environmental microbiology

Environmental Microbiology (2021) **23**(6), 2969–2991 doi:10.1111/1462-2920.15509

# PahT regulates carbon fluxes in *Novosphingobium* sp. HR1a and influences its survival in soil and rhizospheres

**Ana Segura**,[1] **Zulema Udaondo**[2] **and Lázaro Molina**[1*]

[1]*Environmental Protection Department, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas, C/Profesor Albareda 1, Granada, 18008, Spain.*

[2]*Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR, 72205.*

## Summary

***Novosphingobium* sp. HR1a is a good biodegrader of PAHs and aromatic compounds, and also a good colonizer of rhizospheric environments. It was previously demonstrated that this microbe is able to co-metabolize nutrients existing in root exudates together with the PAHs. We have revealed here that PahT, a regulator of the IclR-family, regulates the central carbon fluxes favouring the degradation of PAHs and mono-aromatic compounds, the ethanol and acetate metabolism and the uptake, phosphorylation and further degradation of mono- and oligosaccharides through a phosphoenolpyruvate transferase system (PTS). As final products of these fluxes, pyruvate and acetyl-CoA are obtained. The *pahT* gene is located within a genomic region containing two putative transposons that carry all the genes for PAH catabolism; PahT also regulates these genes. Furthermore, encoded in this genomic region, there are genes that are involved in the recycling of phosphoenolpyruvate, from the obtained pyruvate, which is the motor molecule involved in the saccharide uptake by the PTS system. The co-metabolism of PAHs with different carbon sources, together with the activation of the thiosulfate utilization and an alternative cytochrome oxidase system, also regulated by PahT, represents an advantage for *Novosphingobium* sp. HR1a to survive in rhizospheric environments.**

Received 21 January, 2021; revised 29 March, 2021; accepted 3 April, 2021. *For correspondence. E-mail lazaro.molina@eez.csic.es; Tel. +34 958 181600; Fax +34 958 129600.

## Introduction

Polycyclic aromatic hydrocarbons (PAHs) are released on a daily basis by the incomplete combustion of organic matter in industrial, traffic and heating activities. PAHs are also released in high quantities by natural processes, such as natural fires, volcanic eruptions, or oil spills. In the atmosphere. PAHs form part of the particulate matter, being especially abundant in the small particles (Kong *et al*., 2013). PAH contamination may be found at high concentrations in urban or industrial areas, but, once in the atmosphere, by the action of wind, they can be transported and deposited in places far away from their production sites, leading to diffuse, medium-low concentration, polluted areas (soil and water environments). Because of the toxicity of PAHs, the elimination of these compounds constitutes a priority for different health and environmental agencies (Panagos *et al*., 2013). PAH eradication is complicated, especially in medium–low contaminated sites, because of the high cost of the conventional physical–chemical remediation techniques (Kuppusamy *et al*., 2017).

Rhizoremediation, the utilization of pollutant degrader microorganisms in the plant root environment, constitutes a relatively cheap alternative for the elimination of these diffuse contaminations (Kuiper *et al*., 2004). This strategy can also offer attractive economic benefits (Caplan, 1993) if the used plants are afterwards collected as a source of added value products for the biomass, biofuel or paper industries. Successful examples of rhizoremediation of polychlorinated biphenyls, PAHs and other contaminants have been frequently reported in the literature (Balseiro-Romero *et al*., 2016; Zafra *et al*., 2017; Terzaghi *et al*., 2019; Sandhu *et al*., 2020; Simmer *et al*., 2020; Wolf *et al*., 2020; Zhao *et al*., 2021). However, there are still some key aspects of plant–bacteria interactions during rhizoremediation that remain unknown and may preclude its successful outcome.

One of the key factors in the bioremediation process is the selection of an appropriate microorganism. *Novosphingobium* sp. HR1, the strain used in this study, is an excellent phenanthrene degrader (Segura *et al*., 2017) that was isolated from the rhizosphere of plants growing in sandy soils near an oil refinery in Huelva, in southern Spain (Rodriguez-Conde *et al*., 2016). This microorganism carries the necessary genetic information that codifies for PAH degradation pathways in two putative transposons (Segura *et al*., 2017). The first one contains the genes encoding a ring-hydroxylating PAH dioxygenase (*pahAB*) that belongs to the ring-hydroxylating oxygenases family with broad specificity of substrates (Vila *et al*., 2015); in the second transposon, PahR, a sigma 54 regulator that is the main regulator of the *pahAB* genes (Segura *et al*., 2017), is encoded.

In *Novosphingobium* sp. HR1a, PahAB is required for the degradation of several PAHs (naphthalene, phenanthrene, antracene, chrysene and pyrene) as well as biphenyl (Segura *et al*., 2017). The PAH degradation pathways of *Novosphingobium* sp. HR1a, as described in other microorganisms, consist of the dihydroxylation of one of the aromatic rings that is sequentially reduced and re-aromatized by the action of an aryl-dehydrogenase into the diol form, which is opened by a ring-cleavage dioxygenase (Kanaly and Harayama, 2010; Ghosal *et al*., 2016). After several transformations, the resulting compounds lost one of the aromatic rings that was susceptible to the new attack of a dioxygenase. The -diol intermediates can be cleaved by an intradiol (*ortho*-cleavage) or an extradiol (*meta*-cleavage) ring-cleaving dioxygenase (Fig. 1). Most of the studies carried out with the model PAH, phenanthrene, report the dioxygenation at the 3,4-carbon positions to form phenanthrene-3,-4-dihydrodiol, which is metabolized to 1-hydroxy-2-naphthoic acid (Roy *et al*., 2012). This compound could be degraded via *o*-phthalate and protocatechuate to render succinyl-CoA and acetyl-CoA (Iwabuchi and Harayama, 1997; Gao *et al*., 2013). The decarboxylation of 1-hydroxy-2-naphthoic acid has also been reported and leads to the formation of 1,2-dihydroxynaphthalene which is metabolized by the well-characterized naphthalene degradation pathway via salicylic acid and catechol (or gentisate), to finally produce pyruvate and acetyl-CoA (Evans *et al*., 1965). The initial dioxygenation at the 1,2-position of phenanthrene leading to the formation of 2-hydroxy-1-naphthoic acid has also been reported (Gao *et al*., 2013). However, despite the huge knowledge about the metabolic pathways of degradation PAHs, their regulation is still quite unknown.

A good rhizoremediator agent, in addition to an effective capacity to eliminate the contaminant, should also be able to endure the competitive stress exerted by the native microorganisms and by the host plants (Matilla *et al*., 2007; Sánchez-Cañizares *et al*., 2017). Plants can exude 20%–50% of their photosynthetic production of carbon to root exudates (el Zahar Haichar *et al*., 2014). These nutrients, provided by the host plant, promote the growth of microorganisms and favour contaminant's mobilization (Vandenkoornhuyse *et al*., 2007; Segura and Ramos, 2013; Lu *et al*., 2017; Sasse *et al*., 2018; Rodríguez-Garrido *et al*., 2020). However, the presence of easily utilizable nutrients in the root exudates (sugars, organic acids, amino acids or fatty acids) could also repress the degradation of aromatic compounds (Rentz *et al*., 2004), such as PAHs, in a phenomenon termed catabolic repression (Rojo, 2010).

In this study, we report the important role of *orf1998* (named *pahT*) in the interconnection between PAH degradation and the utilization of root exudates nutrients. This gene codifies a regulator of the IclR family that is located within the second putative transposon. Control over the metabolization of different nutrients is important in the plant rhizosphere during rhizoremediation. We have also demonstrated that *pahT* is expressed in root exudates, even in the absence of PAHs, and plays an important role in the survival of *Novosphingobium* sp. HR1a in the rhizosphere.

## Results

### In silico *identification and organization of PAH degradation pathways in* Novosphingobium *sp. HR1a*

In sphingomonads, PAH degradation genes are normally encoded in plasmids or mobile genetic elements (Stolz, 2014) but the spatial architecture of these genes is very fluid; they can be integrated in different transposons, with different genetic organization and/or in different locations within the same transposon (Suppl. Fig. 1). Although the phylogenetic analysis of the strain *Novosphingobium* sp. HR1a (Suppl. Fig. 2) revealed that it is closely related to *N. resinovorum* (Hegedűs *et al*., 2017) and *N. guangzhouense* (Sha *et al*., 2017), these strains are lacking the genetic information necessary for PAH degradation (Suppl. Fig. 2). *In silico* analysis of the *Novosphingobium* sp. HR1a proteins encoded in this genomic region revealed an average identity of over 80% with similar proteins from *Sphingobium yanoikuyae* B1 (84.18%), *Sphingobium fuliginis* DSM 18781, (83.41%) and *Sphingobium* sp. MP9-4, (81.42%) (Suppl. Table 1). These three *Sphingobium* strains were isolated from contaminated soils, specifically from a PAH polluted stream, from a fly ash dumping site of a thermal power plant and from a petroleum-contaminated soil respectively (Prakash and Lal, 2006; Zhao *et al*., 2015; Zhong *et al*., 2017). *Novosphingobium* sp. PP1Y

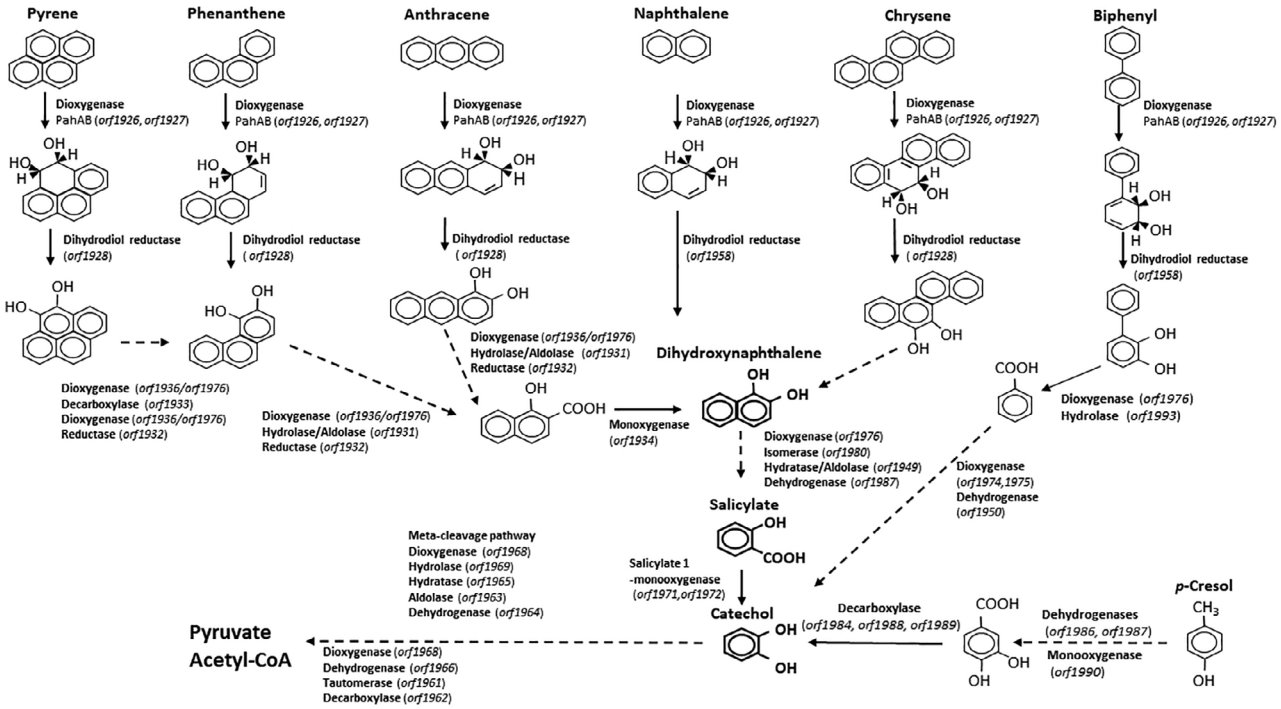


**Fig 1.** Schematic representation of the PAH degradation pathways in *Sphingomonadaceae*: the enzymes responsible for each reaction are in bold; the putative *Novosphingobium* sp. HR1a *orfs* involved in each reaction are below the generic name of the enzyme. The dotted arrow means that different reactions have taken place for the conversion of one compound to another.

(D'Argenio *et al*., 2014), *Sphingomonas paucimobilis* EPA505 (Story *et al*., 2004), *Croceicoccus naphthovorans* PQ2 (Huang *et al*., 2015) and *Novosphingobium pentaromativorans* US6-1 (Choi *et al*., 2015) also presented high sequence homology in this region (Suppl. Table 1), although the protein identity is lower than with the three mentioned *Sphingobium* strains. (around 75%–79%). All these strains are able to metabolize PAHs. Genes putatively associated with PAH degradation pathways are encoded in two contiguous genomic regions flanked by transposases (Segura *et al*., 2017). The genes that encoded for the dioxygenase (*pahAB*) involved in the first step of the degradation of several PAHs are located in the first putative transposon (Fig. 2) and transcribed with other genes encoding putative enzymes with functions involved in the conversion of phenanthrene, anthracene, chrysene and pyrene (*orf1926-orf1936*) to dihydroxy-naphthalene (Segura *et al*., 2017; Figs 1 and 2; Table 1). PahAB are highly conserved in the three mentioned *Sphingobium* strains, in the two *Novosphingobium* strains and in *Sphingomonas paucimobilis* EPA505 and *C. naphthovorans* PQ2 (>90% of identity) (Suppl. Table 1).

Genes involved in the conversion of dihydroxy-biphenyl (*orf1958*, *orf1976* and *orf1993*), *p*-cresol (*orf1986*, *orf1987* and *orf1990*) and dihydroxy-naphthalene (*orf1976*, *orf1980*, *orf1949*, *orf1987*, *orf1971* and *orf1972*) into catechol (Hopper and Taylor, 1975; Schell, 1983; Chakraborty and Das, 2016) are encoded in the putative second transposon (Figs 1 and 2). Catechol is finally converted into pyruvate and acetyl-CoA by the enzymes encoded by *xylEFGHIJK* genes that are clustered together in this putative second transposon (*orf1972-orf1961*). The product of *orf1951-52* has a high sequence homology with XylE, a ferredoxin reductase, an enzyme required for the action of the dioxygenases. The regulatory gene, *pahR*, is encoded in the genomic region between the *nah* and *xyl* genes (Fig. 2).

The genetic information described here, together with previous studies that demonstrated that the induction of *pahAB* genes was higher in the presence of salicylate than in the presence of different PAHs (Segura *et al*., 2017), suggests that PAH degradation in *Novosphingobium* sp. HR1a occurs via salicylate and catechol. Furthermore, we were unable to identify a gene coding for phthalate dioxygenase in *Novosphingobium* sp. HR1a.

### *The regulatory protein encoded by* orf1998 *is involved in PAH degradation*

We had identified the presence of four regulators within the two putative transposons (Fig. 2). We had previously

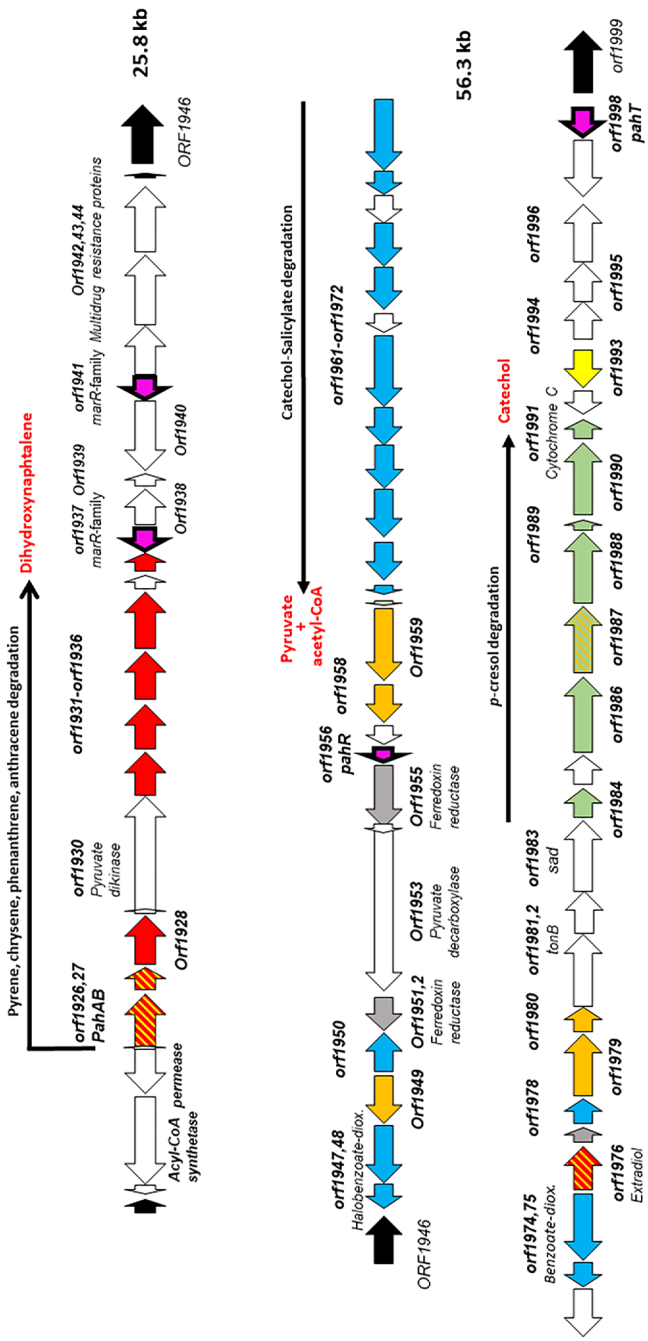


**Fig 2.** Schematic representation of the two putative transposons of *Novosphingobium* sp. HR1a: this genomic region contains the information for the metabolization of different PAHs. Arrows in red represent *orfs* that are involved in PAH conversion into dihydroxynaphthalene. Arrows in dark yellow are *orfs* involved in the conversion of this last compound into salicylate. Arrows in blue indicate *orfs* involved in the conversion of salicylate to catechol and finally to intermediates of the Krebs cycle. Arrows in black represent genes encoding functions related with transposition events. Regulatory genes are depicted as pink arrows. Grey arrows represent *orfs* encoding ferredoxins. Green arrows represent *orfs* involved in *p*-cresol to catechol. Light yellow are genes involved in biphenyl degradation. Striped arrows indicate possible mixed functions. White arrows represent genes that encoded for functions possibly not related with PAH degradation. [Color figure can be viewed at wileyonlinelibrary.com]

**Table 1.** Results of the three transcriptomic analyses.

| ORF | Function | 1998 glc/WT glc | 1998 phe/WT phe | WT phe/WT glc |
|---|---|---|---|---|
| | | | Induction fold | |
| **Degradative transposons genes** | | | | |
| First transponson | | | | |
| 1921 | Mobile element protein | – | – | – |
| 1922 | Hypothetical protein | – | – | 1.4 |
| 1923 | Acetyl-coenzyme A synthetase (EC 6.2.1.1) | – | – | 1.0 |
| 1924 | Permease of the drug/metabolite transporter (DMT) superfamily | – | 0.2 | 1.5 |
| 1925 | Hypothetical protein | 0.2 | 0.5 | 2.4 |
| 1926 | Large subunit naph/bph dioxygenase (PahA) | −1.0 | −0.2 | 7.2 |
| 1927 | Small subunit naph/bph dioxygenase (PahB) | −0.5 | −0.4 | 7.6 |
| 1928 | Benzyl alcohol dehydrogenase (PhnB) | −0.6 | −0.4 | 8.1 |
| 1929 | Hypothetical protein | −0.7 | −0.4 | 6.5 |
| 1930 | Pyruvate phosphate dikinase (EC 2.7.9.1) (PahD) | −0.5 | −0.3 | 8.8 |
| 1931 | Aldolase/hydratase (PahC) | – | −0.4 | 9.1 |
| 1932 | Reductase (NidD) | – | −0.5 | 8.5 |
| 1933 | Oxidase/decarboxylase | – | −0.4 | 8.1 |
| 1934 | Monooxygenase | – | −0.4 | 9.0 |
| 1935 | Hypothetical protein | 0.7 | – | 4.6 |
| 1936 | Dioxygenase | 0.5 | – | 1.1 |
| 1937 | Transcriptional regulator MarR | 0.5 | 0.6 | 0.1 |
| 1938 | Hydrolase-alpha/beta hydrolase fold family | 0.4 | 0.6 | – |
| 1939 | Thioesterase | 0.3 | 0.4 | – |
| 1940 | Phosphotransferase | – | 0.3 | – |
| 1941 | Transcriptional regulator MarR | – | – | – |
| 1942 | Membrane fusion component of tripartite multidrug resistance system | −0.5 | – | −0.8 |
| 1943 | Inner membrane component of tripartite multidrug resistance system | – | 0.5 | −0.6 |
| 1944 | Outer membrane component of tripartite multidrug resistance system | – | 0.6 | −0.6 |
| 1945 | Mobile element protein | – | 0.7 | – |
| Second transposon | | | | |
| 1946 | Mobile element protein | 0.6 | – | 0.9 |
| 1947 | Ortho-halobenzoate 1,2-dioxygenase beta-ISP protein OhbA | −0.7 | −1.5 | 7.7 |
| 1948 | Large subunit aromatic oxigenase OhbB | −0.9 | −0.9 | 5.8 |
| 1949 | 4-hydroxy-tetrahydrodipicolinate synthase (EC 4.3.3.7) NahE | −2.0 | −0.7 | 5.3 |
| 1950 | 1,2-dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenase (EC 1.3.1.25) (XylE) | −0.6 | −0.4 | 2.0 |
| 1951 | Ferredoxin reductase | – | – | – |
| 1952 | Ferredoxin reductase | – | −0.5 | 3.3 |
| 1953 | Pyruvate carboxylase (EC 6.4.1.1) | −0.4 | −0.8 | 3.4 |
| 1954 | Hypothetical protein | – | – | 0.9 |
| 1955 | Ferredoxin reductase | −0.5 | −0.5 | 3.2 |
| 1956 | Regulator aromatic degradative pathways (PahR) | – | −0.4 | 4.0 |
| 1957 | Hypothetical protein | – | −0.3 | 4.8 |
| 1958 | Dihydrodiol dehydrogenase (EC 1.3.1.56) (NahB) | – | −0.5 | 4.3 |
| 1959 | Aldehyde dehydrogenase (EC 1.2.1.3) (PhnF) | – | −0.5 | 3.9 |
| 1960 | Hypothetical protein | −0.5 | −0.8 | 5.4 |
| 1961 | 4-oxalocrotonate tautomerase (EC 5.3.2.-) (XylH) | −0.7 | −0.8 | 4.9 |
| 1962 | 4-oxalocrotonate decarboxylase (EC 4.1.1.77) (xylI) | −0.4 | −0.8 | 4.6 |
| 1963 | 4-hydroxy-2-oxovalerate aldolase (EC 4.1.3.39) (XylK) | −0.6 | −0.7 | 4.8 |
| 1964 | Acetaldehyde dehydrogenase acetylating (EC 1.2.1.10) (DmpF) | −0.5 | −0.7 | 4.6 |
| 1965 | 2-keto-4-pentenoate hydratase (EC 4.2.1.80) (XylJ) | −0.7 | −0.7 | 4.1 |
| 1966 | Putative 5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase oxidoreductase protein (EC 1.2.1.60) (XylG) | −0.4 | −1.0 | 6.9 |
| 1967 | Protein GlcG | −0.6 | −1 | 6.8 |
| 1968 | Catechol 2,3-dioxygenase (EC 1.13.11.2) (XylE) | −0.6 | −0.9 | 5.6 |
| 1969 | 2-hydroxymuconic semialdehyde hydrolase (EC 3.7.1.9) (XylF) | −0.6 | −0.7 | 6.1 |
| 1970 | Glutathione S-transferase (EC 2.5.1.18) | −0.3 | −0.7 | 5.0 |
| 1971 | Ortho-halobenzoate 1,2-dioxygenase beta-ISP protein BphA2d | −0.5 | −0.8 | 4.9 |

*(Continues)*

**Table 1.** Continued

| ORF | Function | 1998 glc/WT glc | 1998 phe/WT phe | WT phe/WT glc |
|---|---|---|---|---|
| | | | Induction fold | |
| 1972 | Large subunit toluate/benzoate dioxygenase (BphA1c) | −0.5 | −0.7 | 4.6 |
| 1973 | 4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262) | −0.8 | −1.0 | 6.1 |
| 1974 | Benzoate 1,2-dioxygenase beta subunit (EC 1.14.12.10) (XylY) | −0.7 | −0.9 | 5.1 |
| 1975 | Large subunit toluate/benzoate dioxygenase (XylX) | −0.7 | −0.5 | 4.0 |
| 1976 | 2,3-dihydroxybiphenyl 1,2-dioxygenase (EC 1.13.11.39) (NahC) | −1.1 | −0.6 | 3.5 |
| 1977 | Ferredoxin subunits of nitrite reductase and ring-hydroxylating dioxygenases | −1.1 | −0.9 | 4.2 |
| 1978 | Ortho-halobenzoate 1,2-dioxygenase beta-ISP protein NagH | −0.9 | −1.2 | 5.2 |
| 1979 | Large subunit toluate/benzoate dioxygenase (NagH) | −0.9 | −1.3 | 6.3 |
| 1980 | 2-hydroxychromene-2-carboxylate isomerase (NahD) | −0.9 | −1.0 | 6.8 |
| 1981 | Outer membrane receptor protein | −0.7 | −0.8 | 4.9 |
| 1982 | Outer membrane receptor protein | −0.8 | −1.0 | 4.2 |
| 1983 | Succinate-semialdehyde dehydrogenase [NAD(P)+] (EC 1.2.1.16) (Sad) | −0.9 | −1.1 | 4.5 |
| 1984 | 3-polyprenyl-4-hydroxybenzoate carboxy-lyase UbiX (EC 4.1.1.-) | −0.9 | −0.8 | 5.6 |
| 1985 | Hypothetical protein | −0.9 | −1.1 | 4.4 |
| 1986 | 4-cresol dehydrogenase [hydroxylating] flavoprotein subunit (EC 1.17.99.1) | −0.9 | −1.1 | 4.8 |
| 1987 | Aldehyde dehydrogenase (EC 1.2.1.3) (NahF) | −1.0 | −1.1 | 4.8 |
| 1988 | Hydroxyaromatic non-oxidative decarboxylase protein C (EC 4.1.1.-) | −0.9 | −1.1 | 5.2 |
| 1989 | Hydroxyaromatic non-oxidative decarboxylase protein D (EC 4.1.1.-) (UbiD) | −0.7 | −1.2 | 5.9 |
| 1990 | 4-cresol dehydrogenase [hydroxylating] flavoprotein subunit (EC 1.17.99.1) | −0.8 | −1.0 | 4.4 |
| 1991 | Cytochrome C553 (soluble cytochrome f) | – | −1.0 | 4.8 |
| 1992 | Hypothetical protein | – | −0.2 | 1.0 |
| 1993 | 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (EC 3.7.1.-) (BphD) | −0.5 | −0.4 | 2.4 |
| 1994 | 2-oxo-hepta-3-ene-1,7-dioic acid hydratase (EC 4.2.-.-) (HpcG) | −0.6 | −0.3 | 4.3 |
| 1995 | 2,4-dihydroxyhept-2-ene-1,C7-dioic acid aldolase (EC 4.1.2.n4) (HpcH) | −0.6 | −0.5 | 2.6 |
| 1996 | L-carnitine dehydratase/bile acid-inducible protein F (EC 2.8.3.16) | – | −1.1 | 3.8 |
| 1997 | MFS family multidrug efflux protein | −0.6 | −0.4 | 1.1 |
| 1998 | Transcriptional regulator IclR family (PahT) | | | 0.3 |
| 1999 | Mobile element protein | – | 0.3 | – |
| **Sugar metabolism** | | | | |
| PTS system | | | | |
| 4145 | RpfN protein | −31.3 | −28.6 | 0.5 |
| 4146 | PTS system fructose-specific IIB component (EC 2.7.1.69) | −25.2 | −26.6 | 0.5 |
| 4147 | 1-phosphofructokinase (EC 2.7.1.56) | −21.5 | −29.1 | 1.0 |
| 4148 | Phosphoenolpyruvate-protein phosphotransferase of PTS system (EC 2.7.3.9) | −15.2 | −26.4 | 0.8 |
| **Oligosaccharide metabolism** | | | | |
| 1277 | Maltodextrin glucosidase (EC 3.2.1.20) | −0.1 | −1.6 | 0.4 |
| 1278 | Maltodextrin glucosidase (EC 3.2.1.20) | −0.5 | −1.1 | – |
| 1279 | Maltodextrin glucosidase (EC 3.2.1.20) | −0.4 | −2.9 | 1.0 |
| 1280 | Predicted maltose-specific TonB-dependent receptor | −0.7 | −1.3 | – |
| **Xylan metabolism** | | | | |
| 4031 | Beta-xylosidase (EC 3.2.1.37) | 2.3 | 2.5 | – |
| 4032 | Putative xylulose kinase | 2.0 | 2.9 | 0.2 |
| 4033 | Membrane protein | 0.4 | 0.2 | – |
| 4034 | Putative xylitol dehydrogenase | 2.1 | 2.5 | 1.4 |
| 4035 | Putative dehydrogenase | 0.9 | 2.5 | 0.7 |
| 4036 | Polysaccharide pyruvyl transferase family protein | 1.5 | 4.0 | 1.4 |
| **Glycan and trehalose metabolism** | | | | |
| 2941 | Trehalose-6-phosphate phosphatase (EC 3.1.3.12) | 2.8 | 1.2 | 0.0 |
| 2942 | Glucoamylase (EC 3.2.1.3) | 2.9 | 1.7 | 0.7 |

*(Continues)*

**Table 1.** Continued

| ORF | Function | 1998 glc/WT glc | 1998 phe/WT phe | WT phe/WT glc |
|---|---|---|---|---|
| | | | Induction fold | |
| 2943 | Alpha, alpha-trehalose-phosphate synthase [UDP-forming] (EC 2.4.1.15) | 3.3 | 1.9 | 1.2 |
| 2944 | Hypothetical protein | 2.5 | 2.9 | 1.2 |
| **Pyruvate metabolism** | | | | |
| 0786 | L-lactate dehydrogenase (EC 1.1.2.3) | – | −3.3 | 1.6 |
| 0787 | Iron-uptake factor PiuC | – | −1.7 | 0.5 |
| 4939 | RidA like protein | – | −1.6 | 0.5 |
| 4940 | D-amino acid dehydrogenase small subunit (EC 1.4.99.1) | – | −6.4 | 3.5 |
| 4941 | Alanine racemase (EC 5.1.1.1) | −0.3 | −3.2 | 0.3 |
| 5791 | Hypothetical protein | 2.4 | 2.6 | – |
| 5792 | Msr0042 protein | 1.3 | 1.0 | 0.1 |
| 5793 | Hypothetical protein | 1.6 | 1.3 | – |
| 5794 | 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) | 2.1 | 1.1 | 0.4 |
| 5795 | Pyruvate oxidase [ubiquinone cytochrome] (EC 1.2.2.2) | 1.6 | 1.0 | 1.7 |
| 5796 | Glutathione S-transferase (EC 2.5.1.18) | 3.2 | 0.9 | 1.2 |
| 5710 | Catalase (EC 1.11.1.6) | 2.8 | 2.4 | 0.9 |
| 5819 | Catalase (EC 1.11.1.6) | 2.3 | 1.4 | 0.1 |
| **Acetate metabolism** | | | | |
| 0282 | Aldehyde dehydrogenase (EC 1.2.1.3) | −8.4 | −25.6 | 2.8 |
| 0283 | Putative dehydrogenase | −10.1 | −24.5 | 2.3 |
| 0284 | GfdT protein | −1.1 | −2.1 | 0.3 |
| 0285 | Sensory box histidine kinase/response regulator | −0.2 | −1.7 | 0.4 |
| 0287 | Transcriptional regulator LuxR family | −1.4 | −5.6 | 0.9 |
| 0288 | Quino(hemo)protein alcohol dehydrogenase PQQ-dependent (EC 1.1.99.8) | −13.3 | −31.1 | 2.4 |
| 0289 | Hypothetical protein | −1.2 | −6.9 | 0.9 |
| 0290 | Extracellular substrate-binding protein associated with quino(hemo)protein alcohol dehydrogenase | −9.0 | −18.1 | 1.5 |
| 0291 | Cytochrome c550, associated with quino(hemo)protein alcohol dehydrogenase (EC 1.1.99.8) | −13.3 | −39.3 | 2.8 |
| 0292 | Hypothetical protein | −0.3 | −4.7 | 1.0 |
| 0293 | PQQ-dependent catabolism-associated beta-propeller protein | −2.7 | −9.9 | 2.0 |
| 1511 | Hypothetical protein | 0.0 | −4.4 | 2.2 |
| 1512 | Acetyl-CoA synthetase (ADP-forming) alpha and beta chains putative | 0.1 | −5.5 | 2.9 |
| 1513 | Acetyl-coenzyme A synthetase (EC 6.2.1.1) | −0.4 | −1.2 | 0.6 |
| 4899 | Acetyl-CoA hydrolase (EC 3.1.2.1) | 29.1 | 25.8 | – |
| **Aromatic compound metabolism** | | | | |
| 5156 | *p*-hydroxybenzoate hydroxylase (EC 1.14.13.2) | −1.2 | −1.5 | −0.5 |
| 5157 | Protocatechuate 3,4-dioxygenase beta chain (EC 1.13.11.3) | −1.9 | −6.5 | 0.3 |
| 5158 | Protocatechuate 3,4-dioxygenase alpha chain (EC 1.13.11.3) | −1.2 | −3.4 | 0.4 |
| 5159 | 3-carboxy-cis,cis-muconate cycloisomerase (EC 5.5.1.2) | −0.6 | −6.0 | 0.6 |
| 5160 | Beta-ketoadipate enol-lactone hydrolase (EC 3.1.1.24) | −0.8 | −2.3 | – |
| 5247 | 3-oxoadipate CoA-transferase subunit A (EC 2.8.3.6) | −3.7 | −1.1 | −1.9 |
| 5248 | 3-oxoadipate CoA-transferase subunit B (EC 2.8.3.6) | −4.2 | −0.7 | −1.3 |
| 5249 | Beta-ketoadipyl CoA thiolase (EC 2.3.1,-) | −2.8 | −0.6 | −1.1 |
| **Phenylalanine metabolism** | | | | |
| 2614 | Maleylacetoacetate isomerase (EC 5.2.1.2) @ Glutathione S-transferase zeta (EC 2.5.1.18) | 4.3 | 2.9 | 0.9 |
| 2615 | Glyoxalase family protein | 4.8 | 2.5 | 0.9 |
| 2616 | 4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27) | 5.1 | 3.3 | 0.5 |
| 2617 | Phenylalanine-4-hydroxylase (EC 1.14.16.1) | 5.2 | 3.8 | 0.2 |
| **Sulfur metabolism** | | | | |
| 0294 | Hypothetical protein | −7.7 | −20.4 | 2.4 |
| 0295 | SoxX/Cytochrome B561 | −1.1 | −6.8 | 0.8 |
| 0296 | Outer membrane receptor protein | −1.8 | −4.3 | 0.8 |
| 0297 | SoxY | −8.6 | −23.8 | 1.3 |
| 0298 | SoxZ | −11.5 | −29.0 | 1.4 |
| 0299 | Rhodanese | −5.5 | −15.7 | 2.2 |
| 0300 | SoxA/Cytochrome c | −11.9 | −29.2 | 1.9 |
| **Respiratory chain** | | | | |
| 3164 | Coproporphyrinogen III oxidase oxygen-independent (EC 1.3.99.22) | – | −4.6 | 14.8 |
| 3165 | Hypothetical protein | – | −5.7 | 31.9 |

*(Continues)*

**Table 1.** Continued

| ORF | Function | 1998 glc/WT glc | 1998 phe/WT phe | WT phe/WT glc |
|---|---|---|---|---|
| | | | Induction fold | |
| 3166 | Type cbb3 cytochrome oxidase biogenesis protein CcoS | – | – | – |
| 3167 | Type cbb3 cytochrome oxidase biogenesis protein CcoI (EC 3.6.3.4) | – | −5.5 | 21.8 |
| 3168 | Type cbb3 cytochrome oxidase biogenesis protein CcoH | −1.3 | −6.3 | 18.0 |
| 3169 | Type cbb3 cytochrome oxidase biogenesis protein CcoG | −1.3 | −7.6 | 15.5 |
| 3170 | Cytochrome c oxidase subunit CcoP (EC 1.9.3.1) | −1.2 | −9.7 | 16.6 |
| 3171 | Cytochrome c oxidase subunit CcoQ (EC 1.9.3.1) | – | −12.3 | 16.9 |
| 3172 | Cytochrome c oxidase subunit CcoO (EC 1.9.3.1) | −0.5 | −11.7 | 16.1 |
| 3173 | Cytochrome c oxidase subunit CcoN (EC 1.9.3.1) | – | −9.9 | 9.4 |
| 3174 | Outer membrane protein W precursor | – | −12.1 | 7.4 |

Only genes of the two putative transposons and genes with $\log_2$fold-change higher/lower than 1.5 have been included. Complete information is given in Suppl. Tables 3, 4 and 5.

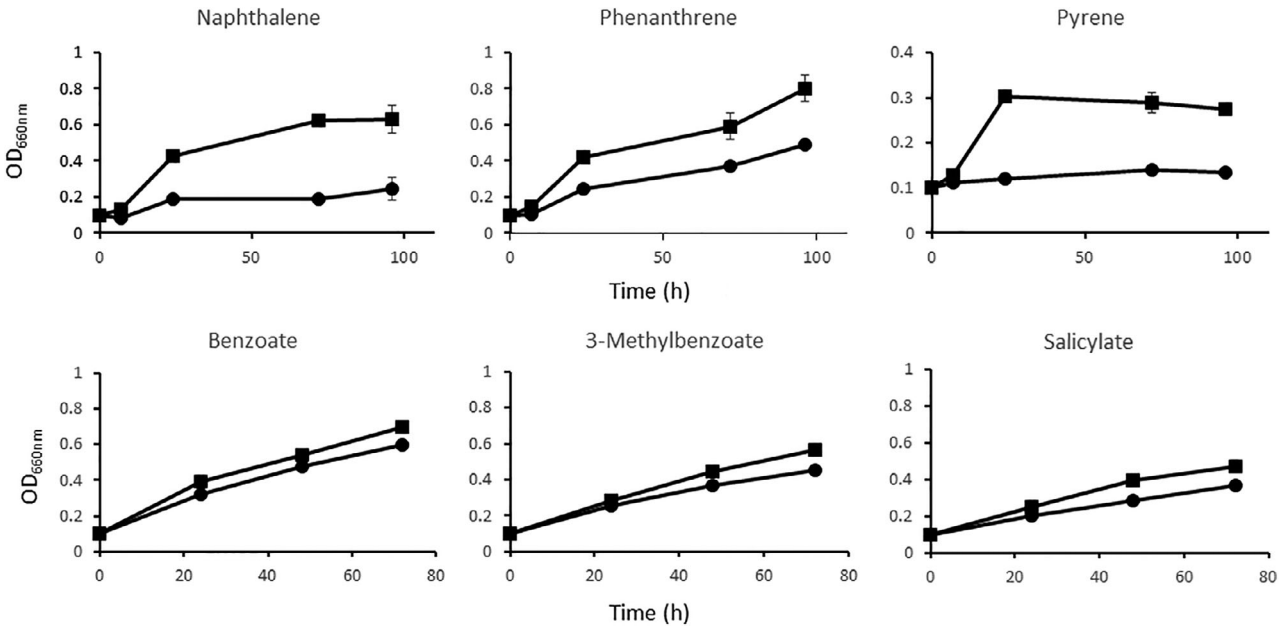


**Fig 3.** Growth of *Novosphingobium* sp. HR1 and *pahT* mutant in different aromatic compounds. Squares: *Novosphingobium* sp. HR1a; circles *pahT* mutant strain. $OD_{660nm}$ differences between both strains at the final time are statistically significant for all the compounds ($p < 0.005$).

determined that PahR plays a crucial role in the regulation of the PAH degradation, while the two regulators belonging to the MarR family (*orf1937* and *orf1941*) were not involved in this regulation (Segura *et al*., 2017). To determine if the fourth regulatory protein encoded by *orf1998* was involved in PAH metabolism in *Novosphingobium* sp. HR1a, we constructed a knockout mutant in this gene and we compared the growth of this strain with that of the wild type in minimal medium M9 plus different PAHs (naphthalene, phenanthrene and pyrene) or monocyclic aromatic compounds (benzoate, 3-methylbenzoate and salicylate) as the sole carbon source (Fig. 3). Growth with the three PAHs tested was significantly slower in the mutant than in the wild-type strain. When tested in monocyclic aromatic compounds the mutant strain also grew slightly more slowly than the wild type but the phenotype was less pronounced than when growing with PAHs. Given the implication of *orf1998* in *Novosphingobium* sp. HR1a growth using PAHs as the sole carbon source, we named this gene *pahT*.

To further confirm the participation of *pahT* in PAH metabolism, we analysed the expression from the *pahA* and *pahR* promoters in the wild type and the mutant strain. β-Galactosidase assays revealed that the absence of a functional *orf1998* led to a significant reduction of the expression from *pahA* and *pahR* promoters in the presence of salicylate, naphthalene and phenanthrene (Fig. 4A and B).

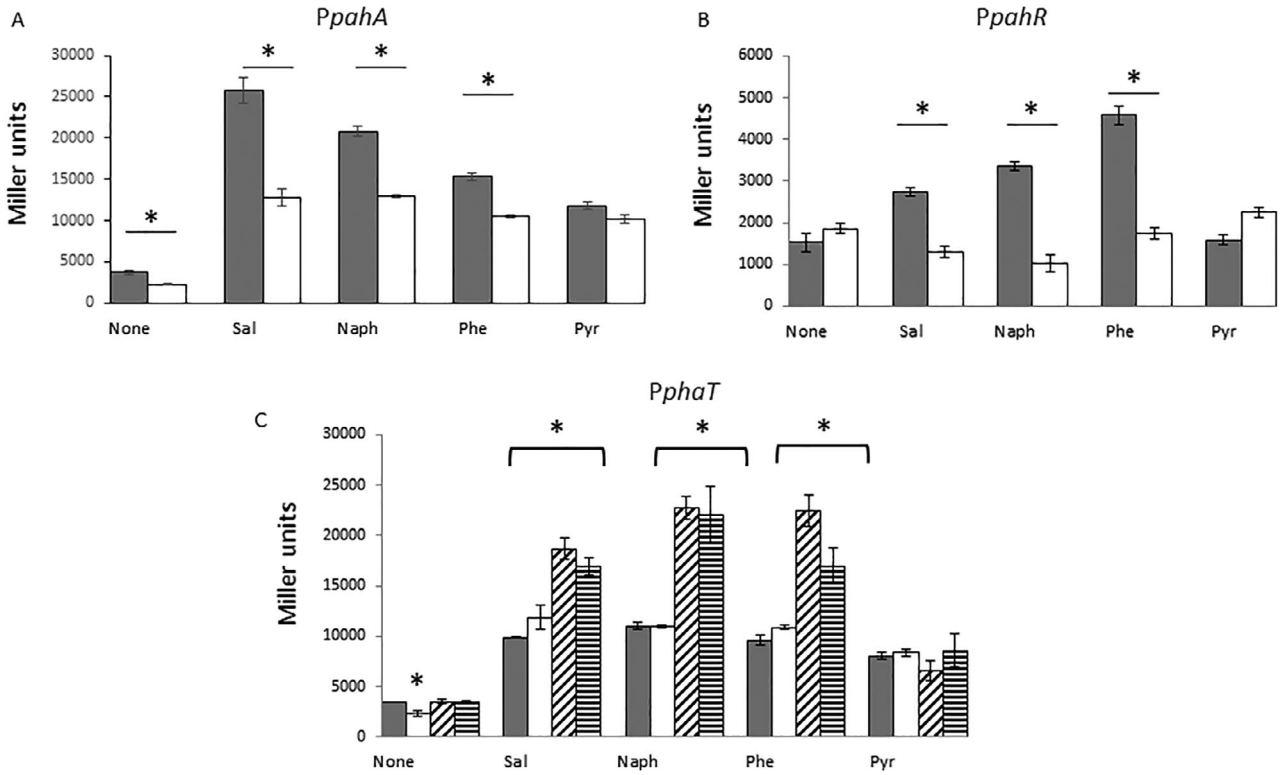


**Fig 4.** Expression from the (A) *phaA* (P*phaA*), (B) the *pahR* (P*pahR*) and (C) *pahT* (P*pahT*) promoters in different genetic backgrounds. Expression was measured as β-galactosidase activity using the wild-type and the mutant strains carrying the appropriate plasmids, pHR1a (*phaA* activity); pPHR (*phaR* activity) and pMP220-Pr1998 (*pahT* activity). Dark grey bars: *Novosphingobium* sp. HR1a; white bars: *Novosphingobium* sp. HR1a *pahT*$^{-}$; texture with oblique lines: *Novosphingobium* sp. HR1a *pahA*$^{-}$ mutant strain; texture with horizontal lines: *Novosphingobium* sp. HR1a *pahR*$^{-}$ mutant strain. *$p < 0.0005$.

We also analysed the expression of the *pahT* promoter in the absence/presence of salicylate, naphthalene and phenanthrene and pyrene (Fig. 4C). The addition of any of these compounds to the growth media induced the expression from the promoter (around threefold induction in presence of salicylate, naphthalene or phenanthrene and twofold induction in the presence of pyrene; Fig. 4C). The expression from the *pahT* promoter was not significantly affected in the knockout mutant strain in presence of PAHs, indicating the lack of self-regulation, at least in the presence of PAHs. However, in the absence of aromatic compounds, a small, but significant decrease of the expression was observed (Fig. 4C). To further analyse the expression from these promoters, we introduced the construction in two mutant strains; one in which the *pahA* gene (dioxygenase) was inactivated and other in which the *pahR* gene was inactivated. In both knockout mutants, the expression level of the *pahT* promoter was higher than in the wild-type strain in the presence of the inducers tested (Fig. 4C), except in the presence of pyrene. The higher expression level in these mutants than in the wild type could be due to the accumulation of PAHs or their degradation products within the mutant cells.

All these results indicate that *pahT* is involved in the regulation of the upper (from PAHs to salicylate) and lower (salicylate toward intermediates of the tricarboxylic acids) metabolic pathways of the degradation of PAHs. The relative low influence of the lack of PahT in the growth rate of the mutant cells when using monocyclic aromatic compounds could be explained because genes encoding alternative pathways for the degradation of benzoate and salicylate were found on the chromosome of *Novosphingobium* sp. HR1a, outside of the two putative transposons in which *pahAB*, *pahR* and *pahT* are encoded (not shown).

### *Genes related with PAH degradation, cytochrome oxidase* cbb$_3$*, iron uptake and* uspA *are strongly regulated in response to phenanthrene*

We carried out transcriptomic studies of the wild type growing on M9 minimal media plus glucose in the presence/absence of phenanthrene, and we compared the level of expression of the transcripts when growing in the presence versus the absence of phenanthrene (Supplementary Material Table 2).

In the wild-type strain, the presence of phenanthrene in the medium produced a significant induction of most of the genes encoded in the two transposons; among them, the previously referred to *pahA* and *pahR* (Table 1). *pahT* showed a modest but statistically significant ($\log_2$-fold change of 0.3, $p_{adj} < 0.05$), induction in response to phenanthrene. These results validated our experimental set-up, as they confirmed the results obtained with the β-galactosidase assays (Fig. 4). The genes within the two transposons whose expression did not change in the presence of phenanthrene were *orf1937* (MarR regulator), *orf1938* and *orf1939* (hypothetical proteins), and *orf1940* (probably encoding a ubiquinone biosynthesis monooxygenase). *orf1941* (MarR regulator) and *orf1942-orf1945* (encoding the components of a tripartite multi-drug resistance system) showed a moderate, but statistically significant, decrease in their expression (Table 1; Suppl. Fig. 3). These results indicate that not all the genes within the two putative transposons are subjected to similar regulation.

The most upregulated genes in response to phenanthrene in the wild-type strain (12-22-fold overexpressed in the presence of phenanthrene) are genes related with the biogenesis of $ccb_3$-type cytochrome *c* oxidase ($cbb_3$-Cox) (Table 1; Suppl. Table 2). Cox enzymes terminate the respiratory chains of aerobic and facultative aerobic organisms and have been implicated in colonization of low $O_2$ containing tissues. In most of the organisms studied, their level of expression increased under low oxygen concentrations (Ekici *et al*., 2012). UspA, the universal stress protein (Kvint *et al*., 2003), was also upregulated (about five times). Among the genes repressed in the presence of phenanthrene, there are several genes involved in iron acquisition and the heat shock proteins GroEL and GroES (Suppl. Table 2).

### pahT *modulates the expression of most, but not all, the genes within the two putative transposons*

We also carried out the transcriptomic analysis with the *pahT* mutant in presence of phenanthrene (Supplementary Material Table 3) and in absence of this compound (Supplementary Material Table 4). When we compared the level of expression of the PAH degradation genes within the putative transposons, we observed that their expression levels were moderately, but nevertheless significantly, reduced in the mutant when compared with the wild-type strain (Table 1) when growing in the presence of phenanthrene (Suppl. Figs 3 and 4). The exception to this general inhibition were the MarR regulators and their neighbouring genes that showed a slightly, but statistically significant, increase in their expression (Table 1, Suppl. Table 3; Suppl. Fig. 3).

In the absence of phenanthrene, a decreased expression of the genes involved in PAH degradation in the mutant compared with the wild-type strain was also observed, although the magnitude of this decrease is lower than in the presence of phenanthrene, except for the *pahA* gene (Suppl. Table 4; Suppl. Fig. 3).

All these results demonstrate that the regulator codified by *pahT* positively modulates the expression of genes involved in phenanthrene degradation, with this modulation being stronger in the genes responsible for the conversion of dihydroxynaphthalene to intermediates of the Krebs cycle than in those encoded in the putative first transposon.

### pahT *is a global regulator*

Our transcriptomic assays revealed that *pahT* also controls the expression of other genes in *Novosphingobium* sp. HR1a that are not located within the two putative transposons. We found that several genes involved in sugar, ethanol and sulfur metabolism, in *p*-hydroxybenzoate and protocatechuate degradation and in $cbb_3$ cytochrome oxidase biosynthesis were strongly downregulated in the *pahT* mutant strain when compared with the wild-type strain (regardless of the presence/absence of phenanthrene in the culture medium) whilst genes involved in polysaccharide and phenylalanine degradation and biosynthesis of compatible solutes were upregulated in the mutant.

### *Genes downregulated in the* pahT *mutant*

i. Sugar metabolism

*orfs 4145–4148* were among the most downregulated in the mutant strain, both in the presence and absence of phenanthrene (Table 1; Suppl. Tables 3 and 4). These genes encoded proteins putatively involved in sugar transport and fructose utilization through a phosphotransferase system (PTS) (Tchieu *et al*., 2001).

*orfs 1277–1280* were also slightly downregulated in the mutant strain compared with the wild type, especially in the presence of phenanthrene. *orf1277*, *orf1278* and *orf1279* encoded three putative maltodextrin glycosidades, whilst *orf1280* encoded for the ExbBD system involved in maltodextrin and disaccharide transport (Neugebauer *et al*., 2005).

These transcriptomic results suggested that the *pahT* mutant is defective in the uptake of mono- and oligosaccharides from the medium. To validate this hypothesis we monitored the growth of the wild type and *pahT* knockout mutant strains in cultures grown in M9 minimal medium with mono- and di-saccharides as the only carbon sources. As seen in Fig. 5, the lag

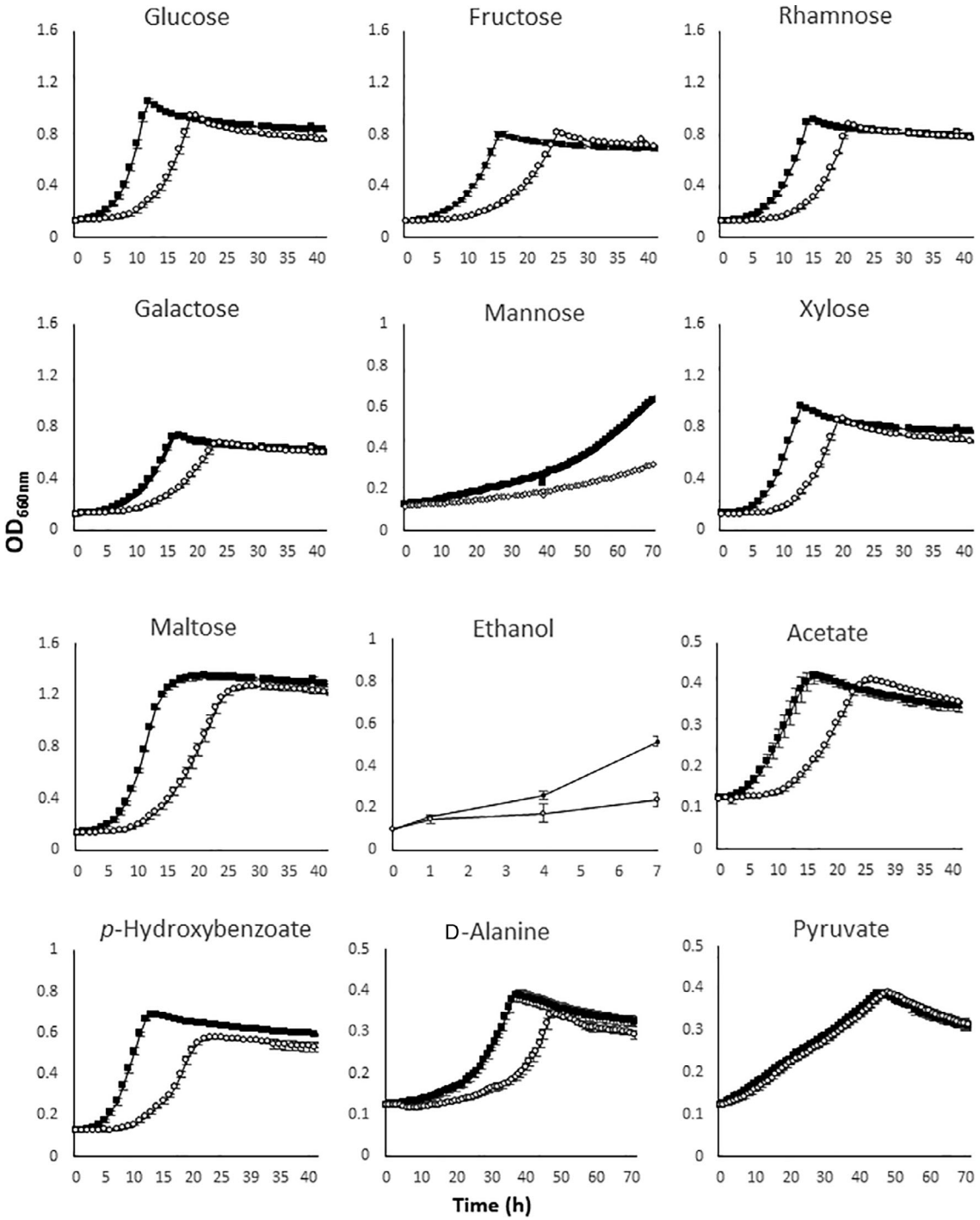


**Fig 5.** Growth curves of the wild-type strain (black squares) and the *pahT* mutant (white circles) in different carbon sources. Ethanol was added to a glass bar every day.

phase of the mutant was significantly longer and the growth rate lower than those of the wild-type strain when these strains were growing in the hexoses (glucose, fructose, rhamnose, galactose and mannose), xylose and the α-disaccharide, maltose. The growth defects of the mutant strain support the results of the transcriptomic analysis.

ii. Ethanol and acetate metabolism

The second group of genes that presented lower levels of transcription in the *pahT* mutant strain than in

the wild type was *orf282-300*. Within this gene cluster, *orf0289* codifies a quino(hemo) alcohol dehydrogenase pyrroloquinoline quinone (PQQ)-dependent, a membrane protein that is involved in the oxidation of ethanol to acetaldehyde. During this oxidation, electrons are transferred to the cytochrome *c* (*orf0291*) that participates in the reduction of the oxidized quinones by cytochrome reductase $bc_1$ complex forming part of the respiratory chain of *Novosphingobium* sp. HR1. The acetaldehyde produced by this dehydrogenase is transformed to acetate by the aldehyde dehydrogenase putatively encoded by *orf0282* and *orf0283* (Toyama *et al*., 1995; Yakushi and Matsushita, 2010). This enzyme is produced in high quantities in acidic and in low-aeration conditions (Matsushita *et al*., 1995). The resulting acetic acid is finally transformed in acetyl-CoA by the action of a chromosomic acetyl-CoA synthetase (*orf1512*, *orf1513*). Most of these genes presented lower levels of expression in the mutant in the presence of phenanthrene than in its absence (Table 1), probably because the level of expression of these genes is higher in the presence of phenanthrene than in its absence in the wild type. Accordingly, with the induction of the ethanol assimilation pathway by PahT, the *phaT*$^-$ mutant grew much more slowly and had a longer lag phase in ethanol than the wild-type strain (Fig. 5). A similar phenotype presented when acetate was the sole carbon source.

iii. Oxidation of thiosulphates

Within the gene cluster involved in the oxidation of ethanol to acetaldehyde, mentioned above (*orf282-300*), a periplasmic sulfur oxidation (SOX) is encoded. This complex is responsible for the oxidation of thiosulphates to sulphate ions (Grabarczyk and Berks, 2017). In this system, SoxZ (*orf297*) and SoxY (*orf298*) form a complex that binds sulfur derivatives, a reaction that is mediated by the action of two cytochromes SoxA (*orf300*) and SoxX (*orf295*). This binding produces the reduction of SoxA and SoxY that can reduce the oxidized quinones by the cytochrome reductase $bc_1$ complex of the respiratory chain. *orf299* codifies a rhodanase-like protein (SoxL) that could act as a sulfurtransferase able to release the sulphate ion from the complex SoxYZ. As with the genes involved in ethanol metabolism, downregulation of these genes was more evident in the presence of phenanthrene, probably because phenanthrene induced the expression of these genes (Table 1).

Growth of the mutant in a minimal medium plus glucose in which the only sulfur source was thiosulphate was indistinguishable from the growth in M9 minimal media plus glucose in which the sulfur source is magnesium sulphate (not shown). *Novosphingobium* sp. HR1a possesses, at least, an alternative mechanism for the utilization of thiosulphates: thiosulfate sulfurtransferase rhodanese (*orf3115*) which transforms the thiosulphate in sulphite and a sulphite reductase (*orf5675,6*) that renders as final product sulphate.

iv. p-hydroxybenzoate metabolism

PahT is also involved in the regulation of genes involved in the degradation of different aromatic compounds. Most of these metabolic pathways render acetyl-CoA as a final product. *orfs* 5156–5160 encode many of the enzymatic activities required for the conversion of *p*-hydroxybenzoate to protocatechuate (*orf5156*), catechol (*orf5157-orf5158*), and further degradation via the *ortho*-pathway (*orf5159*, *orf5160)*. *orf5247* and *orf5248* (3-oxoadipate CoA-transferase) and *orf5249* (β-ketoadipyl-CoA thiolase), the final steps in the degradation of catechol to render succinyl-CoA, were also downregulated in the mutant strain. Phenotypic studies revealed that the *pahT* mutant strain has a longer lag growth phase and slower growth rate than the wild type when they were cultivated in M9 minimal medium with *p*-hydroxybenzoic acid as the sole carbon source (Fig. 5). As seen in Fig. 3, there were only small differences in the growth of the wild type and the mutant strain when grown in benzoate, 3-methylbenzoate or salicylate. This could be explained because the expression of the genes involved in the catechol *meta*-pathway (*orf4703-8*) was not affected in the mutant strain.

v. $cbb_3$ cytochrome genes

The expression of genes involved in the structure, assembly and functioning of the complex of the terminal cytochrome oxidase $cbb_3$ (*orf3164-74*), that were clearly induced in the wild-type strain by the presence of phenanthrene in the media, were also highly downregulated in the *pahT* mutant strain in the presence of phenanthrene, whilst expression changes were quite modest in its absence (Table 1).

vi. Lactate and D-alanine metabolism

*orf0786*, encoding the L-lactate dehydrogenase, which is the enzyme responsible for the transformation of L-lactate in pyruvate, and *orf4939-orf4941* (encoding for a deaminase, and for a D-amino-acid dehydrogenase and alanine racemase respectively) are downregulated in the mutant strain in the presence of phenanthrene (Table 1). The expression of these genes was higher in the wild type when exposed to phenanthrene. The final central metabolite produced in these reactions, as well as in the PAH degradation (Fig. 1) is pyruvate.

Contained in the two putative transposons encoding for the PAH-degradation genes, and regulated by

*pahT* as shown above, there are also three genes related with pyruvate metabolism; in the first transposon, *orf1930* which codifies for a pyruvate phosphate dikinase (PPDK), the enzyme responsible for the recycling of phosphoenolpyruvate from pyruvate, and *orf1933*, a putative D-malic enzyme, transforming malate into pyruvate; in the second transposon, *orf1955* codifies for a pyruvate decarboxylase an enzyme that transforms pyruvate into oxaloacetate, an intermediate of the tricarboxylic acid cycle (Suppl. Figs. 3 and 4). A paralogue of *orf1930* (PPDK) was also identified on the chromosome of *Novosphingobium* sp. HR1a (*orf1458*, 71% identity) and its expression was not regulated by PahT or by the presence of phenanthrene.

Phenotypic analyses were carried out in order to test the different behaviour of *Novosphingobium* sp. HR1 and *pahT* mutant in minimal medium in presence of pyruvate and D-alanine. The mutant strain had a longer lag phase and lower growth rate than the wild type in D-alanine, corroborating the downregulation of genes involved in its transformation into pyruvate observed in the transcriptomic analysis (Fig. 5). No differences in growth were observed in pyruvate, probably because of the multiple pathways by which pyruvate can be metabolized in *Novosphingobium* sp. HR1a, although many of them are not affected by the absence of a functional PahT.

*Genes upregulated in the* pahT *mutant*

i. The most upregulated gene (more than 29 times) in the glucose medium was *orf4899* which codifies an acetyl-CoA hydrolase involved in the conversion of acetyl-CoA into acetate. The expression of this gene was not affected by the presence of phenanthrene.
ii. Phenylalanine degradation genes: Genes involved in the degradation of phenylalanine to fumarate and acetyl-CoA were overexpressed in the mutant strain (*orf2614-7*) *orf2617* codifies a phenylalanine hydroxylase, which transforms phenylalanine into tyrosine; after the action of a transaminase the resulting 4-hydroxyphenylpyruvate is hydroxylated by a dioxygenase (*orf2616*) to produce homogentisate. The aromatic ring of this compound is the substrate of another dioxygenase, and *orf2614* and *orf2615* encoded the enzymes responsible for the next steps in the catabolism to produce acetyl-CoA and fumarate.
iii. Sugar polymers degradation: *orfs* that encoded putative proteins related with the depolymerization of xylan or starch (β-xylosidase [*orf4031*], or glucoamylase [*orf2941*]), were overexpressed in the *phaT* mutant compared with the wild-type strain.
iv. Biosynthesis of osmoprotectants: *orfs* putatively involved in the biosynthesis of osmoprotectants such as Di-*myo*-inositol 1,1′-phosphate, sorbitol or trehalose (inositol 3P synthase [*orf4032*], glucose-fructose oxidoreductase [*orf4035*] or trehalose-6-phosphate-synthase and α-D-trehalose-phosphate synthase [UPD-forming]) and threonine dehydrogenase (*orf4034*), are overexpressed in the mutant (Kingston *et al*., 1996; Meyer *et al*., 2007; Anandham *et al*., 2008).
v. Pyruvate metabolism: Among the reactions that use pyruvate as substrate, we observed an evident overexpression of a PQQ dependent dehydrogenase, the pyruvate oxidase (*orf5795*), in the *pahT* mutant cells in the presence or absence of phenanthrene compared with the wild-type strain (Table 1), with the overexpression being higher in the presence of phenanthrene, probably because phenanthrene induced the expression of these genes. This enzyme is responsible for the oxidation of pyruvate into acetate. The same expression pattern presented the neighbouring genes, *orf5792* and *orf5793* which encode a ferredoxin-like protein and an oxidoreductase respectively. During the oxidation of pyruvate into acetate, oxygen peroxide is produced as a sub-product. The increase in the toxic oxygen peroxide, probably produced by the elevated levels of pyruvate oxidase in the mutant, is detoxified by catalases (the putative products of *orf5710* and *orf5819*) accordingly, these genes are overexpressed in the mutant strain.

pahT *plays an important role in the capacity of root colonization of* Novosphingobium *sp. HR1a*

As PahT was revealed as a regulatory protein of the general metabolism of *Novosphingobium* sp. HR1a, and because of the capacity of this strain to grow and utilize rhizospheric compounds as carbon sources, we compared the growth of the wild type and the mutant strain in gnotobiotic systems with clover. As shown in Fig. 6A, the wild type and mutant strains were able to increase their numbers 3 days after inoculation. In the mutant strain, these numbers were maintained until day 6, then decreasing after 9 days. The wild-type strain followed a similar pattern but the numbers of CFUs were always higher than in the mutant. When clover was not included in the gnotobiotic system and therefore, no carbon or nitrogen sources were available for the bacteria, we observed a rapid decrease in the number of CFUs, with the decrease being faster in the mutant than in the wild-type strain (Fig. 6B).

To further study the lack of fitness of the mutant strain while competing with the wild-type strain for rhizospheric resources, we co-inoculated gnotobiotic systems with

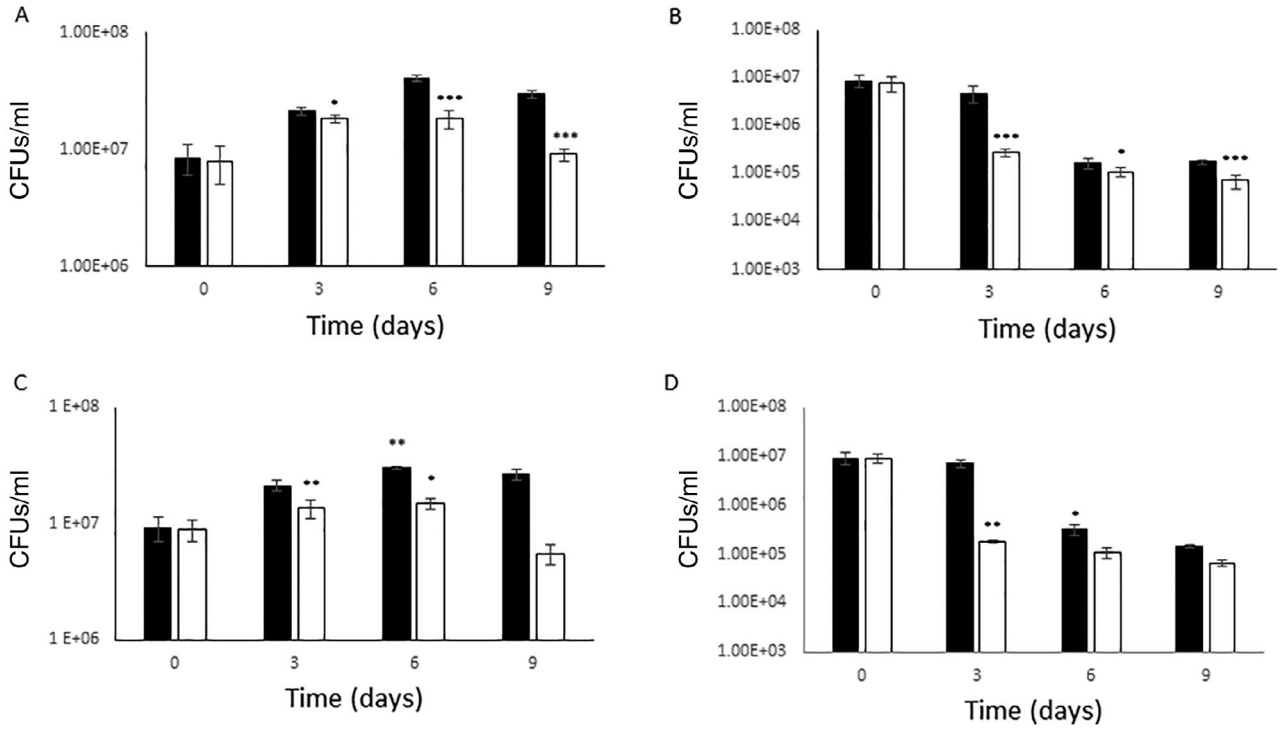


**Fig 6.** Growth of *Novosphingobium* sp. HR1a and *pahT* mutant in gnobiotic systems. Number of CFUs in gnobiotic systems with clover seeds (A) and without seeds (B) when each strain was inoculated separately. Number of CFUs in experiments in which the two strains have been co-inoculated in gnotobiotic systems with clover (C) and control experiments without clover (D). Black bars: *Novosphingobium* sp. HR1a; white bars: *pahT* mutant strain. ***$p < 0.0005$; **$p < 0.005$; *$p < 0.05$.

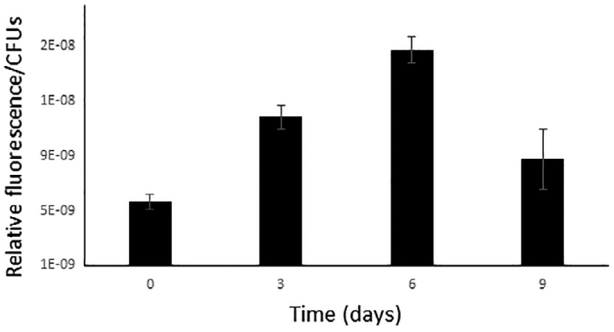


**Fig 7.** Expression from the *pahT* promoter in gnotobiotic systems. The gnotobiotic systems were inoculated with *Novosphingobium* sp. HR1a (pSEVA637-Pr1998). Samples of 200 μl of the liquid media were taken at 3, 6 and 9 days to measure the fluorescence.

both strains at similar initial numbers of CFUs. Once again, the numbers of CFUs in the mutant strain were significantly lower 3 and 6 days after inoculation than the numbers of CFUs in the wild type (Fig. 6C). Co-inoculation with both strains in systems without clover renders a faster decrease in the numbers of CFUs in the mutant than in the wild-type strain (Fig. 6D). These results indicate that the mutant in the *pahT* gene has a reduced capacity of growth in root exudates and a reduced capacity for survival in hostile environments.

To further explore the involvement of *pahT* in the fitness of *Novosphingobium* sp. HR1a in the rhizosphere, we analysed the level of expression of this gene at days 3, 6 and 9 after inoculation, using the strain *Novosphingobium* sp. HR1a (pSEVA637-Pr1998) in which the *pahT* promoter drives the activity of the *gfp* (Suppl. Table 5). We observed an increase of fluorescence per CFUs 3 and 6 days after inoculation (twofold induction after 3 days, and threefold induction after 6 days over the expression of the initial day), and a clear decrease in the expression at day 9 (1.5-fold induction over the expression of the initial day). These results indicate that this regulator is being expressed at least during the first 6 days of growth in the rhizosphere (Fig. 7).

We also tested the expression of the *pahT* promoter whilst growing in carbon sources that have been identified in clover exudates (Molina *et al*., personal communication). The *pahT* gene was expressed in most of the compounds tested; however, the expression pattern was different for each compound (Fig. 8). 1/10 of LB, *p*-hydroxybenzoate and malate (Fig. 8) did not induce the expression of this gene. A significantly lower expression level was observed when using a rich medium (LB) instead of minimal media with single carbon sources or diluted LB media.

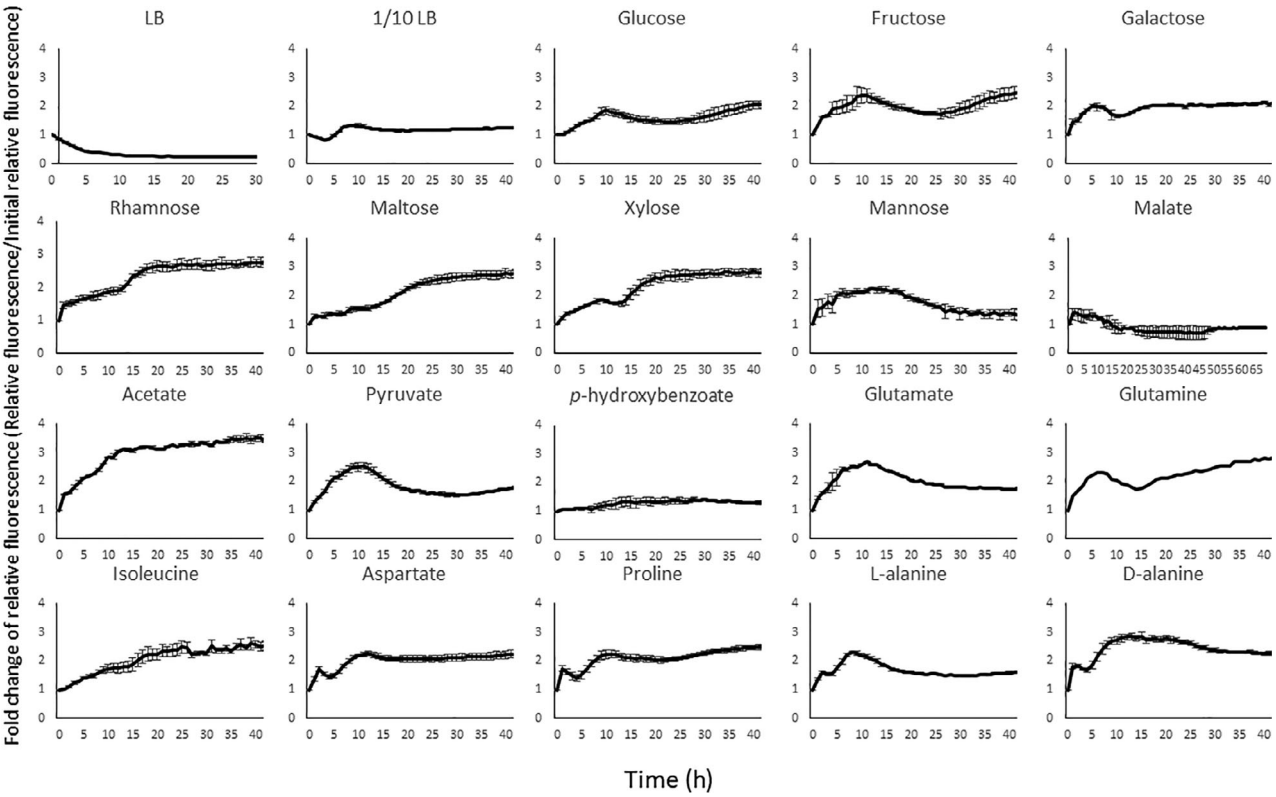


**Fig 8.** Expression from the *pahT* promoter in cultures with different carbon sources. Expression from *pahT* promoter was measured using the reporter strain *Novosphingobium* sp. HR1a (pSEVA637-Pr1998). Fluorescent measurements were weighted by the turbidity of the culture ($OD_{600nm}$) and expressed in the figure as the increase in fluorescence at each time point in comparison with the value at the beginning of the experiment.

These results indicate that gene *pahT* was expressed in the presence of different carbon sources and in the rhizosphere, suggesting that nutrient utilization in the rhizosphere is controlled by this regulatory protein.

## Discussion

*Novosphingobium* sp. HR1 is able to use, in addition to PAHs and other aromatic compounds, a high diversity of primary metabolites (a wide variety of monosaccharides, several amino acids and different organic acids) as the sole carbon source. In many microorganisms, such as *Pseudomonas* and *Enterobacteria*, the presence of easily degraded or preferred carbon sources inhibits the degradation of complex molecules (i.e. aromatic compounds) in a phenomenon called catabolic repression (Rojo, 2010). The expression pattern observed in our transcriptomic analysis in the presence of phenanthrene suggests that one of the roles of PahT in *Novosphingobium* sp. HR1a consists of allowing the co-metabolism of monosaccharides and PAHs (Table 1; Fig. 9). This effect is just the opposite phenomenon of the well-studied catabolic repression.

In the absence of phenanthrene, PahT modulates the sugar flux toward energy production by accelerating the sugar uptake and inhibiting polysaccharide depolymerization (Fig. 9). In the absence of functional PahT alternative carbon sources such as alcohols, formate, glycerol and glutamate are used for energy obtention.

Different metabolic pathways, including sugar metabolization, converge to produce pyruvate in *Novosphingobium* sp. HR1a, as in other microorganisms. PahT regulates the production of this intermediate by upregulating some of these pathways and downregulating others (Fig. 9). Some of these PahT-regulated pathways include the metabolization of PAHs and other aromatic compounds such as *p*-hydroxybenzoate and protocatechuate. Pyruvate is transformed/recycled into phosphoenolpyruvate (PEP) by a PPDK, encoded in the first putative transposon and upregulated by PahT. PEP is the motor molecule that allows the acquisition of sugars by the PTS system (Tjaden *et al*., 2006). Interestingly, this gene is not conserved in all the microbes able to degrade PAHs (Suppl. Table 1), suggesting that its role is not essential in the PAH degradation pathway, although it could be an evolutionary advantage allowing

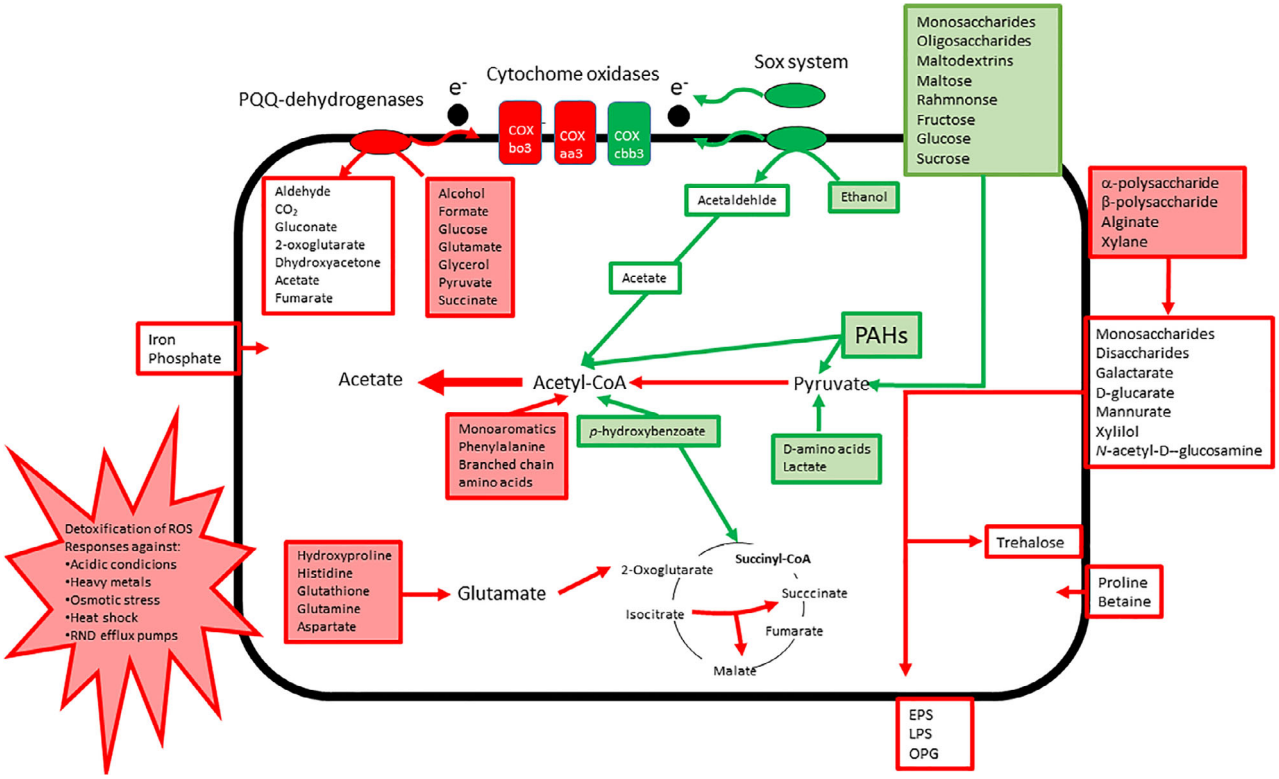


**Fig 9.** Schematic representation of the pathways in which PahT is involved. In green are depicted the pathways that are positively regulated by PahT; in red those that are negatively regulated. [Color figure can be viewed at wileyonlinelibrary.com]

an increase in sugars uptake and PAHs degradation. Encoded in the second transposon, and rarely found in PAH degrading microorganisms (Suppl. Table 1), there is a gene encoding for a putative pyruvate carboxylase. This enzyme, which is activated in presence of high amounts of acetyl-CoA (Bernson, 1976), catalyses the transformation of pyruvate into oxaloacetate (Suppl. Fig. 4) an intermediate of the tricarboxylic acid cycle. This enzyme has been involved in the gluconeogenesis process (Adina-Zada *et al*., 2012). Therefore, a functional PahT is not only necessary for high level of expression of a PTS system but also for the replenishment of the PEP from pyruvate formed during PAH degradation.

Pyruvate is metabolized not only into PEP but also into acetyl-CoA. PahT inhibits the action of the acetyl-CoA hydrolase, an enzyme involved in the conversion of this metabolite into acetic acid (Bernson, 1976), avoiding its accumulation in the cell. Furthermore, PahT activates the production of acetyl-CoA through the oxidation/fermentation of ethanol to acetic acid (Fig. 9). Two acetyl-CoA synthetases are controlled by PahT, *orf1513* and *orf1923* (encoded in the first putative degradative transposon). As with pyruvate dikinase and pyruvate carboxylase, this gene is not conserved in the gene arrangements of the microbes with similar PAH degradation genes (Suppl. Fig. 1; Suppl. Table 1).

When we analysed the genomic region of the PAH degradation genes of different strains, we observed that only *Novosphingobium* sp. PP1Y, *Sphingomonas paucimobilis* EPA505, *C. naphthovorans* PQ2 and *Novosphingobium pentaromativorans* US6-1 encode regulatory proteins with sequence identity higher than 90% with *Novosphingobium* sp. HR1a PahT, while the identity with *S. yanoikuyae* B1, *S. fuliginis* DSM 18781 and *Sphingobium* sp. MP9-4 regulatory protein was around 80% (Suppl. Table 1). All the other strains encoded regulatory proteins of the IclR family with a lower percentage of identity or did not encode regulatory protein. All the strains that encode PahT-like proteins (identity higher than 78%) also encoded the *orf1923* (acetyl-CoA synthetase) while the presence of *orf1930* (pyruvate dikinase) is not correlated with *pahT* (Suppl. Fig. 1). The GC content of the first putative transposon is slightly lower than the CG content of the genome (61.5% versus 64.9%) and whilst *orf1923* and *orf1930* have a GC content of 62% and 63% respectively, *orf1953* (pyruvate decarboxylase) and *pahT* have a GC content 59% and 58% respectively well below the GC

content of the second putative transposon (64%). This lower GC content may indicate the latter acquisition of these genes in the transposon.

The PahT control over the above-mentioned genes is independent of the presence or absence of phenanthrene, although the level of gene expression may change depending on whether these genes are induced by phenanthrene. However, there are several processes controlled by PahT that are mainly dependent on the presence of phenanthrene. These are the activation of the cytochrome $cbb_3$, of the L-lactate dehydrogenase, of iron uptake systems and alanine metabolism. Lactate dehydrogenase and alanine metabolism render pyruvate, a metabolite that is highly regulated by PahT (see above).

In addition to the redirection of the carbon fluxes, thiosulphate assimilation is another process controlled by PahT regardless of the presence of phenanthrene, although this PAH did induce the expression of the genes. Thiosulphate is not the most abundant sulfur source in the environment, but the capacity to assimilate this compound may represent a competitive advantage in certain environments. Many rhizobacteria with this capacity have been isolated and soxB distribution is abundant in the rhizosphere (Meyer *et al*., 2007; Anandham *et al*., 2008).

$cbb_3$ is a proton-pumping respiratory oxidase expressed by bacteria mainly under microaerobic conditions. This oxidase has only been identified in proteobacteria and is characterized by its high affinity toward molecular oxygen. The induction of this system in response to phenanthrene may indicate that growth on phenanthrene consumes plenty of oxygen. Cytochrome $cbb_3$ oxidase may also represent a specialized mechanism for bacterial survival in microaerobic environments; it has previously described been that certain hydrophobic compounds may destabilize bacterial membranes (Duque *et al*., 2004) and this could affect the respiratory chain activating auxiliary systems. Iron uptake could be exacerbated in the presence of phenanthrene because many of the dioxygenases involved in PAH degradation require iron as a co-factor.

*Novosphingobium* sp. HR1a was isolated from a plant rhizosphere (Rodriguez-Conde *et al*., 2016), and therefore it should be adapted to the utilization of the different carbon, nitrogen and sulfur sources that are exudated during plant development. In this complex environment the acquisition of a global regulator, such as PahT, could give a selective advantage to this bacterium. The importance of this regulator in the survival of the strain, not only in the clover rhizosphere but also during starvation in the absence of the plant, has been demonstrated here (Fig. 5).

## Experimental procedures

### *Strains, plasmids, primers and media used*

Strains, plasmids and primers used in this study are described in Suppl. Table 5.

### *Growth experiments*

The strains were grown overnight in LB medium plus the corresponding antibiotic to be used as pre-inocula for the different growth experiments. For the bioscreen experiments, 200 μl M9 minimal medium plus the corresponding carbon source (at a final concentration of 5 mM) were dispensed into wells of honeycomb microplates (OY Growth Curves AB, Raisio, Finland). The wells were inoculated with the overnight culture at an initial optical density ($DO_{660nm}$) of 0.1. The cultures were incubated at 28°C with maximal agitation in an FP-1100-C Bioscreen C MBR analyser system (OY Growth Curves AB). Growth was monitored using a type at 30°C with continuous agitation. Turbidity was measured using a sideband filter at 420–580 nm every 60 min for 42 h.

Growth with aromatic compounds was carried out in 100 ml flasks and incubated at 30°C with agitation of 2000 rpm in an orbital shaker. When benzoate or 3-methylbenzoate were used as the only carbon source, they were added at a final concentration 5 mM in 10 ml of M9 medium. Because of their toxicity, when salicylate was the carbon source, it was added at 1 mM every day during the first 3 days of the experiment; ethanol was added as liquid inside a curve crystal glass cylinder and added every day to prevent complete evaporation. When the PAHs (naphthalene, diphenyl, phenanthrene, anthracene, chrysene or pyrene) were used as the only carbon source, 0.05 mg of crystals was added in 10 ml of M9 medium.

In all cases, the bacterial strains were inoculated to reach an initial optical density at 660 nm of 0.1 and turbidity was recorded at the indicated times in a spectrophotometer.

The inoculated media without a carbon source were used as control in all the cases. Assays were run in duplicate and were repeated for at least three independent experimental rounds.

### *Construction of the mutant in* orf1998

A 340 bp internal fragment of *orf1998* was amplified using oligonucleotides 1998-F and 1998-R (Suppl. Table 5) using genomic DNA of *Novosphingobium* sp. HR1a as template. This fragment was cloned into pMBL™-T plasmid and the resulting plasmid, pMBL1998, was digested with BamHI and ligated with the Ω-Km cassette from plasmid pHP45ΩKm that were previously

digested with BamHI and ScaI and extracted from an agarose gel (Prentki and Krisch, 1984). The resulting plasmid pMBL1998Km was transformed into *Novosphingobium* sp. HR1a by electroporation. Transformants that have integrated the plasmid into the host chromosome via homologous recombination were selected on LB plus kanamycin plates and checked by Southern blot hybridization (not shown).

### *Construction of reported plasmid*

The 340 intergenic region between *orf1998* and *orf1999* of *Novosphingobium* sp. HR1a was amplified with primers incorporating restriction sites (an EcoRI site in the primer designed to meet the 5′ end and a PstI site in the primer designed to meet the 3′ end). Upon amplification, DNA was digested with EcoRI and PstI and ligated into the low-copy-number pMP220 vector (Spaink *et al*., 1987) and the medium-copy number pSEVA637 vector (Silva-Rocha *et al*., 2013), both previously cut with the same enzymes. The resulting plasmids, pMP220-Pr1998 and pSEV637-Pr1998 respectively were sequenced to verify the promoter sequence. Plasmids were individually electroporated into *Novosphingobium* sp. HR1a and/or the *pahT* mutant. Transformants were selected in LB plates plus tetracycline or gentamycin; individual colonies were grown in liquid media and plasmid was extracted and digested with EcoRI and PstI to verify the incorporation of the plasmid.

### *Growth in gnotobiotic systems*

*Novosphingobium* sp. HR1a and the *orf1998* knockout mutant strain were cultivated overnight on M9 minimal medium plus 10 mM glucose. The following day, the cultures were centrifuged and washed three times with 1× M9 minimal media, and finally diluted to an OD at 660 nm of 0.005 (approximately $10^6$ colony-forming units (CFUs) $mL^{-1}$) in 2.5 mM Fe-EDTA solution, and 20 ml of this solution was added to sterilized jars containing 50 ml of glass beads. In control jars, 20 ml of the Fe-EDTA solution, without bacteria, was added. One hundred milligrams of surface-sterilized clover seeds was placed in the corresponding jars and the samples were taken immediately after inoculation and then 3, 6 and 10 days later. The seeds germinated after 2–3 days. The samples were analysed to count the numbers of CFUs in supernatant doing serial dilutions that were plated in LB medium with the corresponding antibiotics. Assays were run in duplicate and were repeated for at least three independent experimental rounds.

To study the colonization of the mutant strain in competition with the wild type, the assays were performed as described above but both strains were inoculated at the optical density in a single jar. The numbers of total CFUs were counted by growing serial dilutions of the samples on LB plus rifampicin (10 μg $ml^{-1}$), and the numbers of the mutant strain were counted on LB plus kanamycin (50 μg $ml^{-1}$). The numbers of the wild-type strain were calculated by the difference between CFUs growing on LB plus rifampicin and CFUs growing on LB plates plus kanamycin. Assays were run in duplicate and were repeated for at least three independent experimental rounds.

### *Gene expression experiments*

To analyse the expression from the *pahAB*, *pahR* and the *orf1998* promoters, plasmids pHR1a, pPHR, pMP220-Pr1998, or pSEVS637-Pr1998 (Suppl. Table 5) were used. Depending on the reporter gene, β-galactosidase or fluorescence analyses were carried out.

*β-Galactosidase.* Selected strains carrying pMP220 derivatives (Suppl. Table 5) were grown overnight in LB medium and the following day were diluted to an initial $OD_{660nm}$ of 0.1 in M9 medium plus glucose. Inducers were added at the beginning of the experiment as follows: Salicylate was added at a final concentration of 1 mM; naphthalene, phenanthrene or pyrene were added as crystals (0.05 mg). β-Galactosidase assays (Miller, 1972) were carried out 7 h after inoculation. Assays were run in duplicate and were repeated for at least three independent experimental rounds.

*Fluorescence.* For the analysis of induction with root exudates, the gnotobiotic assays were prepared as above and inoculated with strain *Novosphingobium* sp. HR1a carrying the corresponding plasmid. Aliquots of 200 μl of the solution contained in the gnotobiotic systems (prepared as above) with and without plants were directly dispensed in the 96-well microplate (Greiner 96 black-welled plates) at the different time points. Fluorescence (excitation 485 nm, emission 520 nm) was immediately measured in a Varioskan LUX Multimode Microplate Reader.

To measure the expression pattern of the *pahT* promoter during growth with individual carbon sources (sugars, organic acids and amino acids), 200 μl of *Novosphingobium* sp. HR1a (pSEVA637-Pr1998) cultures prepared as above (overnight cultures diluted to an $OD_{660nm}$ of 0.1 in M9 minimum media plus the corresponding carbon source [5 mM]), were dispensed in the 96-well plate (Greiner 96 black-welled plates) and cultivated at 28°C with an agitation of 300 rpm in a Varioskan LUX Multimode Microplate Reader measured at the indicated times. Fluorescence (excitation 485 nm,

emission 520 nm) and turbidity at 660 nm of the cultures were measured every hour for 50 h.

Five replicas per carbon source were analysed and were repeated for at least three independent experimental rounds.

### *Statistical analysis*

Statistical differences between the different categories inside each experiment were determined using the one-way analysis of the variance, followed by a Tukey HSD Test ($p < 0.05$) Statistics Kingdom (https://www.statskingdom.com/180Anova1way.html).

### *RNA-seq experiments*

Overnight cultures grown in M9 minimal medium plus glucose of *Novosphingobium* sp. HR1a and the *pahT* mutant were diluted in 10 ml of fresh M9 minimal medium plus glucose or plus glucose and phenanthrene (0.05 mg) to an initial $OD_{660nm}$ of 0.1 and incubated at 30 °C for 7 h at 200 rpm in an orbital shaker. Five millilitres of these cultures was centrifuged (6000 rpm during 10 min) and the pellets were immediately frozen in liquid $N_2$ and stored a −80 °C.

Total RNA was extracted using the Trizol method (TRIzol RNA Isolation Reagents, Thermofisher Scientist) and further DNAse treatment and purification with RNeasy Mini Kit (Qiagen). RNA degradation and contamination were monitored on 1% agarose gels. RNA purity was checked using the NanoPhotometer® spectrophotometer (IMPLEN, CA, USA). RNA integrity and quantification were assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

One microgram RNA per sample was used to prepare the sequencing libraries which were generated by Novogen (Hong Kong) using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's recommendations and index codes were added to attribute sequences to each sample. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using PE Cluster Kit cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina platform and paired-end reads were generated.

Raw data (raw reads) of FASTQ format were first processed through in-house scripts (Novogen). In this step, clean data (clean reads) were obtained by removing reads containing adapters and reads in which uncertain nucleotides (N) were more than the 10% of the read length. Reads with low-quality nucleotides (base quality <20) in more than 50% of the read length were also discarded. Read data were deposited in the Gene Expression Omnibus repository (accession number GSE163593).

Paired-end clean reads were mapped to the reference genome (RAST Genome ID: 6666666.92368) using HISAT2 software v.0.6.1. HISAT2 uses a large set of small GFM indexes that collectively cover the whole genome. These small indexes (called local indexes), combined with several alignment strategies, enable rapid and accurate alignment of sequencing reads. A table with the gene names as in RAST and the equivalent genes in GenBank (JABXWS000000000) is included in Supplementary Table 6.

HTSeq software was used to count the read mapped to each gene, including known and novel genes.

Differential expression analysis between two conditions/groups (three biological replicates per condition) was performed using DESeq2 package v1.20.0 in R. DESeq2 provides statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting *p*-values were adjusted using the Benjamini and Hochberg's approach for controlling the False Discovery Rate. Genes with an adjusted *p*-value ($p_{adj}$) <0.05 found by DESeq2 were assigned as differentially expressed.

GO enrichment analysis of differentially expressed genes was implemented by the GOseq package v2.12 in R, in which gene length bias was corrected. GO terms with corrected $p_{adj}$ value less than 0.05 were considered significantly enriched by differential expressed genes. KOBAS software v3.0 was used to test the statistical enrichment of differential expression genes in KEGG pathways ($p_{adj}$, 0.05).

Quality of RNA reads and exploratory analysis of RNA-seq samples are shown in Supplementary Fig. 5.

## Acknowledgements

This work was supported by the European Regional Development Fund FEDER and a grant from the Spanish Ministry of Science and Innovation (Grant No. BIO2017-85994-P). We thank Angela Tate for improving the use of English in the manuscript.

## References

Adina-Zada, A., Zeczycki, T.N., and Attwood, P.V. (2012) Regulation of the structure and activity of pyruvate carboxylase by acetyl CoA. *Arch Biochem Biophys* **519**: 118–130.

Anandham, R., Indiragandhi, P., Madhaiyan, M., Ryu, K.Y., Jee, H.J., and Sa, T.M. (2008) Chemolithoautotrophic oxidation of thiosulfate and phylogenetic distribution of sulfur

oxidation gene (*soxB*) in rhizobacteria isolated from crop plants. *Res Microbiol* **159**: 579–589.

Balseiro-Romero, M., Gkorezis, P., Kidd, P.S., Vangronsveld, J., and Monterroso, C. (2016) Enhanced degradation of diesel in the rhizosphere of after inoculation with diesel-degrading and plant growth-promoting bacterial strains. *J Environ Qual* **45**: 924–932.

Bernson, V.S.M. (1976) Acetyl-CoA hydrolase; activity, regulation and physiological, significance of the enzyme in brown adipose tissue from hamster. *Eur J Biochem* **67**: 403–410.

Caplan, J.A. (1993) The worldwide bioremediation industry - prospects for profit. *Trends Biotechnol* **11**: 320–323.

Chakraborty, J., and Das, S. (2016) Characterization of the metabolic pathway and catabolic gene expression in biphenyl degrading marine bacterium *Pseudomonas aeruginosa* JP-11. *Chemosphere* **144**: 1706–1714.

Choi, D.H., Kwon, Y.M., Kwon, K.K., and Kim, S.-J. (2015) Complete genome sequence of *Novosphingobium pentaromativorans* US6-1T. *Stand Genomic Sci* **10**: 107.

D'Argenio, V., Notomista, E., Petrillo, M., Cantiello, P., Cafaro, V., Izzo, V., *et al*. (2014) Complete sequencing of *Novosphingobium* sp. PP1Y reveals a biotechnologically meaningful metabolic pattern. *BMC Genomics* **15**: 384.

Duque, E., Garcia, V., de la Torre, J., Godoy, P., Bernal, P., and Ramos, J.-L. (2004) Plasmolysis induced by toluene in a cyoB mutant of *Pseudomonas putida*. *Environ Microbiol* **6**: 1021–1031.

Ekici, S., Pawlik, G., Lohmeyer, E., Koch, H.-G., and Daldal, F. (2012) Biogenesis of *cbb3*-type cytochrome c oxidase in *Rhodobacter capsulatus*. *Biochim Biophys Acta* **1817**: 898–910.

el Zahar Haichar, F., Santaella, C., Heulin, T., and Achouak, W. (2014) Root exudates mediated interactions belowground. *Soil Biol Biochem* **77**: 69–80.

Evans, W., Fernley, H., and Griffiths, E. (1965) Oxidative metabolism of phenanthrene and anthracene by soil *Pseudomonas*. The ring-fission mechanism. *Biochem J* **95**: 819–831.

Gao, S., Seo, J.-S., Wang, J., Keum, Y.-S., Li, J., and Li, Q. X. (2013) Multiple degradation pathways of phenanthrene by *Stenotrophomonas maltophilia* C6. *Int Biodeterior Biodegrad* **79**: 98–104.

Ghosal, D., Ghosh, S., Dutta, T.K., and Ahn, Y. (2016) Current state of knowledge in microbial degradation of polycyclic aromatic hydrocarbons (PAHs): a review. *Front Microbiol* **7**: 1369.

Grabarczyk, D.B., and Berks, B.C. (2017) Intermediates in the Sox sulfur oxidation pathway are bound to a sulfane conjugate of the carrier protein SoxYZ. *PLoS One* **12**: e0173395.

Hegedűs, B., Kós, P.B., Bálint, B., Maróti, G., Gan, H.M., Perei, K., and Rákhely, G. (2017) Complete genome sequence of *Novosphingobium resinovorum* SA1, a versatile xenobiotic-degrading bacterium capable of utilizing sulfanilic acid. *J Biotechnol* **241**: 76–80.

Hopper, D.J., and Taylor, D.G. (1975) Pathways for the degradation of *m*-cresol and *p*-cresol by *Pseudomonas putida*. *J Bacteriol* **122**: 1–6.

Huang, Y., Zeng, Y., Feng, H., Wu, Y., and Xu, X. (2015) *Croceicoccus naphthovorans* sp. nov., a polycyclic aromatic hydrocarbons-degrading and acylhomoserine-lactone-producing bacterium isolated from marine biofilm, and emended description of the genus *Croceicoccus*. *Int J Syst Evol Microbiol* **65**: 1531–1536.

Iwabuchi, T., and Harayama, S. (1997) Biochemical and genetic characterization of 2-carboxybenzaldehyde dehydrogenase, an enzyme involved in phenanthrene degradation by *Nocardioides* sp. strain KP7. *J Bacteriol* **179**: 6488–6494.

Kanaly, R.A., and Harayama, S. (2010) Advances in the field of high-molecular-weight polycyclic aromatic hydrocarbon biodegradation by bacteria. *J Microbial Biotechnol* **3**: 136–164.

Kingston, R.L., Scopes, R.K., and Baker, E.N. (1996) The structure of glucose-fructose oxidoreductase from *Zymomonas mobilis*: an osmoprotective periplasmic enzyme containing non-dissociable NADP. *Structure* **4**: 1413–1428.

Kong, S., Ji, Y., Li, Z., Lu, B., and Bai, Z. (2013) Emission and profile characteristic of polycyclic aromatic hydrocarbons in PM2.5 and PM10 from stationary sources based on dilution sampling. *Atmos Environ* **77**: 155–165.

Kuiper, I., Lagendijk, E.L., Bloemberg, G.V., and Lugtenberg, B.J.J. (2004) Rhizoremediation: a beneficial plant-microbe interaction. *Mol Plant-Microbe Interact* **17**: 6–15.

Kuppusamy, S., Thavamani, P., Venkateswarlu, K., Lee, Y. B., Naidu, R., and Megharaj, M. (2017) Remediation approaches for polycyclic aromatic hydrocarbons (PAHs) contaminated soils: technological constraints, emerging trends and future directions. *Chemosphere* **168**: 944–968.

Kvint, K., Nachin, L., Diez, A., and Nyström, T. (2003) The bacterial universal stress protein: function and regulation. *Curr Opin Microbiol* **6**: 140–145.

Lu, H., Sun, J., and Zhu, L. (2017) The role of artificial root exudate components in facilitating the degradation of pyrene in soil. *Sci Rep* **7**: 7130.

Matilla, M.A., Espinosa-Urgel, M., Rodríguez-Herva, J.J., Ramos, J.L., and Ramos-González, M. (2007) Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. *Genome Biol* **8**: R179.

Matsushita, K., Yakushi, T., Takaki, Y., Toyama, H., and Adachi, O. (1995) Generation mechanism and purification of an inactive form convertible *in vivo* to the active form of quinoprotein alcohol dehydrogenase in *Gluconobacter suboxydans*. *J Bacteriol* **177**: 6552–6559.

Meyer, B., Imhoff, J.F., and Kuever, J. (2007) Molecular analysis of the distribution and phylogeny of the *soxB* gene among sulfur-oxidizing bacteria – evolution of the Sox sulfur oxidation enzyme system. *Environ Microbiol* **9**: 2957–2977.

Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Neugebauer, H., Herrmann, C., Kammer, W., Schwarz, G., Nordheim, A., and Braun, V. (2005) ExbBD-dependent transport of maltodextrins through the novel MalA protein across the outer membrane of *Caulobacter crescentus*. *J Bacteriol* **187**: 8300–8311.

Panagos, P., Van Liedekerke, M., Yigini, Y., and Montanarella, L. (2013) Contaminated sites in Europe: review of the current situation based on data collected

through a European network. *J Environ Public Health* **2013**: 158764.

Prakash, O., and Lal, R. (2006) Description of *Sphingobium fuliginis* sp. nov., a phenanthrene-degrading bacterium from a fly ash dumping site, and reclassification of *Sphingomonas cloacae* as *Sphingobium cloacae* comb. nov. *Int J Syst Evol Microbiol* **56**: 2147–2152.

Prentki, P., and Krisch, H.M. (1984) *In vitro* insertional mutagenesis with a selectable DNA fragment. *Gene* **29**: 303–313.

Rentz, J.A., Alvarez, P.J.J., and Schnoor, J.L. (2004) Repression of *Pseudomonas putida* phenanthrene-degrading activity by plant root extracts and exudates. *Environ Microbiol* **6**: 574–583.

Rodriguez-Conde, S., Molina, L., González, P., García-Puente, A., and Segura, A. (2016) Degradation of phenanthrene by *Novosphingobium* sp. HS2a improved plant growth in PAHs-contaminated environments. *Appl Microbiol Biotechnol* **100**: 10627–10636.

Rodríguez-Garrido, B., Balseiro-Romero, M., Kidd, P.S., and Monterroso, C. (2020) Effect of plant root exudates on the desorption of hexachlorocyclohexane isomers from contaminated soils. *Chemosphere* **241**: 124920.

Rojo, F. (2010) Carbon catabolite repression in *Pseudomonas*: optimizing metabolic versatility and interactions with the environment. *FEMS Microbiol Rev* **34**: 658–684.

Roy, M., Khara, P., and Dutta, T.K. (2012) Meta-cleavage of hydroxynaphthoic acids in the degradation of phenanthrene by *Sphingobium* sp. strain PNB. *Microbiology* **158**: 685–695.

Sánchez-Cañizares, C., Jorrín, B., Poole, P.S., and Tkacz, A. (2017) Understanding the holobiont: the interdependence of plants and their microbiome. *Curr Opin Microbiol* **38**: 188–196.

Sandhu, M., Jha, P., Paul, A.T., Singh, R.P., and Jha, P.N. (2020) Evaluation of biphenyl- and polychlorinated-biphenyl (PCB) degrading *Rhodococcus* sp. MAPN-1 on growth of *Morus alba* by pot study. *Int J Phytoremediation* **22**: 1487–1496.

Sasse, J., Martinoia, E., and Northen, T. (2018) Feed your friends: do plant exudates shape the root microbiome? *Trends Plant Sci* **23**: 25–41.

Schell, M.A. (1983) Cloning and expression in *Escherichia coli* of the naphthalene degradation genes from plasmid NAH7. *J Bacteriol* **153**: 822–829.

Segura, A., Hernández-Sánchez, V., Marqués, S., and Molina, L. (2017) Insights in the regulation of the degradation of PAHs in *Novosphingobium* sp. HR1a and utilization of this regulatory system as a tool for the detection of PAHs. *Sci Total Environ* **590–591**: 381–393.

Segura, A., and Ramos, J.L. (2013) Plant–bacteria interactions in the removal of pollutants. *Curr Opin Biotechnol* **24**: 467–473.

Sha, S., Zhong, J., Chen, B., Lin, L., and Luan, T. (2017) *Novosphingobium guangzhouense* sp. nov., with the ability to degrade 1-methylphenanthrene. *Int J Syst Evol Microbiol* **67**: 489–497.

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., De las Heras, A., *et al*. (2013) The standard European vector architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* **41**: D666–D675.

Simmer, R., Mathieu, J., da Silva, M.L.B., Lashmit, P., Gopishetty, S., Alvarez, P.J.J., and Schnoor, J.L. (2020) Bioaugmenting the poplar rhizosphere to enhance treatment of 1,4-dioxane. *Sci Total Environ* **744**: 140823.

Spaink, H.P., Okker, R.J.H., Wijffelman, C.A., Pees, E., and Lugtenberg, B.J.J. (1987) Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. *Plant Mol Biol* **9**: 27–39.

Stolz, A. (2014) Degradative plasmids from sphingomonads. *FEMS Microbiol Lett* **350**: 9–19.

Story, S.P., Kline, E.L., Hughes, T.A., Riley, M.B., and Hayasaka, S.S. (2004) Degradation of aromatic hydrocarbons by *Sphingomonas paucimobilis* strain EPA505. *Arch Environ Contam Toxicol* **47**: 168–176.

Tchieu, J.H., Norris, V., Edwards, J.S., and Saier, M.H. (2001) The complete phosphotransferase system in *Escherichia coli*. *J Mol Microbiol Biotechnol* **3**: 329–346.

Terzaghi, E., Vergani, L., Mapelli, F., Borin, S., Raspa, G., Zanardini, E., *et al*. (2019) Rhizoremediation of weathered PCBs in a heavily contaminated agricultural soil: results of a biostimulation trial in semi field conditions. *Sci Total Environ* **686**: 484–496.

Tjaden, B., Plagens, A., Dorr, C., Siebers, B., and Hensel, R. (2006) Phosphoenolpyruvate synthetase and pyruvate, phosphate dikinase of *Thermoproteus tenax*: key pieces in the puzzle of archaeal carbohydrate metabolism. *Mol Microbiol* **60**: 287–298.

Toyama, H., Fujii, A., Matsushita, K., Shinagawa, E., Ameyama, M., and Adachi, O. (1995) Three distinct quinoprotein alcohol dehydrogenases are expressed when *Pseudomonas putida* is grown on different alcohols. *J Bacteriol* **177**: 2442–2450.

Vandenkoornhuyse, P., Mahe, S., Ineson, P., Staddon, P., Ostle, N., Cliquet, J.-B., *et al*. (2007) Active root-inhabiting microbes identified by rapid incorporation of plant-derived carbon into RNA. *Proc Natl Acad Sci U S A* **104**: 16970–16975.

Vila, J., Tauler, M., and Grifoll, M. (2015) Bacterial PAH degradation in marine and terrestrial habitats. *Curr Opin Biotechnol* **33**: 95–102.

Wolf, D.C., Cryder, Z., Khoury, R., Carlan, C., and Gan, J. (2020) Bioremediation of PAH-contaminated shooting range soil using integrated approaches. *Sci Total Environ* **726**: 138440.

Yakushi, T., and Matsushita, K. (2010) Alcohol dehydrogenase of acetic acid bacteria: structure, mode of action, and applications in biotechnology. *Appl Microbiol Biotechnol* **86**: 1257–1265.

Zafra, G., Absalón, Á.E., Anducho-Reyes, M.Á., Fernandez, F.J., and Cortés-Espinosa, D.V. (2017) Construction of PAH-degrading mixed microbial consortia by induced selection in soil. *Chemosphere* **172**: 120–126.

Zhao, Q., Hu, H., Wang, W., Peng, H., and Zhang, X. (2015) Genome sequence of *Sphingobium yanoikuyae* B1, a polycyclic aromatic hydrocarbon-degrading strain. *Genome Announc* **3**: e01522-14.

Zhao, X., Miao, R., Guo, M., and Zhou, Y. (2021) Effects of fire Phoenix (a genotype mixture of *Fesctuca arundinecea* L.) and *Mycobacterium* sp. on the degradation of PAHs

and bacterial community in soil. *Environ Sci Pollut Res*. https://doi.org/10.1007/s11356-021-12432-9.

Zhong, J., Luo, L., Chen, B., Sha, S., Qing, Q., Tam, N.F.Y., *et al*. (2017) Degradation pathways of 1-methylphenanthrene in bacterial *Sphingobium* sp. MP9-4 isolated from petroleum-contaminated soil. *Mar Pollut Bull* **114**: 926–933.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Suppl. Fig. 1.** Schematic representation of the distribution of the different genetic 'components' of the PAH degradative transposons of *Novosphingobium* sp. HR1a in other microorganisms. In this scheme, the arrows do not represent transcriptional units and therefore the direction of the arrows is not indicative of the transcription orientation. The arrows represent different and mean genetic 'modules' related with the genes *pahA* (yellow), *pahR* (pink), *pahT* (red), *orf1937* (light orange) and *orf1941* (dark orange). In blue and double arrow transposition elements are indicated. In light green insertion of genes related with lipid metabolism and in dark green insertion of clusters of dioxygenases genes are indicated. In black other type of gene insertions is indicated. Arrows filled with oblique lines indicate that the module is not complete but at least *orf1937* (light orange lines) or *orf1941* (dark orange lines) are included. In *Novosphingobium* sp. PP1Y two set of genes have been reported, one in the chromosome and the second one in a megaplasmid

**Suppl. Fig.2.** Genomic phylogenetic three of *Novosphingobium* strains. Genomic sequences of *Novosphingobium acidiphilum* DSM 19966 (SAMN02440878), *N aromaticivorans* DSM 12444 (SAMN02598432), *N. barchaimii* LL02 (SAMN02727999), *N. capsulatum* NBRC 12533 (SAMD00046742), *N. fuchskuhlense* FNE08-7 (SAMN04193360), *N. ginsenosidimutans* FW-6 (SAMN12419120), *N. guangzhouense* SA925 (SAMN05004390), *N. kunmingense* 18-11HK (SAMN06296206), *N. lentum* NBRC 107847 (SAMD00046709), *N. lindaniclasticum* LE124 (SAMN02471710), *N. malaysiense* MUSC 273 (SAMN03070119), *N. mathurense* SM117 (SAMN06295987), *N. naphthalenivorans* NBRC 102051 (SAMD00046703), *N. nitrogenifigens* DSM 19370 (SAMN02470891), *N. panipatense* SM16 (SAMN06296065), *N. pentaromativorans* US6-1 (SAMN03002180), *N. resinovorum* KF1 (SAMN02676962), *N. rosa* NBRC 15208 (SAMD00046712), *N. sediminis* NBRC 106119 (SAMD00170761), *Novosphingobium* sp. 12–64-8, *Novosphingobium* sp. TW-4 (SAMN14642910), *Novosphingobium* sp. NDB2Meth1 (SAMEA4535142), *Novosphingobium* sp. 9 U (SAMEA6080506), *Novosphingobium* sp. 12–62-10 (SAMN06622390). *Novosphingobium* sp. 12–63-9 (SAMN06622392), *Novosphingobium* sp. 12–64-8 (SAMN06622395), *Novosphingobium* sp. 16–62-11 (SAMN06622287), *Novosphingobium* sp. 17–62-8 (SAMN06622252), *Novosphingobium* sp. 17–62-9 (SAMN06622248), *Novosphingobium* sp. 17–62-19 (SAMN06622246), *Novosphingobium* sp. 28–62-57 (SAMN06622292), *Novosphingobium* sp. 32–60-15 (SAMN06622345), *Novosphingobium* sp. 35–62-5 (SAMN06622330), *Novosphingobium* sp. 63–713 (SAMN05660613), *Novosphingobium* sp. 18,050 (SAMEA6372283), *Novosphingobium* sp. AAP1 (SAMN02925435), *Novosphingobium* sp. AAP83 (SAMN02927142), *Novosphingobium* sp. AAP93 (SAMN02927144), *Novosphingobium* sp. AP12 (SAMN00789124), *Novosphingobium* sp. B 225 (SAMN06335524), *Novosphingobium* sp. B1 (SAMN06272759), *Novosphingobium* sp. B-7 (SAMN02469431), *Novosphingobium* sp. B3058 49 (SAMN09081298), *Novosphingobium* sp. BW1 (SAMN12161804), *Novosphingobium* sp. CCH12-A3 (SAMN04299424), *Novosphingobium* sp. CF614 (SAMN05518801), *Novosphingobium* sp. Chol11 (SAMEA104233113), *Novosphingobium* sp. ERN07 (SAMN14589664), *Novosphingobium* sp. ERW19 (SAMN14589663), *Novosphingobium* sp. FGD1 (SAMN13688776), *Novosphingobium* sp. FSY-8 (SAMN13781904), *Novosphingobium* sp. Fuku2-ISO-50 (SAMN04193359), *Novosphingobium* sp. Gsoil 351 (SAMN13294205), *Novosphingobium* sp. GV002 (SAMN08779352), *Novosphingobium* sp. HII-3 (SAMN08381663), *Novosphingobium* sp. KN65.2 (SAMN02696960), *Novosphingobium* sp. LASN5T (SAMN10475352), *Novosphingobium* sp. Leaf2 (SAMN04151573), *Novosphingobium* sp. M24A2M (SAMN13441224), *Novosphingobium* sp. MBES04 (SAMD00019870), *Novosphingobium* sp. MD-1 (SAMD00027653), *Novosphingobium* sp. NDB2Meth1 (SAMEA4535142), *Novosphingobium* sp. P6W (SAMN03323925), *Novosphingobium* sp. PASSN1 (SAMN07280675), *Novosphingobium* sp. PC22D (SAMN06444841), *Novosphingobium* sp. PhB55 (SAMN10361093), *Novosphingobium* sp. PhB165 (SAMN10361092), *Novosphingobium* sp. PP1Y (SAMEA2272572), *Novosphingobium* sp. Rr 2–17 (SAMN02470867), *Novosphingobium* sp. SCN 63–17 (SAMN03652505), *Novosphingobium* sp. SG707 (SAMN12024102), *Novosphingobium* sp. ST904 (SAMN10361598), *Novosphingobium* sp. SYSU G00007 (SAMN09222672), *Novosphingobium* sp. TCA1 (SAMD00197004), *Novosphingobium* sp. TH158 (SAMN08272654), *Novosphingobium* sp. THN1 (SAMN08741636), *Novosphingobium* sp. TW-4 (SAMN14642910), *Novosphingobium* sp. UBA1939 (SAMN06455794), *N. stygium* ATCC 700280 (SAMN05660666), *N. subterraneum* NBRC 16086 (SAMD00046759), *N. taihuense* DSM 17507 (SAMN13173503), and *N. tardaugens* NBRC 16725 (SAMD00041832) (BioProject: PRJNA224116) were introduced in the free online pipeline Reference sequence Alignment based Phylogeny builder (REALPHY, https://realphy.unibas.ch/realphy/) that can infer phylogenetic trees from whole genome sequence data. These sequences were mapped to each of the references genomes via bowtie2. From these alignments multiple sequence alignments will be reconstructed from which phylogenetic trees are inferred via PhyML.

Bertels F, Silander OK, Pachkov M, Rainey PB, van Nimwegen E. Automated reconstruction of whole-genome phylogenies from short-sequence reads. Mol Biol Evol. 2014 May;31(5):1077–88. doi: 10.1093/molbev/msu088.

**Suppl. Fig 3.** Differential expression of the different gen 'operons' conforming the PAH degradative transposons. Mutant in *pahT* versus WT in minimal medium containing glucose as the only carbon source (A) or glucose plus phenanthrene as carbon sources (B) or in the case of WT growing in glucose plus phenanthrene versus the same strain growing glucose as the only carbon source (C). Red indicated overexpressed genes, green downregulated. The fold change of the expression is indicated in the figure.

**Suppl. Fig 4.** Schematic representation of the main alterations of the central carbon metabolic fluxes 'regulated'

by PahT. In red are indicated the pathways upregulated by this regulator. The indicated genes are present in the degradative transposons but their function is related with central carbon metabolic flux and not with degradative functions.

**Suppl. Fig. 5.** Quality of RNA data and exploratory analysis of RNA-seq samples A) Plot of dispersion estimates for each gene B) Principal component plot (PCA) of expressed genes within three biological replicates of each of the four samples: *Novosphingobium* sp. HR1a grown in glucose (w_glucose), grown in glucose plus phenanthrene (w_phenanthrene) and the isogenic *pahT* mutant grown in glucose (m_glucose) and in glucose plus phenanthrene (m_phenanthrene). C) Dendrogram and heatmap of distances between the four samples (three biological replicates).

**Suppl. Table 1.** Presence, absence and identity percentage of the proteins encoded by the genes of the degradative transposons of *Novosphingobium* sp. HR1a compared with other microorganisms. Numbers indicate the degree of identity, in dark red the identities over 90%, in light red the identities between 80%–89%.

**Suppl. Table 2.** List of genes differentially expressed in *Novosphingobium* sp. HR1 in the presence *vs* the absence of phenanthrene. Only protein encoding genes (peg.) that showed a log2fold change higher or equal to 1.5 (up-regulated) or lower or equal than 1.5 (down-regulated) are shown.

**Suppl. Table 3.** List of genes whose expression levels changed when compared the *phaT* mutant vs the wild-type strain grown in glucose plus phenanthrene. Only genes in which the log2fold-change value was higher than 1.5 or lower than −1.5 are shown.

**Suppl. Table 4.** List of genes whose expression levels changed when compared the *phaT* mutant vs wild-type strain grown in glucose. Only genes in which the log2fold-change value was higher than 1.5 or lower than −1.5 are shown.

**Suppl. Table 5.** List of strains and plasmids used in this study and their principal characteristics.

**Suppl. Table 6.**

**Supplementary Material Table 2.**

**Supplementary Material Table 3.**

**Supplementary Material Table 4.**