β-subunit of a protocatechuate dioxygenase (*LigB*) and one biphenyl-2,3-diol dioxygenase) and one hydrolase involved in aromatic metabolism (Figure 1.2). The gene cluster also included a TetR/AcrR family transcriptional regulator, along a couple of hypothetical proteins. Strikingly, this cluster was found to be enclosed in the genome of the HMW-PAH-degrading *Immundisolibacter cernigliae* TR3.2 (Singleton et al., 2016) with the exact same configuration, both sharing a high nucleotide identity (78-98%). Furthermore, a functional metagenomic screening of fosmid libraries from a PAH-contaminated soil from the Czech Republic subjected to 12 years of *in situ* bioremediation (Duarte et al., 2017), also revealed the presence of gene clusters encoding a vast number of aromatic RHD and ring-cleaving dioxygenases resembling gene clusters found in *I. cernigliae* TR3.2. Remarkably, the detection of conserved genetic organizations encoding enzymes crucial for PAH biodegradation in members of *Immundisolibacter* untaps a common trait for the metabolism of PAHs within this poorly characterized genus.

In fact, a BLAST analysis of the amino acid sequences of the α-subunits of potential PAH-RHD from the *Immundisolibacter* sp. BA1 MAG revealed they are closely related to RHD found in members of *Immundisolibacter* from hydrocarbon-contaminated sites from diverse geographic locations (Table 1.5). The widespread distribution of these RHDs suggests the crucial role of the *Immundisolibacter* genus in the biodegradation of aromatic hydrocarbons. A phylogenetic analysis of the 13 amino acid sequences annotated as α-subunits of naphthalene/biphenyl/benzene dioxygenases shows that most of these enzymes are distantly related to other well-known RHD (Figure 1.3). It is worth noting a separate branch in the phylogenetic tree, comprised of sequences for RHD mainly found in members of *Immundisolibacter* along one of the RHD present in the aforementioned gene cluster (bin1_BAA_00433), suggesting the existence of a novel group of PAH-RHD. These findings highlight the potential of functional screening of metagenomic data for the identification of novel aromatic-degradation genes, distantly related to previously characterized genes from isolates or to genes identified by PCR-mediated functional studies targeting a limited diversity of RHD.

Table 1.5. BLAST analysis of protein sequences of the alpha subunits of potential PAH ring-hydroxylating dioxygenases (annotated as napthtalene 1,2-/biphenyl 2,3-/benzene 1,2-dioxygenases) in the *Immundisolibacter* sp. BA1 MAG.

Bin	Locus tag	Length (aa)	Annotation	BLAST			Id. (%)
				Protein (Accession n.)	Organism	Isolation source	
1	BAA_0138	442	Biphenyl 2,3-dioxygenase subunit alpha	Rieske 2Fe-2S domain-containing protein (MBC7161829.1)	uncultured *Immundisolibacter* sp.	Oil production facility (China)	99
1	BAA_0433	441	Benzene 1,2-dioxygenase subunit alpha	Rieske oxygenase of unknown function alpha-subunit (ART37571.1)	uncultured bacterium clone fosmid E358-360	PAH-contaminated soil (Czech Republic)	97
1	BAA_0933	431	Biphenyl 2,3-dioxygenase subunit alpha	aromatic ring-hydroxylating dioxygenase subunit alpha (MBC7161681.1)	uncultured *Immundisolibacter* sp.	Oil production facility (China)	100
1	BAA_2333	440	Naphthalene 1,2-dioxygenase system, large oxygenase component	aromatic ring-hydroxylating dioxygenase subunit alpha (MBC7163367.1)	uncultured *Immundisolibacter* sp.	Oil production facility (China)	100
1	BAA_2435	446	3-phenylpropionate/cinnamic acid dioxygenase subunit alpha	aromatic ring-hydroxylating dioxygenase subunit alpha (MBC7160873.1)	uncultured *Immundisolibacter* sp.	Oil production facility (China)	99
1	BAA_2437	451	Naphthalene 1,2-dioxygenase system, large oxygenase component	Rieske 2Fe-2S domain-containing protein (MBC7160871.1)	uncultured *Immundisolibacter* sp.	Oil production facility (China)	99
2	BAA_0124	452	Naphthalene 1,2-dioxygenase system, large oxygenase component	aromatic ring-hydroxylating dioxygenase subunit alpha (WP_068805440.1)	*Immundisolibacter cernigliae*	PAH-contaminated soil (USA)	98
3	BAA_0146	450	Naphthalene 1,2-dioxygenase system, large oxygenase component	naphthalene 1,2-dioxygenase (HCO42595.1)	uncultured *Gammaproteobacteria*	Beach sand Deepwater Horizon oil spill (USA)	94
3	BAA_0209	436	Biphenyl 2,3-dioxygenase subunit alpha	Rieske oxygenase of unknown function alpha-subunit (ART37185.1)	uncultured bacterium clone fosmid D474-484	PAH-contaminated soil (Czech Republic)	99
3	BAA_0234	416	Naphthalene 1,2-dioxygenase system, large oxygenase component	aromatic ring-hydroxylating dioxygenase subunit alpha (WP_068802877.1)	*Immundisolibacter cernigliae*	PAH-contaminated soil (USA)	80
4	BAA_0327	439	3-phenylpropionate/cinnamic acid dioxygenase subunit alpha	Rieske 2Fe-2S domain-containing protein (WP_068807673.1)	*Immundisolibacter cernigliae*	PAH-contaminated soil (USA)	92
4	BAA_0465	457	Naphthalene 1,2-dioxygenase system, large oxygenase component	Naphthalene dioxygenase of Proteobacteria alpha-subunit (ART36092.1)	uncultured bacterium clone fosmid B77-89	PAH-contaminated soil (Czech Republic)	99
4	BAA_0488	437	Benzene 1,2-dioxygenase subunit alpha	Rieske 2Fe-2S domain-containing protein (WP_068802386.1)	*Immundisolibacter cernigliae*	PAH-contaminated soil (USA)	96
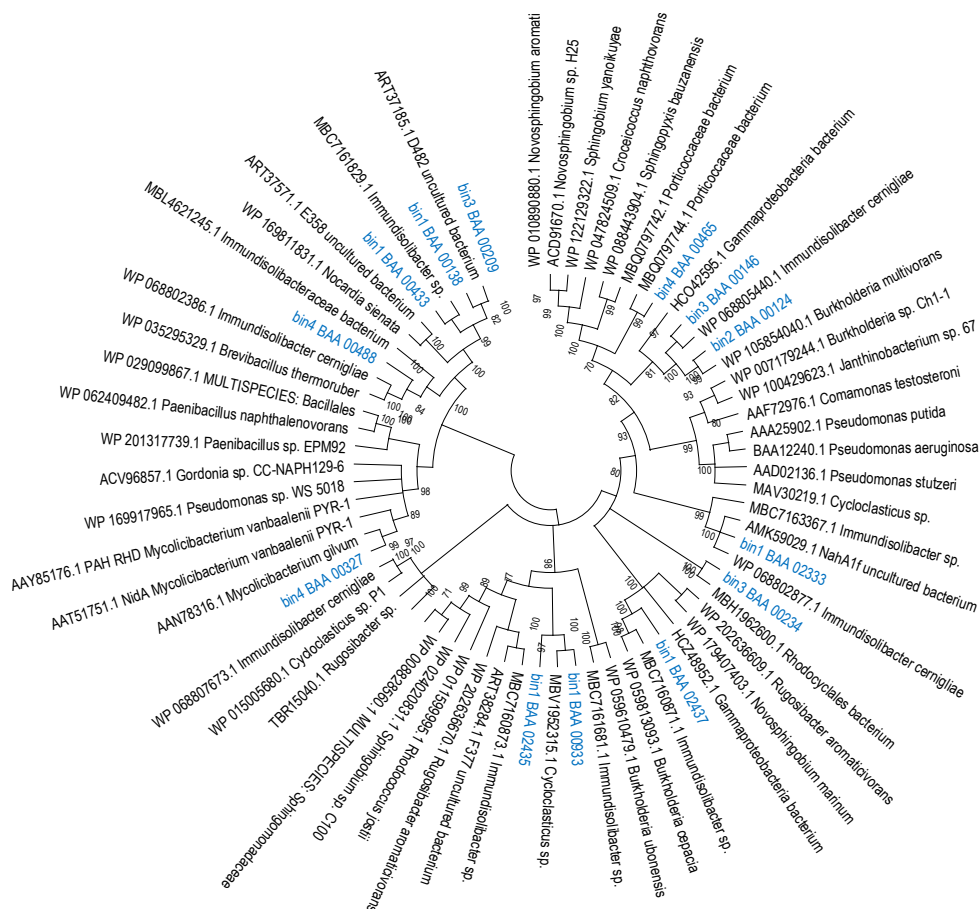
Figure 1.3. Phylogenetic tree of the alpha subunits of aromatic ring-hydroxylating dioxygenases using the maximum-likelihood method of amino acid sequences from the *Immundisolibacter* sp. BA1 MAG (in blue). Reference sequences corresponding to the closest RHD protein sequences available in the GenBank database and well-characterized RHD are also included. The tree was constructed in MEGA X and it involved a total of 66 sequences.

Influence of HMW-PAH co-substrates on B*a*A biodegradation kinetics

The influence of co-occurring PAHs on the biodegradation of B*a*A in a creosote contaminated soil was assessed in soil microcosms spiked either with B*a*A, FT, PY or CHY individually or with B*a*A in combination with one of the other PAHs as co-substrates. Extensive removal of PAHs was observed at the end of the incubation when these were supplied as sole carbon source, being almost complete for B*a*A, FT and PY (95.3%, 96.5%, and 96.2%, respectively), whereas only a 72.8% of CHY was degraded. When co-incubated, B*a*A showed a slightly lower degradation

51

extent (92.1% with FT, 84.3% with PY and 91.8% with CHY) compared to individual BaA incubation, although differences were not significant. As for the PAHs supplied as co-substrates, neither FT nor PY degradation seemed to be affected by the presence of BaA, showing degradation extents comparable to those observed in single substrate incubations (96.1% and 97.3%, respectively). In the case of CHY, however, the presence of BaA seemed to stimulate its degradation, as the final concentration of CHY was significantly lower if compared to the single substrate incubation (80.1% degradation).

Despite modest differences in final degradation percentages, the presence of co-substrates had major effects on degradation kinetics (Figure 1.4). When incubated alone, BaA showed a typical "hockey stick" kinetics with maximum biodegradation rates during the first 30 days (22.1 ± 0.1 nmol·kg^{-1}·day^{-1}). In the presence of FT or PY, BaA degradation was significantly delayed, and only started after 15 days of incubation, when FT or PY had already been removed. Afterwards, BaA concentration was rapidly depleted between 15 and 30 days, showing significantly higher maximum rates compared to incubations with BaA alone (41.4 ± 0.9 nmol·kg^{-1}·day^{-1} with FT and 37.3 ± 2.8 nmol·kg^{-1}·day^{-1} with PY). Biodegradation kinetics of FT and PY remained identical with or without the presence of BaA, being almost completely removed during the first 15 days of incubation with rapid rates (51.3 ± 0.8 nmol·kg^{-1}·day^{-1} FT[+BaA]; 53.8 ± 0.7 nmol·kg^{-1}·day^{-1} FT; 46.3 ± 0.5 nmol·kg^{-1}·day^{-1} PY[+BaA]; 48.3 ± 0.3 nmol·kg^{-1}·day^{-1} PY). When BaA and CHY were co-incubated, the rate of BaA did not suffer significant changes in comparison to the single substrate incubation. On the contrary, CHY utilization was delayed by the presence of BaA, in a similar effect to those observed for FT and PY on BaA degradation. The concentration of CHY did not start to decrease until day 15, but at that point, degradation of CHY started with kinetics significantly higher to those observed for CHY alone (26.4 ± 0.5 nmol·kg^{-1}·day^{-1} and 9.1 ± 1.3 nmol·kg^{-1}·day^{-1}, respectively), that extended until 30 days. The observed sequential biodegradation of PAHs in the co-incubations is consistent with their water solubility, as it is significantly higher for FT (0.26 mg/L) and PY (0.135 mg/L) than for BaA (0.009 mg/L), while BaA is more water soluble than CHY (0.002 mg/L). Then, the microbial populations of the creosote-contaminated soil would first degrade the more readily accessible compound, delaying the degradation of the compound with lower water-solubility.

Interactive effects have been widely observed during the biodegradation of PAH mixtures by pure and mixed cultures. These interactions may be synergistic, resulting in an increase in the degradation of one or more compounds, or inhibitory. Synergistic effects have been attributed to several factors, such as
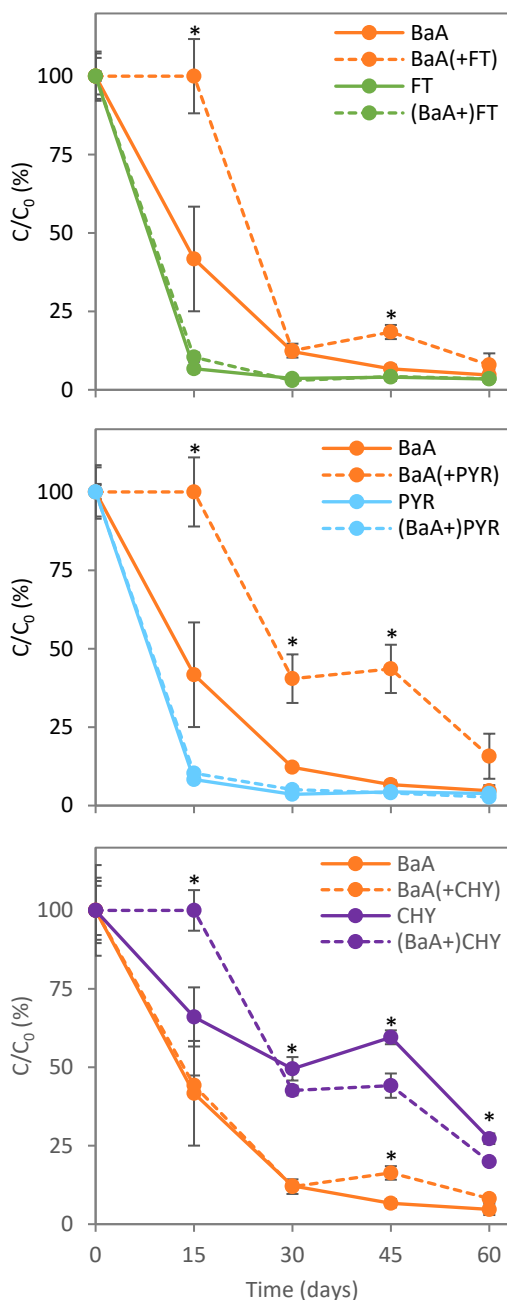
Figure 1.4. Biodegradation kinetics of HMW-PAHs in the soil microcosms. Each data point corresponds to the average of three replicates and error bars represent the standard deviation. Asterisks indicate significant differences (p < 0.05) between individual and combined incubation of a specific substrate at a specific timepoint.

enhancement through biomass growth (Guha et al., 1999) or cross induction (Molina et al., 1999), but more frequently to cometabolic reactions, which are of special relevance in the biodegradation of HMW compounds (Jones et al., 2014). Conversely, inhibitory effects could also occur due to toxicity of the co-substrate (Bouchez et al., 1995), the formation of toxic metabolites (Casellas et al., 1998; Ghosh & Mukherji, 2017), or the relaxed specificity and broad substrate range of the initial enzymes in PAH biodegradation pathways, which may lead to competitive inhibition (Stringfellow & Aitken, 1995; Dean-Ross et al., 2002; Desai et al., 2008). An alternative mechanism would control PAH degradation at the gene expression level via induction or catabolite repression of key enzymes. Facile degradation of a more soluble PAH could either repress enzymes for less accessible PAHs (Juhasz et al., 2002) or release metabolites to induce other pathways leading to cometabolic interactions (Rojo, 2021). All these interactive effects could explain the sequential selective biodegradation of PAHs in mixtures. However, most of these studies have been addressed with pure cultures or defined consortia (Banerjee et al., 1995; Lotfabad & Gray, 2002; Mueller et al., 1989), ignoring other factors operating within complex soil microbial communities. In fact, in the preceding lab-scale biostimulation assay of the creosote-polluted soil (Tauler, 2015), the removal of PAHs within the creosote mixture was also sequential, with the initial disappearance of 2- and 3-ring PAHs and the subsequent biodegradation of HMW-PAHs. These sequential phases, that occurred in accordance with PAHs physicochemical properties, were also associated to important shifts in the active soil microbiome and in the expression level of RHD genes, suggesting a succession in PAH-degrading populations. However, the specific interactions taking place could not be elucidated due to the complexity of the creosote mixture and the simultaneous biodegradation and formation of different polyaromatic compounds.

Metabolomic screens in the B*a*A-spiked soil microcosms

Metabolomic profiles of the polar fractions from the soil microcosms also revealed interactions in the biodegradation of B*a*A when co-incubated with other HMW-PAHs. Throughout the incubation, benz(*a*)anthracene-7,12-dione (B*a*AQ), the ready oxidation product from B*a*A, was detected in all the B*a*A-spiked microcosms (Figure 1.5). After 15 days of incubation, the concentration of this metabolite significantly increased in all conditions, evidencing the formation of B*a*AQ due to the microbial transformation of B*a*A. Considering the almost negligible biodegradation of B*a*A during the first 15 days when co-incubated with other PAHs, the formation of B*a*AQ can be considered as more relevant in relative terms
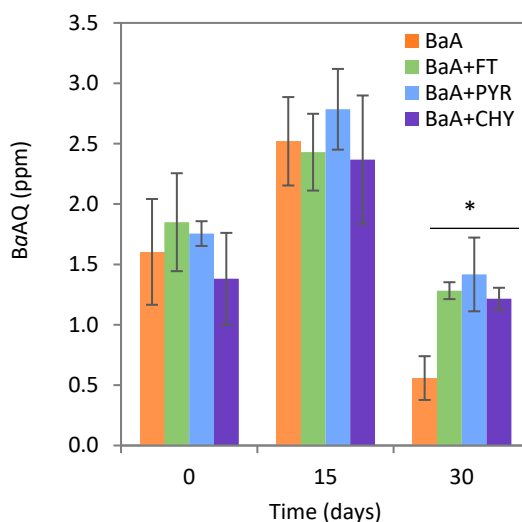
Figure 1.5. Concentration (mg·kg of dry soil⁻¹) of benz(*a*)anthracene-7,12-dione (B*a*AQ) in the B*a*A-spiked soil microcosms. Values are the mean and standard deviation of three replicate. Asterisks indicate significant differences (p<0.05) between the different incubation conditions at a given timepoint.

in the presence of co-substrates, and it suggests that an important fraction of the B*a*A degraded at that point was oxidized by cometabolic reactions. It is also worth noting that the concentration of B*a*AQ significantly declined at 30 days in all conditions, but on a higher extent in the microcosms where B*a*A was the sole carbon source. The slower decrease of B*a*AQ in the co-incubations, indicates that this oxy-PAH was simultaneously formed and degraded, thus corroborating the cometabolic attack of B*a*A when accompanied by other PAHs. Oxy-PAHs are known to have higher water solubility (e.g. B*a*AQ, 0.0289 mg/L) and higher bioavailability than parent-PAHs (Lundstedt et al., 2007), then, it is expected that they are degraded faster than PAHs. Nevertheless, a slower degradation of oxy-PAHs during biodegradation processes in contaminated soils has been attributed to their concurrent microbial formation and degradation (Lundstedt et al., 2003; Wilcke et al., 2014). This phenomenon suggests that oxy-PAHs are not only dead-end products from partial degradation of PAHs or cometabolic reactions, but on the contrary, highlight that microbial communities with the ability to reutilize these compounds may be encountered in soils, using oxy-PAHs as catabolic nodes (Vila et al., 2015). The final fate of oxy-PAHs, however, is still unclear and they rise major concern as recent studies associated the eventual accumulation of oxy-PAHs with an increase in toxicity and genotoxicity of soils after bioremediation (Chibwe et al., 2015; Tian et al., 2017). Therefore, efforts should be made to explore the mechanisms that determine their fate.

Relevant phylotypes in the biodegradation of HMW-PAHs in the soil microcosms

In order to associate substrate interactions to microbial community changes and to identify the bacterial key players involved in the biodegradation of HMW-PAHs in the soil microcosms, 16S rRNA amplicon metagenomic sequencing was carried out for DNA samples after 0, 15 and 30 days of incubation, when maximum degradation rates and major interactions between PAHs were observed. The alpha diversity analysis, expressed as the Simpson's inverse index, revealed noticeable differences throughout the incubation. The highest diversity within the bacterial community was detected at the beginning of the incubation (19.2) followed by a significant decrease for all incubations after 15 days (4.8-10.7) that persisted until 30 days (5.1-9.4). The more uneven distribution of bacterial abundances among taxa, suggests an enrichment of specific bacterial populations associated with the assimilation of the supplied PAHs as carbon sources. Analysis of the bacterial community at the class level (Figure 1.6) revealed that *Gammaproteobacteria* (33.2%), including members of the genera *Immundisolibacter* (11.4%), *Pseudoxanthomonas* (7.8%) and *Pseudomonas* (3.8%), were predominant in the contaminated soil at the beginning of the experiment, followed by
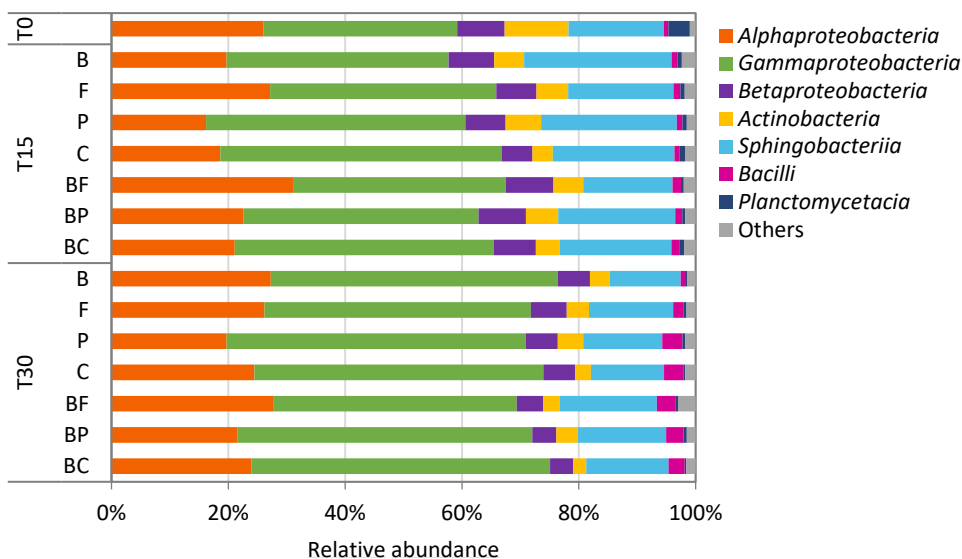
Figure 1.6. Relative abundance of the different bacterial classes in the soil microcosms spiked with B*a*A (B), FT (F), PY (P), CHY (C), B*a*A+FT (BF), B*a*A+PY (BP) or B*a*A+CHY (BC). Others encompass the classes below a 1% of relative abundance and the unclassified sequences.

Alphaproteobacteria (26%), with genera *Sphingobium* (8.6%) and *Phenylobacterium* (1.4%), and members of the *Sphingobacteriia* class (16.3%), mainly belonging to *Olivibacter* (13.9%). Lower relative abundances were accounted for *Betaproteobacteria*, *Actinobacteria*, *Bacilli* and *Planctomycetacia*. In general, no dramatic shifts were observed in microbial community distribution at the class level between conditions or incubation times. Traditionally, Gram negative (GN) bacteria have been associated to the degradation of LMW-PAHs, while the utilization of HMW-PAHs has been related to Gram positives (GP). This perception disagrees with the fact that the microbial communities in all soil microcosms were clearly dominated by GN taxa. The expanding application of culture-independent methods in PAH biodegradation studies is providing increasing evidences of a link between GN bacteria and HMW-PAH disappearance (Vila et al., 2015).

Analysis of the microbial community structure at the OTU-level revealed a similar composition for all conditions, with a predominance of members of *Pseudoxanthomonas* (OTU 1, 24-40%), *Olivibacter* (OTU 2, 8-22%), *Pseudomonas* (OTU 3, 2.5-5%), *Phenylobacterium* (OTU 4, 3.8-6.4%), *Immundisolibacter* (OTU 5, 0.6-10.3%) and *Sphingomonas* (OTU 6, 1.8-5%). However, significant differences were observed for specific OTUs between microcosms (Figures 1.7, 1.8 and 1.9), indicating a differential enrichment associated to PAH assimilation. Analysis focused on those OTUs with >1% relative abundance in any library that experienced a significant increase over time. Statistical analyses of OTU relative abundances were performed in trios comparing the incubations containing B*a*A or the individual PAH co-substrates alone and their co-incubation at each incubation time (e.g. B*a*A alone versus FT alone versus B*a*A and FT together). In this manner, enrichment of each OTU could be associated to the degradation of either B*a*A, the co-substrate or both. Additionally, a LEfSe analysis was performed to differentiate the OTUs explaining differences between all the conditions at each incubation time (Figure 1.10).

OTUs associated to the biodegradation of B*a*A mainly belonged to members of *Olivibacter* (OTUs 2 and 11) at 15 days of incubation, *Immundisolibacter* (OTU 5) at 15 days and especially at 30 days, and to members of unclassified *Sphingomonadaceae* (OTU 12), *Sphingobium* (OTU 18) and *Luteimonas* (OTU 34) at 30 days. The LEfSe analysis revealed that *Olivibacter* (OTU 2) and *Immundisolibacter* (OTU 5) were the main OTUs responsible for the biodegradation of B*a*A, especially at 15 days of incubation. Members of the genus
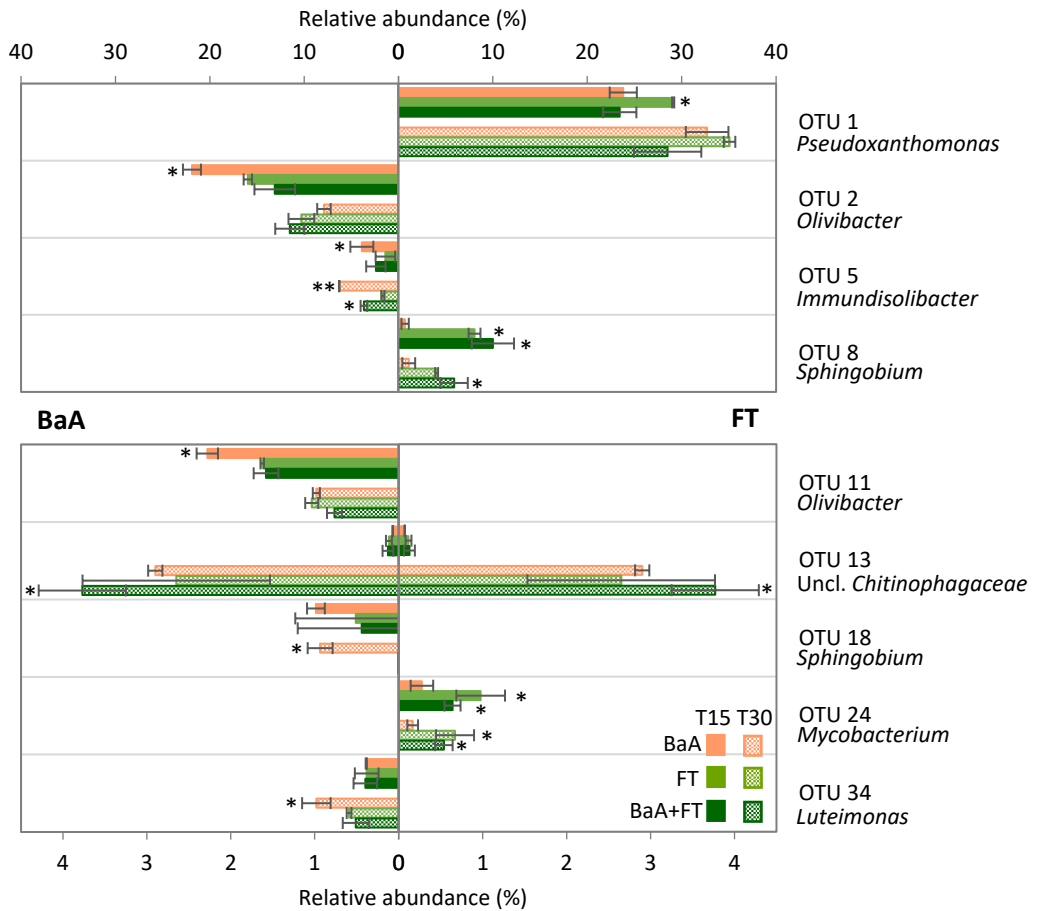
Figure 1.7. Relative abundance of OTUs associated to B*a*A (left) or FT (right) in the soil microcosms containing B*a*A, FT or B*a*A+FT. Values are average and standard deviation for three replicates. Asterisks indicate significant differences (p<0.05) between the three specified conditions at a given timepoint, two asterisks indicate significance in respect to a value that already presented a significant increase.

Olivibacter have been frequently found in contaminated environments (Villaverde et al., 2019), however the reported isolates have been related to the utilization of monoaromatic intermediates in PAH metabolic pathways such as protocatechuate (Ntougias et al., 2007) or catechol (Ntougias et al., 2014), suggesting it could be utilizing metabolites produced by other bacteria. On the other hand, the 16S rRNA sequence of OTU 5 matched at 100% identity with the sequence of *Immundisolibacter* sp. BA1, identified above as the key player in B*a*A degradation by DNA-SIP combined with shotgun metagenomics. For FT-containing microcosms,
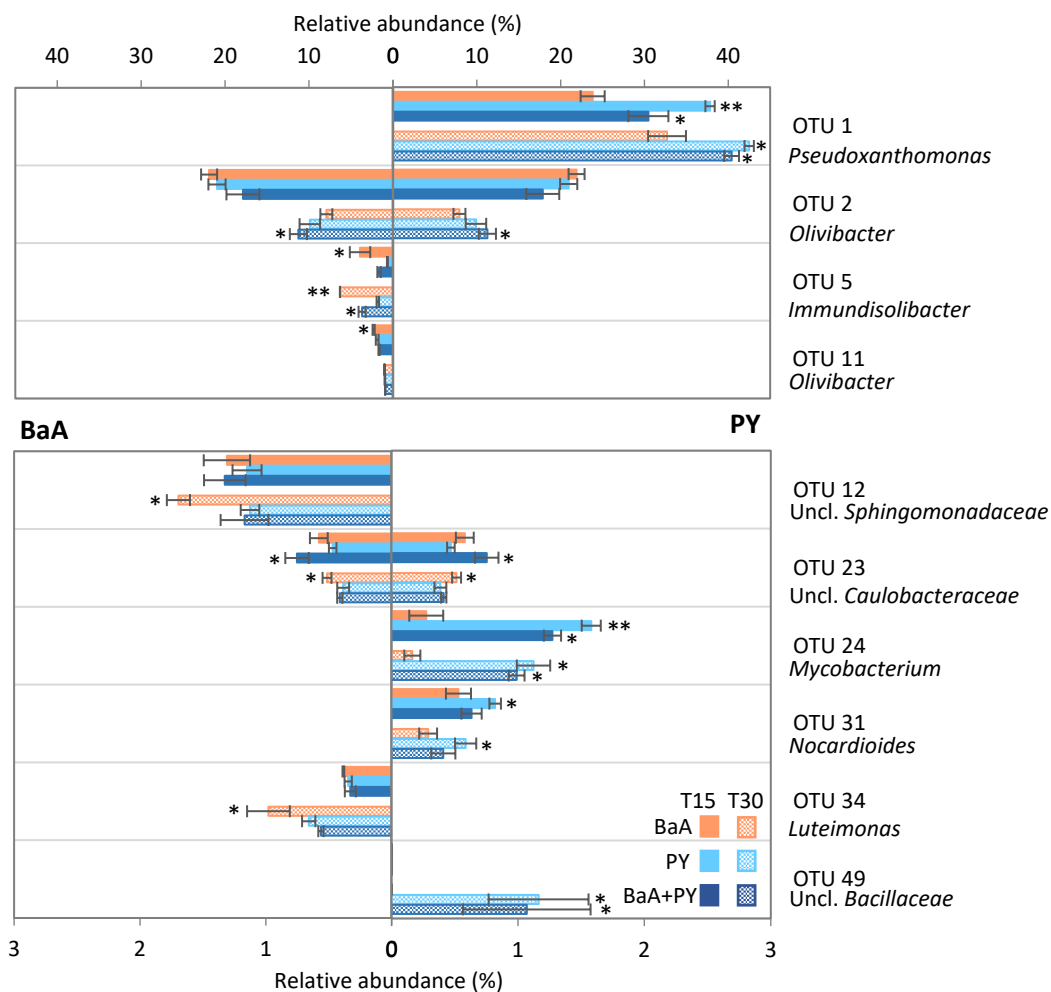
Figure 1.8. Relative abundance of OTUs associated to BaA (left) or PY (right) in the soil microcosms containing BaA, PY or BaA+PY. Values are average and standard deviation for three replicates. Asterisks indicate significant differences (p<0.05) between the three specified conditions at a given timepoint, two asterisks indicate significance in respect to a value that already presented a significant increase.

there was a remarkable increase in the relative abundance of OTU 8, classified as *Sphingobium*, and, in a minor extent, OTU 24, a member of *Mycobacterium*. The presence of *Actinobacteria*, though, was primarily associated to PY, with a prominent increase of *Mycobacterium* (OTU 24) and a slight increment of *Nocardioides* (OTU 31) in microcosms containing this PAH. The relevance of *Mycobacterium* (OTU 24) in the degradation of PY was corroborated by the LEfSe analysis. *Mycobacterium* is a well investigated genus in the biodegradation of HMW-PAHs, particularly associated to pyrene and fluoranthene (Kweon et al.,
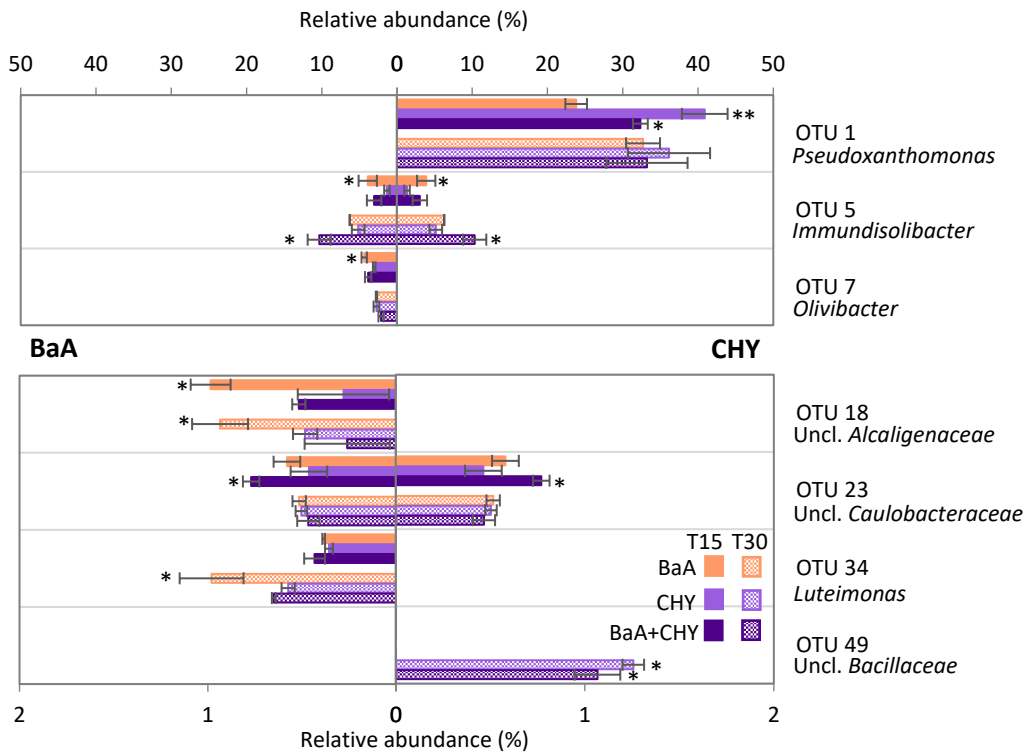
Figure 1.9. Relative abundance of OTUs associated to B*a*A (left) or CHY (right) in the soil microcosms containing B*a*A, CHY or B*a*A+CHY. Values are average and standard deviation for three replicates. Asterisks indicate significant differences (p<0.05) between the three specified conditions at a given timepoint, two asterisks indicate significance in respect to a value that already presented a significant increase.

2010). This genus has been recently divided into four novel genera, including the non-pathogenic genus *Mycolicibacterium* related to PAH utilization (Gupta et al., 2018). *Pseudoxanthomonas* OTU 1, especially at 15 days, and a member of unclassified *Bacillaceae* (OTU 49) at 30 days, presented a significantly higher abundance in the presence of PY and CHY. The LEfSe analysis correlated *Pseudoxanthomonas* (OTU 1) to both PY (at 30 days) and CHY (at 15 days). Actually, the only two *Pseudoxanthomonas* isolates reported for PAH biodegradation were related to the utilization of pyrene (Klankeo et al., 2009) and chrysene (Nayak et al., 2011), indicating a potential role of this genus in the degradation of these HMW-PAHs.
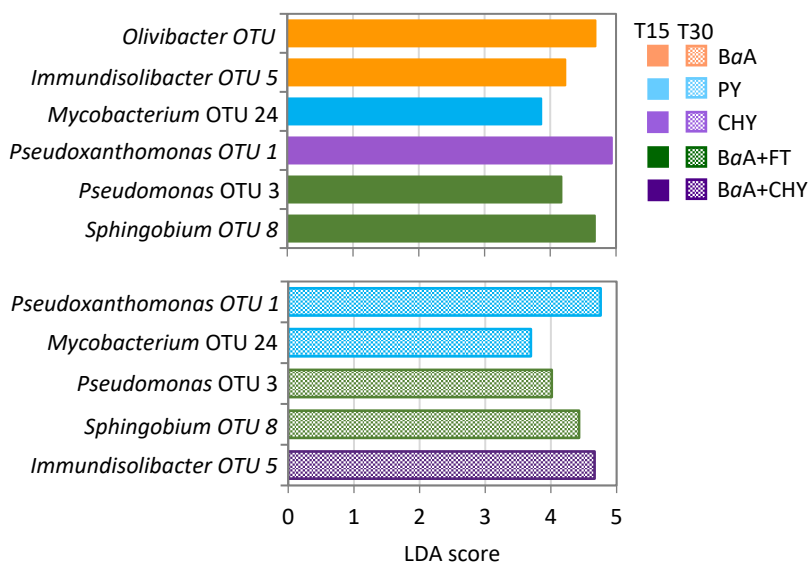
Figure 1.10. LEfSe analysis of the microbial community in the soil microcosms at OTU-level at 15 (top) and 30 (bottom) days of incubation. Linear discriminant analysis (LDA) score is represented for OTUs with a significant increase (p<0.05) over incubation time and a LDA score >3.5.

Microbial interactions during B*a*A degradation in the presence of HMW-PAH co-substrates

The analysis at the OTU-level identified relevant microbial community interactions during the co-incubation of B*a*A with other PAHs as co-substrates. These interactions could probably be explained by cometabolic reactions, as ascertained from the analysis of the biodegradation kinetics and the formation of B*a*AQ. In the presence of FT, *Immundisolibacter* (OTU 5) and *Luteimonas* (OTU 34) experienced a significantly lower increase in abundance at 30 days of incubation compared to the incubations with B*a*A alone. On the contrary, *Sphingobium* (OTU 8) and, in a lesser extent, an unclassified member of *Chitinophagaceae* (OTU 13) were significantly increased in the co-incubation in respect to any of the incubations with single substrates (either FT or B*a*A). In the LEfSe analysis, *Sphingobium* (OTU 8) was associated to the co-incubation of FT and B*a*A. This result indicates that, when FT was present, B*a*A degradation was partially channeled through these other members of the community. Considering that *Sphingobium* OTU 8 was identified as a major FT degrader, this could probably be attributed to a competition phenomenon, where *Sphingobium* OTU 8 would be favored by its

overgrowth at the expenses of FT. This enrichment of OTU 8 at the expenses of FT could also explain the increased degradation rates observed for B*a*A between 15 and 30 days during their co-incubation (Figure 1.4). Sphingomonads are well known for their abilities to degrade a wide range of aromatic compounds, including HMW-PAHs, due to their great enzymatic versatility (Khara et al., 2014; Zhao et al., 2017). In fact, several *Sphingobium* strains have been found to cometabolize B*a*A in the presence of other PAHs, for example, *Sphingobium* sp. KK22, isolated from a phenanthrene enrichment culture of a bacterial consortium able to grow on diesel fuel, was capable to biotransform B*a*A when grown with phenanthrene (Kunihiro et al., 2013). Furthermore, the LEfSe analysis revealed a significant association of a member of *Pseudomonas* (OTU 3) to the co-incubation of B*a*A and FT, suggesting a relevant role in the degradation of these two compounds.

When B*a*A was co-incubated with PY, members of unclassified *Caulobacteraceae* (OTU 23) were enhanced at 15 days and *Olivibacter* (OTU 2) remained dominant at 30 days. Similar to FT, when B*a*A and PY were together, *Immundisolibacter* (OTU 5) presented a significantly lower increase in its relative abundance if compared to B*a*A individually. As suggested before, this could be associated with a possible utilization of B*a*A by other members of the microbial community; however, the extent of B*a*A degradation was lower in the presence of PY at 30 days (Figure 1.4), suggesting a possible inhibition of this specific OTU. The genus *Immundisolibacter* has been associated to the degradation of HMW-PAHs, and in particular to PY (Singleton et al., 2006); nevertheless, OTU 5 did not seem to respond to the presence of PY alone, indicating a specialization of this particular phylotype in the degradation of B*a*A.

Conversely, for the co-incubation of B*a*A with CHY, *Immundisolibacter* (OTU 5) suffered a remarkable increase at 30 days, reaching its maximum relative abundance considering all incubation conditions. This was in agreement with the LEfSe analysis, as OTU 5 was correlated to the co-incubation with B*a*A and CHY. The increased abundance of this OTU coincided with a significantly higher degradation extent of CHY when co-incubated with B*a*A. This synergistic effect of the presence of B*a*A on CHY degradation results from the action of *Immundisolibacter*. Altogether, these findings highlight the multiple interactions that can potentially take place between microbial populations embedded in complex metabolic networks when PAHs are found in mixtures.

The abundance of the total (16S rRNA genes) and active (16S rRNA transcripts) bacterial populations was quantified by qPCR. Overall, bacterial abundance and activity significantly increased after 15 days of incubation for all microcosms,
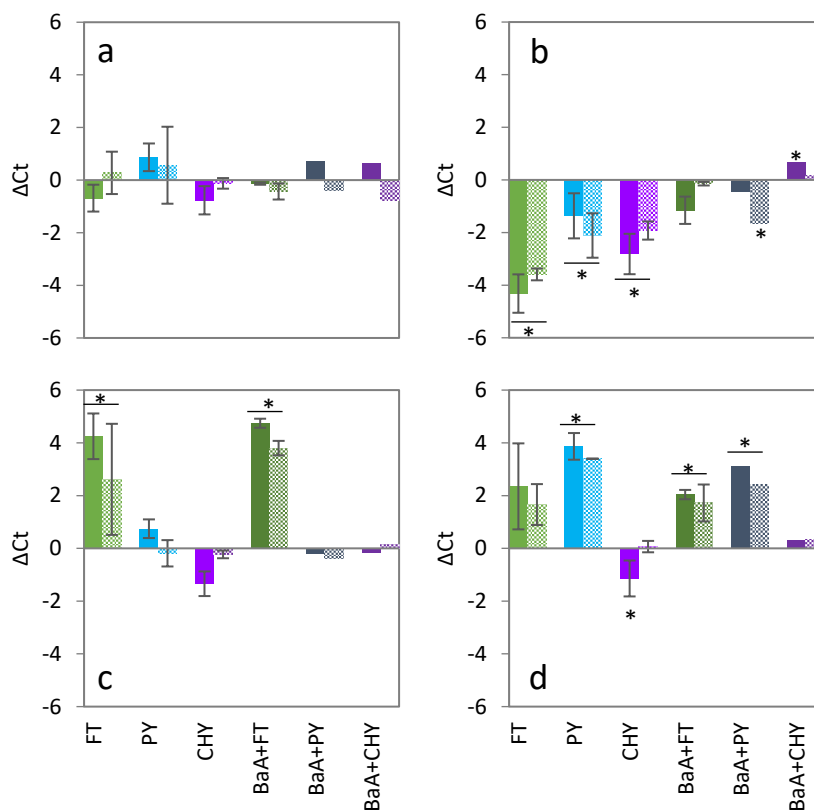
Figure 1.11. Fold-change of phylotype-specific 16S rRNA transcripts in respect to the incubation of B*a*A alone in the soil microcosms quantified by qPCR analysis of cDNA samples. (a) *Pseudoxanthomonas*; (b) *Immundisolibacter*; (c) *Sphingobium*; (d) *Mycobacterium*. Values are average and standard deviation of three replicates. Asterisks indicate significantly (p<0.05) different Ct values compared to the incubation of B*a*A individually.

going from $1.8 \cdot 10^8$ to 10^9 gene copies·g of dry soil^{-1} and from $8.2 \cdot 10^6$ to 10^8 transcript copies·g of dry soil^{-1}, respectively, and remained at the same level after 30 days. The relevance of *Pseudoxantomonas* OTU 1, *Immundisolibacter* OTU 5, *Sphingobium* OTU 8 and *Mycobacterium* OTU 24 was assessed at a transcriptomic level by quantification of their specific 16S rRNA transcripts in the soil microcosms. In order to evaluate their relative contribution to B*a*A degradation, the specific activity of each phylotype in all incubations was compared to their activity in the incubation with B*a*A alone (Figure 1.11). Results were expressed as the average change in Ct values (ΔCt) in respect to the Ct in BaA incubations. *Pseudoxanthomonas* OTU 1 showed minor variations in activity for all incubations, with no differences between conditions. The relative activity of *Immundisolibacter*

OTU 5 in respect to B*a*A was lower in all conditions, except for the co-incubation with B*a*A and CHY. This result confirmed the synergistic effect of this mixture on *Immundisolibacter* OTU 5 activity, and the decreased contribution of this OTU to BaA biodegradation when co-incubated with FT or PY as co-substrates. Analysis of *Sphingobium* OTU 8 16S rRNA transcripts only revealed an increase in activity in those microcosms spiked with FT, either alone or in the presence of B*a*A. This result confirmed the major contribution of *Sphingobium* OTU 8 to FT degradation. Finally, *Mycobacterium* OTU 24 presented the highest activity in microcosms spiked with either FT or PY, especially the latter. The transcriptomic assessment reinforced the findings from the 16S rRNA amplicon sequencing data, demonstrating that increased abundance of specific OTUs in response to either FT, PY or B*a*A correlated with their increased activity in those incubations. Also, transcriptomic data revealed that the presence of other PAHs modulated the activity of *Immundisolibacter* OTU 5 during B*a*A degradation. Activity of OTU 5 decreased in the presence of co-substrates with higher water solubility, while the presence of the less water-soluble CHY resulted in increased activity, thus confirming the existence of competitive inhibition and synergistic phenomena during the degradation of PAH mixtures.

Environmental implications

DNA-SIP combined with metagenomics have identified the key role of members of the order *Immundisolibacterales* in the degradation of the four ring PAH benzo(*a*)anthracene in our model creosote contaminated soil. Our results highlight how this until recently overlooked group of gammaproteobacteria can be a global major player in the fate of HMW-PAHs, as data mining of metagenomic surveys reveal their worldwide distribution in PAH-contaminated soils. Functional analysis of the *Immundisolibacter* MAG detected here revealed an astonishing vast number of aromatic RHD coding genes, suggesting the broad catabolic potential of the phylotype, and the presence of a catabolic cluster with an extremely high degree of conservation with that detected in the type strain TR3.2. This and the high divergence of some *Immunidsolibacterales* RHDs with those of other PAH-degrading bacteria suggest their high degree of specialization, and the need for their consideration when monitoring HMW-PAH degrading communities by means of functional screens (ie. qPCR of RHD genes). Our work also reveals that substrate interactions observed during the degradation of PAH mixtures by complex microbial communities can originate from competitive interactions within the PAH-degrading populations. Those interactions, probably influenced by cometabolic reactions, should be taken into account when exploring the fate of HMW-PAHs in soils.

Chapter 2

Polyomic elucidation of the 9,10-anthraquinone biodegradation pathway by a newly isolated oxy-PAH degrading specialist from PAH-contaminated soil

Polyomic elucidation of the 9,10-anthraquinone biodegradation pathway by a newly isolated oxy-PAH degrading specialist from PAH-contaminated soil

Introduction

Industrial soils related to the production, transport, storage or use of petroleum- or coal-derived products are extensively impacted by contamination with polycyclic aromatic hydrocarbons (PAHs), posing a risk for human health and the environment. PAHs are embedded in complex mixtures that include many other polyaromatic components, including heterocyclic aromatic compounds and oxygenated-PAHs (oxy-PAHs) (Lundstedt et al., 2014). However, measures of risk and remediation effectiveness in PAH-contaminated sites are still based exclusively on concentration levels for the 16 regulated PAHs listed by the US-EPA in 1976, neglecting the co-occurrence and fate of other toxicologically relevant compounds (Andersson & Achten, 2015; Titaley et al., 2020).

Bioremediation is the most cost-efficient technology for the clean-up of hydrocarbon contaminated sites, presenting low environmental footprints and being capable of restoring key natural soil functions (Gillespie & Philp, 2013). Nevertheless, recent studies have revealed that the genotoxicity of soils may eventually increase during bioremediation processes due to the accumulation of oxy-PAHs as a result of the partial oxidation of PAHs (Chibwe et al., 2015; Tian et al., 2017). These polar compounds are generally regarded as more (geno)toxic, mutagenic and carcinogenic (Idowu et al., 2019; Clergé et al., 2019) and present a greater bioavailability and environmental mobility than parent PAHs (Lundstedt et al., 2007; Larsson et al., 2018). Oxy-PAHs, such as aromatic ketones, quinones or lactones, are found along PAHs in the contaminant source, but can also be readily formed due to photo/chemical oxidation (Lampi et al., 2006) or microbial transformation of PAHs (Wilcke et al., 2014; Biache et al., 2017). Several bacterial isolates have been observed to produce oxy-PAHs as dead-end products during aerobic metabolism of PAHs (Moody et al., 2001, 2005). Although their further biodegradation has been reported after stimulation of the microbial activity in soils (Wilcke et al., 2014; Biache et al., 2017), little is known about the microorganisms and mechanisms underlying their fate in the environment.

9,10-Anthraquinone (ANTQ), the ready oxidation product from anthracene (ANT), is one of the most commonly found oxy-PAHs in PAH-contaminated soils and has been classified as possibly carcinogenic to humans (group 2B) by the International Agency for Research on Cancer (IARC, 2013). This oxy-PAH has been reported as a

dead-end product in the metabolic pathway of anthracene transformation by *Mycobacterium vanbaalenii* PYR-1 (Moody et al., 2001) and its formation and subsequent removal has been observed during active biological treatment of contaminated soils (Lundstedt et al., 2003; Hu et al., 2014; Wilcke et al., 2014). In a PAH-contaminated soil from a former manufactured-gas plant, biodegradation of anthraquinone was associated to uncultured *Sphingomonas* and *Phenylobacterium* species (Rodgers-Vieira et al., 2015) detected by DNA-SIP, but the specific metabolic mechanisms involved remain to be elucidated.

In this study, we report the isolation of the first bacterial strain, *Sphingobium* sp. AntQ-1, with the ability to grow on 9,10-anthraquinone as sole carbon and energy source. This strain was isolated from a creosote-contaminated soil sample from a historical wood-treating facility in the South of Spain. AntQ-1 strain has been used here as a model to shed light on the microbial mechanisms involved in oxy-PAH biodegradation in contaminated soils. Combining metabolomic, genomic and transcriptomic analysis we have comprehensively elucidated the metabolic pathway and identified the key genes involved in the biodegradation of anthraquinone. The environmental relevance of the bacterial strain AntQ-1 and its degradative mechanisms during a bioremediation process were ascertained by monitoring the identified genes after biostimulation of the original PAH-contaminated soil.

Materials and methods

Chemicals

Anthracene (ANT, 99% purity), 9,10-anthraquinone (ANTQ, 97%), anthrone (ANTO, 97%) and all other aromatic compounds were purchased from Sigma-Aldrich Chemie (Steinheim, Germany). Media and reagents were purchased from Panreac Química (Barcelona, Spain) or Merck (Darmstadt, Germany). Solvents (organic residue analysis grade) were obtained from J.T. Baker (Deventer, The Netherlands). Diazomethane was generated by alkaline decomposition of Diazald (N-methyl-N-nitroso-p-toluenesulfonamide). All chemicals and solvents were of the highest purity available.

Soil

The creosote-polluted soil sample was collected from a historical wood-treating facility in southern Spain. This soil (silty clay loam-calcaric fulvisoil) presented a very high content in PACs, $\Sigma17$ PAHs=25,791 ± 2,112 ppm, $\Sigma7$ oxy-PAHs =273 ± 9 ppm and $\Sigma7$ N-PACs=2,568 ± 366 ppm (Annex – Table S1.1).

Biodegradation of ANT and ANTQ in sand-in-liquid soil microcosms

A small-scale sand-in-liquid microcosm system (Tauler et al., 2016) was used as a model to assess the microbial community shifts during the biodegradation of ANT and ANTQ in the creosote-contaminated soil. Microcosms consisted of 20 mL-glass vials containing 6 mL of mineral medium and 3 g of thin grain sea sand (Panreac, Barcelona, Spain) coated with 0.1 g·L^{-1} of either ANT or ANTQ. The sand, treated at 400ºC overnight to remove organic residues, was added to the sterile empty vials and spiked with the aforementioned substrates in dichloromethane solution. After complete evaporation of the solvent, vials received 4 mL of sterile mineral medium (MM; Hareland et al., 1975) and 2 mL of inoculum. Identical triplicate abiotic controls only received 6 mL of sterile mineral medium. Series of triplicate microcosms for each substrate were incubated during 0, 5, 10, 15, 20, 25 and 30 days at 150 rpm and 25ºC. At each incubation time, a set of triplicates was extracted for chemical analysis to quantify residual ANT or ANTQ. Another set of triplicates was used to obtain 1 mL samples containing sand and liquid for total DNA and RNA extraction. Details for nucleic acids extraction are explained in Chapter 1. Further 16S rRNA PCR amplification and DGGE analysis was performed as described elsewhere (Tauler et al., 2016), and selected bands were excised for sequencing.

The inoculum for the microcosms derived from the creosote-polluted soil. To minimize the concentration of native PAHs, the soil sample was preincubated for 21 days in aerobic conditions and the removal of PAHs was monitored. 50 g of soil were placed in a 1-L Erlenmeyer flask containing 200 mL of MM and 100 g of sand, and the mixture was incubated at 100 rpm and 25ºC. Every 3 days 10 mL samples of the sand-soil suspension were removed and extracted for chemical analysis. At the end of the incubation, the content in Σ17 PAHs had been reduced by 93% and a 1/10 dilution of the liquid culture was used as inoculum for the microcosms. As a result, the residual concentration of most native PAHs was below detection limits in the miniaturized microcosms.

The liquid and sand phases of the different cultures were separated by filtration through a paper filter (Whatman Grade No.1, GE Healthcare, UK). The aqueous phase was extracted 3-times with 10 mL dichloromethane. 3 g of Na_2SO_4 were added to the solid phase, which was extracted in ultrasonic bath twice with 10 mL dichloromethane/acetone (2:1, v/v) and a third extraction with 10 mL of dichloromethane. Before chemical analysis, the extracts from the pre-incubation culture of the soil were cleaned-up in 2.5 g alumina columns and eluted with 10 mL dichloromethane.

Isolation of the 9,10-anthraquinone-degrading strain *Sphingobium* sp. AntQ-1

Strain AntQ-1 was isolated from the ANTQ-spiked sand-in-liquid microcosms inoculated with the creosote-polluted soil. Serial dilutions from the 30-day ANTQ sand-in-liquid soil microcosms were inoculated on solid mineral medium supplemented with vitamin B12 (50 $\mu g \cdot L^{-1}$) and yeast extract (YE, 0.25 $g \cdot L^{-1}$) and containing 9,10-anthraquinone (0.1 $g \cdot L^{-1}$) as the main carbon source. Solutions of vitamin B12 (filter-sterilized, 0.22μm) and anthraquinone (the latter in acetone) were added to the sterile medium before plating. Plates were incubated at room temperature for 30 days. Colonies surrounded by clearing areas were selected and purified on R2A (Reasoner's 2A) agar plates supplemented with vitamin B12.

Identification of bacterial strain AntQ-1 was carried out according to its 16S rRNA gene sequence. Genomic DNA was purified using InstaGene Matrix (Bio-Rad, Hercules, CA, USA). The nearly complete 16S rRNA gene sequence (approximately 1,350bp) was obtained by sequencing the 16S rRNA gene fragment amplified with primers 27F and 1492R (Weisburg et al., 1991). PCR amplification, purification and sequencing procedures are described below.

Scanning electron microscopy (SEM) was performed at the Scientific and Technological Centres of the University of Barcelona (CCiTUB) on a JEOL JSM7001F microscope. *Sphingobium* sp. AntQ-1 was grown in MM+B12 supplied with ANTQ (0.1 $g \cdot L^{-1}$) and the cells were fixated with 2% glutaraldehyde, subjected to critical point drying and coated with graphite.

Utilization of PAHs and other aromatic compounds by strain *Sphingobium* sp. AntQ-1

ANTQ utilization as sole carbon and energy source was determined by demonstrating cell protein increase concomitant with a decrease of ANTQ concentration in liquid cultures. Cultures were grown in 100-mL Erlenmeyer flasks containing 20 mL of MM+B12 and anthraquinone (0.1 $g \cdot L^{-1}$), at 25ºC under rotary shaking (150 rpm) for 10 days. Uninoculated flasks and flasks without ANTQ served as controls. Every 2 days, entire triplicate cultures were used for measuring protein concentration by a modification of the Lowry method (Daniels et al., 1994). Another set of triplicates were extracted with dichloromethane (5x10 mL) for chemical analysis. ANTQ concentration was determined by GC-FID analysis with a 5-point standard curve as described above. Accumulated metabolites were detected by HPLC analysis of the supernatant after centrifugation (10,000 rpm, 5 min) as described below.

Growth on other aromatic compounds (final concentration 0.1 g·L^{-1}) was tested by quantification of protein in triplicate liquid cultures containing 5 mL of MM+B12 medium. Cultures were incubated for 10 days at 150rpm and 25ºC. The following compounds were assayed: fluorene, anthracene, phenanthrene, benz(*a*)anthracene, 9-fluorenone, anthrone, 9,10-phenanthrenequinone, 7,12-benz(*a*)anthracene-7,12-dione, 5,12-naphthacenequinone, xanthone, indanone, carbazole, acridine, dibenzofuran, dibenzosuberone, benzophenone, 2-methylanthraquinone, bisphenol A, ethinylestradiol, anthraquinone-2-carboxylic acid, benzoic acid, diphenic acid, phthalic acid, 1-hydroxy-2-naphthoic acid, cinnamic acid, protocatechuic acid, carboxybenzaldehyde, benzenetricarboxylic acid, toluic acid, hydroxycoumarin, catechol, salicylic acid, and 1,8-naphthalic anhydride. Growth on naphthalene, camphor and biphenyl was tested in MM+B12 plates supplied with these compounds as crystals on the lid of the plates.

All experiments were inoculated (1%) with a suspension of strain AntQ-1 cells grown in 5 mL of R2A+B12 medium. Cultures were incubated during 48 h at 150 rpm and 25ºC, reaching an A_{600} of approximately 0.9. Cells were centrifuged and washed thrice with MM before their use as inoculum.

Identification of 9,10-anthraquinone metabolites

Metabolites from the degradation of ANTQ by strain AntQ-1 were identified in washed-cell incubations with an excess of ANTQ (0.5 g·L^{-1}). Strain AntQ-1 cells were pregrown in 4x 2-L Erlenmeyer flasks cointaining 400 mL of MM+B12+YE and ANTQ crystals (0.5 g·L^{-1}). At late exponential phase (after 72h) remaining ANTQ crystals were removed by filtration through sterile glass wool and cells were harvested by centrifugation, washed twice and resuspended in 400 mL of phosphate buffer (5 mM, pH 7). The washed-cell suspension was placed in a sterile 2-L Erlenmeyer flask with ANTQ crystals (0.5 g·L^{-1}) and incubated for 48h at 150 rpm and 25ºC, when maximum metabolite accumulation was observed. For the identification of anthrone metabolites, all experiment conditions were the same except for using ANTO as substrate. Controls without cells were included to assess abiotic degradation. After incubation, washed-cell suspensions were centrifuged to remove cells and undissolved ANTQ. Supernatants were first extracted (at neutral pH) with 5x100 mL ethyl acetate and then acidified to pH 2.5 with 6N HCl and extracted again in the same manner. Extracts were analyzed by high pressure liquid chromatography (HPLC) for metabolite detection. Identification of detected metabolites was achieved by gas chromatography coupled to mass spectrometry (GC-MS) and by HPLC coupled to high resolution mass spectrometry with electrospray ionization (HPLC-ESI-HRMS). Structural elucidation of 2-(2-hydroxybenzoyl)-benzoic acid was achieved by nuclear magnetic resonance (NMR).

71

Chemical analyses

Culture extracts were analyzed by gas chromatography with a flame ionization detector (GC-FID) on a TRACE 2000 GC (Thermo Quest, Milan, Italy) as described elsewhere (López et al., 2008). Quantification of PAHs and ANTQ was performed by using a five-point standard calibration curve. GC-MS analyses were performed on an Agilent Technologies 6890N gas chromatograph coupled to a 5975 inert mass spectrometer as described elsewhere (López et al., 2008). Prior to GC-MS analysis, the neutral and acidic extracts were derivatized with diazomethane and the metabolites were identified by comparison with authentic standards, our in-house analytic database or the NIST library.

Reverse-phase HPLC was performed on a Waters W2690/5 chromatograph equipped with a W2996 photo diode-array (PDA) UV-visible detector set at 254 nm. HPLC-ESI-HRMS analysis was conducted on a UHPLC Ultimate 3000 (Thermo Scientific) chromatograph coupled to a PDA Accela (Thermo Fisher Scientific) detector set at 254 nm and a Linear Trap Quadropole (LTQ) Orbitrap Velos (Thermo Scientific) mass spectrometer with electrospray ionization (ESI) and positive polarity (+). Separation was achieved using an Accucore™ C18 column (Thermo Scientific) (100 x 2.1 mm; 2.6 µm particle size) and a linear gradient of methanol (10 to 95% [vol/vol] in 20 min) in acidified water (0.1% CH_2O_2). Flow was kept at 0.2 mL·min^{-1}. The injection volume was 10 µL.

The NMR experiments were performed on a Varian VNMRS500 spectrometer operating at 499.63 MHz in ^{1}H. The ^{1}H spectra were recorded using a standard sequence with a 45° observe pulse, delay recover 1s, 48 transients and at 25 °C temperature controlled. The 2D experiment ^{1}H-^{1}H gCOSY (^{1}H-^{1}H gradient enhanced COrrelation SpectroscopY experiment) was acquired with 16 scans, 128 increments and the spectral width used was 13.7 ppm in the two dimensions. The ^{1}H-^{13}C heterocorrelation experiment was performed using a standard sequence gHSQC (gradient enhanced Heteronuclear Single Quantum Correlation). This spectrum was acquired with spectral windows of 13.7 ppm, for the ^{1}H dimension, and 199.9 ppm, for the ^{13}C dimension. A 1s prescan delay and 24 scans, 128 increments were used. The spectra were processed with the MestReNova program Version 12.03. The ^{13}C chemical shifts were extracted from the gHSQC analysis (only ^{13}C attached to ^{1}H are visible). The ^{1}H chemical shifts and ^{1}H coupling constants were obtained through a multiplet analysis routine included in the MestReNova software.

PCR amplification and Sanger sequencing

16S rRNA gene fragments were amplified by PCR using pureTaq Ready-To-Go PCR bead tubes (GE Healthcare, United Kingdom), in a final volume of 25 µL containing

1 μL DNA extract and 25 pmol of each primer (Sigma-Aldrich, Steinheim, Germany). Amplification was performed with primers GC40-63F and 518R (Muyzer et al., 1993) for DGGE analysis or primers 27F and 1492R (Weisburg et al., 1991) for strain identification. DNA or cDNA extracted from the sand-in-liquid soil microcosms or the DNA of the *Sphingobium* sp. AntQ-1 strain were used as template. cDNA was obtained by retro-transcription of RNA extracts using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA) with random hexamers. PCR amplification was done on an Eppendorf Mastercycler and was validated with agarose gel (TBE 1x; 1% agarose) stained with Nancy 520 (Sigma-Aldrich, Steinheim, Germany). Sequencing of 16S rRNA amplification products was done by Macrogen Europe after purification with ExoSAP-IT (ThermoFisher Scientific, USA). The resulting DNA sequences were manually adjusted using BioEdit v.7.2.5. and analyzed using the *Classifier* and *Sequence Match* tools of the Ribosomal Database Project II (Maidak et al., 2001) and Blastn tool of GenBank.

De novo whole genome sequencing

Genomic DNA of strain AntQ-1 was extracted with the MasterPure Gram Positive DNA Purification kit (Lucigen, Middleton, WI, USA), with O/N lysis, from a 1 mL aliquot of a culture grown in R2A+B12 medium for 48 h. DNA quality was checked by agarose gel electrophoresis and DNA concentration was measured by Qubit™ (Thermo Fisher, USA). Whole genome sequencing was performed at Novogene Co. (Beijing, China) on both an Illumina HiSeq (Illumina, CA, USA) platform, to generate 150 bp paired-end reads, and a PacBio Sequel (Pacific Biosciences, CA, USA) platform. Illumina paired-end reads were analyzed with FastQC (v0.11.8) and Trimmomatic (v0.38) was used for adapter removal and quality filtering. PacBio raw reads were filtered and processed using SMRT Link (v.9.0). Hybrid *de novo* assembly using Illumina and PacBio subreads was performed with Unicycler (v0.4.8), with default parameters. The quality assessment of the genome assembly was done using QUAST (v.5.0.2) and genome assembly graphs were visualized with Bandage (v.0.8.1). The genome sequence of strain AntQ-1 was annotated with Prokka (v.1.14.5). The predicted proteins were assigned to orthologous groups and mapped to KEGG pathways by using the KEGG Automatic Annotation Server (KAAS).

RNA-Seq

Sphingobium sp. AntQ-1 cells grown in R2A+B12 medium were washed twice with MM and used to inoculate 0.5-L Erlenmeyer flasks containing 100 mL of MM+B12 medium with either ANTQ or acetate at 0.1 g·L^{-1}, as sole carbon sources. Each condition was incubated in triplicate. Cultures were grown until mid-exponential phase (54h of incubation), at 150 rpm and 25ºC. Cells were harvested by centrifugation and snap-frozen in liquid nitrogen for subsequent total RNA

extraction. Total RNA was isolated by a method combining the Trizol™ reagent (Invitrogen, Carlsbad, CA, USA) and the RNeasy Mini kit (QIAGEN, Germany). In brief, four microliters of β-mercaptoethanol and 0.4 ml of buffer RLT were added to 1 ml of TrizolTM containing the bacterial pellet. The mixture was transferred to a tube containing 0.8 g of glass beads (diameter 0.1 mm), followed by three times of bead beating for 1 min at 5.5 m·s^{-1} with ice cooling steps in between. Subsequently, 0.2 ml of ice-cold chloroform was added. The solution was mixed gently followed by centrifugation at 12,000 x g for 15 min at 4ºC. RNA isolation was continued with the RNA clean-up using the RNeasy Mini kit according to the manufacturer's instructions. Genomic DNA was removed by an on-column DNase digestion step during RNA purification (DNase I recombinant, RNase-free, Roche Diagnostics, Germany). RNA concentration was measured by Qubit™ (Thermo Fisher, USA) and RNA quality was assessed by Qsep Bioanalyzer (BiOptic Inc., Taiwan).

Ribosomal RNA removal, library preparation and Illumina NovaSeq 6000 2x150bp sequencing was performed by Novogene Co. (Beijing, China). Raw read data quality was checked using FastQC (v0.11.8) and trimmed with Trimmomatic (v0.38). Trimmed reads were mapped against the genome sequence of *Sphingobium* sp. AntQ-1 using bwa-mem (v0.7.17) and quality of the mapping was analyzed with the flagstat command of samtools (v1.9). Mapped reads were quantified with the featureCounts function of the Subread package (v2.0.1). Normalization of counts and differential gene expression analysis was conducted with DESeq2 (Love et al., 2014). All software was used with default settings. Transcript abundance was defined as transcripts per million (TPM), normalized by gene length and sequencing depth.

Quantitative PCR analysis

Validation of the RNA-Seq analysis and quantification of relevant genes in sand-in-liquid and soil samples was achieved by quantitative PCR (qPCR) analysis. RNA samples from the RNA-Seq experiment were retrotranscribed to cDNA with the High Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific). qPCR reactions (20 μL) were carried out on an Applied Biosystems StepOnePlus Real-time PCR system using PowerUp Sybr Green Master Mix (Applied Biosystems, CA, USA), 1 μL of template and 4 pmol of each primer. The primers used for this study are listed in Table 2.1. All primers were designed using Primer-BLAST and experimentally validated by PCR. The relative expression level of the genes was calculated using the $2^{-\Delta\Delta CT}$ method (Schmittgen & Livak, 2008). For quantification, 6-point 10-fold standard plasmid dilution series were used. Standard plasmids were obtained by cloning PCR amplification products with pGEM-T Easy Vector System

(Promega, WI, USA). Plasmids were purified with the GeneJET Plasmid Miniprep Kit (Thermo Scientific, USA), quantified using Qubit and validated by sequencing.

Table 2.1. List of designed primers used for qPCR analysis.

Target	Primer name	Sequence (5'>3')	Amplicon size (bp)	Annealing temp (ºC)
16S rRNA	sphantq-F	GGCCCGCTGGACAAGTATT	281	55
Sphingobium sp. AntQ-1	sphantq-R	GGAAATCGCGATGAGGATGTCA		
sphantq_4473	4473-F	AGTCCCGCACGAATTGTCAT	147	55
	4473-R	ATGGGCACATCGAAAAAGCG		
sphantq_4474	4474-F	ATGAAGCCAACCCTGCACTT	102	55
	4474-R	TCTGCGCTGCCTTATACTCC		
sphantq_4479	4479-F	GAAGCATCGGGCGTCAAAAG	96	55
	4479-R	CGCGAAGAGAGTCATACCCG		
sphantq_4480	4480-F	CGCGTTGCCCTTTATTTCCA	174	55
	4480-R	ATATGCCGCGATGCTGTCTT		
sphantq_4492	4492-F	TCGTCGCGCCTTATACTGTC	110	55
	4492-R	ATGCCACGCAACCCAATAGA		
sphantq_3328	3328-F	GGACTCGCACGGCTTCTATT	167	55
	3328-R	CCTTCATCAGCGTGTCGGAT		
sphantq_4465	4465-F	AACTGTACCGCAAACCTCCC	121	55
	4465-R	GCCACGACGAAATCGAACAC		

Results and discussion

Anthracene and 9,10-anthraquinone biodegradation in sand-in-liquid soil microcosms inoculated with a PAH-contaminated soil

Biodegradation of ANT and ANTQ was assessed in small-scale sand-in-liquid microcosms spiked with either of the two substrates and inoculated with the creosote-contaminated soil. A significant removal of ANT was observed after 25 days of incubation (70%) (Figure 2.7), showing first-order kinetics from day 5 and reaching maximum rates between 15 and 25 days (16.7 ± 0.9 μmol·L^{-1}·day^{-1}). On the other hand, ANTQ was completely removed at the end of the incubation. The oxy-PAH was degraded steadily during the first 10 days (13.5 ± 0.7 μmol·L^{-1}·day^{-1}) and, after a 10-day plateau, its concentration dropped with maximum rates during the last 5 days of incubation (19.3 ± 0.3 μmol·L^{-1}·day^{-1}). No formation of ANTQ was observed in the ANT-spiked microcosms.

Bacterial community response to ANT and ANTQ was evaluated by DGGE analysis of 16S rRNA genes and transcripts after 0, 15 and 25 days of incubation (Annex - Figure S2.1, Table S2.1). Shifts in total (DNA) and active (RNA) bacterial populations

were observed in both conditions. The DGGE profiles from the ANT-microcosms were very similar to those from time 0 samples but presented a substantial increase in the intensity of a specific band identified as a member of the genus *Sphingobium* (99% probability), suggesting its role in the biodegradation of anthracene (band B1). In the ANTQ-microcosms, there was a remarkable change in the DGGE profiles as most bands observed at 0 days or in ANT-microcosms showed much lower intensity, while a new predominant band (band B2) appeared in both total and active populations at all timepoints. Sequence analysis of band B2 also revealed its affiliation to the genus *Sphingobium* (100% probability) but was only 97% similar to that of band B1, indicating that different phylotypes were associated to ANT and ANTQ degradation. None of the major ANTQ degraders were related to ANT, suggesting that anthracene-degrading bacteria might not possess the metabolic capability to grow on anthraquinone, and indicating the existence of populations specialized in the removal of the quinone. This is in agreement with the findings from Rodgers-Vieira et al. (2015), that identified different bacteria involved in the assimilation of anthraquinone or anthracene in a PAH-contaminated soil utilizing DNA-based stable isotope probing (SIP). They associated *Sphingomonas* and *Phenylobacterium* species with anthraquinone degradation and members of *Immundisolibacter* and *Sphingomonadales* were identified as the main anthracene degraders.

Sphingobium sp. AntQ-1, a 9,10-anthraquinone degrading strain

Sphingobium sp. AntQ-1 was isolated from the ANTQ sand-in-liquid microcosms using MM+B12+YE agar plates with ANTQ as main carbon source. The strain was selected based on the observation of clearing areas around its colonies indicating anthraquinone degradation, after 10 days of incubation. Phylogenetic analysis based on the 16S rRNA gene sequence using RDP Seqmatch indicated that strain AntQ-1 was most closely related to the type strain *Sphingobium rhizovicinum* CC-FH12-1 (98.7% identity). BLAST search on GenBank database identified type strains *Sphingobium czechense* LL01 (99.5%) and *Sphingobium cupriresistens* CU4 (99.2%) as its closest relatives. The 16S rRNA gene sequence retrieved from strain AntQ-1 was identical to that of band B2 from the DGGE analysis of the sand-in-liquid soil microcosms spiked with ANTQ, confirming their correspondence. Scanning electron microscopic observations of ANTQ-grown *Sphingobium* sp. AntQ-1 cells revealed it formed rods of approximately 0.9 to 4.5 µm long and 0.4 to 0.7 µm wide, some of them aggregating around ANTQ crystals.
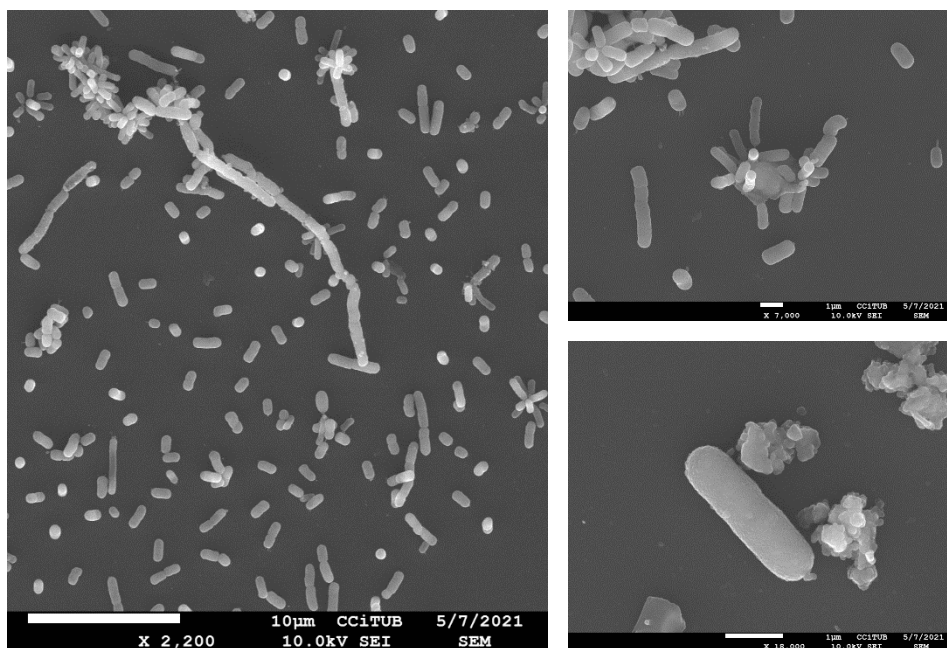
Figure 2.1. Scanning electron microscopy (SEM) images of *Sphingobium* sp. AntQ-1 cells grown with ANTQ as sole carbon and energy source. Crystals of ANTQ are also visualized.

Utilization of 9,10-anthraquinone as sole carbon and energy source by strain AntQ-1 was demonstrated by its removal from liquid MM+B12 cultures along with a concomitant increase in cell protein (Figure 2.2). ANTQ was almost completely removed after 6 days of incubation (97%) when maximum protein concentration was achieved (53 ± 8.1 µg·mL^{-1}). To the best of our knowledge, this is the first reported bacterial isolate with the ability to grow on 9,10-anthraquinone as the sole source of carbon and energy.

A variety of other polyaromatic compounds and known PAH acidic metabolites were tested as growth substrates in MM+B12 cultures. Significant growth was only observed with anthrone (33.8 ± 1.7 µg protein·mL^{-1}), 2-methylanthraquinone (56.8 ± 1.9 µg·mL^{-1}), catechol (12.8 ± 5.3 µg·mL^{-1}), protocatechuic acid (51.8 ± 6.4 µg·mL^{-1}) and benzoic acid (40.2 ± 5.6 µg·mL^{-1}). The strain also grew on MM+B12 with acetate, lactate, pyruvate or glucose. Growth was never observed in MM in the absence of vitamin B12, confirming an auxotrophy for this vitamin. It is worth noting that the strain was not able to grow on ANT, the parent unsubstituted PAH precursor of anthraquinone and anthrone. In fact, this specialization for oxy-PAHs
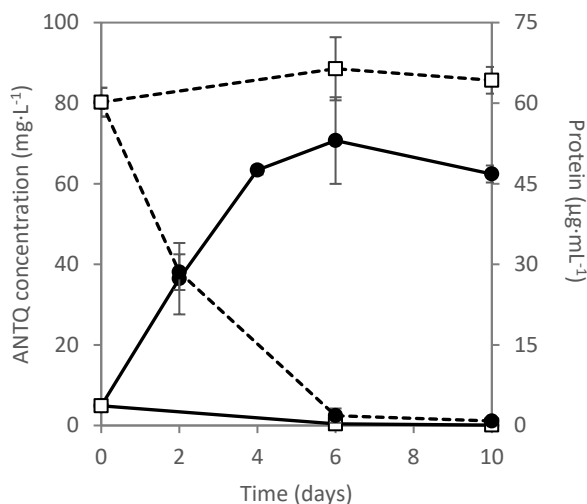
Figure 2.2. Utilization of 9,10-anthraquinone by *Sphingobium* sp. strain AntQ-1 in liquid mineral medium supplemented with vitamin B12 with anthraquinone as the sole source of carbon and energy. Continuous lines correspond to protein concentration with (•) or without (□) ANTQ, and dashed lines correspond to ANTQ concentration in cultures with (•) or without inoculum (□).

was also observed for a 9-fluorenone-degrading strain of *Pseudomonas mendocina,* MC2, isolated from PAH-contaminated river and unable to grow on fluorene or any other PAH (Casellas et al., 1998). This again indicates the existence of populations highly specialized in oxy-PAH utilization in PAH-contaminated soils and sediments.

Detection and identification of 9,10-anthraquinone metabolites

The neutral and acidic extracts from the supernatants of washed-cell suspensions of strain AntQ-1 incubated with ANTQ weighted 2 mg and 3 mg, respectively. Extracts were analyzed by HPLC, and further identification of detected metabolites was achieved by GC-MS and HPLC-ESI(+)-HRMS (Table 2.2). HPLC analysis of both the neutral and acidic extracts revealed a major metabolite at a retention time of 16.5 min, showing a UV-visible spectrum with λ_{max} at 210, 258 and 330 nm. In the HPLC-ESI(+)-HRMS analysis, metabolite I produced the molecular formula $C_{14}H_{11}O_4$ ([M+H]$^+$=243.0694), showing ion fragments at *m/z* 243.0649 (70), 225.0545 (100) and 149.0236 (25). In the GC-MS analysis, metabolite I was detected in the diazomethane treated neutral (42.9%) and acidic (89%) extracts, showing a mass spectrum compatible with a methyl ester derivative of 2-(2-hydroxybenzoyl)-benzoic acid by comparison with the NIST library. The underivatized neutral

Table 2.2. GC-MS Rt and EI mass spectral properties and HPLC-MS Rt and detected mass accuracies of major compounds formed from 9,10-anthraquinone by *Sphingobium* sp. strain AntQ-1 in neutral (N) and acidic (A) extracts.

	Extract	GC-MS				HPLC-MS			
		Rt (min)	Abundance (%)	*m/z* of fragment ions (% relative intensity)	Identification	Rt (min)	[M+H]+	Molecular formula	δ (mmu)
I	N A	26.2	42.9 89.0	256 (M+, 39), 224 (100), 196 (83), 168 (52), 163 (12), 152 (5), 139 (29), 121 (33), 115 (12), 105 (8), 92 (16), 77 (22), 65 (25), 50 (8)	2-(2-hydroxybenzoyl)-benzoic acid ME	15.3	243.0649 243.0651	$C_{14}H_{11}O_4$	-0.2540 -0.1117
II	N	26.7	21.9	224 (M+, 100), 196 (70), 168 (79), 152 (5), 139 (77), 126 (2), 114 (8), 104 (10), 92 (12), 76 (39), 63 (23), 50 (23)	dibenz[*b,e*]oxepin-6,11-dione	n.d.			
III	A	18.1	11.0	194 (M+, 6), 163 (100), 149 (1), 133 (6), 120 (2), 104 (5), 92 (8), 77 (20), 64 (2), 50 (6)	phthalic acid diME	2.1	149.0234	$C_8H_5O_3$	0.3102
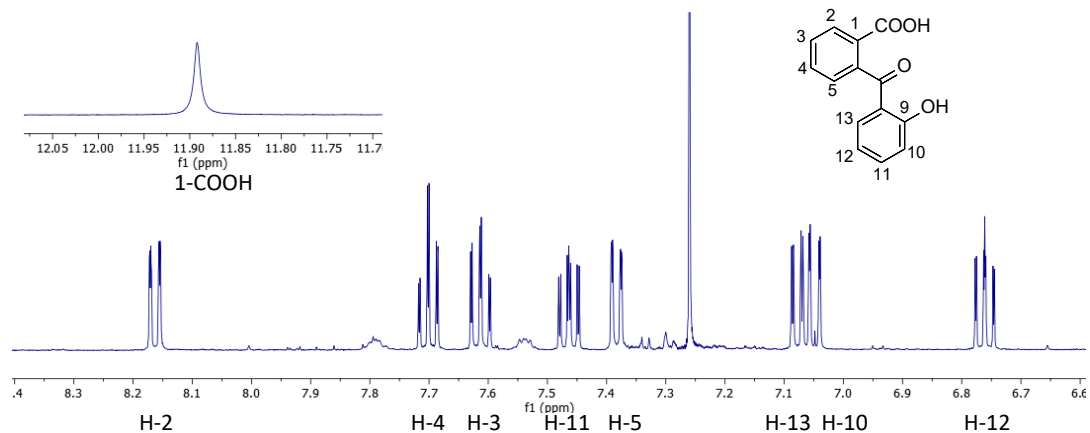
Figure 2.3. Aromatic region of the ^{1}H NMR spectrum of metabolite I.

extract, however, presented a major product (II) with a MS spectrum compatible with dibenz[*b,e*]oxepin-6,11-dione (21.9%). These results suggest either that a portion of metabolite I is extracted in neutral conditions and recircularized in the GC, or that products I and II are in equilibrium in the culture.

To confirm the identification of metabolite I as 2-(2-hydroxybenzoyl)-benzoic acid, the acidic extract was analyzed by NMR, as 89% (2.7 mg) of the total extract corresponded to this product. The monodimensional ^{1}H NMR spectrum of this extract was in perfect agreement with the suggested structure. The coupling pattern shown for all signals permitted the univocal assignment of signals in this spectrum to the diverse proton atoms in the structure (Figure 2.3; Annex - Table S2.2). Nevertheless, additional bidimensional gCOSY and gHSQC spectra were also obtained (Annex - Figures S2.2 and S2.3, respectively). The gCOSY experiment confirmed the assignment of the monodimensional ^{1}H NMR spectra. Given high demand of extract required for the acquisition of a ^{13}C NMR spectra, the gHSQC allowed us to determine the chemical shifts corresponding to the protonated carbon atoms in the molecule. These values were also in good agreement with the suggested structure.

The identified 2-(2-hydroxybenzoyl)-benzoic acid has been previously reported in the biodegradation of anthracene via anthraquinone for ligninolytic fungi and was attributed to the action of manganese peroxidases (Cajthaml et al., 2002; Baborová et al., 2006). In the case of bacteria, the cleavage of the central ring in aromatic quinones has been observed for anthraquinone analogues or other oxy-PAHs. For instance, *Sphingomonas xenophaga* strain QYY converted 1-aminoanthraquinone-2-sulfonic acid, a substituted anthraquinone used as a dye intermediate, to two isomers of the equivalent ring-fission acid and phthalic acid (Lu et al., 2008). A similar pathway was observed for the transformation of benz(*a*)anthracene via benz(*a*)anthracene-7,12-dione by *Mycobacterium vanbaalenii* PYR-1, with further oxidation of the quinone leading to cleavage of the central ring (Moody et al., 2005), and for the degradation of fluorene via 9-fluorenone by *Sphingomonas* sp. LB126 (Wattiau et al., 2001).

In addition to metabolite I, the acidic extract presented a less abundant metabolite (III) that was identified as phthalic acid on the basis of its spectral characteristics and by comparison with authentic material. This product was detected as dimethyl ester derivative of phthalic acid in GC-MS analysis of the diazomethane treated acidic extract (11%), whereas in HPLC-ESI(+)-HRMS analysis the detected accurate mass ([M+H][+]=149.0234) corresponded to phthalic anhydride. This was the only metabolite detected during growth of strain AntQ-1 on anthraquinone, but its accumulation was only transient (data not shown).

Interestingly, when incubated with anthrone, the ketone derivative of anthracene, *Sphingobium* sp. AntQ-1 produced the exact same metabolites in addition to the formation of anthraquinone (Annex - Table S2.3). This indicates that the strain is able to oxidize the central ring of anthrone to produce anthraquinone which would be further metabolized in the same manner. What remains unclear is why the strain is not able to perform the same oxidation on the unsubstituted parent PAH. A plausible explanation would be that the central ring of ANT is aromatic, whereas in anthrone it is not.

Genomic and transcriptomic analysis of *Sphingobium* sp. AntQ-1

The complete genome sequence of the strain *Sphingobium* sp. AntQ-1 was obtained combining Illumina and PacBio sequencing technologies. Illumina sequencing yielded 7,612,734 paired-end reads of 150 bp with a total average Phred quality score of 35, resulting in a sequencing coverage over 200x. For PacBio, a total of 183,593 reads were generated, with an average read length of 8544 bp and N50 9944 bp. The final assembly had a total length of 4,904,521 bp, with an average G+C content of 63.6%, consisting of 8 circular contigs corresponding to 2 chromosomes and 6 plasmids: *chr1* (3,758,784 bp, 64.0% G+C), *chr2* (720,548 bp, 63.0% G+C), *pANTQ-1* (201,917 bp, 60.6% G+C), *pl2* (80,951 bp, 61.6% G+C), *pl3* (53,641 bp, 60.6% G+C), *pl4* (51,841 bp, 63.5% G+C), *pl5* (20,675 bp, 58.9% G+C) and *pl6* (16,164 bp, 57.3% G+C) (Annex - Figure S2.4). The annotated genome contained 4741 coding DNA sequences (CDSs), 63 tRNA genes and 12 rRNA genes.

In pursuance of elucidating the key functional genes associated to the biodegradation pathway of ANTQ, we performed a comparative transcriptomic analysis by RNA-Seq in cultures of strain AntQ-1 on ANTQ versus cultures in acetate as a control condition (Annex - Figure S2.5). RNA-Seq analysis was done on three biological replicates of distinct rRNA-depleted RNA preparations, for each growth condition. The number of clean paired-end reads averaged 7.6 and 9.8 million and the percentage of reads mapping to the genome was over 98.6% and 98.8% for all replicates in ANTQ and acetate conditions, respectively. The percentage of reads assigned to genomic features was over 83.8% for ANTQ samples and over 88.7% for acetate. Differential expression analysis yielded a total of 180 up-regulated genes in the presence of ANTQ whereas only 45 genes were differentially expressed with acetate, considering log2 fold change ≥1.5 or ≤-1.5 and p-value ≤0.01 (Figure 2.4, Annex – Tables S2.4 and S2.5). Up-regulated genes presented a fold change between 1.5 and 5.9 and down-regulated genes between -1.5 and -5.1.

Among the genes that were most highly up-regulated during growth on ANTQ (Table 2.3), a gene encoding a luciferase-like flavin-dependent monooxygenase
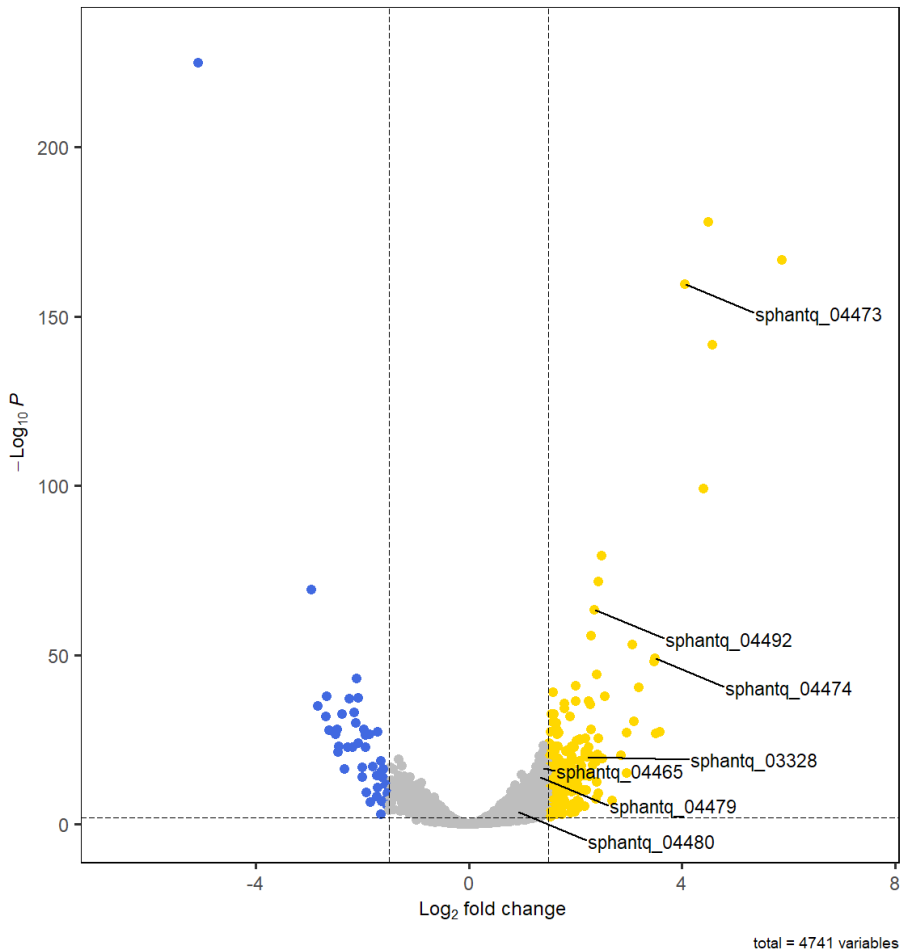
Figure 2.4. Volcano plot displaying the distribution of differentially expressed genes between acetate (blue) and anthraquinone (yellow) grown cultures of *Sphingobium* sp. strain AntQ-1. Positive log2 fold change indicates upregulation in anthraquinone cultures. The horizontal dashed line indicates adjusted p-value threshold = 0.01. The vertical dashed lines indicate log2 fold change threshold = 1.5.

(sphantq_4473) presented the highest transcript per kilobase million (TPM, 21,461) count, with a fold-change of 4.05. Contiguous to it, a gene coding for a hydrolase-like protein (sphantq_4474) presented a 3.49-fold change and 2,570 TPM. Both of these genes were located in the megaplasmid *pANTQ-1* (Figure 2.5). The significant overexpression of this cluster of genes in the presence of ANTQ suggests their essential role in the biodegradation of this oxy-PAH. Based on the identified metabolites and the annotation of this cluster of genes, it seems that the attack on ANTQ would be initiated by a Baeyer-Villiger oxidation and a subsequent

hydrolyzation. Thus, gene sphantq_4473 would probably encode a Baeyer-Villiger monooxygenase (BVMO), as these are flavoenzymes that catalyze the BV oxidation of ketones and cyclic ketones to esters or lactones by inserting one atom of molecular oxygen into a C-C bond adjacent to the carbonyl group, and using NAD(P)H as the reducing agent (Tolmie et al., 2019). The presence of a bacterial luciferase-like protein domain indicates that it would probably be classified as a type II BVMO, that consists of two distinct polypeptides, the oxygenating component that binds FMN as cofactor and the reducing component that uses NADH as cofactor. To date, the only reported type II BVMOs are two diketocamphane monooxygenases from *Pseudomonas putida* ATCC17453, involved in camphor metabolism (Iwaki et al., 2013). Blastp analysis revealed that the amino acid sequence of sphantq_4473 is most closely related to flavin-dependent oxidoreductases belonging to the genus *Ramlibacter* (42-39% identity). The hydrolase encoding gene sphantq_4474 was initially annotated as a hypothetical protein, however, InterProScan analysis showed the presence of an alpha/beta hydrolase_6 protein domain and its amino acid sequence is closely related to an alpha/beta hydrolase of *Bradyrhizobium* sp. WSM3938 (61% identity).

Within plasmid *pANTQ-1*, located only 3kb away from this gene cluster but on the reverse strand, another cluster composed by a gene coding for a BVMO (sphantq_4479) and a gene encoding a monoterpene ε-lactone hydrolase (sphantq_4480) revealed significant upregulation. Although the expression level was considerably lower than that of the first BVMO-hydrolase cluster, upregulation in ANTQ cultures could indicate a relevant role in the biodegradation of this compound. This BVMO displayed type I BVMO hallmarks: two Rossman motifs (GxGxx[G/A]) flanking two BVMO fingerprints ([A/G]GxWxxxx[F/Y]P[G/M]xxxD and FxGxxxHxxxW[P/D]) and shared sequence homology with NAD(P)/FAD-dependent oxidoreductases of diverse sphingomonads (66-67% identity). Type I BVMO is the most extensively studied class of BVMOs for their potential use as biocatalysts in the biosynthesis of secondary metabolites with antibiotic or anti-cancer activity, but also for their ability to catalyze key reactions in metabolic pathways of a wide range of carbon sources such as linear alkanones or alkanals, terpenoids or alicyclic and aromatic ketones (Leisch et al., 2011).

As expected from the detection of phthalate as one of anthraquinone metabolites, genes encoding for the complete phthalic acid metabolic pathway were overexpressed in ANTQ cultures. Significantly higher expression levels were observed for the *pht2345* genes, involved in the initial steps leading to the formation of protocatechuate, with fold changes ranging from 1.59 to 2.36. Most of the genes catalyzing the lower pathway (*ligABCI* and *galD*) were duplicated in the genome of

Table 2.3. Up-regulated genes related to the biodegradation of 9,10-anthraquinone by *Sphingobium* sp. AntQ-1.

ID	Name	Description	Location	Length (bp)	log$_2$ fold change	Adjusted p-value	TPM
Upper ANTq pathway							
4473		Luciferase-like flavin-dependent monooxygenase	*pANTQ-1*	1041	4.05	2.7E-157	21461
4474		alpha/beta hydrolase_6	*pANTQ-1*	1125	3.49	3.0E-47	2570
4479		Baeyer-Villiger monooxygenase	*pANTQ-1*	1620	1.32	2.2E-13	80
4480		Monoterpene epsilon-lactone hydrolase	*pANTQ-1*	912	0.91	9.4E-04	7.2
Phthalate metabolism							
4492	*pht3*	Phthalate 4,5-dioxygenase oxygenase subunit	*pANTQ-1*	1272	2.36	2.1E-61	4782
4509	*pht2*	Phthalate 4,5-dioxygenase reductase subunit	*pANTQ-1*	963	1.59	2.3E-37	3965
4491	*pht4*	Phthalate 4,5-*cis*-dihydrodiol dehydrogenase	*pANTQ-1*	1197	2.30	8.7E-54	276
4512	*pht5*	4,5-dihydroxyphthalate decarboxylase	*pANTQ-1*	993	1.80	4.7E-34	2947
3931	*ligA*	Protocatechuate 4,5-dioxygenase alpha chain	*chr2*	456	2.08	8.2E-24	117
3930	*ligB*	Protocatechuate 4,5-dioxygenase beta chain	*chr2*	897	2.28	6.3E-34	105
4497	*ligA*	Protocatechuate 4,5-dioxygenase alpha chain	*pANTQ-1*	210	0.89	4.7E-02	37
4498	*ligB*	Protocatechuate 4,5-dioxygenase beta chain	*pANTQ-1*	843	0.86	6.9E-05	27
4435	*chqB*	Hydroxyquinol 1,2-dioxygenase	*pANTQ-1*	918	1.05	5.2E-06	47
4436	*macA*	Maleylacetate reductase	*pANTQ-1*	1065	0.72	3.4E-03	22
3929	*ligC*	4-carboxy-2-hydroxymuconate-6-semialdehyde dehydrogenase	*chr2*	942	1.80	1.0E-32	109
4499	*ligC*	4-carboxy-2-hydroxymuconate-6-semialdehyde dehydrogenase	*pANTQ-1*	942	0.18	4.7E-01	69
3937	*ligI*	2-pyrone-4,6-dicarboxylate hydrolase	*chr2*	873	1.58	1.2E-20	62
4495	*ligI*	2-pyrone-4,6-dicarboxylate hydrolase	*pANTQ-1*	888	0.96	1.8E-05	16
3936	*galD*	4-oxalomesaconate tautomerase	*chr2*	1062	1.81	5.3E-17	29
4496	*galD*	4-oxalomesaconate tautomerase	*pANTQ-1*	219	0.72	1.1E-01	7.9
3932	*ligJ*	4-oxalomesaconate hydratase	*chr2*	1026	2.51	1.9E-18	97
3935	*ligK*	4-carboxy-4-hydroxy-2-oxoadipate aldolase	*chr2*	675	1.97	1.6E-21	73

Catechol metabolism

4465	*xylE*	Catechol 2,3-dioxygenase	*pANTQ-1*	924	1.38	1.3E-15	7055
4463	*xylF*	2-hydroxymuconate semialdehyde hydrolase	*pANTQ-1*	843	1.45	1.6E-15	301
4468	*xylG*	2-hydroxymuconate semialdehyde dehydrogenase	*pANTQ-1*	1458	1.41	4.9E-22	2252
4472	*xylH*	2-hydroxymuconate tautomerase	*pANTQ-1*	603	1.57	4.5E-14	1782
4471	*xylI*	4-oxalocrotonate decarboxylase	*pANTQ-1*	771	1.53	6.1E-26	1385
4469	*mhpD*	2-oxopent-4-enoate hydratase	*pANTQ-1*	807	1.50	8.6E-23	2830
4432	*mhpE*	4-hydroxy-2-oxovalerate aldolase	*pANTQ-1*	1047	4.49	2.9E-175	360
4433	*mhpF*	Acetaldehyde dehydrogenase	*pANTQ-1*	945	4.57	2.8E-139	533
3328	*catA*	Catechol 1,2-dioxygenase	*chr1*	927	2.21	8.7E-19	247
3326	*catB*	Muconate cycloisomerase	*chr1*	1158	2.43	4.7E-24	134
3327	*catC*	Muconolactone Delta-isomerase	*chr1*	291	2.08	1.6E-13	118
3335	*pcaD*	3-oxoadipate enol-lactonase	*chr1*	771	2.10	1.9E-13	36
3300	*pcaI*	3-oxoadipate CoA-transferase subunit alfa	*chr1*	687	-0.03	9.3E-01	5.3
4437	*pcaI*	3-oxoadipate CoA-transferase subunit alfa	*pANTQ-1*	585	1.05	2.6E-12	283
3336	*fadA*	acetyl-CoA acyltransferase	*chr1*	1206	1.90	1.3E-12	35
4451	*fadA*	acetyl-CoA acyltransferase	*pANTQ-1*	1257	1.05	2.1E-10	40
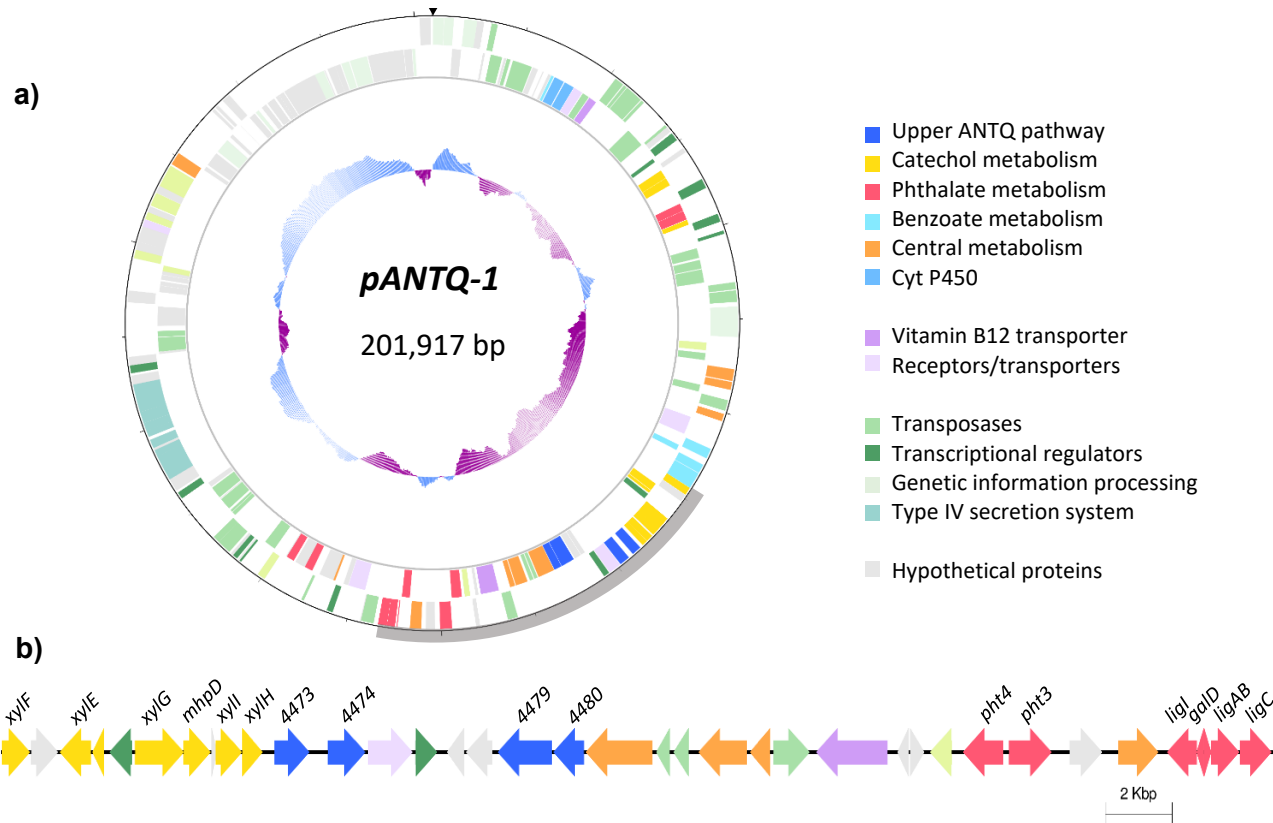
Figure 2.5. (a) Structure of plasmid *pANTQ-1* in *Sphingobium* sp. AntQ-1, encoding most of the genes related to 9,10-anthraquinone biodegradation pathway. Going inwards, the first two circles denote genes in the forward and reverse strands colored based on gene function classification. The inner circle represents GC content (above average, in blue, and below average, in purple). Dashes in the outermost circle are placed every 10 Kbp (b) Arrangement of genes related to 9,10-anthraquinone biodegradation in plasmid *pANTQ-1* of *Sphingobium* sp. strain AntQ-1 between the positions 68745 and 106725 bp, highlighted in grey in the plasmid map.

Sphingobium sp. AntQ-1, one copy being located in the secondary chromosome *chr2* and the other in the megaplasmid *pANTQ-1*. Despite this, overexpression was primarily detected for the genes located in the secondary chromosome, possibly, because of the lack of *ligJK* genes in the plasmid, the rest of the genes involved in the pathway could not be efficiently expressed.

Surprisingly, up-regulation of genes coding for both the complete catechol *meta*- and *ortho*-cleavage pathways was remarkably high in the presence of anthraquinone. The *meta*-cleavage pathway, catalyzed by *xylEFGHI* and *mhpDEF* genes encoded in the plasmid *pANTQ-1*, presented fold-change levels between 1.38 and 4.57 along a significant TPM count. On the other hand, genes coding for the *ortho*-pathway, including *catABCD, pcaDI and fadA*, were found in the primary chromosome *chr1* and showed lower expression levels, although fold-change of the *catABCD* genes was relatively high (2.1-2.4). This would be the second report for the catabolism of catechol by both extra- and intradiol ring cleavage in sphingomonads as Maeda and colleagues (2020) revealed the first evidence for intradiol cleavage of catechol by chemical and genomic analysis of low molecular weight PAHs biodegradation in *Sphingobium barthaii* KK22.

It is worth mentioning that a gene coding for a BtuB vitamin B12 transporter (sphantq_4487) was significantly upregulated in the presence of ANTQ, with a 1.87-fold change and 3,413 TPM. This gene is also located in plasmid *pANTQ-1* in the same region as the two BVMO-hydrolase clusters. Reconstruction of the biosynthetic pathway of vitamin B12 in the genome of *Sphingobium* sp. AntQ-1 revealed its incompleteness as it lacked all the genes for the biosynthesis of the corrin ring, corroborating the auxotrophy of strain AntQ-1 for vitamin B12 and the need for external uptake. The overexpression of the B12 transporter coding gene along the inability of the strain to grow without the presence of this vitamin demonstrates it is an essential co-factor for strain AntQ-1.

Transcription levels of selected differentially regulated genes involved in the biodegradation of anthraquinone were subsequently quantified by RT-qPCR to confirm the RNA-Seq data. The selected genes were: sphantq_4473, sphantq_4474, sphantq_4479, sphantq_4480, *pht3* (sphantq_4492), *xylE* (sphantq_4465) and *catA* (sphantq_3328). The gene expression fold-change between RNA-Seq and RT-qPCR demonstrated a good correlation with almost identical trends for each selected gene (Annex – Figure S2.5), demonstrating the accuracy of the RNA-Seq analysis.

Polyomic reconstruction of the 9,10-anthraquinone degradation pathway

The combination of metabolomic, genomic and transcriptomic data provided a deeper understanding of the ANTQ metabolic pathway in *Sphingobium* sp. AntQ-1
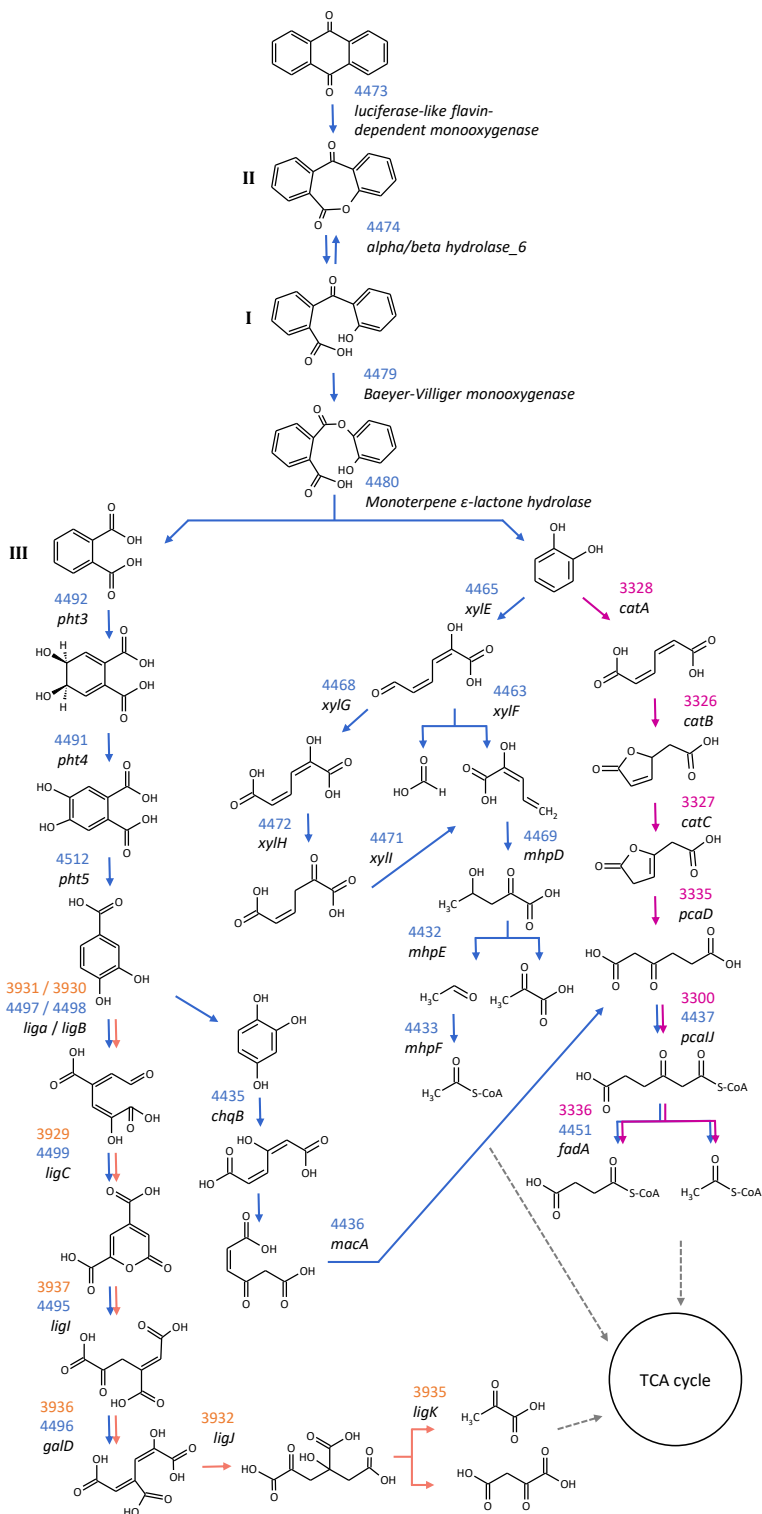
◁ **Figure 2.6.** Schematic pathway proposed for the degradation of 9,10-anthraquinone by *Sphingobium* sp. strain AntQ-1. Pink arrows correspond to reactions catalyzed by enzymes encoded in genes located in the chromosome *chr1*; orange arrows for genes in chromosome chr2; blue arrows for genes in plasmid *pANTQ-1*. Gene IDs are indicated adjacent to the corresponding reactions following the same chromatic criterion.

(Figure 2.6). The identification of dibenz[b,e]oxepin-6,11-dione, 2-(2-hydroxybenzoyl)-benzoic acid and phthalic acid as the major metabolites from anthraquinone degradation suggests that *Sphingobium* sp. AntQ-1 metabolizes anthraquinone by cleavage of the central ring before degradation of the flanking aromatic rings to intermediates of the central metabolism. The attack would be initiated by a Baeyer-Villiger oxidation producing dibenz[b,e]oxepin-6,11-dione, a lactone that would be eventually hydrolyzed resulting in the formation of 2-(2-hydroxybenzoyl)-benzoic acid. Thus, the overexpressed cluster of genes identified in the RNA-Seq analysis would orchestrate the initial attack of ANTQ, catalyzing a Baeyer-Villiger oxidation (sphantq_4473) and an eventual hydrolyzation of the resulting lactone (sphantq_4474).

Evidence of the formation of a lactone by Baeyer-Villiger (BV) oxidation has been documented for *Streptomyces aureofaciens* B-96, that transformed 1,8-dihydroxy-9,10-anthraquinone to give the corresponding lactone (Cudlín et al., 1978). Similar Baeyer-Villiger reactions have also been suggested from the identification of metabolites during the degradation of fluoranthene via acenaphthenone by *Alcaligenes denitrificans* WW1 (Weissenfels et al., 1991) and fluorene via indanone by *Arthrobacter* sp. F101(Grifoll et al., 1992; Casellas et al., 1997) . *Pseudomonas* sp. F274 and *Pseudomonas mendocina* MC2 also produced a similar lactone during the degradation of 9-fluorenone, however its formation was attributed to the cleavage of the central ring following an angular dioxygenation (Grifoll et al., 1994; Casellas et al., 1998).

The presence of a second cluster of BVMO-hydrolase coding genes and overexpression of both phthalate and catechol metabolic pathways, suggests a second BV oxidation and hydrolyzation of the detected 2-(2-hydroxybenzoyl)-benzoic acid that would lead to the formation of phthalate and catechol. In fact, catechol and an aminosulfo-derivative of phthalic acid were detected as biodegradation products from 1-aminoanthraquinone-2-sulfonic acid by a *Rhodococcus pyridinivorans* strain with the ability to transform a variety of anthraquinone derivatives (Lu et al., 2019).

Interestingly, essential genes for the biodegradation pathway of anthraquinone are primarily located in the megaplasmid *pANTQ-1* (Figure 2.5). In particular, plasmid

pANTQ-1 carries the genes related to the upper ANTQ pathway, genes for phthalate metabolism and for the *meta*-cleavage pathway of catechol. However, all of these genes are not organized in coordinately regulated operons as they are scattered across the plasmid and, even, some genes for the phthalate and the catechol *ortho*-cleavage pathways are located in the chromosomes. Megaplasmids (>100 kbp) are commonly found in sphingomonads carrying genes coding for catabolic pathways (Stolz, 2014) and it is usual that these are dispersed through sparse clusters in the genome (Pinyakong et al., 2003). Plasmid *pANTQ-1* also encompassed several genes coding for transposases and type IV secretion system components, raising the possibility for plasmid-mediated gene transfer. Even more, it has also been found that sphingomonads with the ability to degrade compounds that are converted to intermediates common to the naphthalene and biphenyl pathways are able to change the position and orientation of certain conserved gene clusters in their genomes (Basta et al., 2005). All these characteristics give sphingomonads the ability to adapt quickly and efficiently to novel compounds in the environment conferring them exceptional degradative capabilities for a wide range of contaminants. *Sphingobium* sp. AntQ-1, however, utilizes a few aromatic substrates which could be explained by the highly specific enzyme system it possesses. BVMOs are known for their exquisite chemo-, regio- and enantioselectivity for their substrates (Mascotti et al., 2015). If biodegradation of oxy-PAHs relies on highly specialized BVMOs, it could be the reason for their wide environmental distribution and the often higher concentrations compared to their parent hydrocarbons (Bandowe & Wilcke, 2010).

Environmental relevance in PAH-contaminated soils

The relevance of *Sphingobium* sp. AntQ-1 and the described Baeyer-Villiger mechanism for the biodegradation of 9,10-anthraquinone in the creosote-contaminated soil was assessed by specific qPCR quantification of the 16S rRNA gene of the strain and the BVMO-coding gene sphtantq_4473. DNA and cDNA samples from the sand-in-liquid soil microcosms spiked with either ANT or ANTQ at 0, 15 and 25 days of incubation were used to evaluate the suitability of the designed primers for the quantification of the aforementioned genes and their transcripts (Figure 2.7). A significant increase of both gene and transcript copies of *Sphingobium* sp. AntQ-1 16S rRNA and BVMO sphantq_4473 was detected after 15 days of incubation in the presence of ANTQ with respect to time 0. The increment of bacterial community, respectively. In the case of BVMO gene and transcript copies, the levels increased in 4 and 2 orders of magnitude. It is noteworthy that the expression of BVMO during the incubation with ANTQ was in accordance with the biodegradation rates as it decreased after 25 days when ANTQ was completely
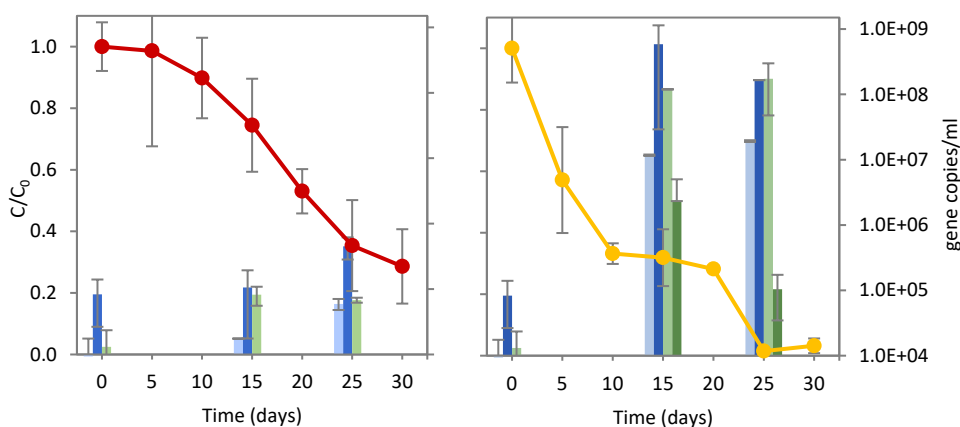
Figure 2.7. Biodegradation of anthracene (red) and 9,10-anthraquinone (yellow) in the sand-in-liquid soil microcosms and quantification of *Sphingobium* sp. AntQ-1 specific 16S rRNA (blue bars) and BVMO 4473 (green bars) gene copies (DNA, light color) and transcripts (cDNA, dark color) in the sand-in-liquid soil microcosms with either anthracene or anthraquinone at 0, 15 and 25 days of incubation.

removed. In the ANT-spiked microcosms, no expression was detected for BVMO. However, a small increase of BVMO gene copies and *Sphingobium* sp. AntQ-1 transcripts was detected, indicating that a fraction of ANT would be transformed to ANTQ, and channeled through the described ANTQ-degradation pathway.

To ascertain the importance of the proposed Baeyer-Villiger oxidation of anthraquinone *in situ*, we monitored the abundance of *Sphingobium* sp. AntQ-1 16S rRNA and BVMO-coding genes during a pilot-scale biostimulation assay of the creosote-polluted soil (Tauler, 2015). Biostimulation of the soil by means of aeration and nutrient amendment for 150 days lead to an important and sequential degradation of PAHs, and transient accumulation of oxy-PAHs. ANTQ was the predominant oxy-PAH formed, accounting for almost a half of the total quantified oxy-PAHs. At its peak of abundance, ANTQ concentration represented a 6% of ANT concentration. Considering the amount of ANTQ produced and the amount of ANT degraded at this timepoint, a 4.3% of ANT carbon was converted to ANTQ (17.8 mmol C·kg^{-1}). Remarkably, quantification of *Sphingobium* sp. AntQ-1 16S rRNA and its BVMO genes revealed a perfect consonance with the biodegradation kinetics of ANTQ in the creosote-contaminated soil, as their respective maxima in gene copy numbers were concomitant with maximum ANTQ biodegradation rates (Figure 2.8). This proves the environmental relevance of Baeyer-Villiger catalyzed biodegradation of oxy-PAHs in PAC-contaminated soils.
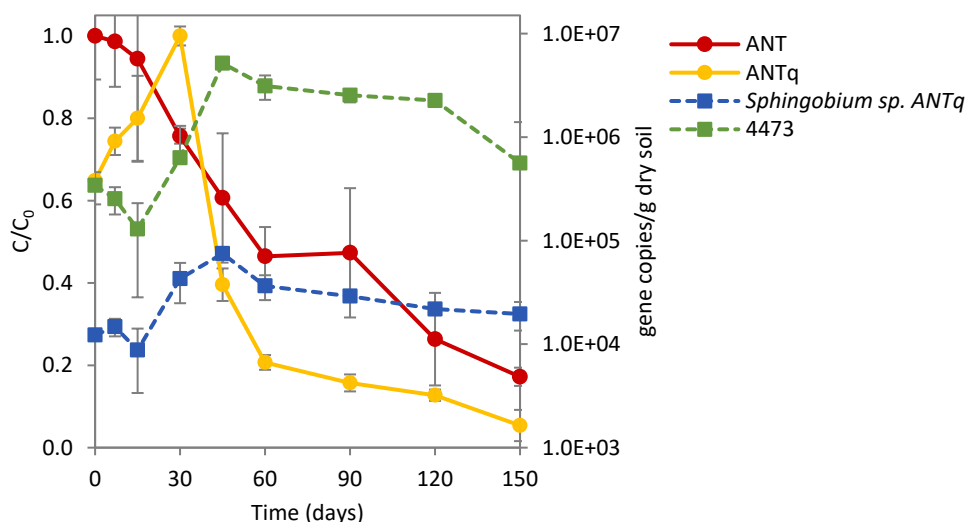
Figure 2.8. Evolution of Sphingobium sp. AntQ-1 specific 16S rRNA (blue) and BVMO 4473 (green) gene copies and the concentration of anthracene (red) and 9,10-anthraquinone (yellow) in the lab-scale biostimulation assay of the creosote-polluted soil.

Recent works have raised increasing awareness on the potential toxicity and persistence of transformation products of organic contaminants (Titaley et al., 2020), and reinforced the need to include such products in risk analysis. Specifically, during the remediation of PAH-contaminated soils, genotoxicity of the soil has been observed to eventually increase despite effective PAH removal, and this increased genotoxicity has been associated to oxy-PAH formation. Our work identifies the existence of highly specialized microbial communities in soils with the ability to degrade PAH transformation products, and elucidates the bacterial mechanisms involved in the biodegradation of 9,10-anthraquinone, a model oxy-PAH frequently identified in PAH-contaminated environments. Quantification of BVMOs by qPCR could be an appropriate method to assess the potential of native microbial communities to cycle PAH-transformation products, therefore providing evidence of the capabilities of native microbial communities to mitigate the risk associated to oxy-PAH formation during remediation.

Chapter 3

Metagenomic insights into the microbial networks of a benz(a)anthracene-7,12-dione degrading community from a creosote-contaminated soil

Metagenomic insights into the microbial networks of a benz(*a*)anthracene-7,12-dione degrading community from a creosote-contaminated soil

Introduction

Oxygenated polycyclic aromatic hydrocarbons (oxy-PAHs) have become a matter of concern as they often present a greater genotoxicity, mutagenicity and carcinogenicity than their parent PAHs (Idowu et al., 2019). In addition, their higher water-solubility increases their environmental mobility and bioavailability in respect to unsubstituted PAHs (Larsson et al., 2018), increasing the risk for human and environmental health. Recent works have revealed that genotoxicity of PAH-contaminated soils can eventually increase after bioremediation despite effective PAH removal, which has been attributed to oxy-PAH formation and accumulation (Chibwe et al., 2015; Tian et al., 2017). The formation of partial oxidation products may be especially relevant during the biodegradation of high molecular weight PAHs, as these compounds often undergo cometabolic reactions (Kanaly & Harayama, 2010). Actually, a bacterial metabolite of the four ring PAH pyrene was recently identified as a major contributor to the increased genotoxicity of a PAH-contaminated soil after bioremediation (Tian et al., 2017).

Benz(*a*)anthracene-7,12-dione (B*a*AQ) is the ready oxidation product of the 4-ring PAH benz(*a*)anthracene (B*a*A). This oxy-PAH, commonly found in PAH-contaminated environments (Lundstedt et al., 2007), is of special concern due to its evidenced genotoxic and mutagenic properties (Clergé et al., 2019). *In vivo* assays using fish embryo have revealed that B*a*AQ presents higher genotoxic potential than its parent PAH and also than the well-known carcinogen benzo(*a*)pyrene (Dasgupta et al., 2014; Mccarrick et al., 2018). Considering its prevalence in soils and toxicological properties, determining the environmental fate of B*a*AQ is of special relevance. In soils, B*a*AQ can be formed by microbial transformation of B*a*A, but also as a result of photooxidation (Lehto et al., 2003) and chemical remediation treatments, such as Fenton oxidation (Lee & Hosomi, 2001) or ozonation (Yao et al., 1998). Several bacterial and fungal isolates have been reported to produce B*a*AQ from the transformation of B*a*A. Enzymatic attack of B*a*A at the C-7,12 positions by *Mycobacterium vanbaalenii* PYR-1 resulted in the formation of B*a*AQ and further ring cleavage products (Moody et al., 2005). Bacterial transformation of B*a*A to B*a*AQ has also been detected for a *Paenibacillus* strain (Deka & Lahkar, 2017), and genomic and transcriptomic analysis of *Rhodococcus* sp. P14 revealed that partial oxidation of B*a*A to B*a*AQ could be

attributed to the action of a cytochrome P450 monooxygenases (Luo et al., 2016). Ligninolytic fungi are also known to produce this oxy-PAH from the oxidation of B*a*A, prompted by laccases or peroxidases, and to further transform it (Cajthaml et al., 2006; Wu et al., 2010; Li et al., 2018). Removal of B*a*AQ has been observed after its transient accumulation during active PAH biodegradation in contaminated soils (Andersson et al., 2003; Lundstedt et al., 2003; Wilcke et al., 2014; Adrion et al., 2016), suggesting its possible reutilization by soil microbial communities. However, the microbial populations and mechanisms arbitrating its environmental fate are still unknown.

In this work, we identified the existence of distinctive microbial populations driving the biodegradation of B*a*A or its ready oxidation product B*a*AQ in a historical creosote-contaminated soil using sand-in-liquid soil microcosms. Based on the identification of a specialized B*a*AQ-degrading subpopulation, we selected a B*a*AQ-degrading microbial consortium by enrichment in sand-in-liquid cultures using this oxy-PAH as the sole carbon source. This consortium served as a model to investigate the metabolic networks involved in oxy-PAH biodegradation in PAH-contaminated soils. We first evaluated the bacterial consortium dynamics during active B*a*AQ-degradation by transcriptomic quantification of relevant phylotypes and by monitoring B*a*AQ removal together with metabolite formation. By integrating data from metabolomics screens and metagenomic functional gene analyses, we ascertained the B*a*AQ metabolic pathway and revealed the metabolic capabilities and potential interactions between the microbial components.

Materials and methods

Chemicals

Benz(*a*)anthracene (B*a*A, 99%) and benz(*a*)anthracene-7,12-dione (B*a*AQ, >97%) were obtained from Sigma-Aldrich Chemie (Steinheim, Germany). Media and reagents were purchased from Panreac Química (Barcelona, Spain) or Merck (Darmstadt, Germany). Solvents (organic residue analysis grade) were obtained from J.T. Baker (Deventer, The Netherlands). All chemicals and solvents were of the highest purity available.

Soil

The creosote-polluted soil sample was originally collected from a historical wood-treating facility in southern Spain. This soil (silty clay loam-calcaric fulvisoil) presented a very high content in PACs, $\Sigma 17$ PAHs=25,791 ± 2,112 ppm, $\Sigma 7$ oxy-PAHs =273 ± 9 ppm and $\Sigma 7$ N-PACs=2,568 ± 366 ppm (Annex – Table S1.1).

B*a*A and B*a*AQ sand-in-liquid soil microcosms

The biodegradation of B*a*A and B*a*AQ by the microbial populations of the creosote-contaminated soil was analyzed in small-scale sand-in-liquid microcosms. Microcosms consisted of 20 mL-glass vials containing 6 mL mineral medium (Hareland et al., 1975) and 3 g of thin grain sea sand (Panreac, Barcelona, Spain) coated with 0.2 g·L^{-1} of either B*a*A or B*a*AQ. The inoculum for the microcosms derived from the creosote-polluted soil. To minimize the concentration of native PAHs, the soil sample was preincubated for 21 days in aerobic conditions as detailed in Chapter 2 of this Thesis. Subsequently, microcosms were inoculated with 2 mL of a 1/10 dilution of a suspension of the pre-incubated soil. For each substrate, series of triplicate microcosms were incubated during 0, 3, 6, 10, 15, 20 and 25 days at 150 rpm and 25ºC. Identical triplicate but non-inoculated flasks were included as abiotic controls. At each incubation time, a set of triplicates was extracted for chemical analysis and another set of triplicates was used to obtain 1 mL samples containing sand and liquid for molecular analysis. Chemical extraction and analysis of residual B*a*A and B*a*AQ in the sand-in-liquid microcosms was performed as described in Chapter 2 of this Thesis. Nucleic acid extraction was performed using the procedures in Chapter 1 and submitted to 16S rRNA amplicon metagenomic sequencing.

16S rRNA amplicon metagenomic sequencing

Paired end reads (2x300bp) of the V1-V3 region of the 16S rRNA from triplicate cDNA samples of selected data points (0 and 20 days) were obtained using Illumina MiSeq equipment and reagents at Molecular Research DNA (Shallowater, Texas, USA). The received raw data was processed using mothur 1.45.3 following the MiSeq standard operating procedure (Kozich et al., 2013). The MiSeq libraries produced a total of 533,705 quality sequences with an average of 59,300 ± 3484 reads per sample. Sequences were aligned using SILVA v138 reference files and clustered into operational taxonomic units (OTUs) with a distance criterion of 0.03. The OTU taxonomic affiliations were assigned to each representative sequence using the RDP (Ribosomal Database Project) database v9. Rarefaction curves, library coverages and α-diversity indexes (Simpson's inverse index) were estimated normalizing the number of reads/sample to that of the least represented sample (54,528 reads). Correspondence analysis (CA) was done using the *vegan* Community Ecology package v.2.5-7 (Oksanen et al., 2020) in R v.4.1.0.

Enrichment cultures to obtain the B*a*AQ-degrading bacterial consortium BQ

The B*a*AQ-degrading community from the creosote-contaminated soil was enriched using a sand-in-liquid culture system, as described elsewhere (Tauler et al., 2016). Briefly, the culture system consisted in 100-mL Erlenmeyer flasks containing mineral medium (20 mL) and washed sea sand (10 g) coated with B*a*AQ (0.2 g·L^{-1}) as sole carbon source. The initial culture was inoculated with a 1 mL sample containing sand and liquid from the B*a*AQ-spiked sand-in-liquid soil microcosms incubated 30 days. Every month, 0.7 mL of sand and liquid were transferred to a newly prepared flask. Incubation was conducted at 25ºC and 150 rpm. Routine DGGE analyses were performed to check the stability of the microbial community composition in the moment of the transfer. All experiments involving the B*a*AQ-degrading consortium BQ were inoculated with a 30 day-culture.

Time-course B*a*AQ biodegradation by bacterial consortium BQ

Series of triplicate sand-in-liquid cultures with B*a*AQ as sole carbon source (0.2 g·L^{-1}) were inoculated with 0.7 mL of consortium BQ and incubated during 0, 3, 5, 10, 15, 20, 25 and 30 days at 150 rpm and 25ºC. Identical uninoculated flasks were included as abiotic controls. At each incubation time, a set of triplicate cultures and controls were removed for chemical analysis. After filtration, the aqueous phase was extracted 5 times with 10 mL ethyl acetate in neutral conditions and 5 times in acidic conditions (pH 2.5). The solid phase was extracted 5 times in ultrasonic bath with 20 mL dichloromethane/acetone (2:1, v/v). Extracts were analyzed by high pressure liquid chromatography (HPLC) as described previously in Chapter 2 of this Thesis. Also, 1 mL samples were collected from an additional set of triplicate flasks and total DNA and RNA were extracted for molecular analysis. Extracts were used as template for qPCR amplifications to monitor the abundance (DNA) and activity/expression (RNA) of relevant phylotypes (16S rRNA) or functional genes (see below).

Identification of B*a*AQ metabolites

Metabolites accumulated during the degradation of B*a*AQ were identified in large-scale sand-in-liquid cultures of consortium BQ. Sand-in-liquid cultures were scaled up to 250-mL Erlenmeyer flasks, containing 50 mL of mineral medium and 25 g of sand coated with B*a*AQ (0.2 g·L^{-1}). Cultures were inoculated with 1.5 mL of consortium BQ and incubated for 0, 5, 10, 15 and 20 days. Controls without cells were included to assess abiotic degradation. At each incubation time, the entire content of the flasks was filtrated, and the liquid phase was extracted as indicated in Chapter 2. Extracts were analyzed by HPLC and, further identification of detected metabolites was achieved by HPLC coupled to a high-resolution mass

spectrometer using electrospray ionization in positive ionization mode (HPLC-ESI(+)-HRMS). Analytical details were described in Chapter 2.

16S rRNA gene clone library of bacterial consortium BQ

 Total DNA from consortium BQ after 18 transfers was PCR-amplified using universal bacterial primers 27F and 1492R (Weisburg et al., 1991) and the GoTaq Master Mix as described in Chapter 2 of this Thesis. The PCR amplification product was visualized using 1% agarose gel, and cloned using the pGEM®-T Easy Vector cloning kit (Promega, Madison, USA). A total of 65 clones were randomly selected, and the corresponding plasmids were purified with the GenJET Plasmid Miniprep kit (Thermo Scientific, USA). All inserts were partially sequenced using universal primer 27F at Macrogen Europe (Madrid, Spain). Selected clones representing clusters of highly similar sequences (>99%) were also sequenced with universal primer 1492R. The resulting DNA sequences were manually adjusted using BioEdit v.7.2.5 and analyzed using the *Classifier* and *Sequence Match* tools of the Ribosomal Database Project II (Maidak, 2001) and Blastn tool of GenBank.

Shotgun metagenomic sequencing and analysis

Total DNA extraction for shotgun metagenomic sequencing of consortium BQ was performed using DNeasy PowerSoil kit (QIAGEN, Germany) with some modifications. Four aliquots of 1.5 mL containing sand and liquid were obtained from a culture of consortium BQ incubated for 15 days. The four aliquots were centrifuged and the pellets containing the sand and bacterial cells were transferred to 4 independent bead tubes. 200 μL of phenol:chloroform:isoamylic (25:24:1) were added to each tube and bead beating was conducted for 45 seconds at 5500 rpm. Lysates were recovered by centrifugation and treated with 200 μL of chloroform:isoamylic (24:1) to remove residual phenol. The four tubes were processed independently until the final purification step, when DNA extracts from the 4 tubes were loaded on a single spin column. Total DNA was eluted in a final volume of 70 μL.

Metagenomic sequencing of the DNA from consortium BQ was performed using an Illumina PE150 NovaSeq 6000 platform at Novogene Europe Co. (Cambridge, UK). Quality filtering and removal of adapters, artifacts and Phix contamination was done on the received paired-end raw reads using BBduk tool from BBmap. Raw read and trimmed read data quality was checked using FastQC (v0.11.8). Trimmed reads were assembled using metaSPAdes v3.11.1 (-meta option) with default options. Assembly statistics was obtained with Quast v.5.1. The assembly was used for binning with the Metawrap v1.2 pipeline running the metaBAT2 and MaxBin2 algorithms, using trimmed reads for mapping and a minimum contig length of

1000 bp. Relative abundance of each bin was inferred by mapping reads to each bin with Bowtie2 v.2.4.2. Bin quality assessment was performed with CheckM v.1.0.18 for contamination and completeness and the bins were referred to as metagenome-assembled genomes (MAGs). Functional annotation of the MAGs was performed using Prokka v.1.14.6. The GhostKOALA annotation server was used to reconstruct metabolic pathways.

Quantitative PCR analysis

The abundance of relevant phylotypes and genes from consortium BQ in sand-in-liquid and soil samples was determined by quantitative PCR (qPCR) analysis. All qPCR reactions were carried out on an Applied Biosystems StepOnePlus Real-time PCR system using PowerUp Sybr Green Master Mix (Applied Biosystems, CA, USA), 1 μL of template and 4 pmol of each primer in a final volume of 20 μL. Before amplification, RNA extracts were reverse transcribed to cDNA using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, USA) as detailed in Chapter 1 of this Thesis. Total bacterial populations (DNA) and total active bacterial populations (cDNA) were quantified by amplification of 16S rRNA gene using universal primers 341F and 534R (Muyzer et al., 1993). Other than these, primers used for this study are listed in Table 3.1. Primers were designed using

Table 3.1. Characteristics of qPCR primer sets used in this study.

Target	Primer name	Sequence (5' > 3')	Amplicon size (bp)	Annealing temp (ºC)
Sphingobium BQ1 16S rRNA gene	sphbq1-F	CTCATGGAGCTTCGGTGGC	168	55
	sphbq1-R	GCGATGAGGATGTCAAACGC		
Stenotrophomonas BQ2 16S rRNA gene	stebq2-F	GCACTTTGGCACGCAGTATC	192	55
	stebq2-R	AGGCACCAATCCATCTCTGG		
Pusillimonas BQ3 16S rRNA gene	pusbq3-F	CCACACTGGGACTGAGACAC	170	55
	pusbq3-R	AGCGGCACCAGGTATTAACC		
Olivibacter BQ4 16S rRNA gene	olibq4-F	ACTGCCTGTGCAAACAGAGA	148	55
	olibq4-R	CGAACTGTGAGGGGCTTTCT		
Pseudomonas BQ5 16S rRNA gene	psebq5-F	GTTGGTGGGGTAATGGCTCA	213	55
	psebq5-R	ACTTACTGCCCTTCCTCCCA		
Achromobacter BQ8 16S rRNA gene	achbq8-F	TAGTTGGTGGGGTAACGGCT	349	55
	achbq8-R	CTTTCCGAACCGCCTGC		
Bradyrhizobium BQ13 16S rRNA gene	brabq13-F	ACGATGAATGCCAGCCGTTA	224	55
	brabq13-R	AGGCTCCGAAGAGAAGGTCA		
ANTQ BVMO of *Sphingobium* sp. AntQ-1 (sphantq_4473)	4473-F	AGTCCCGCACGAATTGTCAT	147	55
	4473-R	ATGGGCACATCGAAAAAGCG		
ANTQ BVMO of *Sphingobium* sp. AntQ-1 (sphantq_4479)	4479-F	GAAGCATCGGGCGTCAAAAG	96	55
	4479-R	CGCGAAGAGAGTCATACCCG		

Primer-BLAST and experimentally validated by PCR. For quantification, 6-point 10-fold linearized standard dilution series were used. Standards were prepared from the corresponding clone in the 16S rRNA clone library, or by cloning the gene amplicons using the pGEM-T Easy Vector System (Promega, WI, USA). Plasmids were purified with the GeneJET Plasmid Miniprep Kit (Thermo Scientific, USA), and validated by sequencing.

Results and discussion

Bacterial populations driving the biodegradation of B*a*A and B*a*AQ in sand-in-liquid soil microcosms

The biodegradation of benz(*a*)anthracene (B*a*A) and benz(*a*)anthracene-7,12-dione (B*a*AQ) by the microbial populations in the creosote-contaminated soil was appraised in small-scale sand-in-liquid soil microcosms spiked with either one of the compounds. After 25 days of incubation the concentration of both substrates had been significantly reduced (B*a*A by 46% and B*a*AQ by 55%; Figure 3.1). B*a*A biodegradation was initated after 3 days of incubation and continued with a slow rate (12.2 ± 1.8 µmol·L^{-1}·day^{-1}) until day 15, when it stopped for 5 days. Subsequently, degradation resumed reaching a maximum rate (35.3 ± 3.3 µmol·L^{-1}·day^{-1}) during the last 5 days. As for B*a*AQ, after a 6 days period of adaptation,
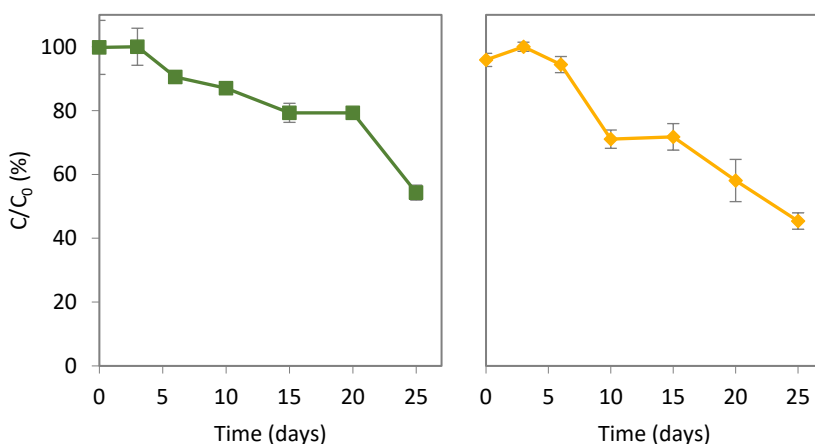
Figure 3.1. Biodegradation kinetics of benz(*a*)anthracene (B*a*A), in green, and benz(*a*)anthracene-7,12-dione (B*a*AQ), in yellow, in the sand-in-liquid soil microcosms. Values are average of three independent microcosms and error bars represent standard deviation.
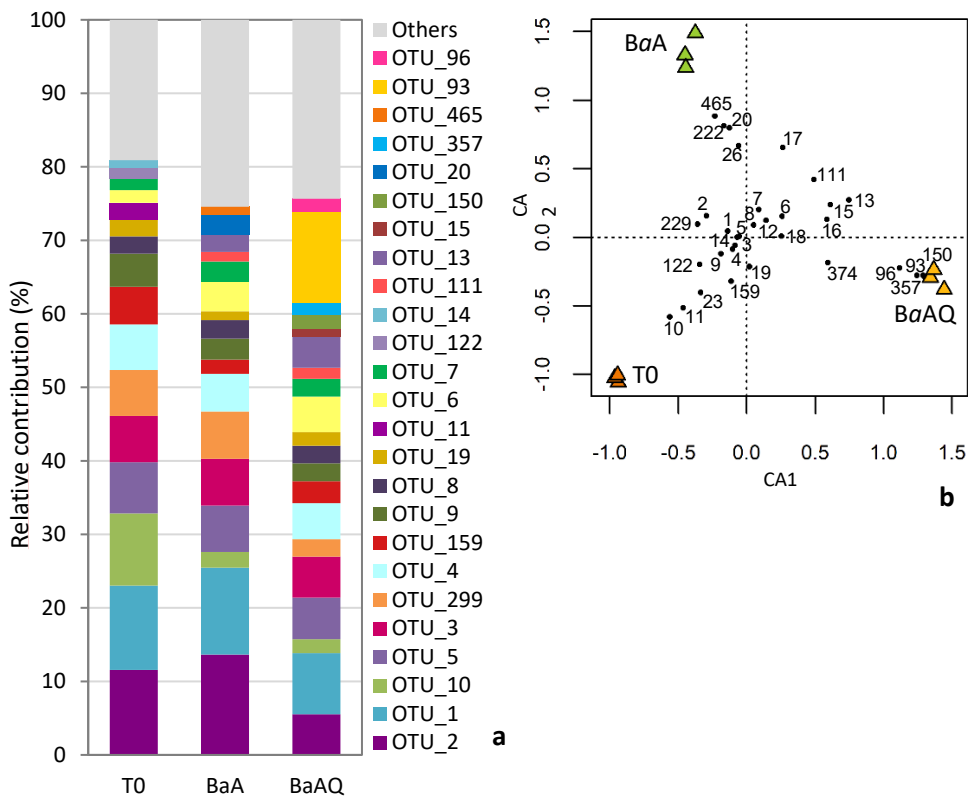
Figure 3.2. **(a)** Relative contribution of the most active OTUs (mean relative contribution >1%) in cDNA samples from the sand-in-liquid soil microcosms at 0 days of incubation (T0) and at 20 days with B*a*A or B*a*AQ. Values are means of three independent microcosms. **(b)** Correspondence Analysis (CA) based on the relative contribution of the most active OTUs (>1%) in cDNA samples from the sand-in-liquid at 0 days (T0) and at 20 days with B*a*A or B*a*AQ. Numbered black dots correspond to the OTUs. The first two axes (CA1 and CA2) of the CA explain 93.3% of the total variance observed.

maximum biodegradation rate was reached between 6 and 10 days (28.1 ± 3.4 µmol·L^{-1}·day^{-1}) and, after a 5-day plateau, the oxy-PAH was steadily consumed until the end of the incubation (12.7 ± 1.2 µmol·L^{-1}·day^{-1}).

The abundance and activity of the bacterial populations, estimated by qPCR amplification of 16S rRNA genes and transcripts respectively, were both relatively high at the beginning of the experiment and did not experience substantial changes over incubation. Total bacteria, initially at $1.3 \cdot 10^8 \pm 4 \cdot 10^7$ gene copies·mL^{-1}, stayed within the same order of magnitude during all the incubation with both

Table 3.2. Taxonomic identification of the most active OTUs (>1%) in the sand-in-liquid soil microcosms, based on the RDP database.

OTU	Taxonomic identification[a]	Class
1, 3, 5, 374	*Sphingobium*	
7, 93, 465	Uncl. *Sphingomonadales*	
26, 96, 150, 222, 357	Uncl. *Sphingomonadaceae*	
4	Uncl. *Rhodobacteraceae*	
6	*Caulobacter*	
12	Uncl. *Acetobacteraceae*	*Alphaprotebacteria*
14, 16	Uncl. *Rhizobiaceae*	
15	*Bosea*	
17	Uncl. *Bradyrhizobiaceae*	
23	*Methylopila*	
8	*Achromobacter*	*Betaproteobacteria*
2, 299	*Immundisolibacter*	
9, 20, 122	Uncl. *Pseudomonadacaeae*	*Gammaproteobacteria*
18, 19, 111, 159	*Pseudomonas*	
10	Uncl. *Xanthomonadaceae*	
13	Uncl. *Isosphaeraceae*	*Planctomycetacia*
11	Uncl. *Weeksellaceae*	*Flavobacteriia*

[a] When genus assignation had a probability lower than 100%, the immediate higher taxon with a 100% probability is shown.

Uncl. = unclassified

substrates, with punctual significant, but slight, increases at 10 days in microcosms with B*a*A ($3.6 \cdot 10^8 \pm 6.3 \cdot 10^7$ gene copies·mL^{-1}) and at 6 days with B*a*AQ ($3.0 \cdot 10^8 \pm 3.5 \cdot 10^7$ gene copies·mL^{-1}). Bacterial activity remained unaltered in both B*a*A and B*a*AQ-spiked microcosms, with values ranging from $1.9 \cdot 10^{10}$ to $3.9 \cdot 10^{10}$ transcripts·mL^{-1}. The high levels of both parameters at the beginning of the experiment can be explained by the preincubation of the soil to reduce its PAHs content, while the fact that the expression level was kept two orders of magnitude higher than the total populations during the entire incubation time indicates the presence of metabolically active populations.

Shifts in the active bacterial community structure were ascertained by 16S rRNA amplicon metagenomic sequencing of cDNA samples from the sand-in-liquid soil microcosms at 0 and 20 days of incubation. A predominance of *Alpha-* and *Gammaproteobacteria* was observed since the beginning of the incubation (47.2 and 47% relative abundance, respectively) and at 20 days for both B*a*A (54.1% and 36.8%) and B*a*AQ (65.6% and 23.1%). Nevertheless, relevant changes occurred in the relative expression of certain OTUs (Figure 3.2, Table 3.2). In the

B*a*A-spiked microcosms, OTU 2 affiliated to the recently described genus *Immundisolibacter* (Corteselli et al., 2017) was the most active OTU, showing a significant increase in relative abundance in cDNA libraries respect time 0 (from 11.5% to 13.7%). The 16S rRNA gene sequence of OTU 2 shared a 100% identity to that of the *Immundisolibacter* phylotype identified as the key player in the biodegradation of B*a*A applying DNA stable isotope probing (SIP) on the same soil (Chapter 1), corroborating its relevance in the degradation of this compound. Other less abundant OTUs that specifically and significantly incremented in B*a*A-microcosms were related to unclassified *Pseudomonadaceae* (OTU 20, from 0.4% to 2.7%) and sphingomonads (OTU 465, from 0.16% to 1.1%; OTU 26, from 0.19% to 0.98%; OTU 222, from 0.14% to 0.94%). Conversely, in the presence of B*a*AQ there was a dramatic increase in the relative abundance of an unclassified member of *Sphingomonadales* (OTU 93, from 0.28% to 12.3%), becoming the most relatively active phylotype in the community. In addition to this OTU, other members of the same order were also significantly more expressed in B*a*AQ-microcosms (OTU 150, 1.9%; OTU 96, 1.8%; OTU 357, 1.6%; OTU 374, 0.85%), suggesting that this compound might be degraded by a specialized uncultured group of *Sphingomonadales*. We recently isolated the strain *Sphingobium* sp. AntQ-1 (Chapter 2), the first bacterial isolate with the ability to degrade anthraquinone, a ready oxidation product from anthracene analogous to B*a*AQ. Altogether, these findings enlarge the broad catabolic range of sphingomonads (Aylward et al., 2013) by adding the utilization of oxy-PAHs.

Correspondence analysis of the most relatively active OTUs (>1%) highlighted the clear divergence between the microbial populations associated to either B*a*A or B*a*AQ degradation (Figure 3.2b). This confirms the existence of bacterial populations specialized in the biodegradation of the aromatic quinone, as none of the B*a*AQ-associated phylotypes were significantly expressed in B*a*A-spiked microcosms. An analogous experiment using sand-in-liquid soil microcosms from the same soil spiked with either anthracene or the corresponding quinone derivative (anthraquinone) revealed a similar divergence, identifying the existence of oxy-PAH-degrading specialists (Chapter 2). Rodgers-Vieira and colleagues (2015) already reported that distinctive bacteria were associated to the assimilation of anthracene or anthraquinone in a PAH-contaminated soil from North Carolina (USA) by DNA-SIP. Thus, it becomes evident that microbial communities from PAH-contaminated soils harbour populations with interconnected metabolic pathways, where dead-end products formed by one population become substrates for others.

B*a*AQ-degrading bacterial consortium BQ

Using the abovementioned B*a*AQ-spiked sand-in-liquid soil microcosms as inoculum, a B*a*AQ-degrading bacterial consortium was stablished by sequential enrichment in sand-in-liquid cultures with B*a*AQ (0.2 g·L^{-1}) as sole carbon source. Consortium BQ has been maintained by monthly transfers during 1.5 years, and routine 16S rRNA gene PCR-DGGE analysis showed a highly stable bacterial community structure (Annex – Figure S3.1). Biodegradation experiments demonstrate that bacterial consortium BQ cultures remove an 80% of the initial B*a*AQ concentration in 30 days. Strikingly, consortium BQ was not able to attack B*a*A, the parent compound of B*a*AQ, evidencing its high degree of specialization. To the best of our knowledge, this is the first oxy-PAH degrading consortium reported so far.

Analysis of the microbial community by 16S rRNA gene clone library revealed that consortium BQ comprised 13 bacterial phylotypes (Table 3.3). The microbial community was dominated by a member of *Sphingobium* (BQ1, 32.5%), whose 16S rRNA gene sequence shared a 98.2% identity to that of OTUs 93 and 96 associated to the utilization of B*a*AQ in the sand-in-liquid soil microcosms. Interestingly, phylotype BQ1 was also closely related to an uncultured *Sphingobium* clone from a DNA-SIP experiment with [U-^{13}C]-anthracene that increased in abundance after the added anthracene had been consumed, which was attributed to the possible growth on anthracene metabolites (Jones et al., 2011). Other abundant members of consortium BQ were related to *Stenotrophomonas* (BQ2, 23.7%; BQ6, 2.6%), *Pusillimonas* (BQ3, 12.3%; BQ9, 1.8%), *Olivibacter* (BQ4, 8.8%) and *Pseudomonas* (BQ5, 7.0%). Sequences of *Stenotrophomonas* BQ2 were very similar (>99%) to 16S rRNA sequences retrieved from other PAH/oil-contaminated soils (Accession numbers: HQ218574, KY962735), but also from plant rhizospheres (HQ912766, FJ493060, KT825842). *Pusillimonas* BQ3 and BQ9 were closely related to isolates of this genus with the ability to utilize hydroxybenzoates (NR_146844), substituted salicylates (NR_043129) or 5-hydroxypicolinic acid (KC602498), suggesting its role in the metabolism of aromatic acids. As for *Olivibacter* BQ4, its sequence resembled (99%) to that of strain *Olivibacter* sp. UBH15, isolated from the same creosote-contaminated soil used here (Tauler et al., 2016). This strain was not able to remove any PAH from a creosote HMW-PAH residue, however, growth has been observed at the expenses of diphenic acid, protocatechuic acid, phthalic acid, 9-fluorenone-4-carboxylic acid and 1,2,3-benzenetricarboxylic acid (data not shown). The 16S rRNA gene of *Pseudomonas* BQ5 showed high homology (>99%) to isolates from PAH-contaminated soils and sediments (MT448942; MF928390), isolates with the ability to degrade xenobiotics such as diethyl

Table 3.3. Analysis of the 16S rRNA sequences recovered in the clone library of consortium BQ.

Phylotype	Freq. (%)	Length (bp)	RDP Classification[a]	Closest relative (Accession number, %)	Closest type strain (Accession number, %)
BQ1	32.5	1305	*Sphingobium*	Uncultured soil bacterium clone SBANT43, SIP anthracene (HM596217, 99%)	*Sphingobium phenoxybenzoativorans* SC_3 (KP257600, 98%)
BQ2	23.7	1370	*Stenotrophomonas*	Uncultured bacterium clone N-132, PAH-contaminated soil (HQ218574, 99%)	*Stenotrophomonas chelatiphaga* LPM-5 (EU573216, 99%)
BQ3	12.3	1325	*Pusillimonas*	*Pusillimonas caeni* strain EBR-8-1, sludge biofilm reactor (NR_146844, 99%)	*Pusillimonas caeni* EBR-8-1 (KF056995, 99%)
BQ4	8.8	1324	*Olivibacter*	*Olivibacter* sp. LS-1, phenanthrene-degrader (CP041643, 99%)	*Olivibacter oleidegradans* TBF2/20.2 (HM021726, 99%)
BQ5	7.0	1353	*Pseudomonas*	*Pseudomonas* sp. strain JWJA2, petroleum-contaminated soil (MT448942, 99%)	*Pseudomonas monteilii* CIP 104883 (AF064458, 99%)
BQ6	2.6	1347	*Stenotrophomonas*	*Stenotrophomonas chelatiphaga* LPM-5, EDTA-degrading (EU573216, 96%)	*Stenotrophomonas chelatiphaga* LPM-5 (EU573216, 96%)
BQ7	2.6	1288	*Sphingobium*	*Sphingobium naphthae* strain K-3-6, oil-contaminated soil (NR_157779, 98%)	*Sphingobium rhizovicinum* CC-FH12-1 (EF465534, 97%)
BQ8	2.6	1313	*Achromobacter*	*Achromobacter spanius* LMG 5911 (AY170848, 99%)	*Achromobacter spanius* LMG 5911 (AY170848, 99%)
BQ9	1.8	1339	*Pusillimonas*	*Pusillimonas caeni* strain EBR-8-1, sludge biofilm reactor (NR_146844, 97%)	*Pusillimonas caeni* EBR-8-1 (KF056995, 97%)
BQ10	1.8	1304	*Sphingomonadaceae*	*Sphingomonas* sp. IC145, carbazole degrader (AB196253, 99%)	*Rhizorhabdus wittichii* DSM 6014 (AB021492, 99%)
BQ11	1.8	1307	*Bosea*	*Bosea* sp. strain SF67, ibuprofen degrader (MN636489, 99%)	*Bosea thiooxidans* DSM 9653 (AJ250796, 99%)
BQ12	1.8	1292	*Rhizobium*	*Rhizobium* sp. 3C6-41, sludge of industrial wastewater treatment (HQ246161, 99%)	*Rhizobium daejeonense* L61T (AY341343, 99%)
BQ13	0.9	1282	*Bradyrhizobium*	*Bradyrhizobium japonicum* USDA 6 DNA (AP012206, 99%)	*Bradyrhizobium subterraneum* 58 2-1 (KP308152, 99%)

[a] When genus assignation had a probability lower than 100%, the immediate higher taxon with a 100% probability is shown.

phthalate (MN964893) or pymetrozine (MW674665), and to the well-characterized *P. putida* KT2440, which harbours a wide metabolic versatility to aromatic compounds (Belda et al., 2016). Less abundant constituents of consortium BQ were affiliated to *Achromobacter* (BQ8, 2.6%), *Sphingomonadaceae* (BQ7, 2.6%; BQ10, 1.8%) and members of the *Hyphomicrobiales* order including *Bosea* (BQ11, 1.8%), *Rhizobium* (BQ12, 1.8%) and *Bradyrhizobium* (BQ13, 0.9%).

Identification of benz(*a*)anthracene-7,12-dione metabolites

Metabolites from the biodegradation of B*a*AQ by consortium BQ were identified by HPLC-ESI(+)-HRMS analysis of neutral and acidic extracts (Table 3.4). Tentative structures of metabolites were based on their exact mass and their MS fragmentation patterns. Metabolite I, with $[M+H]^+$=293.0782 and molecular formula $C_{18}H_{13}O_4$, exhibited ion fragments at *m/z* 293.0782 (94), 249.0887 (100) and 158.0276 (35) and UV/vis spectrum with λ_{max} at 230 and 294 nm. This is consistent with either 2-(2-hydroxynaphthalene-1-carbonyl)benzoic acid or 1-(2-hydroxybenzoyl)-2-naphthoic acid, depending on the position of the carboxylic acid and hydroxyl moieties. Metabolite II produced the molecular formula $C_{18}H_{11}O_3$ ($[M+H]^+$=275.0684), showing ion fragments at *m/z* 275.0684 (100), 247.0727 (1) and 158.0252 (4). The UV-visible spectrum presented λ_{max} at 220, 245, 335 and 370 nm. These properties were compatible with both a hydroxybenz(a)anthracene-7,12-dione and benzo(*e*)naphtha(2,1-*b*)oxepin-8,13-dione. However, tentative identification of metabolite I as an acid indicative of ring cleavage at the C-7 or C-12 positions of B*a*AQ indicated that metabolite II corresponded to the lactone benzo(*e*)naphtha(2,1-*b*)oxepin-8,13-dione. In both acidic and neutral extracts a metabolite with formula $C_{10}H_9O$ ($[M+H]^+$=145.0641) and UV/vis at 226 and 274 nm, compatible with an isomer of naphthol.

The identification of benzo(*e*)naphtha(2,1-*b*)oxepin-8,13-dione or its isomer, 2-(2-hydroxynaphthalene-1-carbonyl)benzoic acid or 1-(2-hydroxybenzoyl)-2-

Table 3.4. HPLC-ESI(+)-HRMS detected metabolites produced from benz(a)anthracene-7,12-dione by consortium BQ in neutral (N) and acidic (A) extracts.

Metabolite	Extract	Rt (min)	$[M+H]^+$	Molecular formula	δ (mmu)	Tentative identification
I	N	13.2	293.0782	$C_{18}H_{13}O_4$	-2.6325	2-(2-Hydroxynaphthalene-1-carbonyl)benzoic acid or 1-(2-hydroxybenzoyl)-2-naphthoic acid
III	N/A	16.7	145.0642	$C_{10}H_9O$	-0.6003	Naphthol
II	N	17.1	275.0684	$C_{18}H_{11}O_3$	-1.8567	Benzo[*e*]naphth[2,1-*b*]oxepin-8,13-dione
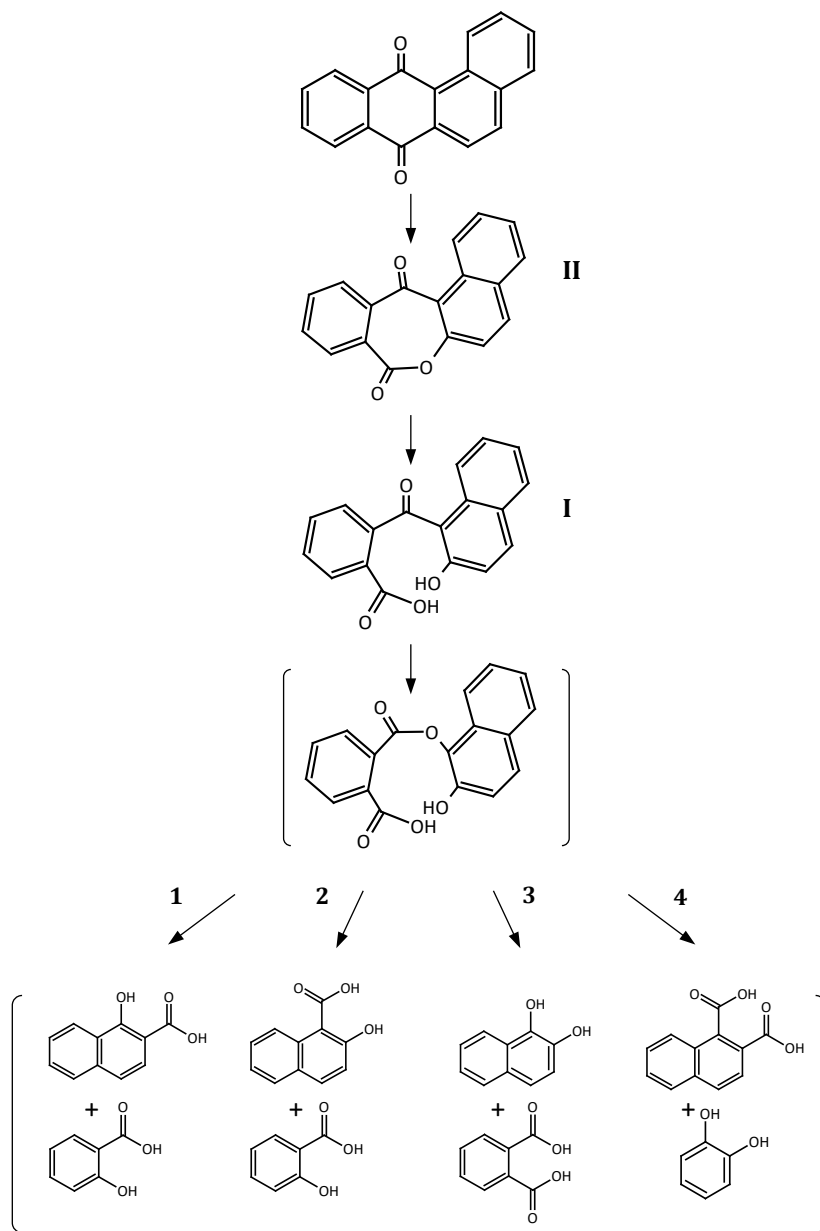
Figure 3.3. Schematic pathway proposed for the degradation of benz(*a*)anthracene-7,12-dione by bacterial consortium BQ. Only one of the possible isomers of metabolites I and II is represented. Compounds in brackets were not detected. Numbered steps correspond to the possible metabolites resulting from the second Baeyer-Villiger reaction. (1) 1-hydroxy-2-naphthoic acid and salicylic acid, (2) 2-hydroxy-1-naphthoic acid and salicylic acid, (3) 1,2-dihydroxynapthtalene and phthalic acid, or (4) naphthalene dicarboxylic acid and catechol.

naphthoic acid and probably naphthol as the major metabolites from B*a*AQ suggests that consortium BQ first attacks the quinone ring and further metabolizes the aromatic rings to intermediates of the central metabolism (Figure 3.3). The attack would be initiated by a Baeyer-Villiger oxidation producing a lactone, benzo(*e*)naphtha(2,1-*b*)oxepin-8,13-dione or its isomer, that would eventually be hydrolyzed resulting in the formation of 2-(2-hydroxynaphthalene-1-carbonyl)benzoic acid or 1-(2-hydroxybenzoyl)-2-naphthoic acid. A similar mechanism has been recently described for the degradation of 9,10-anthraquinone (ANTQ) by *Sphingobium* sp. AntQ-1 (Chapter 2). In that strain, ANTQ degradation was suggested to proceed through a second BV oxidation, based on the presence of two gene clusters coding for a Baeyer-Villiger monooxygenase (BVMO) together with a hydrolase and overexpression of both phthalate and catechol metabolic pathways. If the identification of naphthol is confirmed, it would also suggest that the bacterial consortium BQ performs a second BV attack on the detected acid. Depending on the position of the carboxyl and hydroxyl moieties in metabolite I and on that of the second BV attack, four sets of subsequent metabolites may be formed, as shown in figure 3.3. The formation of 1-(2-hydroxybenzoyl)-2-naphthoic acid from B*a*AQ was also reported for the transformation of B*a*A via B*a*AQ by *Mycobacterium vanbaalenii* PYR-1 (Moody et al., 2005), however, it was not attributed to a Baeyer-Villiger reaction.

Bacterial dynamics during active B*a*AQ-biodegradation by consortium BQ

Insight into the bacterial community dynamics of consortium BQ during degradation of B*a*AQ was accomplished by quantifying the expression of selected phylotypes and functional genes and by tracking the removal of B*a*AQ in parallel to the formation of metabolites (Figures 3.4 and 3.5). After a 3-day lag phase, B*a*AQ was depleted following first-order kinetics with maximum biodegradation rates until the 10th day of incubation (54.2 ± 10.1 µmol·L^{-1}·day^{-1}) that attenuated thereafter (22.1 ± 0.5 µmol·L^{-1}·day^{-1} from day 10 to 20, and 10.5 ± 3.5 µmol·L^{-1}·day^{-1} from day 20 to 30). Coinciding with maximum B*a*AQ degradation rates, there was a significant increase in the abundance of metabolite II, identified as benzo(*e*)naphtha(2,1-*b*)oxepin-8,13-dione. This accumulation was transient, as it progressively decreased after reaching its maximum at 10 days of incubation, indicating its possible reutilization.

The expression level of *Sphingobium* BQ1, *Stenotrophomonas* BQ2, *Pusillimonas* BQ3, *Olivibacter* BQ4, *Pseudomonas* BQ5, *Achromobacter* BQ8 and *Bradyrhizobium* BQ13 was quantified by qPCR specifically targeting their 16S rRNA transcripts. Bacterial activity estimated as total 16S rRNA transcripts experienced a significant
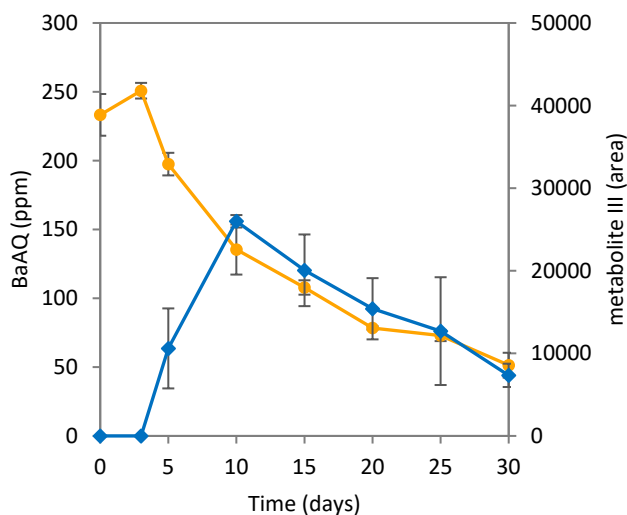
Figure 3.4. Evolution of B*a*AQ and metabolite II in sand-in-liquid cultures of consortium BQ. Abundance of metabolite II is represented as HPLC peak area. Values are average of three independent cultures and error bars represent standard deviation.

increase in expression after 3 days of incubation, when B*a*AQ biodegradation started, reaching the maximum at 10 days. During the period of highest B*a*AQ degradation rates (3 to 10 days), *Sphingobium* BQ1 was the most active phylotype in the consortium, with 1 to 3 orders of magnitude above the rest, and reaching its maximum at 10 days of incubation ($1.1 \cdot 10^{10} \pm 7.2 \cdot 10^9$ transcripts·mL^{-1}). This peak of activity of phylotype BQ1 was concomitant with the peak of concentration of the lactone (metabolite II), suggesting that this phylotype was the main player in the initial attack of B*a*AQ. *Pusillimonas* BQ3, *Pseudomonas* BQ5, *Achromobacter* BQ8 and *Bradyrhizobium* BQ13 showed their maximum expression in the same period (10^7 to 10^9 transcripts·mL^{-1}). *Stenotrophomonas* BQ2 and *Olivibacter* BQ4, on the contrary, were most active from the 10th day, reaching their maximum expression at 20 days (10^8 to 10^9 transcripts·mL^{-1}) when BaAQ degradation was significantly reduced. These results suggest that phylotypes BQ2 and BQ4 were likely thriving on simple compounds produced by other members consortium BQ. The simultaneous expression of different phylotypes within the community suggests that interactions are taking place during active biodegradation of B*a*AQ, and potentially indicate that most phylotypes are probably feeding on BaAQ metabolites formed after the initial attack likely catalyzed by *Sphingobium* BQ1. To shed light on the B*a*AQ degradation pathway and provide insights into the potential cross-feeding interactions between the members of the community, we conducted a metagenomics analysis of consortium BQ.
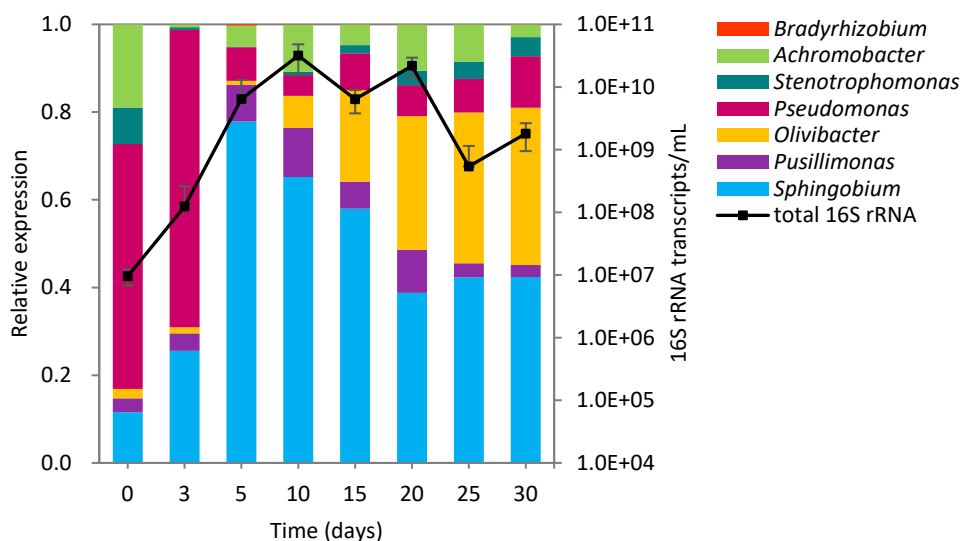
Figure 3.5. Relative expression of representative phylotypes from bacterial consortium BQ and total 16S rRNA transcripts/mL along incubation with B*a*AQ as sole carbon source in sand-in-liquid cultures. Values are average of three independent cultures.

Metabolic functional analysis of bacterial consortium BQ by shotgun metagenomics

The potential metabolic capabilities and interactions of the different members of consortium BQ were assessed by functional analysis of metagenome assembled genomes (MAGs). Shotgun metagenomic sequencing of DNA from a 15-days culture of consortium BQ generated a total of 36.8 million quality-filtered 150 bp paired-end reads that were assembled into 71,695 contigs (largest contig 926,254 bp, N50 64,905 bp) and binned into 15 high-quality MAGs (Table 3.5). The MAGs accounted for an 88.7% of the total reads. Despite no 16S rRNA gene sequences were binned into any of the MAGs, most of the MAGs could be assigned to BQ phylotypes according to their taxonomic identification and their relative abundance in the metagenomic assembly and the 16S rRNA gene clone library. Functional annotation of individual MAGs revealed the presence, to a greater or lesser extent, of genes coding for enzymes driving the metabolism of aromatic compounds in all MAGs (Table 3.6).

As discussed above, B*a*AQ metabolic pathway would be initiated by the action of a Baeyer-Villiger monooxygenase. Currently, BVMOs are considered flavin-dependent monooxygenases (FDMOs) that catalyze BV oxidations (Tolmie et al.,

Table 3.5. Features of the metagenome assembled genomes (MAG) recovered from shotgun metagenomic sequencing of consortium BQ.

MAG	GTDB Classification[a]	Size (Mbp)	GC (%)	Comp[b] (%)	Cont[b] (%)	Rel.[c] abun. (%)	Phylotype[d]
1	*Sphingobium*	3.24	65.8	97.5	0.26	2.2	*Sphingobium* BQ7
2	*Bosea*	5.78	65.4	89.4	2.35	0.5	*Bosea* BQ11
3	*Rhizobiaceae*	6.37	62.2	97.9	2.28	1.1	*Rhizobium*
4	*Shinella*	6.83	63.8	98.7	0.39	1.6	*Rhizobium*
5	*Stenotrophomonas*	4.04	67.2	74.1	0.00	3.2	*Stenotrophomonas* BQ6
6	*Achromobacter*	6.18	64.8	91.3	1.79	1.1	*Achromobacter*
7	*Bradyrhizobium*	8.02	64.0	99.4	1.35	1.5	*Bradyrhizobium* BQ13
8	*Stenotrophomonas*	3.95	66.4	99.5	2.63	19.6	*Stenotrophomonas* BQ2
9	*Sphingomonas*	3.76	67.5	97.6	0.00	2.1	*Sphingomonadaceae* BQ10
10	*Pusillimonas*	3.67	63.2	96.7	0.00	10.5	*Pusillimonas* BQ3
11	*Achromobacter*	6.25	67.6	90.9	0.86	1.2	*Achromobacter*
12	*Sphingobium*	3.86	62.0	99.1	0.69	31.0	*Sphingobium* BQ1
13	*Pigmentiphaga*	4.46	65.3	87.8	0.61	3.5	*Pusillimonas* BQ9
14	*Pseudomonas*	5.33	61.8	96.7	1.15	0.6	*Pseudomonas* BQ5
15	*Pseudosphingobacterium*	6.72	41.1	100	0.95	8.9	*Olivibacter* BQ4

[a] Taxonomic identification was based on the GTDB-Tk tool.

[b] Genome completeness and contamination were calculated with CheckM v.1.0.18

[c] Relative abundance was inferred by mapping paired-end reads to each bin with Bowtie2.

[d] MAGs were assigned to each phylotype according their relative abundance in consortium BQ.

2019). Depending on the type of flavin cofactor, the type of electron donor, the type of reaction catalyzed and the protein folding, FDMOs are classified into eight classes (van Berkel et al., 2006). Only classes A, B and C are known to include BVMOs (Annex – Table S3.1). Class A FDMOs encompass a diverse group of atypical BVMOs (type O) that are mostly multifunctional enzymes catalyzing enzyme cascades that include hydroxylation reactions in addition to the BV oxidation. Class B includes the most extensively studied BVMOs, which are type I. Scanning MAGs for the type I signature protein motif (Fraaije et al., 2002) revealed the presence of type I BVMOs in the MAGs of *Pseudomonas* BQ5 (MAG14_0903, MAG14_0921, MAG14_4357), *Bosea* BQ11 (MAG2_2077) and *Bradyrhizobium* BQ13 (MAG7_0840, MAG7_3680, MAG7_5155, MAG7_7237). Thus, they might be driving the second BV reaction as the BVMO of *Sphingobium* sp. AntQ-1 involved in the analogous step was a type I BVMO (sphantq_4479). The most closely related type I BVMOs to sphantq_4479 were the three CDS of *Pseudomonas* BQ5 (48.2 – 48.8% amino acid sequence identity).

Table 3.6. Number of genes encoding monooxygenase and dioxygenase enzymes related to the metabolism of aromatic compounds in the MAGs of consortium BQ.

	MAG	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		Sphingobium BQ7	*Bosea* BQ11	*Rhizobium*	*Rhizobium*	*Stenotrophomonas* BQ6	*Achromobacter*	*Bradyrhizobium* BQ13	*Stenotrophomonas* BQ2	*Sphingomonadaceae* BQ10	*Pusillimonas* BQ3	*Achromobacter*	*Sphingobium* BQ1	*Pusillimonas* BQ9	*Pseudomonas* BQ5	*Olivibacter* BQ4
	Class A	9	7	4	5	3	11	11	4	1	12	8	2	31	2	1
flavin dependent monooxygenase (FDMOs)	Class B		1		2			7			2				4	
	Class C	7	8	10	7	2	7	19	2	10	5	7	7	7	8	7
biphenyl 2,3-dioxygenase (*bphA*)		1					1	1			1		1	1		
cytochrome P450 (CYP450)		2	2	1	1		1	5		3	2	2	2	1		
1,2-dihydroxynaphthalene dioxygenase (*nahC*)											1	2	1	2		
1-hydroxy-2-naphthoate dioxygenase (*phdI*)											1	1		1		
phthalate-4,5-dioxygenase (*pht23*)			1					1			4	1		6		
protocatechuate 3,4-dioxygenase (*pcaGH*)			2		1		2	2	1		1			1	1	
protocatechuate 4,5-dioxygenase (*ligAB*)		1									1			1		
salicylate 1-hydroxylase (*nahG*)			1	1	1			1			2	1		7		
salicylate 5-hydroxylase (*nagGH*)							1				3	1		1		
catechol 1,2-dioxygenase (*catA*)			2				1			1				1	2	
catechol 2,3-dioxygenase (*xylE*)			3	1	1		1	1						1		1
gentisate 1,2-dioxygenase (*nagI*)		1					2	1			9	2		3		

Legend:
- Sphingomonadaceae (MAGs 1/9/12)
- Olivibacter (MAG 15)
- Pusillimonas (MAGs 10/13)
- Pseudomonas (MAG 14)
- Achromobacter (MAGs 6/11)
- Stenotrophomonas (MAGs 5/8)
- Hyphomicrobiales (MAGs 2/3/4/7)

nahG

phdI
phdJ
phdK

pht23
pht4
pht5

nahC
nahD
nahE
xylC

nahG

pobA

xylXY
xylL

DHBD
cmtC
cmtD

nagGH

ligAB

ligC chqB
ligI
galD
ligJ
ligK

pcaGH pcaB pcaC

macA

catA
catB
catC

pcaD
pcaIJ
pcaF

xylE
xylG
xylH
xylI
xylJ
xylK

xylF

nagI
nagL
nagK

TCA cycle

◁ **Figure 3.6.** Key aromatic hydrocarbon degradation genes enclosed in the MAGs from consortium BQ. MAGs were grouped according their taxonomic identification. Arrows labeled as different colors indicate genes found in each group of MAGs belonging to the same taxon.

On the other hand, the BVMO catalyzing the initial attack of ANTQ by *Sphingobium* sp. AntQ-1 (sphantq_4473) was identified as a probable type II BVMO, that belong to class C FDMOs. To date, the only representatives of type II BVMOs are two diketocamphane monooxygenases (Iwaki et al., 2013). The MAGs of *Sphingobium* BQ1 (MAG12_2582), *Achromobacter* (MAG11_3219), *Bradyrhizobium* BQ13 (MAG7_5918) and *Rhizobium* (MAG3_5954) enclosed genes coding for diketocamphane monooxygenases, however, phylogenetic analysis of amino acid sequences revealed they were distantly related to sphantq_4473. In fact, sphantq_4473 closest relative was a CDS of *Pusillimonas* BQ9 (MAG13_4095) annotated as an alkanal monooxygenase, only sharing a 27.6% identity. Nevertheless, ANTQ-BVMOs of strain AntQ-1 were both part of a cluster composed by two genes, one coding for BVMO and the other for a hydrolase that would catalyze the hydrolyzation of the lactone. From all the potential BVMOs enclosed in the MAGs, only the three type I BVMOs of *Pseudomonas* BQ5 and a type II BVMO of *Pusillimonas* BQ9 (MAG13_3989) were accompanied by a hydrolase-coding gene.

Downstream transformation of B*a*AQ would be channeled to either (1) 1-hydroxy-2-naphthoic acid and salicylic acid, (2) 2-hydroxy-1-naphthoic acid and salicylic acid, (3) 1,2-dihydroxynapthtalene and phthalic acid, or (4) naphthalene dicarboxylic acid and catechol. Comprehensive reconstruction of the metabolic pathways enclosed in the MAGs revealed the metabolic potential of each phylotype to process these aromatic compounds (Table 3.6, Figure 3.6, Annex – Table S3.2). Surprisingly, the most abundant phylotype in consortium BQ, *Sphingobium* BQ1 (MAG12), did not harbor any complete pathway and it only had one gene coding for 1,2-dihydroxynaphthalene dioxygenase (*nahC*) and one for benzoate 1,2-dioxygenase (*xylXYZ*) apart from other genes coding for intermediate or final steps in the catechol or gentisate pathways, respectively (*xylF* and *nagLK*). The less abundant *Sphingobium* BQ7 (MAG1) possessed the complete protocatechuate *meta*-cleavage pathway (*ligABCIJK* and *galD*) and *Sphingomonadaceae* BQ10 had genes coding for the initial steps for catechol *ortho*-cleavage (*catABC* and *pcaD*). However, the metabolic potential of these phylotypes might be underestimated as it is widely known that sphingomonads often harbor aromatic degradative genes in plasmids (Stolz, 2014), which may have escaped from the metagenomic binning process (Antipov et al., 2019).

115

Stenotrophomonas BQ2 (MAG9) and BQ6 (MAG5) mainly harbored genes coding for final steps leading to the tricarboxylic acid (TCA) cycle (*nagLK*, *ligIJK*, *galD*, *pcaF*, *xylJ*). Despite being one of the most abundant phylotypes, only three genes were detected in *Olivibacter* BQ4 (MAG15) from all the metabolic pathways. It included genes encoding for one catechol 2,3-dioxygenase (*xylE*) and the final steps of the gentisate (*nagK*) and the protocatechuate (*pcaF*) pathways. Accordingly, these phylotypes might be thriving on smaller intermediate metabolites produced by other members of the community. *Pseudomonas* BQ5, in contrast, exhibited the ability to catabolize benzoate (*xylXYZ*) to catechol and protocatechuate both via *ortho*-cleavage (*catABC* and *pcaGHBCDIJF*).

The MAGs corresponding to *Pusillimonas* (MAGs 10 and 13) and *Achromobacter* (MAGs 6 and 11) were the ones with the broadest enzymatic range, as they encompassed genes coding for enzymes involved in all the aromatic degradation pathways. These phylotypes shared almost the same metabolic capabilities, being able to completely process 1-hydroxy-2-naphthoic acid, phthalate, catechol and salicylate all the way to the central metabolism. For instance, *Pusillimonas* BQ3 (MAG10) owned the full phthalate pathway (*pht2345*), having 4 gene copies coding for phthalate 4,5-dioxygenase, and both the protocatechuate *ortho*- (*pcaGHBCDIJF*) and *meta*-cleavage pathways (*ligABC*, *galD* and *ligJK*). It also carried the entire salicylate via gentisate pathway (*nagGHILK*), with 3 gene copies encoding salicylate 5-hydroxylase and 9 copies for gentisate 1,2-dioxygenase. This MAG also included the *nahG* gene, acting in the decarboxylation of either salicylate or 1-hydroxy-2-naphthoic acid, and genes coding for the complete cleavage of 1-hydroxy-2-naphthoic acid to phthalate (*phdIJK*). Although it did not have any genes coding for catechol 2,3-dioxygenase, it did have the rest of the genes for the *meta*-cleavage pathway (*xylEFGHIJK*). *Pusillimonas* BQ9, however, did show complete catabolism of catechol by both intra- and extradiol ring cleavage.

Similarly, MAGs from the *Hyphomicrobiales* order, including *Bosea* BQ11 (MAG2), *Rhizobium* (MAGs 3 and 4) and *Bradyrhizobium* BQ13 (MAG7), also displayed extensive metabolic capabilities. They all encompassed *nahG* and *xylE* genes. *Bosea* BQ11 showed the ability to completely metabolize phthalate via *ortho*-cleavage of protocatechuate (*pht2345* and *pcaGHBCDIJF*) and to catalyse the initial steps for *ortho*-cleavage of catechol (*catAB*). *Rhizobium* (MAG4) and *Bradyrhizobium* BQ13 also harbored genes for complete *ortho*-cleavage of protocatechuate. However, both *Rhizobium* MAGs generally possessed genes from intermediate steps from all pathways. In addition, *Bradyrhizobium* BQ13 was the only one of this group able to catabolize gentisate (*nagILK*).

Horizontal gene transfer of plasmid *pANTQ-1*

Since no unequivocal candidate BVMOs were detected in the MAGs for the initial attack of B*a*AQ, mining for the ANTQ BVMO-hydrolase gene cluster of *Sphingobium* sp. AntQ-1 was performed in the whole metagenomic assembly in order to find potential similarities. To our surprise, not only this cluster was found, but also there was a 70.2% (141,746 bp) alignment to the megaplasmid *pANTQ-1* (Chapter 2) with just a 0.39 % mismatch and 0.018 % indels. Plasmid *pANTQ-1* carries the genes coding for the upper ANTQ degradation pathway, phthalate metabolism and catechol *meta*-cleavage pathway. The largest alignment was 26,349 bp long and it corresponded to positions 59,952 to 86,300 of *pANTQ-1*, which include both BVMO-hydrolase coding gene clusters and genes involved in the metabolism of catechol. Alignment to the region encoding phthalate-related genes was also observed (from positions 92,986 to 107,112; and 108,726 to 120,583). However, only a 6.2% alignment, with 2.7% mismatches and 0.17% indels, was detected to the primary chromosome of strain AntQ-1, indicating that this strain was not a member of consortium BQ. Notably, plasmid *pANTQ-1* encompassed several genes coding for transposases and type IV secretion system components. Furthermore, it had a lower GC content (60.6%) than the average genome content (63.6%) of *Sphingobium* sp. AntQ-1. Thus, it is likely that horizontal gene transfer of plasmid *pANTQ-1* to a member of consortium BQ occurred.
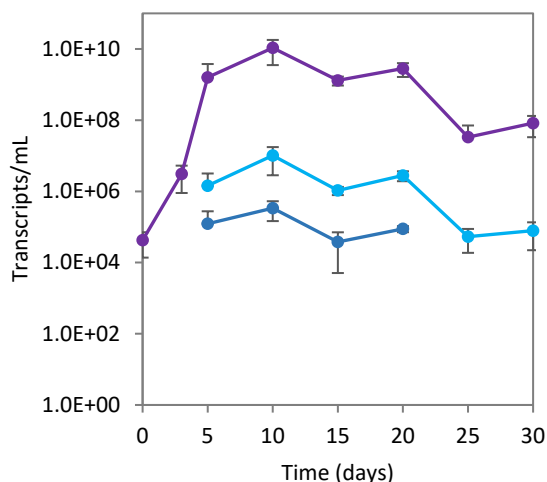
Figure 3.7. Evolution of *Sphingobium* BQ1 16S rRNA (purple) and BVMOs sphantq_4473 (light blue) and sphantq_4479 (dark blue) transcripts in sand-in-liquid cultures of consortium BQ. Values are average of three independent cultures and error bars represent standard deviation.
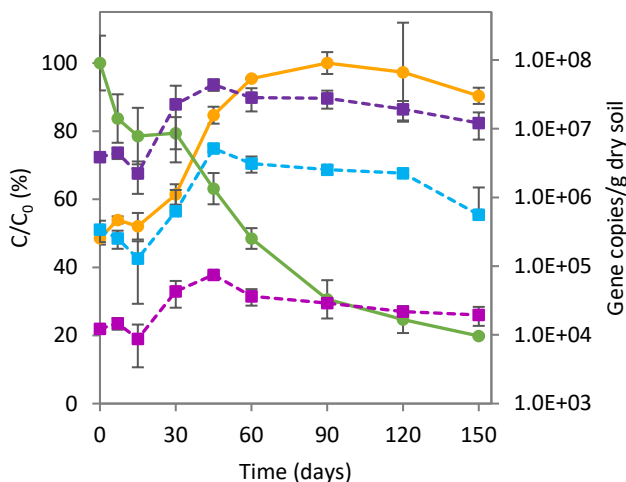
Figure 3.8. Evolution of *Sphingobium* sp. AntQ-1 16S rRNA (magenta), *Sphingobium* BQ1 16S rRNA (purple) and BVMO 4473 (blue) gene copies and the concentration of B*a*A (green) and B*a*AQ (yellow) in the lab-scale biostimulation assay of the creosote-contaminated soil.

Several studies showed that conjugative plasmid transfer from sphingomonads is basically restricted between members of the same family (Stolz, 2014). The most abundant phylotype in the consortium was *Sphingobium* BQ1, however, functional metagenomic analysis of its corresponding MAG did not reveal a clear metabolic function. Quantification of BVMO-coding gene transcripts (sphantq_4473 and 4479) during biodegradation of B*a*AQ by consortium BQ revealed a concurrent expression of these genes and *Sphingobium* BQ1 (Figure 3.7). Altogether, it becomes plausible that the recipient of plasmid *pANTQ-1* could have been *Sphingobium* BQ1, conferring it the ability to degrade B*a*AQ.

To assess the environmental occurrence of this finding, the abundance of phylotype BQ1 and BVMO sphantq_4473 was assessed by qPCR during a pilot-scale biostimulation assay of the original creosote-contaminated soil (Tauler, 2015) (Figure 3.8). In that study, an important and sequential degradation of PAHs was accompanied by a transient accumulation of oxy-PAHs in the soil. B*a*AQ was formed at the same time B*a*A, its parent compound, was being degraded. The increase in *Sphingobium* BQ1 16S rRNA gene copies coincided with a deceleration and the eventual cease in the formation of B*a*AQ. Moreover, BVMO sphantq_4473 gene copies followed identical trends as *Sphingobium* BQ1 16S rRNA gene copies, with concurrent maxima after 45 days of incubation.

Interestingly, the gene copy number of sphantq_4473 was significantly higher (between 1 or 2 orders of magnitude), than that of *Sphingobium* sp. strain AntQ-1

16S rRNA during the entire incubation. Considering that the approximate cellular abundance of plasmid *pANTQ-1*, inferred by the plasmid read depth with respect to the chromosomal read depth in the genome, indicated that strain AntQ-1 probably contains a single copy of *pANTQ-1* per cell (*chr1* 1x depth, *pANTQ-1* 0.87x depth), differences in sphantq_4473 and *Sphingobium* sp. strain AntQ-1 16S rRNA gene copy numbers cannot be explained by the existence of multiple copies of the plasmid in a cell. Therefore, our results indicate that, to a certain extent, horizontal gene transfer of plasmid *pANTQ-1* naturally occurred *in situ* in the contaminated soil between sphingomonads, probably including the uncultured phylotype *Sphingobium* BQ1 detected here.

Exploring the ecological roles of BQ phylotypes

Interestingly, many distinct phylotypes in consortium BQ presumably performed the same metabolic functions, indicating functional redundancy in the microbial community (Louca et al., 2018). Metabolic redundancy seems to be a common aspect of many microbial systems and it has also been observed for PAH degradation (Dunlevy et al., 2013; Guazzaroni et al., 2013). However, the fact that different microorganisms catalyzing the same reactions are maintained in the consortium after multiple transfers with B*a*AQ as the only supplied carbon source, suggests that the coexisting phylotypes have differentiated ecological roles within the community. Some key steps in the aromatic degradation pathways were only detected in a single MAG, in particular, *nahE* gene was only present in the MAG of *Bradyrhizobium* BQ13 or *phdJ* was only possessed by *Pusillimonas* BQ3. Thus, the complete transformation of both 1,2-dihydroxynapthtalene and 1-hydroxy-2-naphthoic acid depended on the action of these singular phylotypes. This indicates the existence of syntrophic associations between the members of consortium BQ, as the rest of the community depends on specific phylotypes for further catabolism of aromatic compounds (Morris et al., 2013). In fact, efforts to isolate members of consortium BQ have failed, likely because of the interdependencies within partner species.

These interdependencies could also be occurring due to exchange of vitamins, amino acids and other cofactors (Seth & Taga, 2014; Zengler & Zaramela, 2018). In an attempt to decipher putative nutritional interactions in consortium BQ, we reconstructed the biosynthetic pathways of essential vitamins and cofactors within the MAGs (Figure 3.9). None of the MAGs encompassed complete pathways for all the selected vitamins and cofactors, indicating that all phylotypes were potentially auxotrophic in at least one compound. Considering the large number of flavin-dependent enzymes encoded in the MAGs, including BVMOs that are key for B*a*AQ degradation, it is surprising that only *Pseudomonas* BQ5, *Olivibacter* BQ4 and
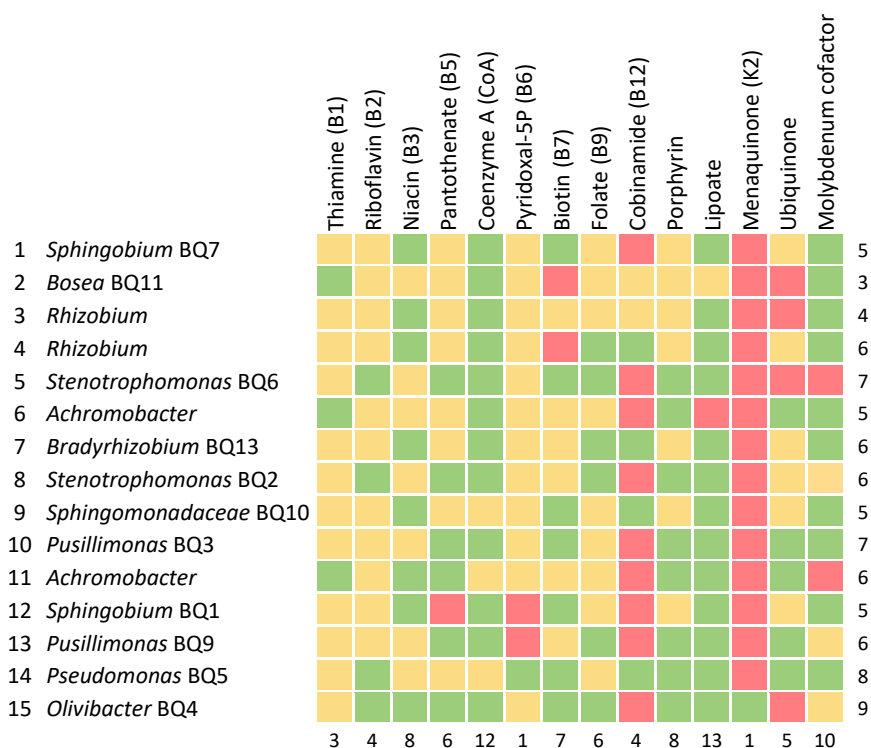
Figure 3.9. Biosynthesis pathways of cofactors and vitamins in the MAGs from consortium BQ. In green, complete pathway; in yellow, one or two steps missing; in red, incomplete or missing pathway. Numbers indicate the number of complete pathways for each MAG or for each cofactor or vitamin. Reconstruction of metabolic pathways was done using KEGG GhostKOALA.

Stenotrophomonas BQ2 and BQ6 were able to completely synthesize riboflavin (vitamin B2). Thiamine (vitamin B1), an essential cofactor for a variety of enzymes such as pyruvate dehydrogenase or pyruvate decarboxylase (Lawhorn et al., 2004), could just be completely produced by *Bosea* BQ11 and *Achromobacter* (MAGs 6 and 11). Solely *Olivibacter* BQ4 showed the ability to synthesize menaquinone, however, this quinone is key in electron transport of anaerobic bacteria and Gram-positive bacteria, whereas aerobic Gram-negative bacteria utilize ubiquinone (Kurosu & Begari, 2010), thus, it does not seem a limiting factor in this community. Pyridoxal-5P (vitamin B6), which plays an important role in amino acid and carbohydrate metabolism (Richts et al., 2019), could only be completely produced by *Pseudomonas* BQ5. Moreover, the lack of cobinamide (vitamin B12) biosynthetic pathway was widespread among most of the consortium members, *Pseudomonas* BQ5, *Sphingomonadaceae* B10, *Rhizobium*

(MAG4) and *Bradyrhizobium* BQ13 were the only phylotypes with the complete pathway. Vitamin B12 is a widely required coenzyme involved in the synthesis of nucleotides and amino acids, in addition to carbon processing and gene regulation, despite this, only a relatively small subset of bacteria are capable of its production (Lu et al., 2019). Actually, ANTQ-degrading *Sphingobium* sp. AntQ-1 is auxotrophic for this vitamin. Given all these potential auxotrophies within the community members of consortium BQ, it seems clear that key interactions were taking place, not only to provide suitable carbon and energy sources, but also to supply essential nutrients and cofactors. Metabolic and nutritional cross-feeding appears to be shaping the community, where each phylotype performs a distinct ecological role despite the metabolic functional redundancy. The metagenomic functional analysis provided an insight into how the consortium members may interact with each other, resulting in the preservation of a stable community structure and an efficient B*a*AQ removal.

General Discussion

General discussion

PAH biodegradation in contaminated soils is a multifaceted process involving a variety of co-occurring contaminants and complex microbial communities harboring different and often interconnected metabolic pathways, influenced by several environmental factors. To achieve a comprehensive understanding of environmental PAH biodegradation it is necessary to apply multidisciplinary approaches that integrate metabolic, taxonomical and functional information from pure cultures, microbial consortia and environmental communities (Vila et al., 2015). In this Thesis, the combination of multi-omic approaches aimed to shed light on the mechanisms driving PAH biodegradation both at the single species and community levels. A previous biostimulation assay (Tauler, 2015) on the creosote-contaminated soil used here served as a basis for further deciphering the specific populations, functions and interactions ruling the fate of the different polyaromatic compounds. Special focus was brought to HMW-PAHs, enriched in the residual fraction remaining after bioremediation, and to oxy-PAHs, whose formation and subsequent degradation in the soil suggested its role as catabolic nodes in the PAH metabolic networks.

In this Thesis, benz(*a*)anthracene was chosen as a model to ascertain the degradative mechanisms for HMW-PAH removal. During bioremediation of PAH-contaminated soils, relatively effective biodegradation of B*a*A is is generally observed , and the microbial communities associated have been the subject of several studies (Lladó et al., 2009; Jones et al., 2011; Singleton et al., 2011; Corona et al., 2017). However, the majority of the B*a*A-degrading bacterial strains reported in the literature are unable to use this compound as a growth substrate, only acting on it cometabolically and thus needing enzymatic induction by co-occurring substrates (Kanaly & Harayama, 2010). Additionally, in comparison with other abundant 4-ring PAHs (including fluoranthene, pyrene and chrysene) BaA has an intermediate water-solubility (0.009 mg·L^{-1}; Table 1 in *Introduction*), which makes it susceptible of interactions with co-ocurring HMW-PAHs with higher and lower water-solubilities. During the mentioned biostimulation study (Tauler et al., 2015), all 4-ring PAHs were depleted simultaneously, B*a*A showing slower biodegradation rates than FT or PY and similar kinetics to CHY, which suggested potential catabolic interactions between them. Thus, it is expected that B*a*A degradation may proceed through a variety of mechanisms when found in the presence of PAH mixtures.

Here, the B*a*A-degrading community within the creosote-polluted soil was assessed by DNA-SIP using ^{13}C-labelled BaA. As previously observed for a PAH-contaminated

soil from a manufactured-gas plant from North Carolina (Jones et al., 2011), the primary player in the biodegradation of this compound (84% in 16S rRNA clone libraries) was identified as a member of the recently described genus *Immundisolibacter* (Corteselli et al., 2017). A functional insight into the BaA degrading community was accomplished by shotgun metagenomic sequencing of ^{13}C-enriched DNA. The metagenome-assembled genome (MAG) corresponding to *Immunisolibacter* encompassed a vast arsenal of genes coding for dioxygenases and monooxygenases with potential activity on aromatic compounds. Interestingly, a gene cluster probably related to HMW-PAH biodegradation, composed of genes coding for 7 α-subunits and 5 β-subunits of aromatic ring-hydroxylating dioxygenases (RHD), 2 aromatic ring-cleaving dioxygenases, 1 hydrolase and a TetR/AcrR family transcriptional regulator, shared the exact same configuration as a cluster enclosed in the genome of *Immundisolibacter cernigliae* TR3.2 with high sequence homology (78-98%). Furthermore, phylogenetic analysis of RHD α-subunits revealed that one of the RHD encoded in this cluster was part of a separate branch including RHDs essentially found in *Immundisolibacter* members from geographically diverse hydrocarbon-contaminated sites. The conserved genetic organization within this genus, encoding enzymes distantly related to well-characterized RHDs, together with its widespread distribution, suggest its pivotal role in the biodegradation of polyaromatic hydrocarbons. Given the potentially untapped PAH-degradation mechanisms they harbor, it becomes crucial to cultivate members of this novel *Immundisolibacterales* order. To date, only one strain belonging to this order has been isolated and its PAH metabolism has not yet been investigated. Numerous efforts were made to isolate *Immundisolibacter* members from the creosote-polluted soil of study, but they were unfruitful. Now, the information extracted from the MAG on their nutritional requirements, will facilitate directed isolation (Gutleben et al., 2018). Culture-based methods will allow unequivocal elucidation of the BaA metabolic pathway and the key functional genes and enzymes involved within this BaA-degrading *Immundisolibacter*.

Relevant interactions between BaA and co-occurring HMW-PAHs were ascertained in soil microcosms spiked with binary mixtures of BaA and FT, PY or CHY. Within these mixtures, a sequential removal of these PAHs, consistent with their water solubility, was observed. The presence of FT or PY delayed the biodegradation of BaA, while CHY utilization was delayed by the presence of BaA. These preferences in substrate utilization were associated to community interactions within the microbial populations of the creosote-contaminated soil. The BaA primary degrader, *Immundisolibacter*, was outcompeted by members of the community associated to the degradation of the more water-soluble FT and PY, mainly affiliated to the genera *Sphingobium* and *Mycobacterium*, respectively. In fact, relative

abundance of *Sphingobium* members was even enhanced when B*a*A and FT were supplied together. These results suggest that, in the presence of FT, part of B*a*A degradation was probably channeled through cometabolic reactions by members of *Sphingobium.* The ability of members of *Sphingobium* to cometabolically transform B*a*A has been previously described (Kunihiro et al., 2013). In the case of PY, despite some PY-degrading mycobacteria have shown the ability to transform B*a*A (Moody et al., 2005), the abundance of *Mycobacterium* did not seem to increase significantly in the presence of BaA. Therefore, the delay in BaA utilization could be probably associated to competitive interactions between members of *Mycobacterium* and *Immundisolibacter*. Conversely, the combination of B*a*A and CHY produced a synergic effect on *Immundisolibacter*, that resulted in an increased biodegradation of CHY. As a consequence of these interactions, there was a greater formation of benz(*a*)anthracene-7,12-dione, the ready oxidation product of B*a*A. However, this oxy-PAH was only transiently accumulated as it was subsequently degraded, confirming the presence of populations that could efficiently remove it. Altogether, these findings highlight the important role of the interactive effects that may take place when PAHs are found in mixtures within a complex soil microbiome, as these will determine their biodegradability and fate. In this context, cometabolism and partial degradation processes gain special relevance in funneling PAH carbon fluxes.

The biodegradation of oxy-PAHs was investigated for the 3-ring 9,10-anthraquinone (ANTQ) and the 4-ring benz(*a*)anthracene-7,12-dione (B*a*AQ). These compounds were chosen as representatives of LMW and HMW-PAH partial oxidation products, respectively. Several reports have identified the common presence of these two compounds in PAH contaminated sites with significant concentrations (Lundstedt et al., 2014). In the biostimulation assay by Tauler, both compounds were transiently accumulated during biodegradation of ANT and B*a*A. ANTQ has been classified as possibly carcinogenic to humans by the IARC (IARC, 2013) and B*a*AQ has been proposed as a priority compound for ecotoxicological risk assessment due to its genotoxic potential (Clergé et al., 2019). On this basis, recent works have raised increasing concern on the potential toxicity and persistence of oxy-PAHs (Titaley et al., 2020), emphasizing the need to include them in risk analysis and post-remediation effectivity assessment. In this Thesis, we aimed to decipher the bacterial mechanisms that determine their fate in PAH-contaminated soils, providing insights into the capabilities of native microbial communities to mitigate the risk associated to oxy-PAH formation during bioremediation.

We conducted two analogous experiments to compare the microbial populations driving the biodegradation of these two oxy PAHs and their corresponding parent compounds (ANT and B*a*A, respectively). For each PAH/oxy-PAH pair we identified the presence of divergent community structures, indicative of the existence of

populations specialized in the utilization of the oxy-PAHs. The biodegradation of ANT and ANTQ was associated to members of *Sphingobium*, however different phylotypes were enriched in response to each substrate. As mentioned above, B*a*A was utilized by a member of *Immundisolibacter*, while the degradation of the corresponding quinone (B*a*AQ) was associated to unclassified members of the order *Sphingomonadales*. These results are in agreement with previous observations from a PAH-contaminated soil from North Carolina (USA), where distinctive populations were also associated to the assimilation of ANT and ANTQ by DNA-SIP (Rodgers-Vieira et al., 2015). Furthermore, the obtained ANTQ-degrading isolate (*Sphingobium* sp. AntQ-1) and the B*a*AQ-degrading consortium were not able to grow or transform the respective parent PAHs. This confirms the existence of highly specialized microbial populations within the soil community with the ability to degrade PAH transformation products. Oxy-PAHs have been generally regarded as dead-end products in PAH metabolic and cometabolic pathways from pure cultures (Grifoll et al., 1992; Moody et al., 2001; van Herwijnen et al., 2003; Roy, 2013). Thus, the dead-end metabolites from PAH-degrading populations become substrates for the oxy-PAH degrading microorganisms, acting as catabolic nodes within the soil metabolic network.

The isolation of the first ANTQ-degrading bacterial isolate, *Sphingobium* sp AntQ-1, allowed the comprehensive reconstruction of the ANTQ metabolic pathway by combining multi-omic approaches. The integration of data from metabolomic screens and metagenomic functional analysis of the B*a*AQ-degrading bacterial consortium BQ also permitted the decipherment of the metabolism of B*a*AQ. Interestingly, both oxy-PAHs were metabolized through Baeyer-Villiger oxidations to give the corresponding lactones that were then subsequently hydrolyzed resulting in the cleavage of the quinone ring. The initial attack is orchestrated by a common gene encoding a Baeyer-Villiger monooxygenase (BVMO). These enzymes are known for their exquisite chemo-, regio- and enantioselectivity (Mascotti et al., 2015), which could explain the extremely specialized oxy-PAH degrading populations encountered in the PAH-contaminated soil. For instance, the well-studied cyclohexanone monooxygenase readily accepts a wide range of cyclic aliphatic ketones but will not convert aromatic or simple aliphatic compounds (Torres Pazmiño et al., 2010). BV monooxygenation of oxy-PAHs has been suggested for intermediates in the metabolism of fluorene and fluoranthene (Grifoll et al., 1992, 1994; Casellas et al., 1997, 1998; Weissenfels et al., 1991), however, the enzymes catalyzing these oxidations had never been elucidated. The environmental relevance of the Baeyer-Villiger mechanism in determining the fate of oxy-PAHs *in situ* was confirmed by monitoring the identified genes in the biostimulation assay of the creosote-contaminated soil, as the BVMO-coding gene abundance revealed a

perfect consonance with the biodegradation kinetics of ANTQ and B*a*AQ. Our results suggest that BV monooxygenases could be a key step in the processing of oxy-PAHs in contaminated soils, thus untapping their potential key role in decreasing the risk posed by the formation of genotoxic transformation products. Further research is needed in order to assess their potential contribution for the processing of other PAH-oxidation products, such as 2H-naphtho[2,1,8-*def*]chromen-2-one, a genotoxic lactone formed during the transformation of pyrene by different bacterial isolates, including members of *Immundisolibacter* (Tian et al., 2017).

In addition to performing key reactions in the metabolic pathways of a wide range of carbon sources such as linear alkanones or alkanals, terpenoids, steroids or alicyclic and aromatic ketones (Leisch et al., 2011), BVMOs are specially known for their potential use as biocatalysts in the biosynthesis of polyketides with antibiotic or antitumoral activity (Fürst et al., 2019). These natural products share some structure analogy to oxy-PAHs (Zhang et al., 2018); thus, it seems that the identified oxy-PAH-oxidizing BVMOs could potentially be used as biocatalytic enzymes to obtain compounds of interest. Also, the presence of these BVMOs in the PAH-contaminated soil could be explained by their involvement in processing secondary metabolites commonly encountered in soils from the action of microorganisms or plants.

Essential genes coding for the enzymes channeling the biodegradation of ANTQ were principally located in the megaplasmid *pANTQ-1* belonging to the genome of *Sphingobium* sp. AntQ-1. Plasmid *pANTQ-1* carried the genes for the upper ANTQ-metabolic pathway and the complete phthalate and catechol metabolism to intermediates of the TCA cycle. Strikingly, a major part of this degradative plasmid was also found in the metagenomic library from the B*a*AQ-degrading bacterial consortium. The presence of genes coding for transposases and type IV secretion system components within the plasmid supported the occurrence of horizontal gene transfer to a member of consortium BQ, probably to the *Sphingobium* BQ1 phylotype. Sphingomonads commonly encompass their degradative genes on plasmids or megaplasmids, which contributes to the extraordinary catabolic flexibility of this group (Aylward et al., 2013). Conjugative plasmid transfer in sphingomonads mainly occurs among members of the same family (Stolz, 2014), conferring them a selective advantage in front of other bacteria as they can rapidly adapt to new potential carbon sources in the environment (Top & Springael, 2003). Interestingly, a wide diversity of *Sphingobium* phylotypes were found in the creosote-polluted soil, each one associated to the biodegradation of different polyaromatic compounds, i.e., ANT, B*a*A, FT, ANTQ or B*a*AQ. This indicates that each *Sphingobium* phylotype would preferentially utilize a specific substrate for growth, however, they would likely also be catalyzing partial oxidation processes of other

compounds. For instance, the member of *Sphingobium* associated to FT utilization was enhanced when co-incubated with B*a*A, indicating that it could benefit by the cometabolization of the latter. Different *Sphingobium* strains have been reported to only grow on a limited number of PAHs but, at the same time, the relaxed specificity of their RHD confers them the ability to transform a variety of other compounds (Vila et al., 2015). For example, *Sphingobium* sp. PNB is only capable of growing with phenanthrene, naphthalene and biphenyl as sole carbon and energy sources, but can cometabolize FT, B*a*A, PY and benzo(a)pyrene (Roy, 2013). The promiscuity of their enzymes together with the dissemination of degradative genes by horizontal gene transfer, enables this group to evolve and acquire new biodegradation capacities and pathways (Kolvenbach et al., 2014).

Despite the advantageous catabolic versatility of sphingomonads, the ANTQ-degrading *Sphingobium* sp. AntQ-1 and the *Sphingobium* phylotypes within the B*a*AQ-degrading consortium displayed vitamin auxotrophies, indicating they depend on other members of the soil community to obtain these vitamins. The reconstruction of biosynthetic pathways of essential vitamins and cofactors within the MAGs of the B*a*AQ-degrading consortium BQ not only revealed potential auxotrophies for this taxonomic group but for all the consortium constituents identified. This suggests the existence of interspecies dependencies among the consortium members to obtain the required cofactors for growth. Most microorganisms in the environment are auxotrophs (Zengler & Zaramela, 2018), thus relying on external nutrients for thriving. These nutritional requirements have profound implications for microbial interactions and thus the overall microbial network, as they create interdependencies that shape the composition and function of the microbial community (Seth & Taga, 2014). Besides these potential nutritional dependencies, bacterial consortium BQ also showed syntrophic associations, as genes coding for key steps in the catabolism of aromatic compounds were only harbored by unique phylotypes. Biodegradation of B*a*AQ appeared to rely, thus, on metabolic and nutritional cross-feeding, where each member of the community performs a distinct ecological role. The differentiated ecological roles explain the apparent metabolic functional redundancy inferred from the metagenomic analysis. Many essential microbial partnerships occur in nature, such as the cooperative biodegradation of perchloroethylene (PCE) or trichloroethylene (TCE) through syntrophic associations and cofactor exchanges (Field & Sierra-Alvarez, 2004). A recent study also showed potential metabolic exchanges and amino acid and vitamin auxotrophies in members of hydrocarbon-degrading aerobic and anaerobic communities from an oil reservoir (Liu et al., 2018). Microbial consortium BQ served as a simplified model for multispecies mechanisms and interactions within the soil microbial community and highlighted their importance for efficient removal of

polyaromatic compounds in soils, opening the door to the necessity of novel cofactors.

Altogether, the results gathered in this Thesis evidence that biodegradation of PAH mixtures in contaminated soils is channeled through interactive metabolic networks composed of (1) primary PAH-degraders that completely mineralize PAHs or (2) produce dead-end products, such as oxy-PAHs, as a result of branched metabolic pathways or cometabolic reactions; (3) specialized oxy-PAHs degraders; and (4) populations feeding on simple aromatic metabolites producing amino acids, vitamins and cofactors required by other community members (Figure 1).
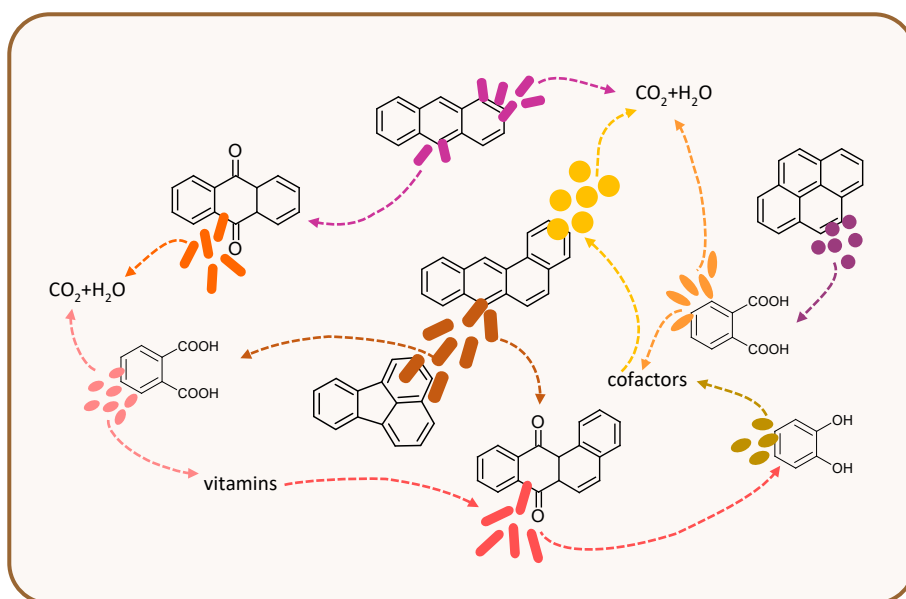
Figure 1. Schematic model for the bacterial metabolic networks channeling PAH biodegradation in contaminated soils.

Conclusions

Conclusions

I.   DNA stable isotope probing, using U^{13}C-benz(*a*)anthracene, in combination with shotgun metagenomics of ^{13}C-labeled DNA revealed that the main B*a*A-degrader in the creosote-polluted soil was a member of the genus *Immundisolibacter*.

II.   Functional analysis of the *Immundisolibacter* metagenome-assembled genome untapped a conserved unique genetic organization within this genus, harboring novel aromatic ring-hydroxylating dioxygenases.

III.   Co-incubation of B*a*A with other HMW-PAHs in soil microcosms influenced their biodegradation, causing a major delay in the removal of the less soluble PAHs and an increased formation of B*a*AQ, the ready oxidation product of B*a*A.

IV.   Biodegradation of the HMW-PAHs was associated to four key players: *Pseudoxanthomonas*, for PY and CHY; *Immundisolibacter*, for B*a*A and CHY; *Sphingobium,* for FT; and *Mycobacterium*, for PY.

V.   Microbial interactions occurred during B*a*A degradation in the presence of co-substrates. Members of *Immundisolibacter* were outcompeted by FT and PY degraders, while their action was enhanced in the presence of CHY. Similarly, *Sphingobium* was favored by the co-incubation of FT and B*a*A. This demonstrates that microbial interactive processes within soil communities operate in preferential utilization of PAHs in mixtures.

VI.   Different *Sphingobium* phylotypes were linked to the removal of ANT and ANTQ in sand-in-liquid soil microcosms, suggesting specialized oxy-PAH degrading populations.

VII.   *Sphingobium* sp. AntQ-1 was isolated from the ANTQ-spiked sand-in-liquid soil microcosms. The strain had the ability to grow on ANTQ as sole carbon and energy source, but not on ANT.

VIII.   Polyomic reconstruction of the ANTQ metabolic pathway by *Sphingobium* sp. AntQ-1 revealed that this oxy-PAH is metabolized by two consecutive Baeyer-Villiger oxidations followed by hydrolyzations of the C-O bond. This results in the cleavage of the central ring and the formation of phthalate and catechol conveying to intermediates of the central metabolism.

IX.   Essential genes for the ANTQ biodegradation pathway were located in megaplasmid *pANTQ-1*, including the genes coding for enzymes in the upper ANTQ pathway, and the lower routes for the metabolism of phthalate and *meta*-cleavage of catechol.

X.   The environmental relevance of the identified Baeyer-Villiger monooxygenase was confirmed by monitoring the BVMO-coding gene in a previous pilot-scale biostimulation assay of the creosote-contaminated soil.

XI.   Biodegradation of B*a*AQ in sand-in-liquid soil microcosms was attributed to a specialized uncultured group of *Sphingomonadales*, confirming again the existence of oxy-PAH degrading specialists in the soil community.

XII.   A B*a*AQ-degrading bacterial consortium was established by enrichment in sand-in-liquid cultures using this oxy-PAH as the sole carbon source. The consortium was constituted by members of *Sphingobium*, *Stenotrophomonas*, *Pusillimonas*, *Olivibacter*, *Pseudomonas*, *Achromobacter*, *Sphingomonadaceae*, *Bosea*, *Rhizobium* and *Bradyrhizobium*.

XIII.   The integration of data from metabolomic screens and metagenomic functional gene analyses revealed that the B*a*AQ metabolic pathway was also initiated by a Baeyer-Villiger oxidation.

XIV.   The presence of plasmid *pANTQ-1* in the metagenomic library from consortium BQ suggested the occurrence of a horizontal gene transfer event, probably between sphingomonads.

XV.   Metagenomic functional analysis indicated the existence of metabolic redundancy within consortium BQ, which could be explained by syntrophic associations and nutritional interdependencies that confer each phylotype a distinct ecological role.

References

References

Achten, C., & Andersson, J. T. (2015). Overview of Polycyclic Aromatic Compounds (PAC). *Polycyclic Aromatic Compounds*, *35*(2–4), 177–186. https://doi.org/10.1080/10406638.2014.994071

Adrion, A. C., Singleton, D. R., Nakamura, J., Shea, D., & Aitken, M. D. (2016). Improving Polycyclic Aromatic Hydrocarbon Biodegradation in Contaminated Soil Through Low-Level Surfactant Addition after Conventional Bioremediation. *Environmental Engineering Science*, *33*(9), 659–670. https://doi.org/10.1089/ees.2016.0128

Ahmad, F., Zhu, D., & Sun, J. (2020). Bacterial chemotaxis: a way forward to aromatic compounds biodegradation. In *Environmental Sciences Europe* (Vol. 32, Issue 1). https://doi.org/10.1186/s12302-020-00329-2

Alexander, J., Benford, D., Cockburn, A., Cravedi, J.-P., Dogliotti, E., Di Domenico, A., Luisa Fernández-Cruz, M., Fink-Gremmels, J., Fürst, P., Galli, C., Grandjean, P., Gzyl, J., Heinemeyer, G., Johansson, N., Mutti, A., Schlatter, J., van Leeuwen, R., Van Peteghem, C., & Verger, P. (2008). Polycyclic Aromatic Hydrocarbons in Food - Scientific Opinion of the Panel on Contaminants in the Food Chain. *EFSA Journal*, *6*(8), 1–114. https://doi.org/10.2903/j.efsa.2008.724

Andersson, B. E., Lundstedt, S., Tornberg, K., Schnürer, Y., Öberg, L. G., & Mattiasson, B. (2003). Incomplete degradation of polycyclic aromatic hydrocarbons in soil inoculated with wood-rotting fungi and their effect on the indigenous soil bacteria. *Environmental Toxicology and Chemistry*, *22*(6), 1238–1243. https://doi.org/10.1002/etc.5620220608

Andersson, J. T., & Achten, C. (2015). Time to Say Goodbye to the 16 EPA PAHs? Toward an Up-to-Date Use of PACs for Environmental Purposes. *Polycyclic Aromatic Compounds*, *35*(2–4), 330–354. https://doi.org/10.1080/10406638.2014.991042

Antipov, D., Raiko, M., Lapidus, A., & Pevzner, P. A. (2019). Plasmid detection and assembly in genomic and metagenomic data sets. *Genome Research*, *29*(6), 961–968. https://doi.org/10.1101/GR.241299.118

Aylward, F. O., McDonald, B. R., Adams, S. M., Valenzuela, A., Schmidt, R. A., Goodwin, L. A., Woyke, T., Currie, C. R., Suen, G., & Poulsen, M. (2013). Comparison of 26 Sphingomonad Genomes Reveals Diverse Environmental Adaptations and Biodegradative Capabilities. *Applied and Environmental Microbiology*, *79*(12), 3724–3733. https://doi.org/10.1128/AEM.00518-13

Baborová, P., Möder, M., Baldrian, P., Cajthamlová, K., & Cajthaml, T. (2006). Purification of a new manganese peroxidase of the white-rot fungus Irpex lacteus, and degradation of polycyclic aromatic hydrocarbons by the enzyme. *Research in Microbiology*, *157*, 248–253. https://doi.org/10.1016/j.resmic.2005.09.001

Baboshin, M., Akimov, V., Baskunov, B., Born, T. L., Khan, S. U., & Golovleva, L. (2008). Conversion of polycyclic aromatic hydrocarbons by Sphingomonas sp. VKM B-2434. *Biodegradation*, *19*(4), 567–576. https://doi.org/10.1007/s10532-007-9162-2

Bamforth, S. M., & Singleton, I. (2005). Bioremediation of polycyclic aromatic hydrocarbons: Current knowledge and future directions. *Journal of Chemical Technology and Biotechnology*, *80*(7), 723–736. https://doi.org/10.1002/jctb.1276

Bandowe, B. A. M. M., & Wilcke, W. (2010). Analysis of Polycyclic Aromatic Hydrocarbons and Their Oxygen-Containing Derivatives and Metabolites in Soils. *Journal of Environmental Quality*, *39*(4), 1349–1358. https://doi.org/10.2134/jeq2009.0298

Banerjee, D. K., Fedorak, P. M., Hashimoto, A., Masliyah, J. H., Pickard, M. A., & Gray, M. R. (1995). Monitoring the biological treatment of anthracene-contaminated soil in a rotating-drum bioreactor. *Applied Microbiology and Biotechnology*, *43*(3), 521–528. https://doi.org/10.1007/BF00218460

Basta, T., Buerger, S., & Stolz, A. (2005). Structural and replicative diversity of large plasmids from sphingomonads that degrade polycyclic aromatic compounds and xenobiotics. *Microbiology*, *151*(6), 2025–2037. https://doi.org/10.1099/mic.0.27965-0

Bastiaens, L., Springael, D., Wattiau, P., Harms, H., DeWachter, R., Verachtert, H., & Diels, L. (2000). Isolation of adherent polycyclic aromatic hydrocarbon (PAH)-degrading bacteria using PAH-sorbing carriers. *Applied and Environmental Microbiology*, *66*(5), 1834–1843. https://doi.org/10.1128/AEM.66.5.1834-1843.2000

Belda, E., Heck, R. G. A. van, Lopez-Sanchez, M. J., Cruveiller, S., Barbe, V., Fraser, C., Klenk, H.-P., Petersen, J., Morgat, A., Nikel, P. I., Vallenet, D., Rouy, Z., Sekowska, A., Santos, V. A. P. M. dos, Lorenzo, V. de, Danchin, A., & Médigue, C. (2016). The revisited genome of Pseudomonas putida KT2440 enlightens its value as a robust metabolic chassis. *Environmental Microbiology*, *18*(10), 3403–3424. https://doi.org/10.1111/1462-2920.13230

Biache, C., Ouali, S., Cébron, A., Lorgeoux, C., Colombano, S., & Faure, P. (2017). Bioremediation of PAH-contamined soils: Consequences on formation and degradation of polar-polycyclic aromatic compounds and microbial community abundance. *Journal of Hazardous Materials*, *329*, 1–10. https://doi.org/10.1016/j.jhazmat.2017.01.026

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. *Applied Microbiology and Biotechnology*, *43*(1), 156–164. https://doi.org/10.1007/BF00170638

Boulangé, M., Lorgeoux, C., Biache, C., Michel, J., Michels, R., & Faure, P. (2019). Aging as the main factor controlling PAH and polar-PAC (polycyclic aromatic compound) release mechanisms in historically coal-tar-contaminated soils. *Environmental Science and Pollution Research*, *26*(2), 1693–1705. https://doi.org/10.1007/s11356-018-3708-1

Brezna, B., Kweon, O., Stingley, R. L., Freeman, J. P., Khan, A. A., Polek, B., Jones, R. C., & Cerniglia, C. E. (2006). Molecular characterization of cytochrome P450 genes in the polycyclic aromatic hydrocarbon degrading Mycobacterium vanbaalenii PYR-1. *Applied Microbiology and Biotechnology*, *71*(4), 522–532. https://doi.org/10.1007/s00253-005-0190-8

Cajthaml, T., Erbanová, P., Šašek, V., & Moeder, M. (2006). Breakdown products on metabolic pathway of degradation of benz[a]anthracene by a ligninolytic fungus. *Chemosphere*, *64*(4), 560–564. https://doi.org/10.1016/j.chemosphere.2005.11.034

Cajthaml, T., Möder, M., Kačer, P., Šašek, V., & Popp, P. (2002). Study of fungal degradation products of polycyclic aromatic hydrocarbons using gas chromatography with ion trap mass spectrometry detection. *Journal of Chromatography A*, *974*(1–2), 213–222. https://doi.org/10.1016/S0021-9673(02)00904-4

Casellas, M., Grifoll, M., Bayona, J. M., & Solanas, A. M. (1997). New metabolites in the degradation of fluorene by Arthrobacter sp. strain F101. *Applied and Environmental Microbiology*,

63(3), 819–826. https://doi.org/10.1128/aem.63.3.819-826.1997

Casellas, M., Grifoll, M., Sabaté, J., & Solanas, A. M. (1998). Isolation and characterization of a 9-fluorenone-degrading bacterial strain and its role in synergistic degradation of fluorene by a consortium. *Canadian Journal of Microbiology*, *44*(8), 734–742. https://doi.org/10.1139/w98-066

Cébron, A., Norini, M. P., Beguiristain, T., & Leyval, C. (2008). Real-Time PCR quantification of PAH-ring hydroxylating dioxygenase (PAH-RHDα) genes from Gram positive and Gram negative bacteria in soil and sediment samples. *Journal of Microbiological Methods*, *73*(2), 148–159. https://doi.org/10.1016/j.mimet.2008.01.009

Cerniglia, C. E. (1992). Biodegradation of polycyclic aromatic hydrocarbons. *Biodegradation*, *3*(2–3), 351–368. https://doi.org/10.1007/BF00129093

Chakraborty, J., Ghosal, D., Dutta, A., & Dutta, T. K. (2012). An insight into the origin and functional evolution of bacterial aromatic ring-hydroxylating oxygenases. *Journal of Biomolecular Structure and Dynamics*, *30*(4), 419–436. https://doi.org/10.1080/07391102.2012.682208

Chibwe, L., Geier, M. C., Nakamura, J., Tanguay, R. L., Aitken, M. D., & Simonich, S. L. M. (2015). Aerobic Bioremediation of PAH Contaminated Soil Results in Increased Genotoxicity and Developmental Toxicity. *Environmental Science and Technology*, *49*(23), 13889–13898. https://doi.org/10.1021/acs.est.5b00499

Clergé, A., Le Goff, J., Lopez, C., Ledauphin, J., & Delépée, R. (2019). Oxy-PAHs: occurrence in the environment and potential genotoxic/mutagenic risk assessment for human health. *Critical Reviews in Toxicology*, *49*(4), 302–328. https://doi.org/10.1080/10408444.2019.1605333

Corona, L., Dendooven, L., Chicken, A., Hernández, O., & Iturbe, R. (2017). Removal of Two High Molecular Weight PAHs from Soils with Different Water Content. *Bulletin of Environmental Contamination and Toxicology 2017 99:5*, *99*(5), 619–624. https://doi.org/10.1007/S00128-017-2168-5

Corteselli, E. M., Aitken, M. D., & Singleton, D. R. (2017a). Description of Immundisolibacter cernigliae gen. Nov., sp. nov., a high-molecular-weight polycyclic aromatic hydrocarbondegrading bacterium within the class Gammaproteobacteria, and proposal of Immundisolibacterales ord. nov. and Immundisolibacteraceae fa. *International Journal of Systematic and Evolutionary Microbiology*, *67*(4), 925–931. https://doi.org/10.1099/ijsem.0.001714

Corteselli, E. M., Aitken, M. D., & Singleton, D. R. (2017b). Rugosibacter aromaticivorans gen. nov., sp. nov., a bacterium within the family Rhodocyclaceae, isolated from contaminated soil, capable of degrading aromatic compounds. *International Journal of Systematic and Evolutionary Microbiology*, *67*(2), 311–318. https://doi.org/10.1099/ijsem.0.001622

Coyotzi, S., Pratscher, J., Murrell, J. C., & Neufeld, J. D. (2016). Targeted metagenomics of active microbial populations with stable-isotope probing. *Current Opinion in Biotechnology*, *41*, 1–8. https://doi.org/10.1016/J.COPBIO.2016.02.017

Cudlín, J., Steinerová, N., Sedmera, P., & Vokoun, J. (1978). Microbial analogy of Baeyer-Villiger reaction with an anthraquinone derivative. *Collection of Czechoslovak Chemical Communications*, *43*, 1808–1810.

Daniels, L., Handson, R., & Philips, J. (1994). Chemical analysis. In A. Gerhardt, R. Murray, W.

Wood, & N. Krieg (Eds.), *Methods for general and molecular bacteriology* (pp. 512–554). ASM Press.

Dasgupta, S., Cao, A., Mauer, B., Yan, B., Uno, S., & McElroy, A. (2014). Genotoxicity of oxy-PAHs to Japanese medaka (Oryzias latipes) embryos assessed using the comet assay. *Environmental Science and Pollution Research*, *21*(24), 13867–13876. https://doi.org/10.1007/s11356-014-2586-4

Davies, J. I., & Evans, W. C. (1964). Oxidative metabolism of naphthalene by soil pseudomonads. The ring-fission mechanism. *The Biochemical Journal*, *91*(2), 251–261. https://doi.org/10.1042/bj0910251

Dean-Ross, D., Moody, J., & Cerniglia, C. E. (2002). Utilization of mixtures of polycyclic aromatic hydrocarbons by bacteria isolated from contaminated sediment. *FEMS Microbiology Ecology*, *41*(1), 1–7. https://doi.org/10.1016/S0168-6496(02)00198-8

DeBruyn, J. M., Mead, T. J., & Sayler, G. S. (2012). Horizontal Transfer of PAH Catabolism Genes in Mycobacterium : Evidence from Comparative Genomics and Isolated Pyrene-Degrading Bacteria. *Environmental Science & Technology*, *46*(1), 99–106. https://doi.org/10.1021/es201607y

Deka, H., & Lahkar, J. (2017). Biodegradation of Benzo(a)anthracene Employing Paenibacillus sp. HD1PAH: A Novel Strain Isolated from Crude Oil Contaminated Soil. *Polycyclic Aromatic Compounds*, *37*(2–3), 161–169. https://doi.org/10.1080/10406638.2016.1253593

Desai, A. M., Autenrieth, R. L., Dimitriou-Christidis, P., & McDonald, T. J. (2008). Biodegradation kinetics of select polycyclic aromatic hydrocarbon (PAH) mixtures by Sphingomonas paucimobilis EPA505. *Biodegradation*, *19*(2), 223–233. https://doi.org/10.1007/s10532-007-9129-3

Duarte, M., Nielsen, A., Camarinha-Silva, A., Vilchez-Vargas, R., Bruls, T., Wos-Oxley, M. L., Jauregui, R., & Pieper, D. H. (2017). Functional soil metagenomics: elucidation of polycyclic aromatic hydrocarbon degradation potential following 12 years of in situ bioremediation. *Environmental Microbiology*, *19*(8), 2992–3011. https://doi.org/10.1111/1462-2920.13756

Dunlevy, S. R., Singleton, D. R., & Aitken, M. D. (2013). Biostimulation reveals functional redundancy of Anthracene-degrading bacteria in polycyclic aromatic hydrocarbon-contaminated soil. *Environmental Engineering Science*, *30*(11), 697–705. https://doi.org/10.1089/ees.2013.0067

Field, J. A., & Sierra-Alvarez, R. (2004). Biodegradability of chlorinated solvents and related chlorinated aliphatic compounds. In *Reviews in Environmental Science and Biotechnology* (Vol. 3, Issue 3, pp. 185–254). https://doi.org/10.1007/s11157-004-4733-8

Fraaije, M. W., Kamerbeek, N. M., Van Berkel, W. J. H., & Janssen, D. B. (2002). Identification of a Baeyer-Villiger monooxygenase sequence motif. *FEBS Letters*, *518*(1–3), 43–47. https://doi.org/10.1016/S0014-5793(02)02623-6

Fuentes, S., Barra, B., Gregory Caporaso, J., & Seeger, M. (2016). From rare to dominant: A fine-tuned soil bacterial bloom during petroleum hydrocarbon bioremediation. *Applied and Environmental Microbiology*, *82*(3), 888–896. https://doi.org/10.1128/AEM.02625-15

Fürst, M. J. L. J., Gran-Scheuch, A., Aalbers, F. S., & Fraaije, M. W. (2019). Baeyer-Villiger Monooxygenases: Tunable Oxidative Biocatalysts. *ACS Catalysis*, *9*(12), 11207–11241. https://doi.org/10.1021/acscatal.9b03396

Furuno, S., Päzolt, K., Rabe, K., Neu, T. R., Harms, H., & Wick, L. Y. (2010). Fungal mycelia allow chemotactic dispersal of polycyclic aromatic hydrocarbon-degrading bacteria in water-unsaturated systems. *Environmental Microbiology*, *12*(6), 1391–1398. https://doi.org/10.1111/j.1462-2920.2009.02022.x

Garcia-Junco, M., Gomez-Lahoz, C., Niqui-Arroyo, J. L., & Ortega-Calvo, J. J. (2003). Biosurfactant- and biodegradation-enhanced partitioning of polycyclic aromatic hydrocarbons from nonaqueous-phase liquids. *Environmental Science and Technology*, *37*(13), 2988–2996. https://doi.org/10.1021/es020197q

Ghosal, D., Ghosh, S., Dutta, T. K., & Ahn, Y. (2016). Current state of knowledge in microbial degradation of polycyclic aromatic hydrocarbons (PAHs): A review. In *Frontiers in Microbiology* (Vol. 7, Issue AUG). Frontiers Media S.A. https://doi.org/10.3389/fmicb.2016.01369

Ghosh, I., & Mukherji, S. (2017). Substrate interaction effects during pyrene biodegradation by Pseudomonas aeruginosa RS1. *Journal of Environmental Chemical Engineering*, *5*(2), 1791–1800. https://doi.org/10.1016/j.jece.2017.03.016

Gillespie, I. M. M., & Philp, J. C. (2013). Bioremediation, an environmental remediation technology for the bioeconomy. *Trends in Biotechnology*, *31*(6), 329–332. https://doi.org/10.1016/j.tibtech.2013.01.015

Green, M. R., & Sambrook, J. (2016). Precipitation of DNA with ethanol. *Cold Spring Harbor Protocols*, *2016*(12), 1116–1120. https://doi.org/10.1101/pdb.prot093377

Grifoll, M., Solanas, A.M. & Bayona, J.M. (1990) Characterization of genotoxic components in sediments by mass spectrometric techniques combined with *Salmonella*/microsome test. *Arch. Environ. Contam. Toxicol.* 19, 175–184. https://doi.org/10.1007/BF01056084

Grifoll, M., Casellas, M., Bayona, J. M., & Solanas, A. M. (1992). Isolation and characterization of a fluorene-degrading bacterium: Identification of ring oxidation and ring fission products. *Applied and Environmental Microbiology*, *58*(9), 2910–2917. https://doi.org/10.1128/aem.58.9.2910-2917.1992

Grifoll, M., Selifonov, S. A., & Chapman, P. J. (1994). Evidence for a novel pathway in the degradation of fluorene by Pseudomonas sp. strain F274. *Applied and Environmental Microbiology*, *60*(7), 2438–2449. https://doi.org/10.1128/aem.60.7.2438-2449.1994

Grifoll, M., Selifonov, S. A., Gatlin, C. V., & Chapman, P. J. (1995). Actions of a versatile fluorene-degrading bacterial isolate on polycyclic aromatic compounds. *Applied and Environmental Microbiology*, *61*(10), 3711–3723. https://doi.org/10.1128/aem.61.10.3711-3723.1995

Guazzaroni, M.-E., Herbst, F.-A., Lores, I., Tamames, J., Peláez, A. I., López-Cortés, N., Alcaide, M., Del Pozo, M. V., Vieites, J. M., von Bergen, M., Gallego, J. L. R., Bargiela, R., López-López, A., Pieper, D. H., Rosselló-Móra, R., Sánchez, J., Seifert, J., & Ferrer, M. (2013). Metaproteogenomic insights beyond bacterial response to naphthalene exposure and bio-stimulation. *The ISME Journal*, *7*(1), 122–136. https://doi.org/10.1038/ismej.2012.82

Guha, S., Peters, C. A., & Jaffé, P. R. (1999). Multisubstrate biodegradation kinetics of naphthalene, phenanthrene, and pyrene mixtures. *Biotechnology and Bioengineering*, *65*(5), 491–499. https://doi.org/10.1002/(SICI)1097-0290(19991205)65:5<491::AID-BIT1>3.0.CO;2-H

Gupta, R. S., Lo, B., & Son, J. (2018). Phylogenomics and Comparative Genomic Studies Robustly Support Division of the Genus Mycobacterium into an Emended Genus Mycobacterium and Four Novel Genera. *Frontiers in Microbiology*, *0*(FEB), 67.

https://doi.org/10.3389/FMICB.2018.00067

Gutierrez, T., Singleton, D. R., Aitken, M. D., & Semple, K. T. (2011). Stable Isotope Probing of an Algal Bloom To Identify Uncultivated Members of the Rhodobacteraceae Associated with Low-Molecular-Weight Polycyclic Aromatic Hydrocarbon Degradation. *Applied and Environmental Microbiology*, *77*(21), 7856–7860. https://doi.org/10.1128/AEM.06200-11

Gutleben, J., Chaib De Mares, M., van Elsas, J. D., Smidt, H., Overmann, J., & Sipkema, D. (2018). The multi-omics promise in context: from sequence to microbial isolate. *Critical Reviews in Microbiology*, *44*(2), 212–229. https://doi.org/10.1080/1040841X.2017.1332003

Hareland, W. A., Crawford, R. L., Chapman, P. J., & Dagley, S. (1975). Metabolic function and properties of 4-hydroxyphenylacetic acid 1-hydroxylase from Pseudomonas acidovorans. *Journal of Bacteriology*, *121*(1), 272–285. https://doi.org/10.1128/jb.121.1.272-285.1975

Haritash, A. K., & Kaushik, C. P. (2009). Biodegradation aspects of Polycyclic Aromatic Hydrocarbons (PAHs): A review. *Journal of Hazardous Materials*, *169*(1–3), 1–15. https://doi.org/10.1016/j.jhazmat.2009.03.137

He, Z., Deng, Y., Van Nostrand, J. D., Tu, Q., Xu, M., Hemme, C. L., Li, X., Wu, L., Gentry, T. J., Yin, Y., Liebich, J., Hazen, T. C., & Zhou, J. (2010). GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. *The ISME Journal 2010 4:9*, *4*(9), 1167–1179. https://doi.org/10.1038/ismej.2010.46

Head, I. M., Jones, D. M., & Röling, W. F. M. (2006). Marine microorganisms make a meal of oil. *Nature Reviews Microbiology 2006 4:3*, *4*(3), 173–182. https://doi.org/10.1038/NRMICRO1348

Hu, J., Adrion, A. C., Nakamura, J., Shea, D., & Aitken, M. D. (2014). Bioavailability of (geno)toxic contaminants in polycyclic aromatic hydrocarbon-contaminated soil before and after biological treatment. *Environmental Engineering Science*, *31*(4), 176–182. https://doi.org/10.1089/ees.2013.0409

IARC. (2013). Anthraquinone. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans - Vol. 101* (pp. 41–70). International Agency for Research on Cancer. https://monographs.iarc.fr/wp-content/uploads/2018/06/mono101-001.pdf

Idowu, O., Semple, K. T., Ramadass, K., O'Connor, W., Hansbro, P., & Thavamani, P. (2019). Beyond the obvious: Environmental health implications of polar polycyclic aromatic hydrocarbons. *Environment International*, *123*(December 2018), 543–557. https://doi.org/10.1016/j.envint.2018.12.051

Iwai, S., Johnson, T. A., Chai, B., Hashsham, S. A., & Tiedje, J. M. (2011). Comparison of the Specificities and Efficacies of Primers for Aromatic Dioxygenase Gene Analysis of Environmental Samples. *Applied and Environmental Microbiology*, *77*(11), 3551. https://doi.org/10.1128/AEM.00331-11

Iwaki, H., Grosse, S., Bergeron, H., Leisch, H., Morley, K., Hasegawa, Y., & Lau, P. C. K. (2013). Camphor pathway redux: Functional recombinant expression of 2,5-and 3,6-diketocamphane monooxygenases of Pseudomonas putida ATCC 17453 with their cognate flavin reductase catalyzing Baeyer-Villiger reactions. *Applied and Environmental Microbiology*, *79*(10), 3282–3293. https://doi.org/10.1128/AEM.03958-12

Jerina, D. M., Van Bladeren, P. J., Yagi, H., Gibson, D. T., Mahadevan, V., Neese, A. S., Koreeda, M., Sharma, N. D., & Boyd, D. R. (1984). Synthesis and absolute configuration of the bacterial cis-1,2-, cis-8,9-, and cis-10,11-dihydro diol metabolites of benz[a]anthracene formed by

strain of Beijerinckia. *The Journal of Organic Chemistry*, *49*(19), 3621–3628. https://doi.org/10.1021/jo00193a033

Jones, M. D., Singleton, D. R., Carstensen, D. P., Powell, S. N., Swanson, J. S., Pfaender, F. K., & Aitken, M. D. (2008). Effect of incubation conditions on the enrichment of pyrene-degrading bacteria identified by stable-isotope probing in an aged, PAH-contaminated soil. *Microbial Ecology*, *56*(2), 341–349. https://doi.org/10.1007/s00248-007-9352-9

Jones, M. D., Singleton, D. R., Sun, W., & Aitken, M. D. (2011a). Multiple DNA Extractions Coupled with Stable-Isotope Probing of Anthracene-Degrading Bacteria in Contaminated Soil. *Applied and Environmental Microbiology*, *77*(9), 2984–2991. https://doi.org/10.1128/AEM.01942-10

Jones, M. D., Crandell, D. W., Singleton, D. R., & Aitken, M. D. (2011b). Stable-isotope probing of the polycyclic aromatic hydrocarbon-degrading bacterial guild in a contaminated soil. *Environmental Microbiology*, *13*(10), 2623–2632. https://doi.org/10.1111/j.1462-2920.2011.02501.x

Jones, M. D., Rodgers-Vieira, E. A., Hu, J., & Aitken, M. D. (2014). Association of Growth Substrates and Bacterial Genera with Benzo[a]pyrene Mineralization in Contaminated Soil. *Environmental Engineering Science*, *31*(12), 689. https://doi.org/10.1089/EES.2014.0275

Jouanneau, Y., Micoud, J., & Meyer, C. (2007). Purification and Characterization of a Three-Component Salicylate 1-Hydroxylase from Sphingomonas sp. Strain CHY-1. *Applied and Environmental Microbiology*, *73*(23), 7515. https://doi.org/10.1128/AEM.01519-07

Jouanneau, Y., Meyer, C., & Duraffourg, N. (2016). Dihydroxylation of four- and five-ring aromatic hydrocarbons by the naphthalene dioxygenase from Sphingomonas CHY-1. *Applied Microbiology and Biotechnology*, *100*(3), 1253–1263. https://doi.org/10.1007/s00253-015-7050-y

Juhasz, A. L., & Naidu, R. (2000). Bioremediation of high molecular weight polycyclic aromatic hydrocarbons: a review of the microbial degradation of benzo[a]pyrene. *International Biodeterioration & Biodegradation*, *45*(1–2), 57–88. https://doi.org/10.1016/S0964-8305(00)00052-4

Juhasz, A., Stanley, G., & Britz, M. (2002). Metabolite repression inhibits degradation of benzo[a]pyrene and dibenz[a,h]anthracene by Stenotrophomonas maltophilia VUN 10,003. *Journal of Industrial Microbiology and Biotechnology*, *28*(2), 88–96. https://doi.org/10.1038/SJ/JIM/7000216

Kanaly, R. A., & Harayama, S. (2000). Biodegradation of High-Molecular-Weight Polycyclic Aromatic Hydrocarbons by Bacteria. *Journal of Bacteriology*, *182*(8), 2059–2067. https://doi.org/10.1128/JB.182.8.2059-2067.2000

Kanaly, R. A., & Harayama, S. (2010). Advances in the field of high-molecular-weight polycyclic aromatic hydrocarbon biodegradation by bacteriambt-130 136..164. *Microbial Biotechnology*, *3*(2), 136–164. https://doi.org/10.1111/j.1751-7915.2009.00130.x

Keith, L. H., & Telliard, W. A. (1979). Priority pollutants. I. A perspective view. *Environmental Science and Technology*, *13*(4), 416–423. https://doi.org/10.1021/es60152a601

Khara, P., Roy, M., Chakraborty, J., Ghosal, D., & Dutta, T. K. (2014). Functional characterization of diverse ring-hydroxylating oxygenases and induction of complex aromatic catabolic gene clusters in Sphingobium sp. PNB. *FEBS Open Bio*, *4*, 290–300. https://doi.org/10.1016/J.FOB.2014.03.001

Kim, K. H., Jahan, S. A., Kabir, E., & Brown, R. J. C. (2013). A review of airborne polycyclic aromatic hydrocarbons (PAHs) and their human health effects. *Environment International*, *60*, 71–80. https://doi.org/10.1016/J.ENVINT.2013.07.019

Kim, S. J., Kweon, O., Freeman, J. P., Jones, R. C., Adjei, M. D., Jhoo, J. W., Edmondson, R. D., & Cerniglia, C. E. (2006). Molecular cloning and expression of genes encoding a novel dioxygenase involved in low- and high-molecular-weight polycyclic aromatic hydrocarbon degradation in Mycobacterium vanbaalenii PYR-1. *Applied and Environmental Microbiology*, *72*(2), 1045–1054. https://doi.org/10.1128/AEM.72.2.1045-1054.2006

Kim, S. J., Kweon, O., Jones, R. C., Freeman, J. P., Edmondson, R. D., & Cerniglia, C. E. (2007). Complete and integrated pyrene degradation pathway in Mycobacterium vanbaalenii PYR-1 based on systems biology. *Journal of Bacteriology*, *189*(2), 464–472. https://doi.org/10.1128/JB.01310-06

Kim, Y.-H., Moody, J. D., Freeman, J. P., Brezna, B., Engesser, K.-H., & Cerniglia, C. E. (2004). Evidence for the existence of PAH-quinone reductase and catechol-O-methyltransferase in Mycobacterium vanbaalenii PYR-1. *Journal of Industrial Microbiology & Biotechnology*, *31*(11), 507–516. https://doi.org/10.1007/s10295-004-0178-x

Klankeo, P., Nopcharoenkul, W., & Pinyakong, O. (2009). Two novel pyrene-degrading Diaphorobacter sp. and Pseudoxanthomonas sp. isolated from soil. *Journal of Bioscience and Bioengineering*, *108*(6), 488–495. https://doi.org/10.1016/j.jbiosc.2009.05.016

Kolvenbach, B. A., Helbling, D. E., Kohler, H.-P. E., & Corvini, P. F.-X. (2014). Emerging chemicals and the evolution of biodegradation capacities and pathways in bacteria. *Current Opinion in Biotechnology*, *27*, 8–14. https://doi.org/10.1016/j.copbio.2013.08.017

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Applied and Environmental Microbiology*, *79*(17), 5112–5120. https://doi.org/10.1128/AEM.01043-13

Kunihiro, M., Ozeki, Y., Nogi, Y., Hamamura, N., & Kanaly, R. A. (2013). Benz[a]anthracene biotransformation and production of ring fission products by Sphingobium sp. strain KK22. *Applied and Environmental Microbiology*, *79*(14), 4410–4420. https://doi.org/10.1128/AEM.01129-13

Kuppusamy, S., Thavamani, P., Venkateswarlu, K., Lee, Y. B., Naidu, R., & Megharaj, M. (2016). *Remediation approaches for polycyclic aromatic hydrocarbons (PAHs) contaminated soils: Technological constraints, emerging trends and future directions*. https://doi.org/10.1016/j.chemosphere.2016.10.115

Kurosu, M., & Begari, E. (2010). Vitamin K2 in Electron Transport System: Are Enzymes Involved in Vitamin K2 Biosynthesis Promising Drug Targets? *Molecules*, *15*(3), 1531. https://doi.org/10.3390/MOLECULES15031531

Kweon, O., Kim, S. J., Freeman, J. P., Song, J., Baek, S., & Cerniglia, C. E. (2010). Substrate specificity and structural characteristics of the novel Rieske nonheme iron aromatic ring-hydroxylating oxygenases NidAB and NidA3B3 from Mycobacterium vanbaalenii PYR-1. *MBio*, *1*(2). https://doi.org/10.1128/MBIO.00135-10

Kweon, O., Kim, S. J., Holland, R. D., Chen, H., Kim, D. W., Gao, Y., Yu, L. R., Baek, S., Baek, D. H., Ahn, H., & Cerniglia, C. E. (2011). Polycyclic aromatic hydrocarbon metabolic network in Mycobacterium vanbaalenii PYR-1. *Journal of Bacteriology*, *193*(17), 4326–4337. https://doi.org/10.1128/JB.00215-11

Kweon, O., Kim, S. J., Blom, J., Kim, S. K., Kim, B. S., Baek, D. H., Park, S. I., Sutherland, J. B., & Cerniglia, C. E. (2015). Comparative functional pan-genome analyses to build connections between genomic dynamics and phenotypic evolution in polycyclic aromatic hydrocarbon metabolism in the genus Mycobacterium. *BMC Evolutionary Biology*, *15*(1), 1–23. https://doi.org/10.1186/s12862-015-0302-8

Lampi, M. A., Gurska, J., McDonald, K. I. C., Xie, F., Huang, X. D., Dixon, D. G., & Greenberg, B. M. (2006). Photoinduced toxicity of polycyclic aromatic hydrocarbons to Daphnia magna: Ultraviolet-mediated effects and the toxicity of polycyclic aromatic hydrocarbon photoproducts. *Environmental Toxicology and Chemistry*, *25*(4), 1079–1087. https://doi.org/10.1897/05-276R.1

Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N. (Nil), Baldé, A. B., Bertollini, R., Bose-O'Reilly, S., Boufford, J. I., Breysse, P. N., Chiles, T., Mahidol, C., Coll-Seck, A. M., Cropper, M. L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., … Zhong, M. (2018). The Lancet Commission on pollution and health. *The Lancet*, *391*(10119), 462–512. https://doi.org/10.1016/S0140-6736(17)32345-0

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G., & Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, *31*(9), 814–821. https://doi.org/10.1038/nbt.2676

Larsson, M., Lam, M. M., Van Hees, P., Giesy, J. P., & Engwall, M. (2018). Occurrence and leachability of polycyclic aromatic compounds in contaminated soils: Chemical and bioanalytical characterization. *Science of the Total Environment*, *622–623*, 1476–1484. https://doi.org/10.1016/j.scitotenv.2017.12.015

Lawhorn, B. G., Mehl, R. A., & Begley, T. P. (2004). Biosynthesis of the thiamin pyrimidine: the reconstitution of a remarkable rearrangement reaction. *Organic & Biomolecular Chemistry*, *2*(17), 2538–2546. https://doi.org/10.1039/B405429F

Leisch, H., Morley, K., & Lau, P. C. K. (2011). Baeyer-villiger monooxygenases: More than just green chemistry. *Chemical Reviews*, *111*(7), 4165–4222. https://doi.org/10.1021/cr1003437

Li, X., Pan, Y., Hu, S., Cheng, Y., Wang, Y., Wu, K., Zhang, S., & Yang, S. (2018). Diversity of phenanthrene and benz[a]anthracene metabolic pathways in white rot fungus Pycnoporus sanguineus 14. *International Biodeterioration and Biodegradation*, *134*, 25–30. https://doi.org/10.1016/j.ibiod.2018.07.012

Liu, J., Techtmann, S. M., Woo, H. L., Ning, D., Fortney, J. L., & Hazen, T. C. (2017). Rapid Response of Eastern Mediterranean Deep Sea Microbial Communities to Oil. *Scientific Reports 2017 7:1*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-05958-x

Liu, Y. F., Galzerani, D. D., Mbadinga, S. M., Zaramela, L. S., Gu, J. D., Mu, B. Z., & Zengler, K. (2018). Metabolic capability and in situ activity of microorganisms in an oil reservoir. *Microbiome*, *6*(1). https://doi.org/10.1186/s40168-017-0392-1

Lladó, S., Jiménez, N., Viñas, M., & Solanas, A. M. (2009). Microbial populations related to PAH biodegradation in an aged biostimulated creosote-contaminated soil. *Biodegradation*, *20*(5), 593–601. https://doi.org/10.1007/s10532-009-9247-1

López, Z., Vila, J., Ortega-Calvo, J. J., & Grifoll, M. (2008). Simultaneous biodegradation of creosote-polycyclic aromatic hydrocarbons by a pyrene-degrading Mycobacterium. *Applied Microbiology and Biotechnology*, *78*(1), 165–172.

https://doi.org/10.1007/s00253-007-1284-2

Lotfabad, S. K., & Gray, M. R. (2002). Kinetics of biodegradation of mixtures of polycyclic aromatic hydrocarbons. *Applied Microbiology and Biotechnology*, *60*(3), 361–365. https://doi.org/10.1007/s00253-002-1104-7

Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., O'Connor, M. I., Ackermann, M., Hahn, A. S., Srivastava, D. S., Crowe, S. A., Doebeli, M., & Parfrey, L. W. (2018). Function and functional redundancy in microbial systems. *Nature Ecology and Evolution*, *2*(6), 936–943. https://doi.org/10.1038/s41559-018-0519-1

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lu, H., Zhou, J., Wang, J., Liu, G., & Zhao, L. (2008). Decolorization of 1-aminoanthraquinone-2-sulfonic acid by Sphingomonas xenophaga. *World Journal of Microbiology and Biotechnology*, *24*(7), 1147–1152. https://doi.org/10.1007/s11274-007-9586-1

Lu, H., Wang, X., Zang, M., Zhou, J., Wang, J., & Guo, W. (2019). Degradation pathways and kinetics of anthraquinone compounds along with nitrate removal by a newly isolated Rhodococcus pyridinivorans GF3 under aerobic conditions. *Bioresource Technology*, *285*. https://doi.org/10.1016/j.biortech.2019.121336

Lu, X., Heal, K. R., Ingalls, A. E., Doxey, A. C., & Neufeld, J. D. (2019). Metagenomic and chemical characterization of soil cobalamin production. *The ISME Journal 2019 14:1*, *14*(1), 53–66. https://doi.org/10.1038/s41396-019-0502-0

Lundstedt, Staffan, Haglund, P., & Öberg, L. (2003). Degradation and formation of polycyclic aromatic compounds during bioslurry treatment of an aged gasworks soil. *Environmental Toxicology and Chemistry*, *22*(7), 1413–1420. https://doi.org/10.1002/etc.5620220701

Lundstedt, Staffan, White, P. A., Lemieux, C. L., Lynes, K. D., Lambert, I. B., Öberg, L., Haglund, P., & Tysklind, M. (2007). Sources, fate, and toxic hazards of oxygenated Polycyclic Aromatic Hydrocarbons (PAHs) at PAH-contaminated sites. *Ambio*, *36*(6), 475–485. https://doi.org/10.1579/0044-7447(2007)36[475:SFATHO]2.0.CO;2

Lundstedt, S., Bandowe, B. A. M., Wilcke, W., Boll, E., Christensen, J. H., Vila, J., Grifoll, M., Faure, P., Biache, C., Lorgeoux, C., Larsson, M., Frech Irgum, K., Ivarsson, P., & Ricci, M. (2014). First intercomparison study on the analysis of oxygenated polycyclic aromatic hydrocarbons (oxy-PAHs) and nitrogen heterocyclic polycyclic aromatic compounds (N-PACs) in contaminated soil. In *TrAC - Trends in Analytical Chemistry* (Vol. 57, pp. 83–92). Elsevier B.V. https://doi.org/10.1016/j.trac.2014.01.007

Luo, A., Wu, Y. R., Xu, Y., Kan, J., Qiao, J., Liang, L., Huang, T., & Hu, Z. (2016). Characterization of a cytochrome P450 monooxygenase capable of high molecular weight PAHs oxidization from Rhodococcus sp. P14. *Process Biochemistry*, *51*(12), 2127–2133. https://doi.org/10.1016/J.PROCBIO.2016.07.024

Madsen, E. L. (2006). The use of stable isotope probing techniques in bioreactor and field studies on bioremediation. *Current Opinion in Biotechnology*, *17*(1), 92–97. https://doi.org/10.1016/j.copbio.2005.12.004

Maeda, A. H., Nishi, S., Hatada, Y., Ohta, Y., Misaka, K., Kunihiro, M., Mori, J. F., & Kanaly, R. A. (2020). Chemical and genomic analyses of polycyclic aromatic hydrocarbon biodegradation in Sphingobium barthaii KK22 reveals divergent pathways in soil

sphingomonads. *International Biodeterioration and Biodegradation*, *151*. https://doi.org/10.1016/j.ibiod.2020.104993

Mahaffey, W. R., Gibson, D. T., & Cerniglia, C. E. (1988). Bacterial oxidation of chemical carcinogens: Formation of: Polycyclic aromatic acids from benz[a]anthracene. *Applied and Environmental Microbiology*, *54*(10), 2415–2423. https://doi.org/10.1128/AEM.54.10.2415-2423.1988

Maidak, B. L., Cole, J. R., Lilburn, T. G., Parker, C. T., Saxman, P. R., Farris, R. J., Garrity, G. M., Olsen, G. J., Schmidt, T. M., & Tiedje, J. M. (2001). The RDP-II (Ribosomal Database Project). *Nucleic Acids Research*, *29*(1), 173–174. https://doi.org/10.1093/nar/29.1.173

Mallick, S., Chakraborty, J., & Dutta, T. K. (2011). Role of oxygenases in guiding diverse metabolic pathways in the bacterial degradation of low-molecular-weight polycyclic aromatic hydrocarbons: A review. *Critical Reviews in Microbiology*, *37*(1), 64–90. https://doi.org/10.3109/1040841X.2010.512268

Mascotti, M. L., Lapadula, W. J., & Ayub, M. J. (2015). The origin and evolution of Baeyer - Villiger Monooxygenases (BVMOs): An ancestral family of flavin monooxygenases. *PLoS ONE*, *10*(7), 1–16. https://doi.org/10.1371/journal.pone.0132689

Matsumura, Y., Hosokawa, C., Sasaki-Mori, M., Akahira, A., Fukunaga, K., Ikeuchi, T., Oshiman, K. I., & Tsuchido, T. (2009). Isolation and characterization of novel bisphenol - A-degrading bacteria from soils. *Biocontrol Science*, *14*(4), 161–169. https://doi.org/10.4265/bio.14.161

McCarrick, S., Cunha, V., Zapletal, O., Vondráček, J., & Dreij, K. (2019). In vitro and in vivo genotoxicity of oxygenated polycyclic aromatic hydrocarbons. *Environmental Pollution*, *246*, 678–687. https://doi.org/10.1016/j.envpol.2018.12.092

Meynet, P., Head, I. M., Werner, D., & Davenport, R. J. (2015). Re-evaluation of dioxygenase gene phylogeny for the development and validation of a quantitative assay for environmental aromatic hydrocarbon degraders. *FEMS Microbiology Ecology*, *91*, 49. https://doi.org/10.1093/femsec/fiv049

Molina, M., Araujo, R., & Hodson, R. E. (1999). Cross-induction of pyrene and phenanthrene in a Mycobacterium sp. isolated from polycyclic aromatic hydrocarbon contaminated river sediments. *Canadian Journal of Microbiology*, 45(6), 520-529. https://doi.org/10.1139/cjm-45-6-520

Moody, J. D., Freeman, J. P., Doerge, D. R., & Cerniglia, C. E. (2001). Degradation of Phenanthrene and Anthracene by Cell Suspensions of Mycobacterium sp. Strain PYR-1. *Applied and Environmental Microbiology*, *67*(4), 1476–1483. https://doi.org/10.1128/AEM.67.4.1476-1483.2001

Moody, J. D., Freeman, J. P., & Cerniglia, C. E. (2005). Degradation of benz[a]anthracene by Mycobacterium vanbaalenii strain PYR-1. *Biodegradation*, *16*(6), 513–526. https://doi.org/10.1007/s10532-004-7217-1

Morris, B. E. L., Henneberger, R., Huber, H., & Moissl-Eichinger, C. (2013). Microbial syntrophy: interaction for the common good. *FEMS Microbiology Reviews*, *37*(3), 384–406. https://doi.org/10.1111/1574-6976.12019

Mueller, James G., Chapman, P. J., & Pritchard, P. H. (1989a). Creosote-contaminated sites their potential for bioremediation. *Environmental Science and Technology*, *23*(10), 1197–1201. https://doi.org/10.1021/es00068a003

Mueller, J G, Chapman, P. J., & Pritchard, P. H. (1989b). Action of a Fluoranthene-Utilizing Bacterial Community on Polycyclic Aromatic Hydrocarbon Components of Creosote. *Applied and Environmental Microbiology,* 55(12), https://doi.org/3085-3090. 10.1128/aem.55.12.3085-3090.1989

Muyzer, G., De Waal, E. C., & Uitterlinden, A. G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology*, *59*(3), 695–700. https://doi.org/10.1128/aem.59.3.695-700.1993

Nayak, A. S., Sanjeev Kumar, S., Santosh Kumar, M., Anjaneya, O., & Karegoudar, T. B. (2011). A catabolic pathway for the degradation of chrysene by Pseudoxanthomonas sp. PNK-04. *FEMS Microbiology Letters*, *320*(2), 128–134. https://doi.org/10.1111/j.1574-6968.2011.02301.x

Nostrand, J. D. Van, He, Z., & Zhou, J. (2012). Use of functional gene arrays for elucidating in situ biodegradation. *Frontiers in Microbiology*, *3*(SEP), 339. https://doi.org/10.3389/fmicb.2012.00339

Ntougias, S., Fasseas, C., & Zervakis, G. I. (2007). Olivibacter sitiensis gen. nov., sp. nov., isolated from alkaline olive-oil mill wastes in the region of Sitia, Crete. *International Journal of Systematic and Evolutionary Microbiology*, *57*(2), 398–404. https://doi.org/10.1099/ijs.0.64561-0

Ntougias, S., Lapidus, A., Han, J., Mavromatis, K., Pati, A., Chen, A., Klenk, H.-P., Woyke, T., Fasseas, C., Kyrpides, N. C., & Zervakis, G. I. (2014). High quality draft genome sequence of Olivibacter sitiensis type strain (AW-6T), a diphenol degrader with genes involved in the catechol pathway. *Standards in Genomic Sciences 2014 9:3*, *9*(3), 783–793. https://doi.org/10.4056/SIGS.5088950

Nzila, A. (2013). Update on the cometabolism of organic pollutants by bacteria. *Environmental Pollution*, *178*, 474–482. https://doi.org/10.1016/j.envpol.2013.03.042

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P. R., O'hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E., & Maintainer, H. W. (2020). Package "vegan": Community Ecology Package Version 2.5-7. In *cran.ism.ac.jp*. http://cran.ism.ac.jp/web/packages/vegan/vegan.pdf

Oostingh, G. J., Schmittner, M., Ehart, A. K., Tischler, U., & Duschl, A. (2008). A high-throughput screening method based on stably transformed human cells was used to determine the immunotoxic effects of fluoranthene and other PAHs. *Toxicology in Vitro*, *22*(5), 1301–1310. https://doi.org/10.1016/j.tiv.2008.03.003

Ortega-Calvo, J. J., Tejeda-Agredano, M. C., Jimenez-Sanchez, C., Congiu, E., Sungthong, R., Niqui-Arroyo, J. L., & Cantos, M. (2013). Is it possible to increase bioavailability but not environmental risk of PAHs in bioremediation? *Journal of Hazardous Materials*, *261*, 733–745. https://doi.org/10.1016/j.jhazmat.2013.03.042

Ortega-Calvo, J.-J., Harmsen, J., Parsons, J. R., Semple, K. T., Aitken, M. D., Ajao, C., Eadsforth, C., Galay-Burgos, M., Naidu, R., Oliver, R., Peijnenburg, W. J. G. M., Römbke, J., Streck, G., & Versonnen, B. (2015). From Bioavailability Science to Regulation of Organic Chemicals. *Environmental Science & Technology*, *49*(17), 10255–10264. https://doi.org/10.1021/acs.est.5b02412

Panagos, P., Van Liedekerke, M., Yigini, Y., & Montanarella, L. (2013). Contaminated sites in Europe: Review of the current situation based on data collected through a European

network. *Journal of Environmental and Public Health*, *2013*. https://doi.org/10.1155/2013/158764

Pinyakong, O., Habe, H., & Omori, T. (2003). The unique aromatic catabolic genes in sphingomonads degrading polycyclic aromatic hydrocarbons (PAHs). *J. Gen. Appl. Microbiol.*, 1–19.

Posada-Baquero, R., Grifoll, M., & Ortega-Calvo, J. J. (2019). Rhamnolipid-enhanced solubilization and biodegradation of PAHs in soils after conventional bioremediation. *Science of the Total Environment*, *668*, 790–796. https://doi.org/10.1016/j.scitotenv.2019.03.056

Richardson, S. D., Lebron, B. L., Miller, C. T., & Aitken, M. D. (2011). Recovery of Phenanthrene-Degrading Bacteria after Simulated in Situ Persulfate Oxidation in Contaminated Soil. *Environmental Science & Technology*, *45*(2), 719–725. https://doi.org/10.1021/es102420r

Richts, B., Rosenberg, J., & Commichau, F. M. (2019). A Survey of Pyridoxal 5′-Phosphate-Dependent Proteins in the Gram-Positive Model Bacterium Bacillus subtilis. *Frontiers in Molecular Biosciences*, *6*(MAY), 32. https://doi.org/10.3389/FMOLB.2019.00032

Rodgers-Vieira, E. A., Zhang, Z., Adrion, A. C., Gold, A., & Aitken, M. D. (2015). Identification of anthraquinone-degrading bacteria in soil contaminated with polycyclic aromatic hydrocarbons. *Applied and Environmental Microbiology*, *81*(11), 3775–3781. https://doi.org/10.1128/AEM.00033-15

Rojo, F. (2021). A new global regulator that facilitates the co-metabolization of polyaromatic hydrocarbons and other nutrients in Novosphingobium. *Environmental Microbiology*, *23*(6), 2875–2877. https://doi.org/10.1111/1462-2920.15527

Roy, M. (2013). Catabolic Versatility of Sphingobium sp. Strain PNB Capable of Degrading Structurally Diverse Aromatic Compounds. *Journal of Bioremediation and Biodegradation*, *04*(01), 1–6. https://doi.org/10.4172/2155-6199.1000173

Sakshi, & Haritash, A. K. (2020). A comprehensive review of metabolic and genomic aspects of PAH-degradation. In *Archives of Microbiology* (Vol. 202, Issue 8, pp. 2033–2058). https://doi.org/10.1007/s00203-020-01929-5

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols*, *3*(6), 1101–1108. https://doi.org/10.1038/nprot.2008.73

Schuler, L., Jouanneau, Y., Ní Chadhain, S. M., Meyer, C., Pouli, M., Zylstra, G. J., Hols, P., & Agathos, S. N. (2009). Characterization of a ring-hydroxylating dioxygenase from phenanthrene-degrading Sphingomonas sp. strain LH128 able to oxidize benz[a]anthracene. *Applied Microbiology and Biotechnology*, *83*(3), 465–475. https://doi.org/10.1007/s00253-009-1858-2

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biology*, *12*(6), R60. https://doi.org/10.1186/gb-2011-12-6-r60

Seth, E. C., & Taga, M. E. (2014). Nutrient cross-feeding in the microbial world. *Frontiers in Microbiology*, *0*(JULY), 350. https://doi.org/10.3389/FMICB.2014.00350

Singh, A., & Lal, R. (2009). Sphingobium ummariense sp. nov., a hexachlorocyclohexane (HCH)-degrading bacterium, isolated from HCH-contaminated soil. *International Journal of Systematic and Evolutionary Microbiology*, *59*(1), 162–166. https://doi.org/10.1099/ijs.0.65712-0

Singleton, D. R., Powell, S. N., Sangaiah, R., Gold, A., Ball, L. M., & Aitken, M. D. (2005). Stable-isotope probing of bacteria capable of degrading salicylate, naphthalene, or phenanthrene in a bioreactor treating contaminated soil. *Applied and Environmental Microbiology*, *71*(3), 1202–1209. https://doi.org/10.1128/AEM.71.3.1202-1209.2005

Singleton, D. R., Sangaiah, R., Gold, A., Ball, L. M., & Aitken, M. D. (2006). Identification and quantification of uncultivated Proteobacteria associated with pyrene degradation in a bioreactor treating PAH-contaminated soil. *Environmental Microbiology*, *8*(10), 1736–1745. https://doi.org/10.1111/j.1462-2920.2006.01112.x

Singleton, D. R., Richardson, S. D., & Aitken, M. D. (2011). Pyrosequence analysis of bacterial communities in aerobic bioreactors treating polycyclic aromatic hydrocarbon-contaminated soil. *Biodegradation*, *22*(6), 1061–1073. https://doi.org/10.1007/s10532-011-9463-3

Singleton, D. R., Jones, M. D., Richardson, S. D., & Aitken, M. D. (2013). Pyrosequence analyses of bacterial communities during simulated in situ bioremediation of polycyclic aromatic hydrocarbon-contaminated soil. *Applied Microbiology and Biotechnology*, *97*(18), 8381–8391. https://doi.org/10.1007/s00253-012-4531-0

Singleton, D. R., Dickey, A. N., Scholl, E. H., Wright, F. A., & Aitken, M. D. (2016). Complete genome sequence of a bacterium representing a deep uncultivated lineage within the Gammaproteobacteria associated with the degradation of polycyclic aromatic hydrocarbons. *Genome Announcements*, *4*(5), 4–5. https://doi.org/10.1128/genomeA.01086-16

Sohn, J. H., Kwon, K. K., Kang, J. H., Jung, H. B., & Kim, S. J. (2004). Novosphingobium pentaromativorans sp. nov., a high-molecular-mass polycyclic aromatic hydrocarbon-degrading bacterium isolated from estuarine sediment. *International Journal of Systematic and Evolutionary Microbiology*, *54*(5), 1483–1487. https://doi.org/10.1099/ijs.0.02945-0

Srogi, K. (2007). Monitoring of environmental exposure to polycyclic aromatic hydrocarbons: a review. *Environmental Chemistry Letters 2007 5:4*, *5*(4), 169–195. https://doi.org/10.1007/S10311-007-0095-0

Stolz, A. (2009). Molecular characteristics of xenobiotic-degrading sphingomonads. In *Applied Microbiology and Biotechnology* (Vol. 81, Issue 5, pp. 793–811). https://doi.org/10.1007/s00253-008-1752-3

Stolz, A. (2014). Degradative plasmids from sphingomonads. *FEMS Microbiology Letters*, *350*(1), 9–19. https://doi.org/10.1111/1574-6968.12283

Stringfellow, W. T., & Aitken, M. D. (1995). Competitive metabolism of naphthalene, methylnaphthalenes, and fluorene by phenanthrene-degrading pseudomonads. *Applied and Environmental Microbiology*, *61*(1), 357–362. https://doi.org/10.1128/aem.61.1.357-362.1995

Tauler, M. (2015). Bacterial populations and functions driving the decontamination of polycyclic aromatic compounds polluted soils. *Doctoral Thesis.* Universitat de Barcelona.

Tauler, M., Vila, J., Nieto, J. M., & Grifoll, M. (2016). Key high molecular weight PAH-degrading bacteria in a soil consortium enriched using a sand-in-liquid microcosm system. *Applied Microbiology and Biotechnology*, *100*(7), 3321–3336. https://doi.org/10.1007/s00253-015-7195-8

Tian, Z., Gold, A., Nakamura, J., Zhang, Z., Vila, J., Singleton, D. R., Collins, L. B., & Aitken, M. D.

(2017a). Nontarget Analysis Reveals a Bacterial Metabolite of Pyrene Implicated in the Genotoxicity of Contaminated Soil after Bioremediation. *Environmental Science & Technology*, *51*(12), 7091–7100. https://doi.org/10.1021/acs.est.7b01172

Tian, Z., Vila, J., Wang, H., Bodnar, W., & Aitken, M. D. (2017b). Diversity and Abundance of High-Molecular-Weight Azaarenes in PAH-Contaminated Environmental Samples. *Environmental Science and Technology*, *51*(24), 14047–14054. https://doi.org/10.1021/ACS.EST.7B03319

Tian, Z., Vila, J., Yu, M., Bodnar, W., & Aitken, M. D. (2018). Tracing the Biotransformation of Polycyclic Aromatic Hydrocarbons in Contaminated Soil Using Stable Isotope-Assisted Metabolomics. *Environmental Science and Technology Letters*, *5*(2), 103–109. https://doi.org/10.1021/acs.estlett.7b00554

Titaley, I. A., Simonich, S. L. M., & Larsson, M. (2020). Recent Advances in the Study of the Remediation of Polycyclic Aromatic Compound (PAC)-Contaminated Soils: Transformation Products, Toxicity, and Bioavailability Analyses. In *Environmental Science and Technology Letters* (Vol. 7, Issue 12, pp. 873–882). American Chemical Society. https://doi.org/10.1021/acs.estlett.0c00677

Tolmie, C., Smit, M. S., & Opperman, D. J. (2019). Native roles of Baeyer-Villiger monooxygenases in the microbial metabolism of natural compounds. In *Natural Product Reports* (Vol. 36, Issue 2, pp. 326–353). Royal Society of Chemistry. https://doi.org/10.1039/c8np00054a

Top, E. M., & Springael, D. (2003). The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds. In *Current Opinion in Biotechnology* (Vol. 14, Issue 3, pp. 262–269). https://doi.org/10.1016/S0958-1669(03)00066-1

Torres Pazmiño, D. E., Dudek, H. M., & Fraaije, M. W. (2010). Baeyer-Villiger monooxygenases: recent advances and future challenges. In *Current Opinion in Chemical Biology* (Vol. 14, Issue 2, pp. 138–144). https://doi.org/10.1016/j.cbpa.2009.11.017

van Berkel, W. J. H., Kamerbeek, N. M., & Fraaije, M. W. (2006). Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *Journal of Biotechnology*, *124*(4), 670–689. https://doi.org/10.1016/j.jbiotec.2006.03.044

van der Oost, R., Beyer, J., & Vermeulen, N. P. E. (2003). Fish bioaccumulation and biomarkers in environmental risk assessment: a review. *Environmental Toxicology and Pharmacology*, *13*(2), 57–149. https://doi.org/10.1016/S1382-6689(02)00126-6

van Herwijnen, R., Wattiau, P., Bastiaens, L., Daal, L., Jonker, L., Springael, D., Govers, H. A. J., & Parsons, J. R. (2003). Elucidation of the metabolic pathway of fluorene and cometabolic pathways of phenanthrene, fluoranthene, anthracene and dibenzothiophene by Sphingomonas sp. LB126. *Research in Microbiology*, *154*, 199–206. https://doi.org/10.1016/S0923-2508(03)00039-1

Verma, H., Kumar, R., Oldach, P., Sangwan, N., Khurana, J. P., Gilbert, J. A., & Lal, R. (2014). Comparative genomic analysis of nine Sphingobium strains: Insights into their evolution and hexachlorocyclohexane (HCH) degradation pathways. *BMC Genomics*, *15*(1), 1–19. https://doi.org/10.1186/1471-2164-15-1014

Vila, J., López, Z., Sabaté, J., Minguillón, C., Solanas, A. M., & Grifoll, M. (2001). Identification of a Novel Metabolite in the Degradation of Pyrene by Mycobacterium sp. Strain AP1: Actions of the Isolate on Two- and Three-Ring Polycyclic Aromatic Hydrocarbons. *Applied and Environmental Microbiology*, *67*(12), 5497–5505. https://doi.org/10.1128/AEM.67.12.5497-5505.2001

Vila, J., & Grifoll, M. (2009). Actions of Mycobacterium sp. strain AP1 on the saturated- and aromatic-hydrocarbon fractions of fuel oil in a marine medium. *Applied and Environmental Microbiology*, *75*(19), 6232–6239. https://doi.org/10.1128/AEM.02726-08

Vila, J., Nieto, J. M., Mertens, J., Springael, D., & Grifoll, M. (2010). Microbial community structure of a heavy fuel oil-degrading marine consortium: Linking microbial dynamics with polycyclic aromatic hydrocarbon utilization. *FEMS Microbiology Ecology*, *73*(2), 349–362. https://doi.org/10.1111/j.1574-6941.2010.00902.x

Vila, J., Tauler, M., & Grifoll, M. (2015). Bacterial PAH degradation in marine and terrestrial habitats. *Current Opinion in Biotechnology*, *33*, 95–102. https://doi.org/10.1016/j.copbio.2015.01.006

Villaverde, J., Láiz, L., Lara-Moreno, A., González-Pimentel, J. L., & Morillo, E. (2019). Bioaugmentation of PAH-Contaminated Soils With Novel Specific Degrader Strains Isolated From a Contaminated Industrial Site. Effect of Hydroxypropyl-β-Cyclodextrin as PAH Bioavailability Enhancer. *Frontiers in Microbiology*, *0*, 2588. https://doi.org/10.3389/FMICB.2019.02588

Wang Z, Fingas M, Blenkinsopp S, Sergy G, Landriault M, Sigouin L, Foght J, Semple K, Westlake DW. (1998) Comparison of oil composition changes due to biodegradation and physical weathering in different oils. *J Chromatogr A.* 5;809(1-2):89-107. doi: 10.1016/s0021-9673(98)00166-6. PMID: 9677713.

Wattiau, P., Bastiaens, L., Van Herwijnen, R., Daal, L., Parsons, J. R., Renard, M.-E., Springael, D., & Cornelis, G. R. (2001). Fluorene degradation by Sphingomonas sp. LB126 proceeds through protocatechuic acid: a genetic analysis. In *Res. Microbiol* (Vol. 152).

Weisburg, W. G., Barns, S. M., Pelletier, D. A., & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology*, *173*(2), 697–703. https://doi.org/10.1128/jb.173.2.697-703.1991

Weissenfels, W. D., Beyer, M., Klein, J., & Rehm, H. J. (1991). Microbial metabolism of fluoranthene: isolation and identification of ring fission products. *Applied Microbiology and Biotechnology*, *34*(4), 528–535. https://doi.org/10.1007/BF00180583

Wilcke, W., Kiesewetter, M., & Musa Bandowe, B. A. (2014). Microbial formation and degradation of oxygen-containing polycyclic aromatic hydrocarbons (OPAHs) in soil during short-term incubation. *Environmental Pollution*, *184*, 385–390. https://doi.org/10.1016/j.envpol.2013.09.020

Wu, Y.-R., Luo, Z.-H., & Vrijmoed, L. L. P. (2010). *Biodegradation of anthracene and benz[a]anthracene by two Fusarium solani strains isolated from mangrove sediments.* https://doi.org/10.1016/j.biortech.2010.07.049

Yen, K. M., & Gunsalus, I. C. (1985). Regulation of naphthalene catabolic genes of plasmid NAH7. *Journal of Bacteriology*, *162*(3), 1008–1013. https://doi.org/10.1128/jb.162.3.1008-1013.1985

Yun, S. H., Choi, C.-W., Lee, S.-Y., Lee, Y. G., Kwon, J., Leem, S. H., Chung, Y. H., Kahng, H.-Y., Kim, S. J., Kwon, K. K., & Kim, S. Il. (2014). Proteomic Characterization of Plasmid pLA1 for Biodegradation of Polycyclic Aromatic Hydrocarbons in the Marine Bacterium, Novosphingobium pentaromativorans US6-1. *PLoS ONE*, *9*(3), e90812. https://doi.org/10.1371/journal.pone.0090812

Zengler, K., & Zaramela, L. S. (2018). The social network of microorganisms — how auxotrophies

shape complex communities. *Nature Reviews Microbiology 2018 16:6*, *16*(6), 383–390. https://doi.org/10.1038/S41579-018-0004-5

Zhang, C., Sun, C., Huang, H., Gui, C., Wang, L., Li, Q., & Ju, J. (2018). Biosynthetic Baeyer-Villiger Chemistry Enables Access to Two Anthracene Scaffolds from a Single Gene Cluster in Deep-Sea-Derived Streptomyces olivaceus SCSIO T05. *Journal of Natural Products*, *81*(7), 1570–1577. https://doi.org/10.1021/acs.jnatprod.8b00077

Zhang, Z., Sangaiah, R., Gold, A., & Ball, L. M. (2011). *Organic & Biomolecular Chemistry Synthesis of uniformly 13 C-labeled polycyclic aromatic hydrocarbons †*. https://doi.org/10.1039/c0ob01107j

Zhao, Q., Yue, S., Bilal, M., Hu, H., Wang, W., & Zhang, X. (2017). Comparative genomic analysis of 26 Sphingomonas and Sphingobium strains: Dissemination of bioremediation capabilities, biodegradation potential and horizontal gene transfer. *Science of the Total Environment*, *609*, 1238–1247. https://doi.org/10.1016/j.scitotenv.2017.07.249

Annex

Annex

Supplementary information of Chapter 1
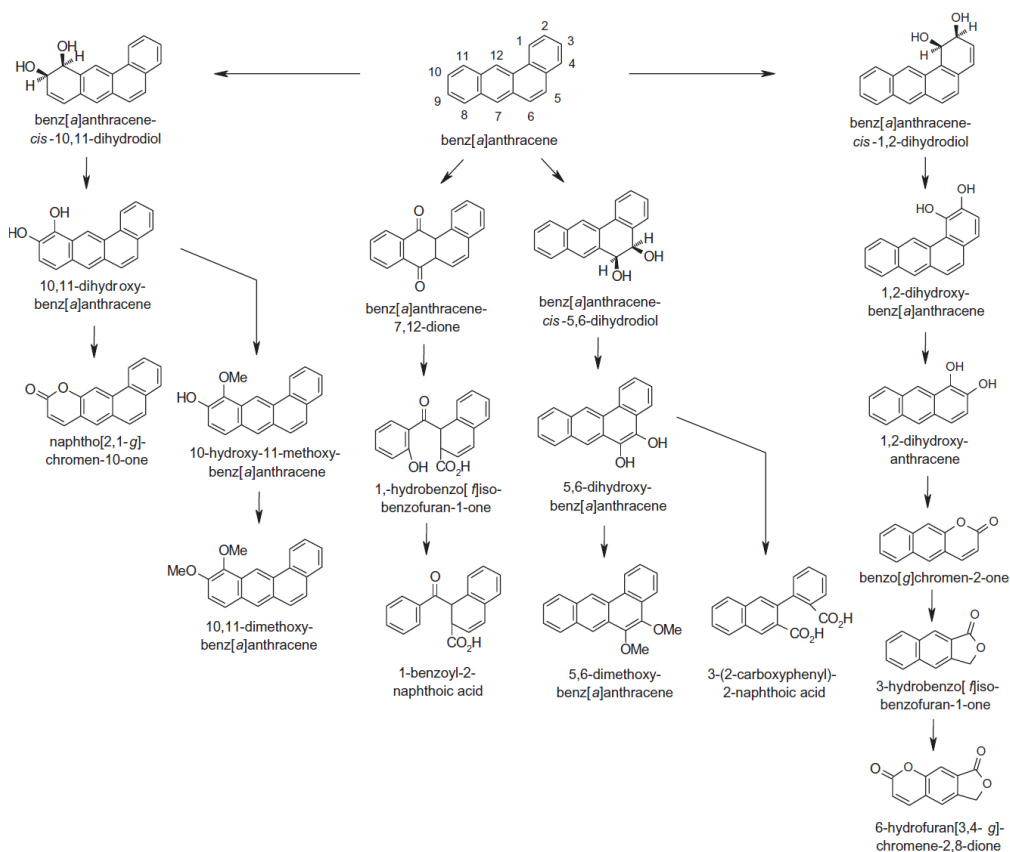
Figure S1.1. Pathways proposed for the biodegradation of benz[*a*]anthracene by mycobacteria (Kanaly & Harayama, 2010).

Table S1.1. Concentration of polycyclic aromatic compounds in the creosote-contaminated soil before and after biostimulation (Tauler, 2015).

Compound	Original soil			Treated soil		
PAHs						
Naphthalene	563.4	±	12.5	4.8	±	1.4
Acenaphthylene	119.3	±	6.8	1.5	±	0.5
Acenaphthene	2276.9	±	4.5	4.5	±	1.2
Fluorene	2744.3	±	94.7	16.1	±	4.2
Phenanthrene	8865.4	±	65.5	69.8	±	3.4
Anthracene	3747.5	±	396.5	215.0	±	27.7
Fluoranthene	4288.8	±	32.9	85.3	±	29.3
Pyrene	2294.2	±	11.3	65.8	±	21.5
Benzo[a]anthracene	694.6	±	67.1	48.4	±	11.7
Chrysene	705.4	±	62.8	61.0	±	14.3
Benzo[b]fluoranthene	219.9	±	25.2	48.2	±	12.5
Benzo[k]fluoranthene	207.0	±	20.2	36.0	±	10.4
Benzo[a]pyrene	158.1	±	16.6	34.1	±	8.7
Benzo[e]pyrene	144.7	±	17.5	25.0	±	7.0
Indeno[cd]pyrene	46.2	±	6.1	10.8	±	3.4
Dibenzo[ah]anthracene	13.1	±	2.0	3.1	±	1.3
Benzo[ghi]perylene	84.2	±	8.1	17.1	±	4.0
∑17 PAHs	27173.0	±	600.0	747	±	95
LMW	18316.7	±	574.2	312	±	3
HMW	8856.2	±	167.7	435	±	123
N-PACs						
Methylquinoline	33.5	±	2.6	0.44	±	0.23
Dimethylquinoline	44.4	±	8.1	1.87	±	0.62
Benzo[h]quinoline	32.8	±	1.5	2.78	±	0.19
Acridine	339.7	±	30.5	19.61	±	2.54
Methylacridine	107.3	±	11.8	9.06	±	2.21
Carbazole	1839.3	±	414.0	88.59	±	22.53
Methylcarbazole	169.2	±	7.9	5.01	±	0.43
∑7 N-PACS	2568.0	±	366.0	127	±	17
Oxy-PAHs						
9-Fluorenone	40.9	±	1.1	1.47	±	0.35
Anthracene-9,10-dione	122.2	±	1.3	3.38	±	2.38
4H-Cyclopenta[def]phenanthrenone	17.7	±	1.3	10.55	±	2.41
2-Methylanthracene-9,10-one	20.0	±	2.0	1.54	±	0.50
Benz[a]fluorenone	22.7	±	0.7	4.47	±	1.65
Benz[a]anthracene-7,12-dione	6.6	±	0.3	3.43	±	1.13
Naphthacene-5,12-dione	43.8	±	2.7	18.04	±	5.82
∑7 Oxy-PAHs	273.0	±	9.0	43	±	14
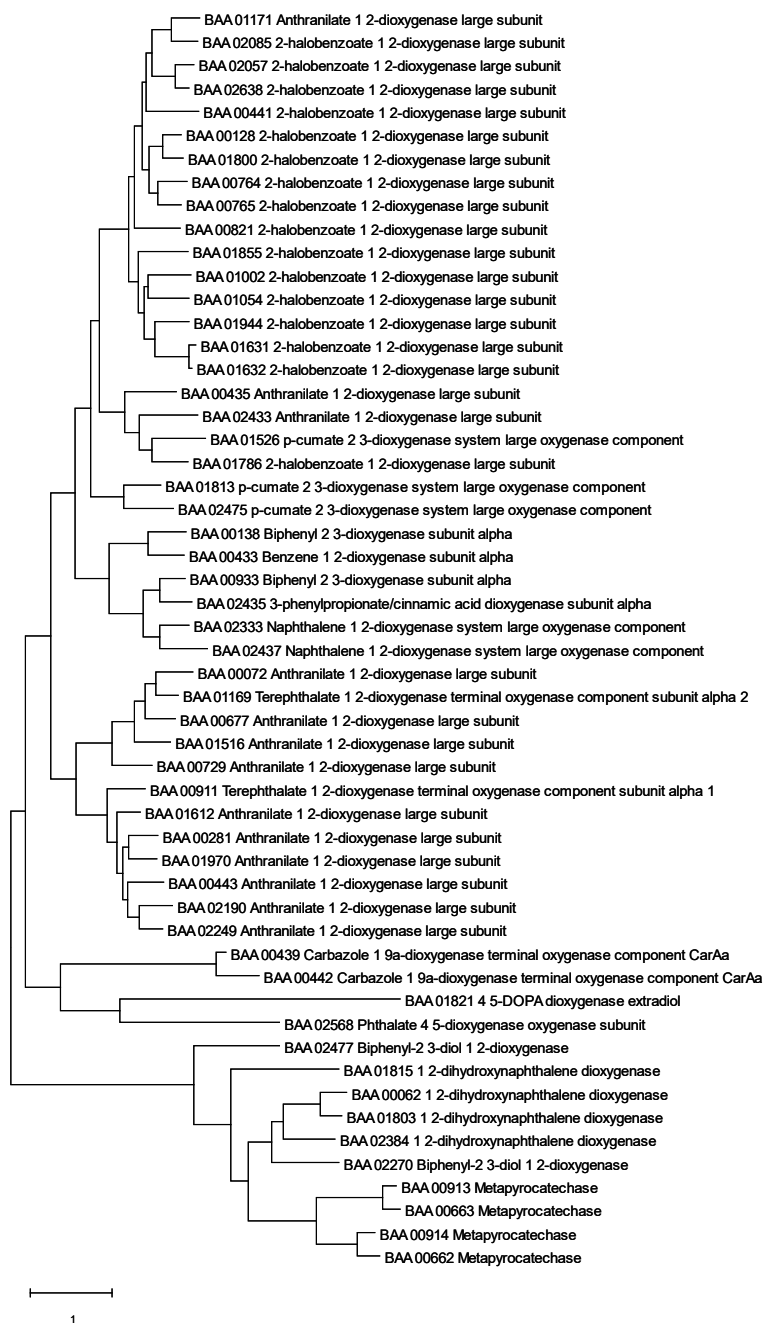
Figure S1.2. Phylogenetic tree of dioxygenases using the maximum-likelihood method of amino acid sequences from the *Immundisolibacter* sp. BA1 MAG. The tree was constructed in MEGA X.

Supplementary information of Chapter 2
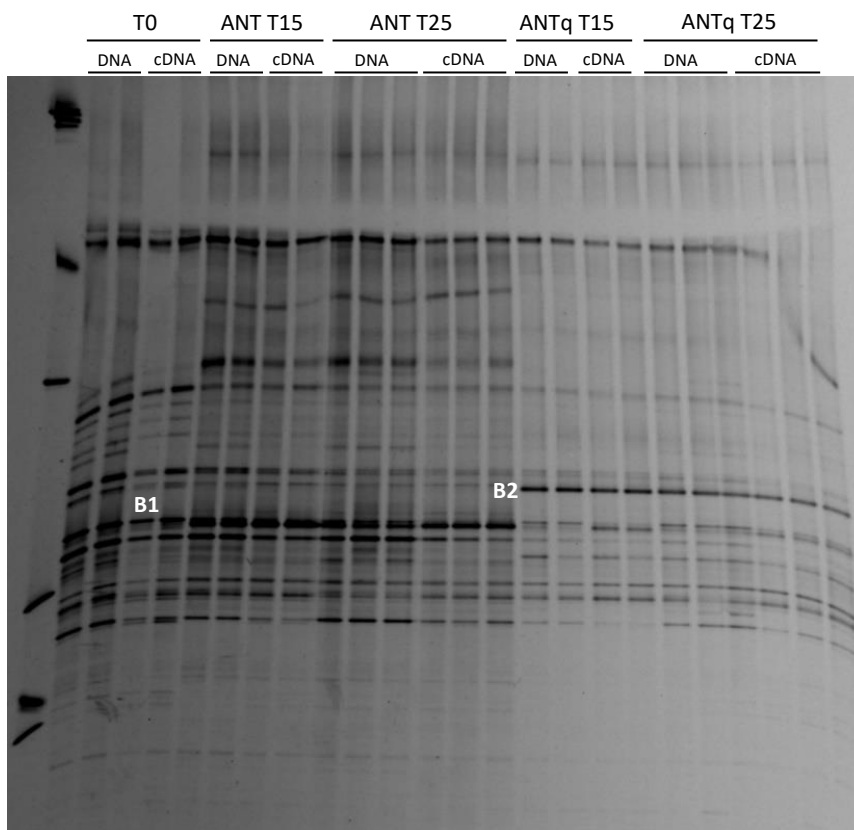
Figure S2.1. DGGE profiles of the 16S rRNA gene fragments from DNA and cDNA samples from the ANT and ANTQ-spiked sand-in-liquid soil microcosms at 0, 15 and 25 days of incubation.

Table S2.1. Phylogenetic affiliation of the most relevant excised bands from the DGGE profiles of the ANT and ANTQ-spiked sand-in-liquid soil microcosms based on RDP database.

Band	Length (bp)	RDP Classification	Prob (%)	Blast closest match (Accession number)	Ident (%)
B1	321	*Sphingobium*	99	*Sphingobium fontiphilum* strain Chen16-4 (NR_109304)	99.7
B2	318	*Sphingobium*	100	*Sphingomonas bisphenolicum* strain AO1 (NR_153665)	100

Table S2.2. ^{1}H and ^{13}C NMR chemical shifts (δ) and coupling constants (J) for metabolite I in CDCl$_3$ (500 MHz).
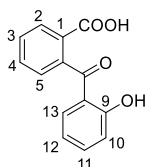
Assignment	δ (multiplicity; J)*
1-COOH	11.89
H-2	8.16 (dd; 7.8, 1.3 Hz)
H-3	7.61 (ddd; 7.8, 7.6, 1.3 Hz)
H-4	7.70 (ddd; 7.6, 7.6, 1.3 Hz)
H-5	7.38 (dd; 7.6, 1.3 Hz)
9-OH	2.01
H-10	7.05 (dd; 8.4, 1.1 Hz)
H-11	7.46 (ddd; 8.4, 7.2, 1.7 Hz)
H-12	6.76 (ddd; 8.1, 7.2, 1.1 Hz)
H-13	7.08 (dd, 8.1, 1.7 Hz)
C-2	131.1
C-3	129.8
C-4	133.2
C-5	127.4
C-10	118.2
C-11	136.4
C-12	118.9
C-13	132.6

*Chemical shifts (δ) are given in ppm. Multiplicity is indicated as dd (double doublet) or ddd (double double doublet) considering the appearance of signals in the spectrum. Coupling constants (J) are given in Hz.
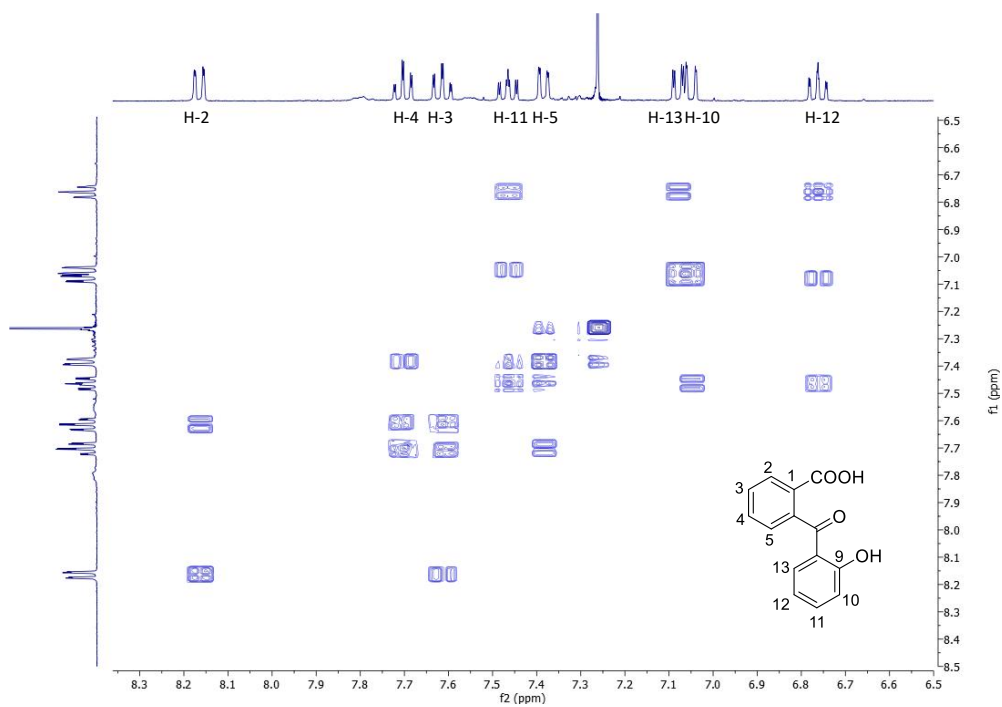
Figure S2.2. Aromatic region of the gCOSY spectrum of metabolite I.
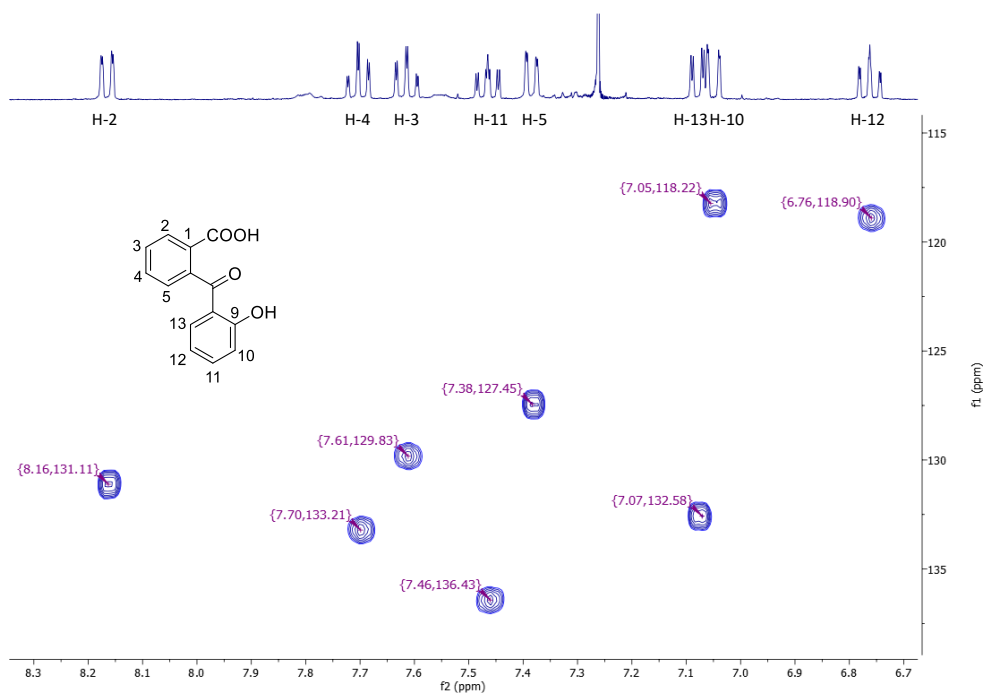
Figure S2.3. Aromatic region of the gHSQC spectrum of metabolite I.

Table S2.3. GC-MS Rt and EI mass spectral properties and HPLC-MS Rt and detected mass accuracies of major compounds formed from anthrone by *Sphingobium* sp. strain AntQ-1 in neutral (N) and acidic (A) extracts.

	Extract	GC-MS				HPLC-ESI(+)-HRMS			
		Rt (min)	Abundance (%)	*m/z* of fragment ions (% relative intensity)	Identification	Rt (min)	[M+H]+	Molecular formula	δ (mmu)
IV	N A	26.57	7.5 30.4	208 (M+, 100), 180 (100), 165 (4), 152 (81), 126 (10), 111 (2), 98 (4), 90 (4), 76 (33), 63 (8), 50 (14)	9,10-anthraquinone	18.5	209.0602	$C_{14}H_9O_2$	0.2134
I	A	26.74	22.3	256 (M+, 25), 223 (100), 196 (79), 168 (44), 163 (10), 152 (6), 139 (25), 121 (33), 115 (10), 104 (6), 92 (14), 77 (25), 65 (27), 50 (8)	2-(2-hydroxybenzoyl)-benzoic acid ME	15.4	243.0643	$C_{14}H_{11}O_4$	-0.8725
III	A	18.07	15.7	194 (M+, 6), 163 (100), 149 (1), 133 (6), 120 (2), 104 (5), 92 (8), 77 (20), 64 (2), 50 (6)	phtalatic acid diME	2.15	149.0234	$C_8H_5O_3$	0.1054
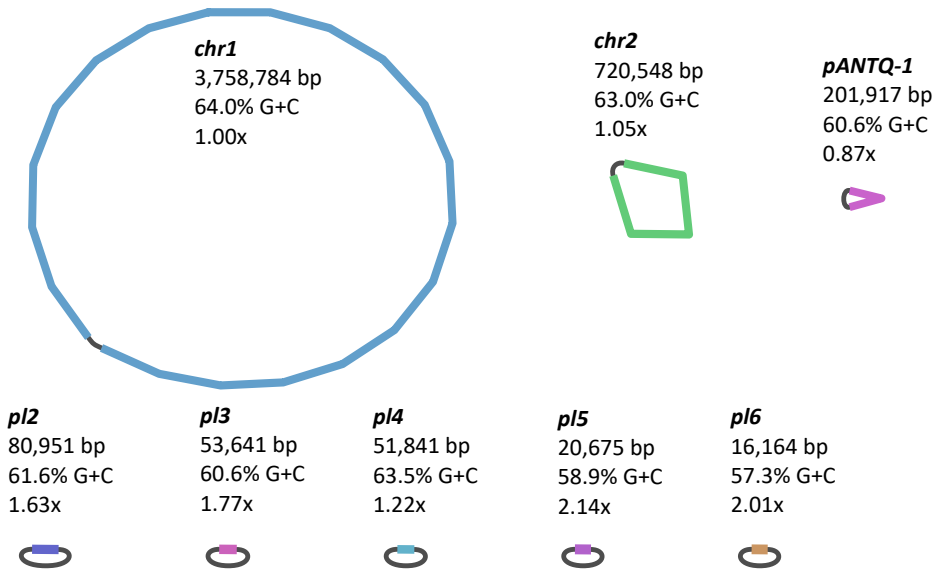
Figure S2.4. Assembly graph of the *Sphingobium* sp. AntQ-1 genome.
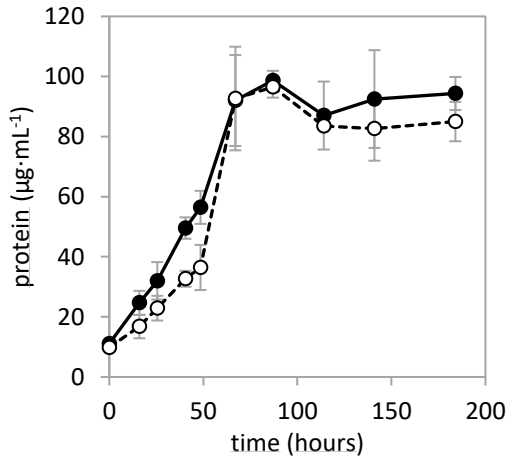
Figure S2.5. Growth of *Sphingobium* sp. strain AntQ-1 in liquid mineral medium supplemented with vitamin B12 either with acetate (○) or 9,10-anthraquinone (•) as the sole carbon and energy source.

Table S2.4. Differentially expressed genes during growth of *Sphingobium* sp. AntQ-1 on acetate, considering p-value <0.01 and log2 fold change ≤-1.5.

ID	Annotation	Length (bp)	log2 fold change	Adjusted p-value	TPM
1235	Isocitrate lyase	1596	-5.08	5.88E-222	509.86
4069	Diaminobutyrate--2-oxoglutarate transaminase	1311	-2.95	2.19E-67	118.91
2547	Flagellin	801	-2.83	2.20E-33	1762.69
1436	hypothetical protein	177	-2.67	2.51E-30	272.27
4070	L-2,4-diaminobutyric acid acetyltransferase	339	-2.67	3.08E-36	126.35
4665	putative metallo-hydrolase	918	-2.62	2.01E-26	2856.84
4662	hypothetical protein	867	-2.49	2.09E-25	116.50
4668	hypothetical protein	504	-2.46	1.26E-26	641.93
515	hypothetical protein	420	-2.44	2.38E-20	171.16
514	hypothetical protein	357	-2.43	7.22E-22	1138.31
4667	hypothetical protein	408	-2.38	4.05E-31	333.24
4663	hypothetical protein	426	-2.33	2.44E-15	2013.60
4666	Peroxiredoxin	618	-2.27	1.51E-21	459.32
4758	hypothetical protein	867	-2.23	1.99E-35	82.43
4084	ECF RNA polymerase sigma factor SigF	537	-2.18	1.46E-21	88.88
4390	putative oxidoreductase	762	-2.15	1.91E-31	678.28
510	Sulfurtransferase	342	-2.12	1.59E-28	2063.52
497	Peroxiredoxin	663	-2.10	2.91E-41	666.30
4795	hypothetical protein	867	-2.07	9.43E-36	65.28
2537	Chemotaxis protein PomA	669	-2.06	9.08E-23	68.05
498	hypothetical protein	372	-2.00	3.63E-13	285.07
507	Biofilm growth-associated repressor	321	-2.00	5.39E-16	471.62
4391	Acetyl-CoA acetyltransferase	1185	-1.97	1.52E-26	691.00
4067	Ectoine dioxygenase	921	-1.94	4.39E-25	43.48

4728	hypothetical protein	867	-1.94	1.51E-21	91.06
2123	hypothetical protein	1401	-1.92	8.98E-09	207.47
4068	L-ectoine synthase	414	-1.86	2.09E-25	116.84
4664	hypothetical protein	438	-1.84	2.73E-06	864.04
4392	Acyl-CoA dehydrogenase	1125	-1.80	3.38E-16	666.15
4083	hypothetical protein	270	-1.72	1.54E-13	55.30
4563	hypothetical protein	654	-1.72	9.13E-08	155.51
508	Multidrug resistance protein MdtC	3219	-1.70	3.49E-10	495.25
400	hypothetical protein	672	-1.70	5.68E-26	655.58
2525	Motility protein A	864	-1.65	1.95E-14	134.74
2526	Flagellar hook-associated protein 3	846	-1.65	1.24E-17	253.74
271	hypothetical protein	3216	-1.64	4.39E-03	128.99
4562	Thiol-disulfide oxidoreductase ResA	783	-1.63	2.16E-06	63.28
512	hypothetical protein	204	-1.61	3.02E-13	1135.94
4099	hypothetical protein	210	-1.59	1.96E-15	1141.42
1643	Diguanylate cyclase VdcA	1167	-1.59	8.07E-13	82.96
2528	hypothetical protein	330	-1.56	3.70E-11	154.37
4082	hypothetical protein	456	-1.56	7.42E-07	26.57
4814	hypothetical protein	354	-1.54	1.16E-05	43.24
4564	Trifunctional nucleotide phosphoesterase protein YfkN	1443	-1.52	9.82E-09	83.61
4100	hypothetical protein	141	-1.50	1.30E-06	1161.12

Table S2.5. Differentially expressed genes during growth of *Sphingobium* sp. AntQ-1 on 9,10-anthraquinone, considering p-value <0.01 and log2 fold change ≤-1.5.

ID	Annotation	Length (bp)	log2 fold change	Adjusted p-value	TPM
420	hypothetical protein	141	1.50	1.37E-08	282.34
4469	2-oxopent-4-enoate hydratase	807	1.50	8.60E-23	2829.91
3856	Glycerol dehydrogenase large subunit	2340	1.51	3.17E-11	9.90
468	Arsenate reductase	528	1.52	5.19E-16	57.54
25	Outer membrane protein OprM	1437	1.52	5.55E-19	29.11
637	hypothetical protein	168	1.52	6.74E-12	102.37
33	hypothetical protein	1029	1.53	1.23E-04	4.97
3316	Hexuronate transporter	1329	1.53	6.11E-10	13.55
4471	4-oxalocrotonate decarboxylase	771	1.53	6.11E-26	1384.82
2957	hypothetical protein	198	1.54	8.54E-05	29.37
431	hypothetical protein	633	1.54	1.28E-03	2.80
20	hypothetical protein	1224	1.54	2.62E-06	3.59
644	Polyamine aminopropyltransferase	942	1.54	2.74E-12	71.27
3892	(-)-trans-carveol dehydrogenase	768	1.54	1.52E-09	19.47
1846	hypothetical protein	318	1.54	5.06E-20	955.21
1611	hypothetical protein	174	1.54	4.34E-14	369.04
3	hypothetical protein	252	1.54	1.60E-02	1.98
4464	hypothetical protein	795	1.55	5.36E-31	1157.27
4594	hypothetical protein	636	1.55	1.14E-04	4.29
74	hypothetical protein	1251	1.56	5.97E-06	3.46
3769	Antibiotic efflux pump outer membrane protein ArpC	1413	1.56	1.37E-08	6.50
926	hypothetical protein	351	1.56	1.08E-15	129.24
4597	hypothetical protein	4218	1.56	1.96E-06	2.83
4424	IS5 family transposase ISMex35	351	1.57	5.96E-22	166.04

4	hypothetical protein	519	1.57	4.78E-04	4.38
4448	Tyrosine recombinase XerD	1017	1.57	2.92E-10	9.16
4472	2-hydroxychromene-2-carboxylate isomerase	603	1.57	4.53E-14	1781.99
438	hypothetical protein	1125	1.58	1.59E-04	2.68
3937	2-pyrone-4,6-dicarboxylate hydrolase	873	1.58	1.16E-20	62.24
4017	Putative phosphoserine phosphatase 2	570	1.58	1.78E-03	2.78
719	Glycogen synthase	1062	1.58	2.28E-12	17.54
4514	Nicotinate-nucleotide pyrophosphorylase [carboxylating]	867	1.58	7.78E-29	159.53
4509	Phthalate dioxygenase reductase	963	1.59	2.27E-37	3965.44
3817	HTH-type transcriptional activator RhaR	864	1.59	7.66E-10	14.29
4510	hypothetical protein	543	1.59	4.55E-31	7916.29
1773	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	768	1.59	5.22E-16	24.55
1694	putative oxidoreductase	786	1.60	4.18E-17	591.73
4521	IS21 family transposase ISPpu7	1527	1.60	1.17E-07	4.56
3024	Fatty acid oxidation complex subunit alpha	2196	1.60	8.67E-07	839.39
4461	2-halobenzoate 1,2-dioxygenase large subunit	1266	1.60	7.71E-15	517.79
2329	hypothetical protein	312	1.61	7.57E-09	12328.34
4443	hypothetical protein	417	1.61	1.40E-04	5.20
393	hypothetical protein	207	1.61	5.46E-04	8.41
3709	hypothetical protein	366	1.62	5.28E-10	764.64
4541	Type IV secretion system protein VirB11	1005	1.62	6.31E-04	2.84
841	hypothetical protein	264	1.62	1.99E-20	419.24
3839	Putative acyltransferase	1203	1.62	1.43E-14	35.45
2842	HTH-type transcriptional repressor BdcR	594	1.63	1.90E-13	67.16
4538	Type IV secretion system protein virB8	690	1.63	6.42E-05	4.96
4534	Type IV secretion system protein virB4	2418	1.63	1.60E-05	3.78
26	Fatty acid resistance protein FarA	1164	1.64	1.14E-13	25.78
1679	hypothetical protein	297	1.64	2.71E-12	54.23
3163	hypothetical protein	198	1.64	5.30E-03	4.26

4459	hypothetical protein	660	1.64	1.61E-28	231.01
3620	Multidrug efflux pump subunit AcrB	3174	1.65	5.14E-17	147.21
4502	Pesticin receptor	2013	1.65	7.95E-22	18.05
4446	Tyrosine recombinase XerC	1239	1.65	4.55E-08	5.48
4581	hypothetical protein	747	1.65	9.59E-27	119.67
2322	hypothetical protein	570	1.66	2.09E-25	134.13
1708	hypothetical protein	174	1.66	1.09E-18	264.20
4511	hypothetical protein	705	1.67	2.80E-15	11885.33
3150	hypothetical protein	405	1.68	8.29E-05	11.23
3833	Alcohol dehydrogenase	1002	1.68	6.93E-22	33.37
4444	hypothetical protein	354	1.68	2.90E-04	5.13
781	hypothetical protein	693	1.68	3.02E-18	56.25
4485	Alcohol dehydrogenase	600	1.68	1.10E-10	14.15
27	Multidrug export protein EmrB	1548	1.69	1.11E-08	23.82
883	hypothetical protein	945	1.69	1.12E-25	213.67
718	Catalase C	2187	1.69	9.27E-17	51.98
1742	hypothetical protein	609	1.69	4.08E-05	5.13
4636	hypothetical protein	192	1.70	1.07E-05	25.43
4449	Narbonolide/10-deoxymethynolide synthase PikA2, modules 3 and 4	906	1.71	5.48E-16	18.87
4287	Multifunctional non-homologous end joining protein LigD	2502	1.71	1.04E-08	4.33
4590	Tyrosine recombinase XerC	1143	1.71	2.15E-17	20.06
4599	Single-stranded DNA-binding protein	321	1.72	4.48E-10	24.87
1677	hypothetical protein	249	1.72	1.87E-12	302.10
4508	hypothetical protein	291	1.73	3.09E-15	312.37
5	hypothetical protein	546	1.74	3.13E-04	4.00
2	Chromosome-partitioning ATPase Soj	639	1.75	8.17E-05	3.72
3155	hypothetical protein	162	1.76	5.90E-03	2.81
3768	Toluene efflux pump periplasmic linker protein TtgA	1194	1.76	5.83E-08	8.41
4520	hypothetical protein	291	1.76	1.26E-03	3.90

4569	hypothetical protein	453	1.78	2.89E-05	7.90
161	hypothetical protein	798	1.78	2.52E-08	27.23
3929	4-carboxy-2-hydroxymuconate-6-semialdehyde dehydrogenase	942	1.80	1.00E-32	109.17
4512	hypothetical protein	993	1.80	4.71E-34	2947.29
4595	DNA primase TraC	924	1.80	1.10E-06	3.74
3936	(4E)-oxalomesaconate Delta-isomerase	1062	1.81	5.33E-17	28.66
4570	hypothetical protein	1131	1.81	1.62E-10	9.93
4542	hypothetical protein	2013	1.82	9.96E-12	3.90
3126	hypothetical protein	549	1.82	2.35E-16	2850.08
1962	p-hydroxybenzoic acid efflux pump subunit AaeB	2067	1.83	1.72E-12	33.56
3861	Ferredoxin-2	207	1.83	1.44E-20	224.62
3047	hypothetical protein	294	1.85	1.31E-20	175.25
375	hypothetical protein	522	1.86	1.75E-05	5.89
4487	Vitamin B12 transporter BtuB	2136	1.87	2.48E-17	3413.24
4486	IS5 family transposase ISMex35	1077	1.90	1.72E-14	99.42
3336	Beta-ketoadipyl-CoA thiolase	1206	1.90	1.29E-12	34.77
2359	hypothetical protein	858	1.90	2.86E-30	3369.03
4513	hypothetical protein	174	1.90	2.40E-17	1424.59
690	hypothetical protein	168	1.91	9.59E-17	79.18
3767	Cobalt-zinc-cadmium resistance protein CzcA	3198	1.91	3.01E-17	14.69
12	hypothetical protein	333	1.91	1.48E-03	2.88
4572	Nucleoid occlusion protein	936	1.93	8.52E-05	2.70
14	Type IV secretion system protein virB4	2457	1.93	3.65E-09	3.24
3933	hypothetical protein	888	1.93	7.40E-22	26.35
3843	hypothetical protein	852	1.93	3.36E-16	192.75
1864	hypothetical protein	609	1.94	4.43E-19	23.22
4505	hypothetical protein	234	1.96	1.03E-05	14.99
10	hypothetical protein	435	1.97	1.52E-05	4.81
3935	4-carboxy-4-hydroxy-2-oxoadipate aldolase	675	1.97	1.55E-21	73.43

656	hypothetical protein	300	1.98	2.66E-10	111.87
4587	hypothetical protein	459	1.98	2.67E-10	13.28
4740	Nodulation protein D 2	579	1.98	1.26E-10	42.51
4574	hypothetical protein	180	1.99	9.94E-04	4.53
4588	hypothetical protein	1809	2.01	1.02E-34	22.46
4476	HTH-type transcriptional regulator CymR	648	2.01	4.92E-39	311.67
1963	hypothetical protein	216	2.02	1.83E-09	26.91
4779	FMNH(2)-dependent dimethylsulfone monooxygenase	1143	2.03	1.80E-23	32.46
4539	Type IV secretion system protein virB9	858	2.03	1.28E-06	3.21
1750	hypothetical protein	636	2.04	4.31E-11	9.90
4420	hypothetical protein (transposase)	1014	2.04	5.03E-14	9.34
3164	hypothetical protein	555	2.05	1.60E-05	5.06
16	hypothetical protein	390	2.05	4.85E-05	6.31
3337	Putative aminoacrylate hydrolase RutD	702	2.06	2.76E-13	62.91
4575	hypothetical protein	396	2.06	8.66E-05	4.22
4584	hypothetical protein	948	2.06	3.03E-09	5.09
4573	Iron-sulfur cluster carrier protein	765	2.07	2.34E-06	3.14
3931	Protocatechuate 4,5-dioxygenase alpha chain	456	2.08	8.19E-24	117.50
3332	putative succinyl-CoA:3-ketoacid coenzyme A transferase subunit A	708	2.08	9.17E-18	137.76
3327	Muconolactone Delta-isomerase	291	2.08	1.64E-13	118.23
4488	hypothetical protein	330	2.09	8.81E-24	8797.83
3335	3-oxoadipate enol-lactonase 2	771	2.10	1.90E-13	35.55
3334	Chloroacetanilide N-alkylformylase 2, ferredoxin component	318	2.11	2.98E-16	76.11
3330	2-halobenzoate 1,2-dioxygenase small subunit	516	2.13	8.92E-15	138.19
2390	hypothetical protein	396	2.16	1.61E-09	119.99
372	hypothetical protein	648	2.17	3.02E-05	3.64
657	hypothetical protein	369	2.19	1.89E-19	174.21
3871	hypothetical protein	735	2.19	4.67E-24	49.93
4540	hypothetical protein	1218	2.19	1.26E-13	6.46

3152	hypothetical protein	177	2.20	2.33E-20	19526.78
1	hypothetical protein	888	2.20	1.27E-09	6.82
3328	Catechol 1,2-dioxygenase 2	927	2.21	8.72E-19	246.66
3329	2-halobenzoate 1,2-dioxygenase large subunit	1365	2.25	4.79E-14	143.47
4418	IS5 family transposase ISAzs28	822	2.25	1.11E-21	45.89
4589	hypothetical protein	1818	2.26	8.39E-35	14.31
3331	3-beta-hydroxycholanate 3-dehydrogenase (NADP(+))	786	2.26	3.01E-19	129.87
1701	Multifunctional non-homologous end joining protein LigD	582	2.27	2.34E-13	15.08
3930	Protocatechuate 4,5-dioxygenase beta chain	897	2.28	6.26E-34	104.86
4489	hypothetical protein	423	2.29	1.45E-26	1296.36
4491	Putative 4,5-dihydroxyphthalate dehydrogenase	1197	2.30	8.70E-54	276.73
4523	ISNCY family transposase ISRel10	447	2.32	1.67E-16	27.07
3333	putative succinyl-CoA:3-ketoacid coenzyme A transferase subunit B	648	2.35	2.23E-18	128.69
4492	Phthalate 4,5-dioxygenase oxygenase subunit	1272	2.36	2.07E-61	4782.39
4586	hypothetical protein	600	2.39	3.94E-07	7.34
4453	IS5 family transposase ISAzs28	822	2.39	1.71E-17	37.01
4442	ISL3 family transposase ISMex26	1563	2.40	8.01E-12	4.90
4431	hypothetical protein	513	2.41	2.13E-42	154.14
4320	hypothetical protein	219	2.42	1.25E-19	398.68
3326	Muconate cycloisomerase 1	1158	2.43	4.70E-24	133.82
4490	FMN-dependent NADH-azoreductase	645	2.43	1.27E-69	1612.12
8	hypothetical protein	1740	2.43	9.76E-09	3.52
4458	2,3-dihydroxyphenylpropionate/2,3-dihydroxicinnamic acid 1,2-dioxygenase	1089	2.49	3.47E-77	245.18
3932	2-keto-4-carboxy-3-hexenedioate hydratase	1026	2.51	1.86E-18	97.00
469	hypothetical protein	333	2.55	3.08E-36	156.04
4593	hypothetical protein	612	2.70	1.10E-06	2.44
777	hypothetical protein	693	2.85	2.07E-19	103.67
645	hypothetical protein	201	2.96	7.37E-26	1350.77
17	hypothetical protein	1365	2.97	3.54E-14	9.01

4475	Hexuronate transporter	1338	3.08	4.33E-51	1134.67
4780	Methionine aminopeptidase 1, mitochondrial	531	3.10	6.61E-29	72.15
4524	hypothetical protein (transposase)	282	3.19	1.05E-38	244.00
4741	FMNH(2)-dependent dimethylsulfone monooxygenase	1173	3.47	2.84E-46	87.91
4474	hypothetical protein (alpha/beta hydrolase_6)	1125	3.49	3.05E-47	2570.66
778	ECF RNA polymerase sigma factor SigR	489	3.51	1.98E-25	89.57
4743	Riboflavin biosynthesis protein RibBA	1104	3.58	5.68E-26	39.12
4473	Limonene 1,2-monooxygenase	1041	4.05	2.71E-157	21461.85
4742	Pesticin receptor	2409	4.40	7.70E-97	280.62
4432	4-hydroxy-2-oxovalerate aldolase	1047	4.49	2.88E-175	359.67
4433	Acetaldehyde dehydrogenase	945	4.57	2.77E-139	532.73
780	hypothetical protein	702	5.87	3.12E-164	4929.61
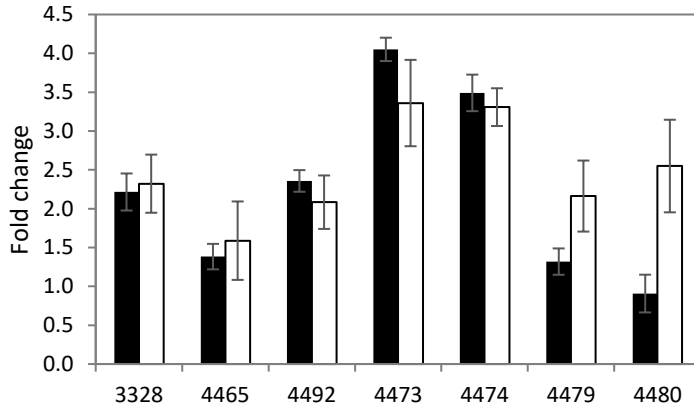
Figure S2.5. Fold change of selected DEGs by RNA-Seq, black bars, and RT-qPCR analysis, white bars. Values are means of three replicates and error bars represent standard deviation.

Supplementary information of Chapter 3
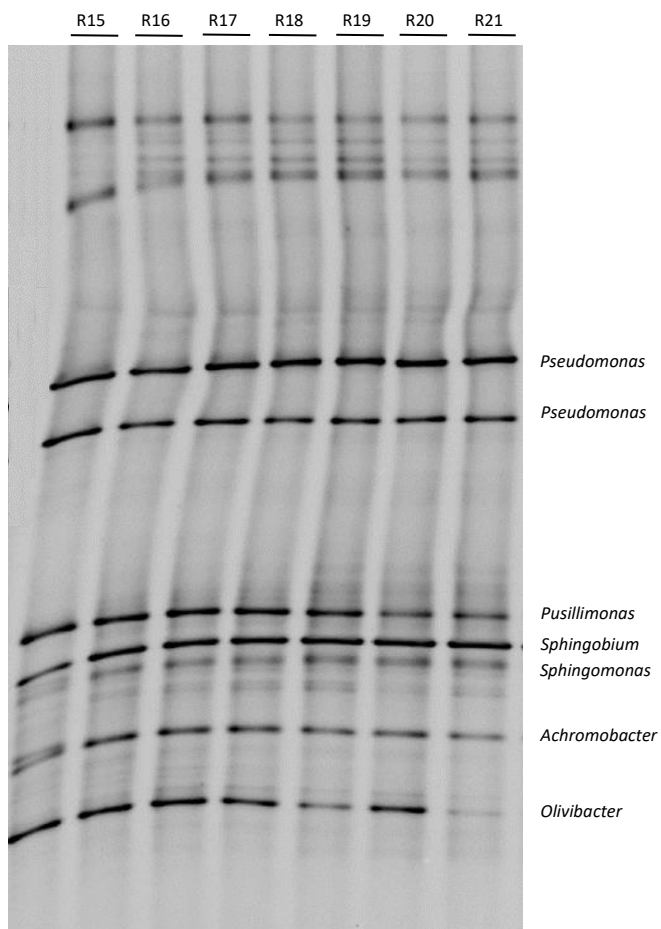
Figure S3.1. DGGE profiles of the 16S rRNA gene fragments from DNA samples of the B*a*AQ-degrading consortium BQ monthly transfers. Taxonomic identification of the bands was based on RDP database (probability 100%).

Table S3.1. List of flavin-dependent monooxygenases found in the metagenome-assembled genomes from the bacterial consortium BQ.

Class A

2,4-dichlorophenol 6-monooxygenase
2,6-dihydroxypyridine 3-monooxygenase
2-octaprenyl-6-methoxyphenol hydroxylase
3-hydroxybenzoate 4-monooxygenase
3-hydroxybenzoate 6-hydroxylase
4-hydroxybenzoate 3-monooxygenase (NAD(P)H)
4-hydroxyphenylacetate 3-monooxygenase
4-nitrophenol 4-monooxygenase/4-nitrocatechol 2-monooxygenase
5-methylphenazine-1-carboxylate 1-monooxygenase
6-hydroxy-3-succinoylpyridine 3-monooxygenase
6-hydroxynicotinate 3-monooxygenase
6-methylpretetramide 4-monooxygenase
Anhydrotetracycline monooxygenase
Anthranilate 3-monooxygenase
Benzoyl-CoA oxygenase
Kynurenine 3-monooxygenase
Para-nitrophenol 4-monooxygenase
Pentachlorophenol 4-monooxygenase
Phenol hydroxylase P5 protein
p-hydroxybenzoate hydroxylase
p-hydroxyphenylacetate 3-hydroxylase

Class B

4-hydroxyacetophenone monooxygenase
Baeyer-Villiger monooxygenase
Cyclohexanone 1,2-monooxygenase
FAD-containing monooxygenase EthA
L-ornithine N(5)-monooxygenase
Pentalenolactone D synthase
Phenylacetone monooxygenase

Class C

2,5-diketocamphane 1,2-monooxygenase
3,6-diketocamphane 1,6 monooxygenase
Alkanal monooxygenase
Alkanesulfonate monooxygenase
Flavin-dependent monooxygenase
Limonene 1,2-monooxygenase
Methanesulfonate monooxygenase
Nitrilotriacetate monooxygenase
Nitronate monooxygenase
Putative monooxygenase MoxC
Pyrimidine monooxygenase RutA
Tryptophan 2-monooxygenase

Table S3.2. Aromatic hydrocarbon degradation genes enclosed in the MAGs from consortium BQ.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		Sphingobium BQ7	*Bosea* BQ11	*Rhizobiaceae*	*Rhizobiaceae*	*Stenotrophomonas* BQ6	*Pusillimonas*	*Bradyrhizobium* BQ13	*Stenotrophomonas* BQ2	*Sphingomonadaceae* BQ10	*Pusillimonas* BQ3	*Achromobacter* BQ8	*Sphingobium* BQ1	*Pusillimonas* BQ9	*Pseudomonas* BQ5	*Olivibacter* BQ4
nahG	salicilate 1-hydroxylase		x	x	x			x			x	x		x		
nahC	1,2-dihydroxynaphthalene dioxygenase										x	x	x	x		
nahD	2-hydroxychromene-2-carboxylate isomerase					x		x					x	x		
nahE	trans-o-hydroxybenzylidenepyruvate hydratase-aldolase							x								
nahF	salicylaldehyde dehydrogenase					x							x	x		
nagGH	salicylate 5-hydroxylase						x				x	x		x		
nagI	gentisate 1,2-dioxygenase	x					x	x			x		x	x		
nagL	maleylpyruvate isomerase	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
nagK	fumarylpyruvate hydrolase		x	x	x	x	x	x	x		x	x	x	x	x	x
phdI	1-hydroxy-2-naphthoate dioxygenase										x	x		x		
phdJ	4-(2-carboxyphenyl)-2-oxobut-3-enoate aldolase										x					
phdK	2-formylbenzoate dehydrogenase		x								x	x	x	x		
pht23	phthalate 4,5-dioxygenase		x					x			x	x		x		
pht4	phthalate 4,5-*cis*-dihydrodiol dehydrogenase		x								x	x		x		
pht5	4,5-dihydroxyphthalate decarboxylase		x					x			x			x		
ligAB	protocatechuate 4,5-dioxygenase alpha chain	x									x			x		
ligC	4-carboxy-2-hydroxymuconate-6-semialdehyde dehydrogenase	x									x			x		
ligI	2-pyrone-4,6-dicarboxylate hydrolase	x				x	x	x					x	x		
galD	4-oxalomesaconate tautomerase	x	x			x					x					

Gene	Enzyme	1	2	3	4	5	6	7	8	9	10	11	12	13
ligJ	4-oxalomesaconate hydratase	x	x			x		x		x		x		
ligK	4-carboxy-4-hydroxy-2-oxoadipate aldolase	x	x		x	x	x	x		x		x		
chqB	hydroxyquinol 1,2-dioxygenase		x			x		x		x	x	x		
macA	maleylacetate reductase		x			x		x		x	x	x		
pcaGH	protocatechuate 3,4-dioxygenase		x			x	x	x	x	x		x	x	
pcaB	3-carboxy-cis,cis-muconate cycloisomerase		x			x	x	x				x	x	
pcaC	4-carboxymuconolactone decarboxylase	x	x	x	x		x	x		x	x	x	x	
pobA	3-hydroxybenzoate 4-monooxygenase	x	x		x		x	x		x		x	x	
DHBD	2,3-dihydroxybenzoate decarboxylase									x				
cmtC	2,3-dihydroxy-p-cumate/2,3-dihydroxybenzoate 3,4-dioxygenase		x							x	x	x		
cmtD	HCOMODA/2-hydroxy-3-carboxy-muconic semialdehyde decarboxylase		x							x	x	x		
xylXYZ	benzoate/toluate 1,2-dioxygenase					x			x			x	x	x
xylL	dihydroxycyclohexadiene carboxylate dehydrogenase					x			x				x	
xylE	catechol 2,3-dioxygenase		x	x	x		x	x					x	x
xylF	2-hydroxymuconate semialdehyde hydrolase									x		x	x	
xylG	2-hydroxymuconate semialdehyde dehydrogenase		x				x			x	x	x		
xylH	2-hydroxymuconate tautomerase	x	x	x	x		x	x		x	x	x	x	
xylI	4-oxalocrotonate decarboxylase									x		x		
xylJ	2-oxopent-4-enoate hydratase		x			x	x	x	x	x	x	x		
xylK	4-hydroxy-2-oxovalerate aldolase		x	x					x	x	x	x	x	
catA	catechol 1,2-dioxygenase		x			x			x	x		x	x	
catB	muconate cycloisomerase		x			x			x		x	x	x	
catC	muconolactone Delta-isomerase				x	x			x		x	x	x	
pcaD	3-oxoadipate enol-lactonase		x		x		x	x	x	x	x	x	x	
pcaIJ	3-oxoadipate CoA-transferase				x		x	x		x	x	x	x	
pcaF	acetyl-CoA acyltransferase	x	x	x	x	x	x	x	x	x	x	x		x

Publications

In the course of her PhD studies, the candidate participated in the following publications:

Vila, J., **Jiménez-Volkerink, S. N**. & Grifoll, M. (2019). Biodegradability of Recalcitrant Aromatic Compounds. *Comprehensive Biotechnology (Third Edition)*, *6*, 15–28. Pergamon. https://doi.org/10.1016/B978-0-444-64046-8.00343-8

Abstract

Polycyclic aromatic hydrocarbons (PAHs) are major environmental pollutants affecting soils and sediments in a number of point source polluted sites. Of all the biogeochemical processes involved in their environmental fate, microbial biodegradation by the naturally occurring microbial communities is the only one that guarantees the destruction of the PAH molecules. Bioremediation, exploiting these natural biodegradation processes, is considered the most cost-efficient technology for the clean-up of polluted sites. However, the success of bioremediation is still constrained by its unpredictable outcome, and the lack of tools for the accurate risk assessment and monitoring of the process. To overcome these limitations, it is essential to have a wide understanding of the microbial processes that determine the fate of PAHs in situ, moving forward from the traditional "black box" perspective to an actual environmental biotechnology. To understand the complex catabolic networks involved in environmental PAH biodegradation, a multiphasic research approach is required that expands from single cell–single substrate studies to the application of -omics tools in situ. This article aims to give a general view of the state-of-the-art knowledge on the metabolic networks that drive PAH removal (encompassing genomic, transcriptomic, proteomic and metabolomic studies of single cells and complex microbial communities), and on the environmental factors that affect them during the *in situ* processes.

Posada-Baquero, R., **Jiménez-Volkerink, S. N.**, García, J. L., Vila, J., Cantos, M., Grifoll, M., & Ortega-Calvo, J. J. (2020). Rhizosphere-enhanced biosurfactant action on slowly desorbing PAHs in contaminated soil. *Science of The Total Environment*, 720, 137608. https://doi.org/10.1016/j.scitotenv.2020.137608

Abstract

We studied how sunflower plants affect rhamnolipid biosurfactant mobilization of slowly desorbing fractions of polycyclic aromatic hydrocarbons (PAHs) in soil from a creosote-contaminated site. Desorption kinetics of 13 individual PAHs revealed that the soil contained initially up to 50% slowly desorbing fractions. A rhamnolipid biosurfactant was applied to the soil at the completion of the sunflower cycle (75 days in greenhouse conditions). After this period, the PAHs that remained in the soil were mainly present in a slowly desorbing form as a result of the efficient biodegradation of fast-desorbing PAHs by native microbial populations. The rhamnolipid enhanced the bioavailable fraction of the remaining PAHs by up to 30%, as evidenced by a standardized desorption extraction with Tenax, but the enhancement occurred with only planted soils. The enhanced bioavailability did not decrease residual PAH concentrations under greenhouse conditions, possibly due to ecophysiological limitations in the biodegradation process that were independent of the bioavailability. However, biodegradation was enhanced during slurry treatment of greenhouse planted soils that received the biosurfactant. The addition of rhamnolipids caused a dramatic shift in the soil bacterial community structure, which was magnified in the presence of sunflower plants. The stimulated groups were identified as fast-growing and catabolically versatile bacteria. This new rhizosphere microbial biomass possibly interacted with the biosurfactant to facilitate intra-aggregate diffusion of PAHs, thus enhancing the kinetics of slow desorption. Our results show that the usually limited biosurfactant efficiency with contaminated field soils can be significantly enhanced by integrating the sunflower ontogenetic cycle into the bioremediation design