3 Biotech (2020) 10:140
https://doi.org/10.1007/s13205-020-2133-6

ORIGINAL ARTICLE

Genes and novel sRNAs involved in PAHs degradation in marine bacteria *Rhodococcus* sp. P14 revealed by the genome and transcriptome analysis

Tao Peng[1] · Jie Kan[1] · Jing Hu[2] · Zhong Hu[1]

Received: 23 October 2019 / Accepted: 11 February 2020 / Published online: 26 February 2020

Abstract
Rhodococcus sp. P14 is able to degrade various polycyclic aromatic hydrocarbons (PAHs). In this study, 6 ring-hydroxylating dioxygenases and 24 monooxygenases genes related to PAHs degradation were identified in its genome. Moreover, various genes, like serine hydrolase, hydratase, alcohol dehydrogenase, protocatechuate 3,4-dioxygenase, β-ketoadipate CoA transferase and β-Ketoadipyl CoA thiolase, which were supposed to be involved in PAHs degradation were also identified. Based on the genome analysis, the proposed PAHs degradation pathway was constructed in P14 strain, which showed that PAHs was degraded into the acetyl CoA and succinyl CoA, then mineralized to CO_2 via the TCA cycle. Furthermore, several genes, including cytochrome P450 (RS16725; RS16695; RS12220), catalase (RS15825), dehydrogenase (RS15755; RS18420) and hydrolase (RS16460; RS24665), showed increased expression level during PAHs degradation according to the transcriptome data. In addition, 12 novel sRNAs which were supposed to have the regulation function in PAHs degradation were identified. This study gives us the outlook of PAHs degradation pathway in *Rhodococcus* sp. P14. Moreover, it first demonstrates that sRNAs may harbor the regulation function in PAHs degradation.

Keywords *Rhodococcus* · PAHs · Biodegradation · sRNA · Transcriptome

Introduction

Polycyclic aromatic hydrocarbons (PAHs) are organic pollutes which are widespread in the environment over the world. They are toxic and can be bioaccumulation in the organism (Mandalakis et al. 2005), therefore, they are harmful to human. Although adsorption, volatilization, photolysis and chemical degradation are able to remove PAHs from the contaminated environment, biodegradation of PAHs is considered as the major pathway for bioremediation, due to its environment friendly and cost less. Biodegradation can breakdown PAHs to produce CO_2 on aerobic or anaerobic condition (Aitken et al. 1998; Moody et al. 2001). Many bacteria, which have the PAHs degradation ability, have been isolated, including *Pseudomonas*, *Sphingomonas*, *Rhodococcus*, *Bacillus*, *Flavobacterium*, *Corynebacterium*, *Aeromonas*, *Mycobacterium*, *Nocardia*, *VIbrio*, *Micrococcus*, *Cyanobacteria* and *Beijernckia* (Aitken et al. 1998).

In bacteria, PAHs degradation pathway are divided into three main steps, including RCP (ring cleavage process), SCP (side chain process) and CAP (central aromatic process) before the metabolites are degraded by TCA cycle to produce CO_2 and H_2O (Kweon et al. 2007). In RCP, the degradation of PAHs is initialed with aromatic ring hydroxylation. In this step, One or two atoms from dioxygen are introduced into PAHs, which is catalyzed by two kinds of oxygenases, cytochrome P450 monooxygenases (CYPs) and ring-hydroxylating dioxygenases (RHDs). Next, the dihydrodiol is transformed to produce ortho or meta dehydroxylated, which is performed by dihydrodiol dehydrogenase (Kweon et al. 2007). The carbon–carbon bonds of the aromatic ring were broken by the ring-cleavage dioxygenase which can introduce the molecular oxygen. Afterwards, the aromatic

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13205-020-2133-6) contains supplementary material, which is available to authorized users.

✉ Zhong Hu
hzh@stu.edu.cn

[1] Department of Biology, Shantou University, Shantou 515063, Guangdong, China

[2] Affiliated Hospital 1, College of Medical, Shantou University, Guangdong 515063, China

compound was degraded by series reactions to one aromatic intermediate (SCP), which was degraded through CAP later. Various enzymes are involved in SCP and CAP, such as protocatechuate 3,4-dioxygenase which can introduce two atoms of oxygen to protocatechuate (Wolgel et al. 1993).

The *Rhodococcus* genus are widely distributed in soil, water and marine sediments. *Rhodococcus* genus are resistant to various environmental stress, due to the catabolic diversity, therefore, they have potential applications in biocatalysts, biotransformation and bioremediation (Aitken et al. 1998; Martínková et al. 2009). *Rhodococcus* sp. P14 was isolated from the marine sediment which was contaminated with crude oil (Song et al. 2011). It can degrade various PAHs, especially high molecule weight PAHs on aerobic condition, which suggests that it has the potential application for bioremediation. For further research, the genome of P14 was sequenced and submitted to GenBank database (Zhang et al. 2012). In the genome, a cytochrome P450 (CYP450) was characterized in the oxidization of various PAHs (Luo et al. 2016). Although some genes related to PAHs degradation were identified in the genome of P14, PAHs degradation pathway in the strain P14 is still unclear. Recently, many studies demonstrated that sRNAs harbored regulation functions in various biological processes, such as stress responses, secretion, biofilm formation, quorum sensing and virulence in bacteria (Mann et al. 2012; Schmidtke et al. 2013). However, there is no study showed that sRNAs had the regulation function in PAHs degradation.

In this study, the genes which were supposed to be involved in PAHs degradation were identified and the possible PAHs degradation pathway was constructed in P14 strain. Moreover, this study attempted to identify the sRNAs which may have the regulation function in PAHs degradation. This study provides a blueprint and instruction for the application of P14 on PAHs bioremediation in the environments.

Materials and methods

Genome sequencing and analysis

The complete genome of *Rhodococcus* sp. P14 was sequenced by the high-throughput Solexa technology (Illumina GA2x) as described before (Zhang et al. 2012). The whole-genome sequence of the strain P14 was deposited in the GenBank database with accession number NZ_CP024315.

Alignments and phylogenetic analyses

The pairwise and multiple alignments were performed by CLUSTALX version 1.83. All parameters were settled as default values. The phylogenetic tree was constructed by the Neighbor-Joining method and visualized with TREE-VIEW later. The reliability of the tree was evaluated by 1,000 bootstrap replications.

RNA isolation

The *Rhodococcus* sp. P14 cells were grown in 2216E liquid medium (Song et al. 2011) at 25 °C. The cells were harvested by centrifugation at 8000 g when the OD_{600} reached at 0.4.and washed by MSM medium three times. The cells were separated into two parts. The first part was cultured with pyrene as the only carbon source. The second part was cultured with glucose as the single carbon source (control group). After 4 h, the cells were collected by centrifugation at 8000 g. Hot phenol method was performed to isolate the total RNA (Peng et al. 2018a, b) and gel electrophoresis was used to check the quality of RNAs.

RNA sequencing

After RNA isolation, RNA sequencing was performed as mentioned earlier (Peng et al. 2018b). In short, the rRNAs were removed by the Illumina TruSeq Stranded Kit (Epicentre, Madison, WI, USA) according to the manufacturer's instructions. Afterwards, the random primers PCR were used to construct the cDNA library. The Illumina sequencing platform (Illumina HiSeq X10) was used to construct the cDNA library. The clean reads were obtained by removing the low-quality sequences reads with more than 5% N bases (unknown bases) and adaptor sequences. Afterwards, the SOAP2 was used for mapping the clean reads to the reference genome (accession NO. CP024315) (Li et al. 2009). The transcriptome data were deposited in the NCBI Sequence Read Archive (SRA) as accession number (PRJNA558057).

Bioinformatics analysis

The new transcripts were analyzed by Coding Potential Calculator in the antisense to mRNA (AM) and intergenic region (IGR) (Kong et al. 2007). The new transcript was considered as sRNAs if no coding sequence was identified. The sRNAs were annotated according to BSRD (Li et al. 2013) and Rfam database (Nawrocki et al. 2015) using infernal (Nawrocki and Eddy 2013). RNAfold was used to predict the secondary structures of sRNAs (Denman 1993). RSEM (Li and Dewey 2011) was applied to calculate the expression of genes and sRNAs. The interactions between sRNAs and mRNAs were predicted by RNAplex (Tafer and Hofacker 2008).

Results

Genome analysis related to PAH degradation

Ring cleavage process (RCP)

The P14 strain can use PAHs for only energy source for growth, which suggests that it has a complete PAHs degradation pathway. To find the genes involved in PAHs degradation, the genome analysis was performed. For the first and important process RCP, both CYPs and RHDs, which can initial the degradation, have been found in *Rhodococcus* sp. P14. 6 ring-hydroxylating dioxygenases (RS10945; RS12875; RS14045; RS17650; RS20575; RS24215) have been identified in the genome. Moreover, A total of 24 genes encoding CYP (RS02600; RS04390; RS04500; RS10880; RS10920; RS12220; RS14035; RS14215; RS14335; RS14340; RS14365; RS16695; RS16705; RS16725; RS16820; RS17565; RS19350; RS19735; RS19800; RS19865, RS20975; RS20985; RS20995; RS23175) were predicted by analyzing the genome sequence. The CYP proteins were named according to the CYP450 homepage (https://drnelson.uthsc.edu/CytochromeP450.html) (Broderick 1999). The phylogenetic tree for the CYP proteins from strain P14 was constructed as shown in Fig. 1. The result showed that CYP proteins from P14 strain were widely distributed in the phylogenetic tree.

After ring hydroxylation, dihydrodiol dehydrogenase was responsible for generating ortho or meta dehydroxylated. Only one paralog (RS24230) to dihydrodiol dehydrogenase was identified by analyzing the genome sequence. In

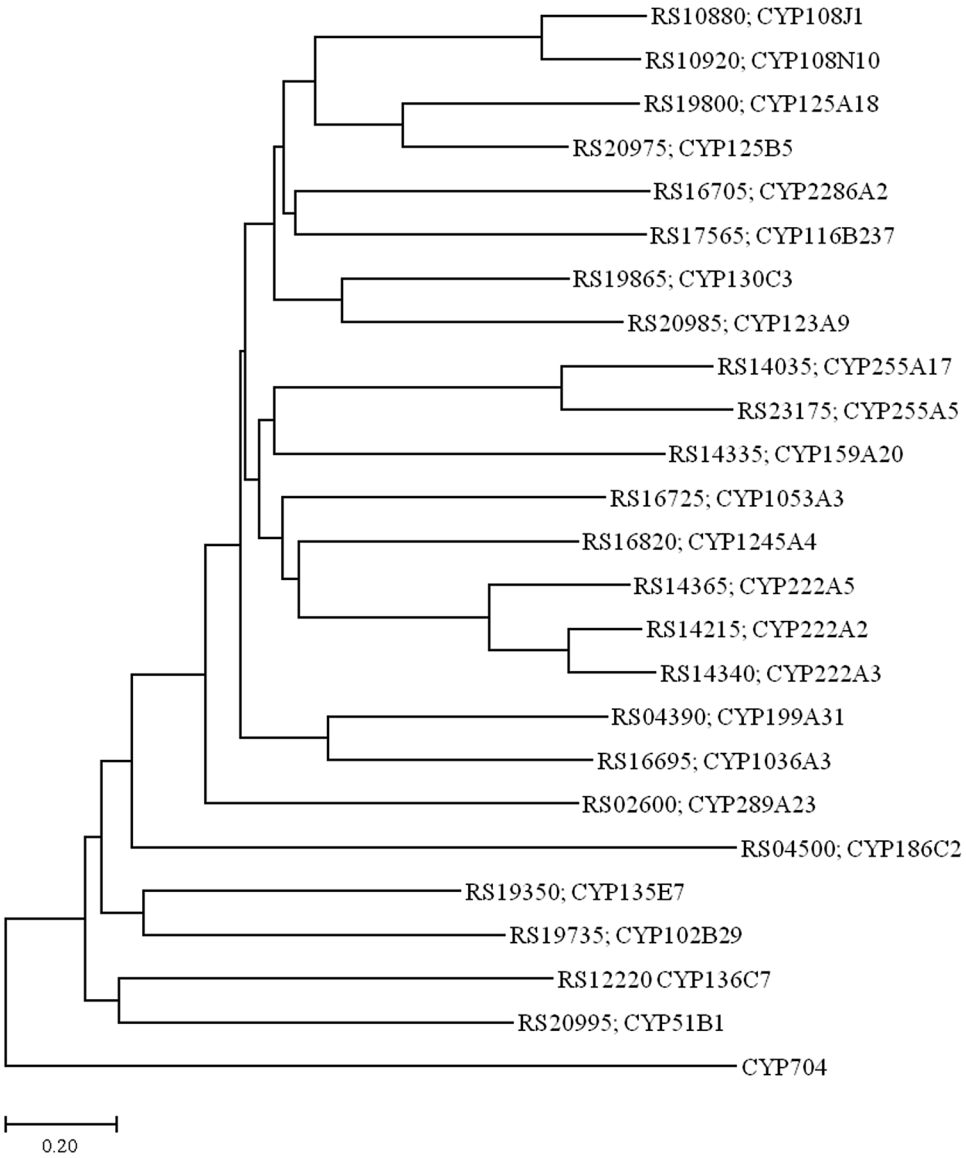

Fig. 1 Phylogenetic tree of 24 CYP proteins from *Rhodococcus* sp. P14. The names of these proteins were shown according to the CYP450 homepage (https://drnelson.uthsc.edu/CytochromeP450.html). These proteins were widely distributed according to the phylogenetic tree

the case of ring-cleavage dioxygenase, which is responsible for breaking carbon–carbon bonds of the aromatic ring by addition of molecular oxygen, the genome appeared to encode 2 paralogs of ring-cleavage dioxygenase (RS10975; RS13920).

Side-chain process (SCP) and central aromatic process (CAP)

After the ring was cleaved, the side chain process was performed. Several studies showed that many aromatic-degrading microorganisms were able to transform various aromatic compounds to a common aromatic intermediate like catechol, protocatechuate or gentisic acid (Kweon et al. 2007; Kim et al. 2006). By analyzing the genome, several genes which encode for the enzymes, serine hydrolase (RS02720; RS05540; RS20255; RS22810), hydratase (RS11875; RS19685; RS23150; RS24245), aldolase (RS11865; RS19595; RS23220; RS24235), alcohol dehydrogenase (RS04465; RS05545; RS07220; RS07915; RS08135; RS09255; RS10890; RS10940; RS11195; RS13670; RS14240; RS15785; RS15835; RS21075; RS23115; RS24090), aldehyde dehydrogenase (RS11870; RS19590; RS23225; RS24240), carboxymuconolactone decarboxylase (RS00675; RS02700; RS07310; RS09325; RS13825; RS14925; RS20950; RS21225; RS22800) were identified. They were responsible for a series of reactions to transfer ring-cleavage metabolites to protocatechuate. After that, protocatechuate was degraded through CAP. Protocatechuate 3,4-dioxygenase plays an important role in this process (Wolgel et al. 1993). In P14, 2 prologs to protocatechuate 3,4-dioxygenase (RS01070; RS13845) were identified. Moreover, β-Carboxy-cis,cis-muconate cycloisomeras (RS13850). γ-Carboxymuconolactone decarboxylase (RS22800) and β-ketoadipate CoA transferase (RS13865; RS24985; RS17305) were responsible for generating β-ketoadipyl CoA, then β-ketoadipyl CoA was transformed to succinyl CoA and acetyl CoA by β-Ketoadipyl CoA thiolase. Only one probable paralog (RS00030) for this enzyme was predicted. Based on the genome analysis, the degradation pathway for PAHs was constructed as shown in Fig. 2.

Tricaboxylic acid cycle (TCA)

After CAP, the acetyl CoA and succinyl CoA were introduced into TCA to generate CO_2 and ATP for bacteria growth. The genome P14 was analyzed to find the genes which encode enzymes related to TCA cycle. Based on the KEGG database, a complete step for the conversion of citrate to oxaloacetate was constructed (Kanehisa et al. 2006) as shown in Fig. 3, which showed a total of 18 genes possibly ware in relation to the TCA cycle.

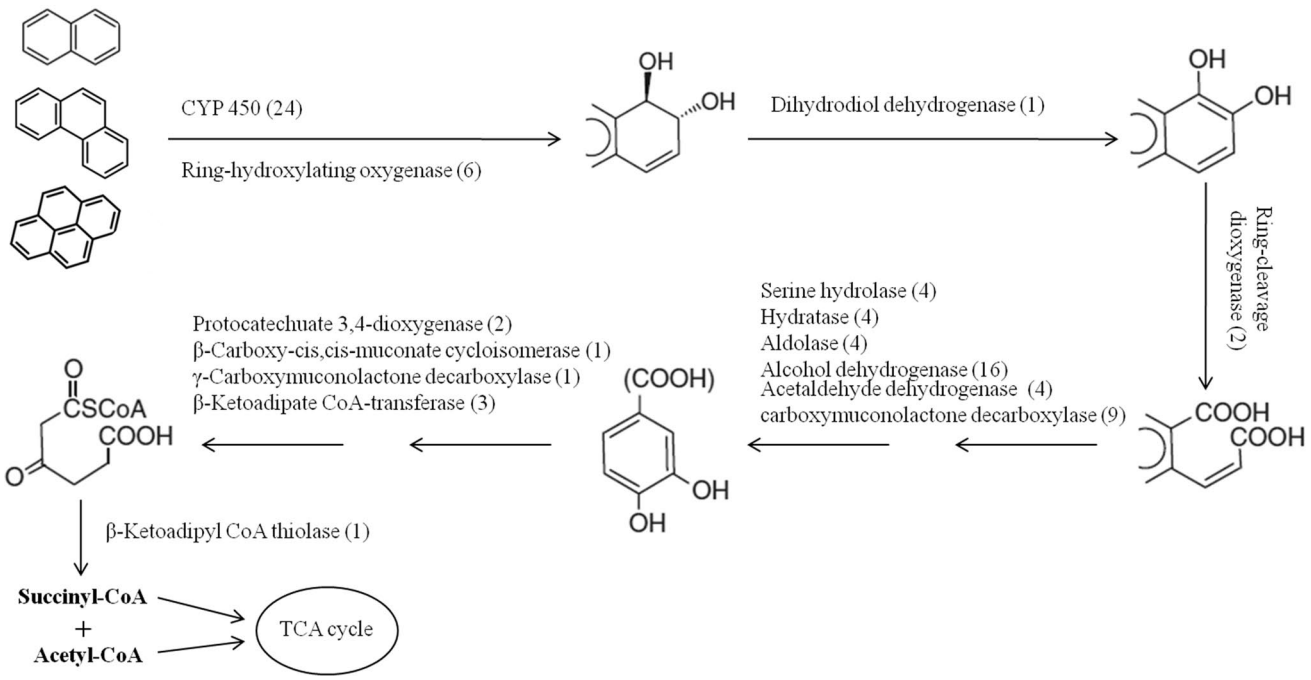

Fig. 2 Overview of PAHs degradation pathway in P14. One solid arrow indicates one-step reaction and two solid arrows indicates two or more steps. Enzyme names were showed with the number of genes identified in the P14 genome. The PAHs were transformed to and acetyl CoA and succinyl CoA, then mineralized to CO_2 via the TCA cycle

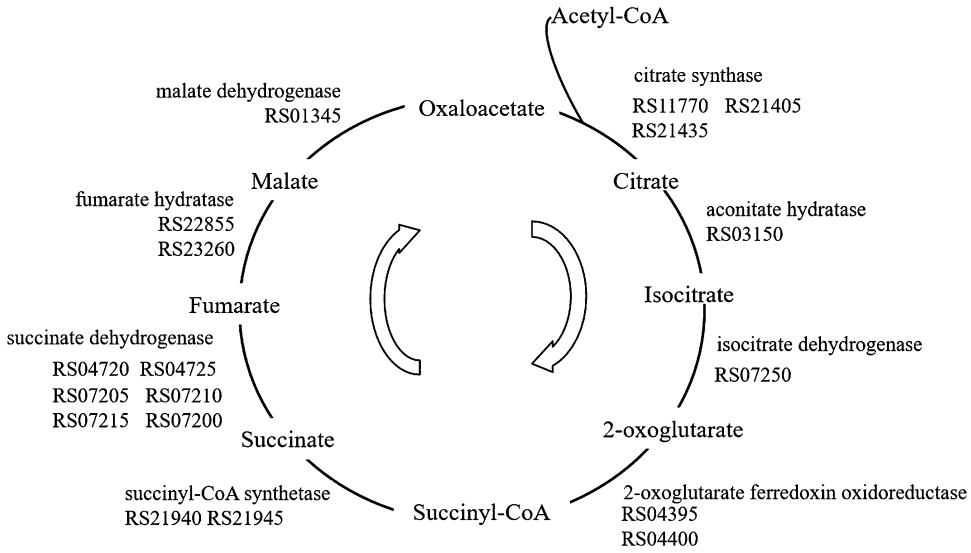

Fig. 3 TCA cycle of *Rhodococcus* sp. P14 as predicted from genome sequence. Enzyme names were showed with genes' number identified in the P14 genome. Total of 18 genes involved in TCA cycle were identified

Genes involved in detoxification during PAHs degradation

PAHs oxidation can produce a number of reactive oxygen species (ROS). In PYR-1, a detoxification function of catalase peroxidase has been revealed (Wang et al. 2000). In the strain P14, several paralogs to catalase peroxidase (RS09880; RS15825; RS15990; RS18725) and superoxide dismutase (RS09540; RS17200; RS18160) were also identified in the genome.

Genes and sRNAs involved in PAHs degradation analyzed by transcriptome data

To confirm the genes which are responsible for PAHs degradation in the cell, the RNA sequencing was performed during PAHs degradation in P14 strain. The results showed that the mRNA level of 101 genes were up-regulated during PAHs degradation ($\log2 \geq 0.6$-fold change and $P \leq 0.05$) (Table 1). Among them, several genes, including cytochrome P450 (RS16725; RS16695; RS12220), catalase (RS15825), dehydrogenase (RS15755; RS18420) and hydrolase (RS16460; RS24665) which were identified based on the genome analysis, showed increased expression level.

Based on the transcriptome data, 75 novel sRNAs were identified. Among them, 12 sRNAs showed a variation on expression level during PAHs degradation ($P \leq 0.05$ and $\log2 \geq 0.6$ or ≤ -0.6 fold change) (Table 2). Their sequences were listed in the supplement data. The trans-encoded sRNAs can regulate gene's expression by directly binding to target mRNAs. In this study, we used RNAplex to predict the target genes which may be regulated by these sRNAs (Table 3). Interestingly, these target genes also showed the change in expression level during PAHs degradation according to the transcriptome data, which implies that these genes may be regulated by the sRNAs during PAHs degradation. These results indicate that sRNAs may have regulation function in the PAHs degradation in P14.

Discussion

Rhodococcus sp. P14 is able to utilize numerous PAH compounds, especially high molecular weight PAHs, as its sole carbon sources. Several hydroxylation metabolites were detected in strain P14 which indicates that both CYPs and RHDs are possible for the hydroxylation of PAHs in P14 (Luo et al. 2016). Serval RHD genes have been characterized in *Rhodococcus* genus. One RHD NarA showed the ability in the degradation of a wide range of aromatic hydrocarbons, including the PAHs (Kimura et al. 2006). In P14 strain, one of the ring-hydroxylating dioxygenases identified as BaaA (RS14054) was proved to have the oxidization activity on anthracene and benz[a]anthracene, resulting in 9,10-dihydroxyanthracene and 7,12-dihydroxybenz[a]anthracene (Peng et al. 2018a). Ring-hydroxylating dioxygenase is a multi-component enzyme system containing a terminal oxygenase and an electron transfer component (Gibson and Parales 2000; Mason and Cammack 1992). Several electron transfer components, such as NAD(P)-dependent oxidoreductase (RSP10955) and ferredoxin reductase (RS14055) were identified near to oxygenase by genome analysis. There were 24 CYP proteins identified and widely distributed according to the phylogenetic tree, which demonstrates a remarkable degree of sequence diversity. Among them, the function of CYP108J1 was studied and it showed various PAHs oxidization activity including biphenyl, phenanthrene, anthracene and benz[a]anthracene (Luo et al. 2016). Various dioxygenase and monooxygenases can increase the PAHs degradation ability of *Rhodococcus* sp. P14, while only 1 paralog of dihydrodiol dehydrogenase was identified in the

Table 1 List of genes with log2 ≥ 0.6 fold change according to the transcriptome data

Gene	log2 (FC)	Description
RS09705	13.15	Hypothetical protein
RS09725	12.98	Hydantoinase
RS09710	12.74	Acetone carboxylase subunit alpha
RS09715	12.67	5-Oxoprolinase (ATP-hydrolyzing)
RS09720	12.63	Acetone carboxylase subunit gamma
RS15800	11.01	Monooxygenase component MmoB/DmpM
RS15815	10.97	Propane monooxygenase hydroxylase alpha subunit
RS15810	10.90	CDP-6-deoxy-delta-3,4-glucoseen reductase, partial
RS15695	10.72	Membrane protein
RS15805	10.48	Toluene hydroxylase
RS15790	10.47	Metal-sulfur cluster biosynthetic enzyme
RS15795	10.47	Amidohydrolase family protein
RS15785	10.37	Alcohol dehydrogenase
RS15780	10.18	Multispecies: molecular chaperone GroEL
RS15775	10.08	Antibiotic biosynthesis monooxygenase
RS15700	8.61	Multispecies: AMP-dependent synthetase
RS15690	8.39	FAD-dependent oxidoreductase
RS18150	7.89	3-Hydroxybutyrate dehydrogenase
RS20005	7.85	Succinyl-CoA–3-ketoacid CoA-transferase beta subunit
RS18145	7.84	MFS transporter
RS20010	7.62	Acetyl-CoA:acetoacetyl-CoA transferase, alpha subunit
RS15685	7.49	PucR family transcriptional regulator
RS09695	7.04	Multispecies: hypothetical protein
RS00030	5.78	Acetyl-CoA acetyltransferase
RS15825	5.31	Catalase
RS16720	4.92	Hypothetical protein
RS16725	4.68	Cytochrome P450
RS09690	4.66	Hypothetical protein
RS15680	4.24	Haloacid dehalogenase
RS09700	4.22	Transcriptional regulator
RS15760	4.06	Glutaryl-CoA dehydrogenase, mitochondrial
RS09685	3.90	RDD domain-containing protein
RS15755	3.77	Aldehyde dehydrogenase
RS16695	3.75	Cytochrome P450
RS16690	3.70	Ferredoxin
RS22070	3.48	50S ribosomal protein L28
RS16615	3.42	Acylase
RS15765	3.34	4-Aminobutyrate aminotransferase
RS06970	3.16	L-Lysine 6-transaminase
RS16730	3.08	Enoyl-CoA hydratase
RS15750	3.06	L-Carnitine dehydrogenase
RS15820	2.89	Multispecies: Fis family transcriptional regulator
RS22065	2.81	50S ribosomal protein L33
RS22085	2.74	50S ribosomal protein L32
RS16610	2.68	Acylase
RS16680	2.63	Cupin, partial
RS08175	2.61	SPW repeat-containing protein
RS02625	2.55	Membrane protein
RS16685	2.51	2,4-Dichlorophenol 6-monooxygenase
RS24175	2.45	AMP-dependent synthetase
RS01780	2.35	1-Phosphofructokinase

Table 1 (continued)

Gene	log2 (FC)	Description
RS24185	2.30	IclR family transcriptional regulator
RS14110	2.20	Lipoprotein
RS16735	2.19	Acyl-CoA synthetase
RS16605	2.13	Hypothetical protein
RS24005	2.13	IclR family transcriptional regulator
RS20015	2.00	IclR family transcriptional regulator
RS17195	1.99	Peptidase
RS16460	1.98	Alpha/beta hydrolase
RS05360	1.95	Hypothetical protein
RS22075	1.94	Cobalamin biosynthesis protein CobW
RS11100	1.89	Conserved hypothetical protein
RS22080	1.88	50S ribosomal protein L31 type B
RS17380	1.83	Membrane protein
RS00190	1.79	SAM-dependent methyltransferase
RS15655	1.78	Apocarotenoid-15,15′-oxygenase
RS24190	1.77	TetR family transcriptional regulator
RS18420	1.75	Aldehyde dehydrogenase
RS09085	1.67	Membrane protein
RS00335	1.60	L-Lactate permease
RS22680	1.60	Membrane protein
RS21690	1.60	Erythromycin esterase-like enzyme
RS24825	1.58	Carboxylesterase
RS12220	1.57	cytochrome P450
RS00355	1.52	Multispecies: hypothetical protein
RS22060	1.46	30S ribosomal protein S14
RS08180	1.43	Transcriptional regulator
RS16465	1.41	Fasciclin
RS17385	1.41	Ion channel family protein
RS15145	1.37	Hypothetical protein
RS03655	1.33	Aminodeoxychorismate synthase, component I
RS12190	1.24	Hypothetical protein
RS08325	1.23	Ring-hydroxylating dioxygenase-like protein
RS24170	1.20	MaoC family dehydratase
RS21880	1.19	LuxR family transcriptional regulator
RS17095	1.18	Domain of Uncharacterised Function (DUF1540)
RS18085	1.17	Pyridoxamine 5′-phosphate oxidase
RS24180	1.16	FAD-binding monooxygenase
RS12370	1.16	Dienelactone hydrolase
RS15150	1.14	Membrane protein
RS12285	1.14	Alkane 1-monooxygenase
RS24665	1.13	Alpha/beta hydrolase fold
RS18415	1.12	L-Aspartate oxidase
RS04135	1.12	Alpha/beta fold family hydrolase
RS24970	1.10	Pyruvate dehydrogenase E1 component subunit beta
RS24165	1.05	Phenol 2-monooxygenase
RS17930	1.05	Beta-lactamase
RS02580	1.04	Hypoxic response protein 1
RS16200	1.04	Thiamine biosynthesis protein ThiJ
RS04440	1.03	Pyruvate dehydrogenase subunit beta
RS24910	1.02	Hypothetical protein

Table 2 List of sRNAs with log2 ≥ 0.6 or ≤ − 0.6 fold change according to the transcriptome data

sRNA	log_2(FC)	Location
predRNA10	− 6.70	AM[a]
predRNA11	− 4.13	AM
predRNA14	− 5.38	IGR[b]
predRNA17	− 19.73	IGR
predRNA2	− 4.85	AM
predRNA22	− 4.57	IGR
predRNA31	− 5.95	IGR
predRNA52	− 21.97	IGR
predRNA55	3.78	AM
predRNA57	− 4.10	IGR
predRNA7	− 2.69	IGR
predRNA74	− 20.79	IGR

[a]Antisense to mRNA(AM)

[b]Intergenic region(IGR)

genome, which indicates that dihydrodiol dehydrogenase is possible the rate-limiting enzyme during PAHs degradation. The dihydrodiol dehydrogenase identified in *Sphingomonas* strain CHY-1 can prevent the accumulation of PAH catechols, which is important to the bacteria cells (Jouanneau and Meyer 2006).

Moreover, RNA-seq was performed to confirm that the mRNA level of several genes, such as cytochrome P450 (RS16725; RS16695; RS12220), catalase (RS15825), dehydrogenase (RS15755; RS18420) and hydrolase (RS16460; RS24665) were up-regulated during PAHs degradation. Interestingly, various genes encoding the sigma factors and response regulators showed increased expression levels during PAHs degradation, may due to the PAHs degradation is a complex pathway, which needs many genes involved and these genes need a strict regulation. One NarL-like regulator has been proved to have negative effect on the expression of *cyp108j1* in PAHs degradation in P14 (Kan et al. 2020). The BaaA showed no change on expression level during PAHs degradation, may due to only one-time point was checked by RNA-seq during this complex pathway. Many studies showed that sRNAs harbor various functions in bacteria (Mann et al. 2012). In our study, the ring-hydroxylating dioxygenases RS14045 (BaaA) was predicted as the target gene of the presRNA 57, suggesting that sRNAs may be involved in the regulation of genes which have a function in PAHs degradation. Moreover, the sRNAs can likely regulate some regulators activity during PAHs degradation in P14 which was reported in other mechanisms in bacteria (Klein et al. 2016). No research showed that sRNAs were involved in PAHs degradation before. In this study, although the functions of sRNAs in PAHs degradation were not confirmed, we first point out the possibility that sRNAs can be involved in PAHs degradation.

In this study, genomic analysis was performed to determine the genes, which were proposed to be responsible for the PAHs degradation in P14. RHD, CYP450 and other genes like hydratase–aldolase, hydrolase, alcohol dehydrogenase, protocatechuate 3,4-dioxygenase and β-ketoadipate CoA transferase were identified in P14′ genome. The pathway in which PAHs were degraded into the acetyl CoA and succinyl CoA, then mineralized to CO_2 via the TCA cycle

Table 3 The target genes of the sRNAs predicted by RNAplex

sRNA	Target gene	Energy (kcal/mol)	Description
predRNA14	RS01615	− 11.25	XRE family transcriptional regulator
predRNA17	RS01915	− 18.52	RNA polymerase sigma factor
predRNA17	RS03140	− 14.95	Transcriptional regulator
predRNA17	RS21880	− 9.40	LuxR family transcriptional regulator
predRNA22	RS12160	− 10.32	MerR family DNA-binding transcriptional regulator
predRNA22	RS22800	− 7.75	4-Carboxymuconolactone decarboxylase
predRNA31	RS09700	− 13.15	Transcriptional regulator
predRNA52	RS05045	− 8.47	Transcriptional regulator MraZ
predRNA52	RS10705	− 11.76	LuxR family transcriptional regulator
predRNA52	RS12730	− 5.37	Glyoxalase
predRNA57	RS14045	− 6.58	1,2-Dioxygenase large subunit
predRNA7	RS05045	− 8.33	Transcriptional regulator MraZ
predRNA7	RS10830	− 12.45	FadR family transcriptional regulator
predRNA74	RS01615	− 6.99	XRE family transcriptional regulator
predRNA74	RS10830	− 9.41	FadR family transcriptional regulator
predRNA74	RS10865	− 11.15	LuxR family transcriptional regulator
predRNA74	RS14235	− 13.67	FadR family transcriptional regulator
predRNA74	RS22740	− 11.65	Transcription elongation factor GreA

was constructed based on the genome analysis. Although the PAHs degradation pathway needs more studies for confirmation, this study would provide a blueprint and instruction for the application of P14 on PAHs bioremediation in contaminated environments. For example, the dihydrodiol dehydrogenase is considered as a rate-limiting enzyme during PAHs degradation in P14, we can increase the copies of dihydrodiol dehydrogenase to enhance the PAHs degradation ability of P14. Overall, this study is helpful for further research to illustrate the PAHs degradation pathway in *Rhodococcus* genus by clarifying the function of genes identified.

Acknowledgements This research was supported by NFSC (31670117 and 31700109) (National Nature Science Foundation of China).

Author contributions TP, JK and JH analyzed the data and wrote the paper, ZH designed the experiment.

Compliance with ethical standards

Conflict of interest No conflicts of interest in the authorship or publication of this contribution.

References

Aitken MD, Stringfellow WT, Nagel RD, Kazunga C, Chen S (1998) Characteristics of phenanthrene-degrading bacteria isolated from soils contaminated with polycyclic aromatic hydrocarbons. Can J Microbiol 44:743–752

Broderick JB (1999) Catechol dioxygenases. Essays Biochem 34:173–189

Denman RB (1993) Using RNAFOLD to predict the activity of small catalytic RNAs. Biotechniques 15:1090–1095

Gibson DT, Parales RE (2000) Aromatic hydrocarbon dioxygenases in environmental biotechnology. Curr Opin Biotechnol 11:236–243

Jouanneau Y, Meyer C (2006) Purification and characterization of an arene *cis*-dihydrodiol dehydrogenase endowed with broad substrate specificity toward polycyclic aromatic hydrocarbon dihydrodiols. Appl Environ Microbiol 72:4726–4734

Kan J, Peng T, Huang T, Xiong G, Hu Z (2020) NarL, a Novel Repressor for CYP108j1 Expression during PAHs Degradation in Rhodococcus sp. P14. Int J Mol Sci 21:983

Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M (2006) From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res 34:354–357

Kim SJ, Kweon O, Freeman JP, Jones RC, Adjei MD, Jhoo JW, Edmondson RD, Cerniglia CE (2006) Molecular cloning and expression of genes encoding a novel dioxygenase involved in low- and high-molecular-weight polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. Appl Environ Microbiol 72:1045–1054

Kimura N, Kitagawa W, Mori T, Nakashima N, Tamura T, Kamagata Y (2006) Genetic and biochemical characterization of the dioxygenase involved in lateral dioxygenation of dibenzofuranfrom Rhodococcus opacus strain SAO101. Appl Microbiol Biotechnol 73:474–484

Klein G, Stupak A, Biernacka D, Wojtkiewicz P, Lindner B, Raina S (2016) Multiple transcriptional factors regulate transcription of the rpoE gene in *Escherichia coli* under different growth conditions and when the lipopolysaccharide biosynthesis is defective. J Biol Chem 291:22999–23019

Kong L, Zhang Y, Ye ZQ, Liu X, Zhao S, Wei L (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res 35:W345–349

Kweon O, Kim SJ, Jones RC, Freeman JP, Cerniglia CE (2007) Complete and integrated pyrene degradation pathway in *Mycobacterium vanbaalenii* PYR-1 based on systems biology. J Bacteriol 189:464–472

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323

Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K (2009) SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25:1966–1967

Li L, Huang D, Cheung MK, Nong W, Huang Q, Kwan HS (2013) BSRD: a repository for bacterial small regulatory RNA. Nucleic Acids Res 41:D233–238

Luo A, Wu Y, Xu Y, Kan J, Qiao J, Huang T, Hu Z (2016) Characterization of a cytochrome P450 monooxygenase capable of high molecular weight PAHs oxidization from *Rhodococcus* sp. P14. Process Biochem 51:2127–2133

Mandalakis M, Gustafsson O, Alsberg T, Egebäck AL, Reddy CM, Xu L, Klanova J, Holoubek I, Stephanou EG (2005) Contribution of biomass burning to atmospheric polycyclic aromatic hydrocarbons at three European background sites. Environ Sci Technol 39:2976–2982

Mann B, Van OT, Wang J, Obert C, Wang YD, Carter R et al (2012) Control of virulence by small RNAs in *Streptococcus pneumoniae*. PLoS Pathog 8:e1002788

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

Mason JR, Cammack R (1992) The electron-transport proteins of hydroxylating bacterial dioxygenases. Annu Rev Microbiol 46:277–305

Moody JD, Freeman JP, Doerge DR, Cerniglia CE (2001) Degradation of phenanthrene and anthracene by cell suspensions of *Mycobacterium* sp. Strain PYR-1. Appl Environ Microbiol 67:1476–1483

Nawrocki EP, Eddy SR (2013) Infernal 1.1: 100-fold faster RNA homology searches. Bioinformatics 29:2933–2935

Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR (2015) Rfam 12.0: updates to the RNA families database. Nucleic Acids Res 43:D130–137

Peng T, Luo A, Kan J, Liang L, Huang T, Hu Z (2018a) Identification of a ring-hydroxylating dioxygenases baa capable of benz[a] anthracene oxidization from *Rhodococcus* sp. P14. J Mol Microb Biotech 28:183–189

Peng T, Kan J, Lun J, Hu Z (2018b) Identification of novel sRNAs involved in oxidative stress response in fish pathogen *Vibiro alginolyticus* by transcriptome analysis. J Fish Dis 42:277–292

Schmidtke C, Abendroth U, Brock J, Serrania J, Becker A, Bonas U (2013) Small RNA sX13: a multifaceted regulator of virulence in the plant pathogen Xanthomonas. PLoS Pathog 9:e1003626

Song X, Xu Y, Li G, Zhang Y, Huang T, Hu Z (2011) Isolation, characterization of *Rhodococcus* sp. P14 capable of degrading high-molecular-weight polycyclic aromatic hydrocarbons and aliphatic hydrocarbons. Mar Pollut Bull 62:2122–2128

Tafer H, Hofacker IL (2008) RNAplex: a fast tool for RNA-RNA interaction search. Bioinformatics 24:2657–2563

Wang R, Wennerstrom D, Cao WW, Khan AA, Cerniglia CE (2000) Cloning, expression, and characterization of the *katG* gene, encoding catalase-peroxidase, from the poly-cyclic aromatic hydrocarbon-degrading bacterium *Mycobacterium* sp. strain PYR-1. Appl Environ Microbiol 66:4300–4304

Wolgel SA, Dege JE, Perkins-Olson PE, Jaurez-Garcia CH, Crawford RL, Münck E, Lipscomb JD (1993) Purification and characterization of protocatechuate 2, 3-dioxygenase from Bacillus macerans: a new extradiol catecholic dioxygenase. J Bacteriol 175:4414–4426

Zhang Y, Qin F, Qiao J, Li G, Shen C, Huang T, Hu Z (2012) Draft genome sequence of *Rhodococcus* sp. P14, a high molecular weight polycyclic aromatic hydrocarbons biodegrader. J Bacteriol 194:3546