

Full Length Research Paper

Evidence for horizontal gene transfer of a tetrahydrofuran monooxygenase: Cloning and analysis of a gene cluster for tetrahydrofuran degradation in *Rhodococcus* sp. YYL

Yanlai Yao^{1,2}, Zhenmei Lu^{2*}, Fengxiang Zhu¹ and Hang Min²

¹Institute of Environment, Resource, Soil and Fertilizer, Zhejiang Academy of Agricultural Sciences, 198 Shiqiao Road, Hangzhou 310021, China.

²Institute of Microbiology, College of Life Science, Zhejiang University, 388 Yuhangtang Road, Hangzhou 310058, China.

Accepted 19 April, 2013

A gene cluster encoding tetrahydrofuran (THF) degrading related enzymes was first isolated from *Rhodococcus* sp. YYL, a strain of the genus *Rhodococcus*, which was first reported to degrade tetrahydrofuran, by combination of CConsensus-DEgenerate Hybrid oligonucleotide Primer PCR and thermal asymmetric interlaced PCR. Sequence analysis of the gene cluster revealed 9 open reading frames containing genes encoding a *Rhodococcus* sp. YYL multiple component tetrahydrofuran monooxygenase. The tetrahydrofuran monooxygenase genes had similarities higher than 92% with these in *Pseudonocardia* sp. K1, but had low similarities with other monooxygenase genes using BLAST search in GenBank. The high specific of tetrahydrofuran monooxygenase genes was further proved by PCR amplification of no product using genomic DNA of other bacteria or even activated sludge as template with primers targeting at the α -subunit of this monooxygenase. Phylogenetic analysis based on amino acid sequences and 16S rDNA strongly suggested that this monooxygenase occurred by horizontal gene transfer. Base composition analysis and the existence of a transposase gene in the gene cluster further proposed that this monooxygenase originated from the genus *Rhodococcus* and it might be transformed by a transposon between different species.

Key words: Consensus-degenerate hybrid oligonucleotide primer polymerase chain reaction (CODEHOP PCR), thermal asymmetric interlaced PCR (TAIL-PCR), tetrahydrofuran monooxygenase, horizontal gene transfer.

INTRODUCTION

Horizontal gene transfer is considered one of the most important and rapid evolutionary forces, which can result in strains with beneficial mutations from more than one parent (Thomas and Nielsen, 2005). Compositional analysis such as GC content is often used in identification

of horizontal gene transfer now and after the genomes of many organisms have been sequenced (Nakamura et al., 2004). However, the origin of horizontally transferred genes was often unknown because of the huge number of microorganisms but relative smaller genome database.

Searching for the origin of horizontally transferred genes always attracts much interest, as it would help to understand the biological mechanisms of this process.

Tetrahydrofuran (THF) is a widely used compound which is considered as not being easily biodegradable and can be degraded by only a few strains mainly included in the genus *Rhodococcus* and *Pseudonocardia* (Bernhardt and Diekmann, 1991; Parales et al., 1994; Kohlweyer et al., 2000; Mahendra and Alvarez-Cohen, 2005; Nakamiya et al., 2005; Zhou et al., 2011a; Zhou et al., 2011b; Tajima et al., 2012). A gene cluster responsible for THF degradation in *Pseudonocardia* strain K1 was reported by Thiemer et al. (2003) and Oppenheimer et al. (2010). The genus *Rhodococcus* was the first one published to have the capability of utilizing THF as sole carbon source. However, no information about genes for THF degradation in any strain of the genus *Rhodococcus* has been reported up today.

The gene cluster responsible for THF degradation in *Pseudonocardia* strain K1 had been shown on a plasmid and containing the THF degrading key enzyme, multiple tetrahydrofuran monooxygenase (Thm), which was composed of four subunits including THF-monooxygenase α -subunit (ThmA), NADH:acceptor oxidoreductase (ThmD), THF-monooxygenase β -subunit (ThmB) and a coupling protein (ThmC) (Thierner et al., 2003). Analysis of this *thm* gene cluster illustrated that, it was distinguished with the genome of *Pseudonocardia* strain K1 in GC content. Since it is often found that horizontally transferred DNA have anomalous nucleotide compositions differing from the rest parts of the genome and two most common ways for horizontal gene transfer was through homologous recombination or plasmid transfer (Thomas and Nielsen, 2005; Monier et al., 2007), it is speculated that the *thm* genes were laterally inherited (Thierner et al., 2003). However, no further evidence was available, since no other THF-degrading related genes have been reported in any other strains.

Cloning of unknown functional genes is always the task attempt that attracts much interest of researchers. Genomic sequencing is a popular way to uncover the unknown genes. However, the relative high costs still could not be afforded by all laboratories, since there are so a large number of microorganisms. A lot of methods mainly relying on degenerate primer have been developed and available for cloning unknown gene, especially for homogenous genes. COnsensus-DEgenerate Hybrid oligonucleotide Primer (CODEHOP) PCR is one of the efficient way for identifying new members of protein families, which may be useful in cloning homogenous genes from distantly related organisms and has been utilized for identifying and characterizing new gene orthologs and paralogs in plants, animals and bacterial species (Rose et al., 1998; Rose et al., 2003; Lourenco et al., 2004; Baines et al., 2005). However, the amplified fragments using degenerate primer are always not longer than 500 bp,

because of the unspecific amplification resultings from degenerate primers. Genome walking techniques are often adopted for completing the whole ORF cloning of functional genes along with degenerate PCR. Genome walking techniques based on the PCR, such as thermal asymmetric interlaced PCR and self-formed adaptor PCR, are more successfully and widely used as they are fast and less labor intenseness (Liu and Whittier, 1995; Wang et al., 2007). These strategies use nested sequence-specific primers together with degenerate primers in a multistep thermal cycling program and have been proved to be effective in cloning of gene-specific flanking regions (Li and Gray 2005; Liu and Chen, 2007; Wang et al., 2007).

In this paper, combination of CODEHOP PCR and TAIL PCR was adopted to isolate the unknown gene cluster for THF degrading from *Rhodococcus* sp.YYL capable of tolerating THF at high concentration and degrading THF with a high rate (Yao et al., 2009). Cloned fragments were assembled and analyzed for open reading frames. The specific of this monooxygenase was tested. Multiple sequence alignment and phylogenetic analysis were used to discuss the occurrence of the THF-degrading monooxygenase genes subsequently.

MATERIALS AND METHODS

Bacteria and media

Rhodococcus sp. YYL capable of utilizing THF at high concentration was used in this study (Yao et al., 2009). The culture was carried out in BMM (basal mineral medium) with THF or glucose as the sole carbon source at 30°C at a shaking speed of 140 rpm as described by Yao et al. (2009).

Isolation of genomic DNA

Isolation of genomic DNA from *Rhodococcus* sp.YYL always failed with several methods until some modifications were made according to the method introduced by Elisabeth and Patrick (2000). The modified way was shown as following: Cells of *Rhodococcus* sp.YYL were collected at the exponential phase by centrifugation at 8 000 g (Sigma 3K15, made in Germany) for 10 min and washed twice with washing buffer (0.1 M NaCl, 10 mM Tris-HCl, 1.0 mM EDTA, pH8.0). The cells were re-suspended in 567 μ L lysozyme buffer (50 mM Tris-HCl, 5 mM EDTA, pH 8.0) in a 2 ml tube, and lysozyme was added to a relative high concentration of 5 μ g/ μ L followed by incubation at 37°C for 2 h with shaking at intervals. Then 15 μ L of 20% SDS and 3 μ L of 20 mg/ml protease K were added with incubation at 55°C for 30 min. One hundred μ L of 5 mol/L NaCl was added and mixed gently and sufficiently after the treatment with SDS and protease K, followed by addition of 100 μ L of 3 mol/L potassium acetate and sufficient mixture. After laid in ice for 10 min to precipitate proteins, the tube was centrifuged for 5 min at 13 000 g. The supernate was transferred to a new tube and extracted with equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) and next again with equal volume of chloroform: isoamyl alcohol (24:1).

Finally, the supernate was put into a new sterilized 1.5 ml tube and equal volume of isopropyl alcohol was added. The tube was standstilled in refrigerator at -20°C for 20 min, followed by

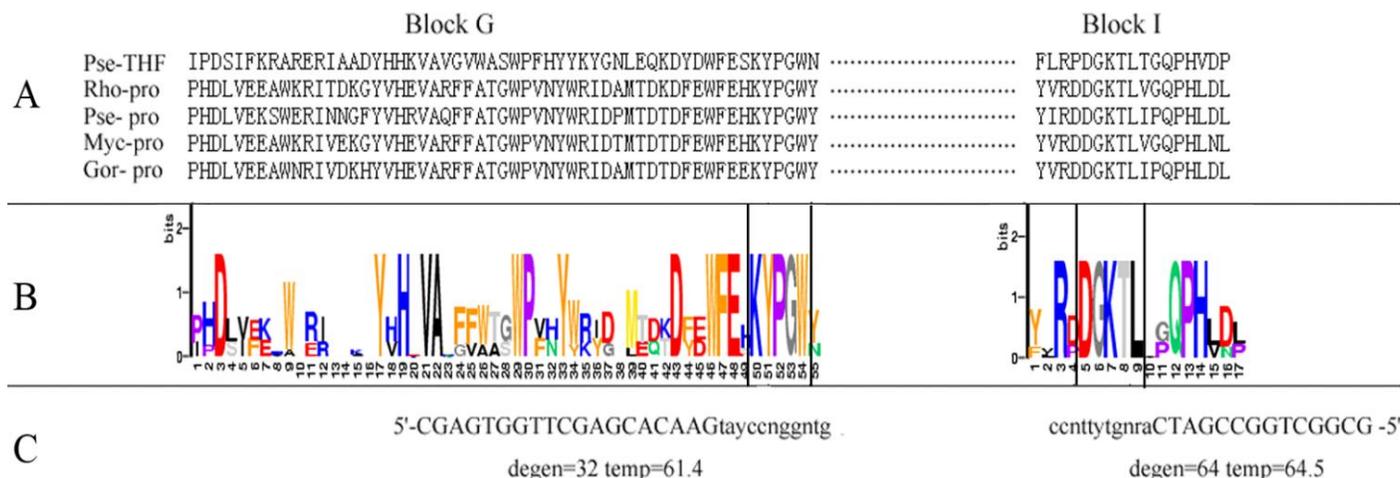


Figure 1. Primer design for CODEHOP PCR. **A**, A portion of the multiple alignment of five-monooxygenase amino acid sequences generated from the program Block Marker. The GenBank accession numbers of the five monooxygenases along with the stain names were shown as below: Pse-THF, AJ296087 (*Pseudonocardia* sp.K1); Rho-pro, YP_700435 (*Rhodococcus* sp.RHA1); Pse-pro, BAF34308 (*Pseudonocardia* sp.TY-7); Myc-pro, U34405 (*Mycobacterium smegmatis* str.MC2 155); Gor-pro, BAD03956(*Gordonia* sp.TY-5); **B**, Conserved amino acid blocks generated from the multiple alignment. The height of each amino acid is proportional to its degree of conservation. The motif in the square was used for designing degenerate 3' core of the CODEHOP primer in this study; **C**, Nucleotide sequence of the CODEHOP primers shown with the 5' consensus clamp in capitalization and a degenerate 3' core region. The mean of degenerate letters in the degenerate 3' core region was present as below: N=A/C/G/T, R=A/G, Y=C/T.

centrifuging for 15 min to precipitate the genomic DNA. The DNA was washed gently with 75% alcohol and then 100% alcohol. After total evaporation of the alcohol at room temperature, suitable volume of TE solution was added to resolve the DNA.

Primer design and PCR program for CODEHOP PCR

Consensus-Degenerate Hybrid Oligonucleotide Primers were designed with the aid of the web-based program CODEHOP (Rose et al., 1998, 2003), which was accessed at <http://bioinformatics.weizmann.ac.il/blocks/codehop.html>. Amino acid sequence of the α -subunit of a multicomponent tetrahydrofuran monooxygenase in *Pseudonocardia* strain K1 was blasted with these in GenBank. Five sequences with highly similarity from different genera under Nocardioform Actinomycetes were retrieved and submitted to the Block marker program on line for multiple align and block making (http://bioinformatics.weizmann.ac.il/blocks/blockmkr/www/make_blocks.html). The generated blocks were used for primer design. Each primer consisted of a degenerate 3' core region and a non-degenerate 5' consensus clamp (Figure 1). The Tm values of primers were set higher than 60°C. PCR amplification was performed in a 50 μ L reaction volume with 1 μ L genomic DNA, 1 μ L of 20 mM each primer, 4 μ L of 0.25 mM each dNTP, and 2.5 units Taq polymerase (Takara Biotechnology (Dalian) Co., Ltd, China). Reactions were carried out in a MyGene Series Gradient Thermal Cycle (LongGene scientific Instruments Co.,Ltd, China) using hot start and touchdown protocol recommended for the CODEHOP PCR. The thermal cycling conditions were: 1 cycle at 95°C for 5 min; addition of Taq polymerase when the temperature decreased to 80°C; 10 cycles of 94°C for 30 s, annealing at temperatures from 65 to 55°C for 1 min with a stepwise decreasing gradient of 1°C per cycle, and 72°C for 45 s; 20 cycles of 94°C for 30 s, 55°C for 1 min, and 72°C for 45 s; and 1 final additional cycle at 72°C for 10 min. The resulting PCR product was checked on 1.0% (w/v) agarose gels and sequenced directly by Shanghai Sangon Biological

Engineering Technology and Services Co., Ltd, China.

Cloning of the tetrahydrofuran monooxygenase gene cluster

TAIL PCR as a gene walking method was adopted to complete the monooxygenase α -subunit gene cloning. Special primers for TAIL PCR were designed using the fragment amplified by CODEHOP PCR as a template according to the principle recommended by the literature as listed in Table 1 (Liu and Whittier, 1995). The degenerate primer was designed according to the conserved motif of succinate semialdehyde dehydrogenase and NADH:acceptor oxidoreductase, which were reported to flank monooxygenase α -subunit in *Pseudonocardia* sp.K1(Kohlweyer et al., 2000). The program of PCR amplification was carried out according to the standard procedure, except that the Tm value in the high stringency cycles was set based on Tms of special primers (Liu et al., 1995, Liu and Whittier, 1995).

Continuous gene walking was further performed to retrieve the unknown adjacent sequence of the THF degrading monooxygenase alpha-subunit gene in *Rhodococcus* sp.YYL. All the amplified fragments had overlapped 5' or 3' end of about 100 bp with the adjacent sequences for sequence assembly.

All primers were synthesized by Shanghai Sangon Biological Engineering Technology and Services Co.,Ltd, China. All the resulting PCR products were checked on 1.0% (w/v) agarose gels and sequenced directly by Shanghai Sangon Biological Engineering Technology and Services Co., Ltd, China. All the DNA marker used in this study were a DL2000 DNA marker (Takara Biotechnology (Dalian) Co., Ltd, China) with molecular length from top to bottom as 2 000, 1 000, 750, 500, 250 and 100 bp.

Northern blot analysis

Northern blot analysis was done to further prove the amplified gene as that targeted at. *Rhodococcus* sp.YYL was cultured in BSM with

Table 1. Sequences of the primers used in the study.

Primer name	Sequence
Primers for TAIL PCR	
tdmAlsp1	5'-AGCCGAGGTACGCACGAT-3'
tdmAlsp2	5'-AGCGGTCCAACCAAACACG-3'
tdmAlsp3	5'-TGCGGCGAGCAATAGAACC-3'
TdmIAD1*	5'-GARGARATHHTTYGG-3'
tdmArsp1	5'-ACTCCGAGATGGTGC GTGAC-3'
tdmArsp2	5'-CTGTGTCGCTGACCGTGT TTG-3'
tdmArsp3	5'-CTCGCTGGAGTGGAAAGTGA-3'
TdmArADr1*	5'-GTYTCDATCCAYTG-3'
Primers for northern blot probe amplifying	
Probsense	5'-ACGCTCTGCCCACTGTCA-3'
ProbAntisenseout	5'-CGGTGTATTCCACTCGCTAT -3'
ProbAsensein	5'-CCGAGGTACGCACGAT-3'

*Arbitrary degenerate primer.

0.5% (w/v) glucose or 12.5 mM THF as sole carbon source. Total RNA was extracted at the exponential phase with AxyPrep Multisource Total RNA Miniprep Kit (Axygen Biosciences 33210 Central Avenue, Union City, CA 94587, USA). RNA was subjected to electrophoresis in a 1.2% agarose gel containing 2.2 M formaldehyde, and transferred onto a nylon membrane (Hybond N+; GE Healthcare UK Ltd., Buckinghamshire, UK) (Sambrook and Russell, 2001). The filter was hybridized with DIG-labeled DNA probe for 12 h at 42°C using DIG-High Prime DNA Labeling and Detection Starter Kit (Roch, Germany). The probe with a sequence length of 755 bp located in the THF degrading monooxygenase α -subunit was amplified by nested PCR. The primers used for probe preparing were Probsense (5'-ACGCTCTGCCCACTGTCA-3'), ProbAntisenseout (5'-CGGTGTATTCCACTCGCTAT-3'), and ProbAsensein (5'-CCGAGGTACGCACGAT-3').

Specificity test of the THF monooxygenase genes

Genomic DNA of six kinds of other bacteria and two samples of activated sludge before or after bioaugmentation with *Rhodococcus* sp. YYL were extracted by the method mentioned above and used as the template for specificity test of these monooxygenase genes by PCR amplification of partial α -subunit gene with the primers of Probsense and ProbAntisenseout. The thermal cycling conditions were: 1 cycle at 95°C for 5 min; 30 cycles of 94°C for 30 s, 60°C for 1 min, and 72°C for 1 min; and 1 final additional cycle at 72°C for 10 min. The resulting PCR product was checked on 1.0% (w/v) agarose gels.

Sequence analysis

Sequences were assembled using the program ContigExpress (a component of Vector NTI Advance 9.1.0). Sequence management and ORF finding were done by program AnnHyb (version 4.935). The ORFs and deduced amino acid sequences were submitted to GenBank for similarity comparison. Sequences with high similarity and 16S rDNA strains related to these sequences were retrieved from the NCBI nucleic acid and protein database. The program ClustalX (version 1.8) with default parameters was run for multiple

sequence alignment. Phylogenetic and evolutionary distance analysis of the aligned sequences was performed by the program MEGA (version 2.1).

Nucleotide accession number

The nucleotide sequence data reported in this study are available at GenBank under accession number EU732588.

RESULTS

Amplification of partial fragment of the tetrahydrofuran monooxygenase gene cluster by CODEHOP PCR

The Block Making program recognized 10 regions (designated as blocks A through J) of conserved protein sequences within oxygenase subunits of the 5 entered monooxygenases. The 10 conserved blocks were submitted to the CODEHOP program for identifying possible PCR primers, and the resulting primers were screened for degeneracies of 128 or less (program default). Functional primers were identified in 9 of the 10 blocks (except Block F). Several pairs of primers were provided. A pair of primers from Block G and complement of Block I were chosen for the PCR amplification, following the basic principle of primer design and effective amplification with degenerate primers under the fragment length of 500 bp. Cycling conditions were optimized using hot start and touchdown PCR (Don et al., 1991).

A target sequence of 331 bp was finally obtained by CODEHOP PCR (Figure 2A). Through database homology searching, the fragment was found to be part of an oxygenase gene and had similarities of 96 and 93% with the monooxygenases in *Rhodococcus* sp. THF 100

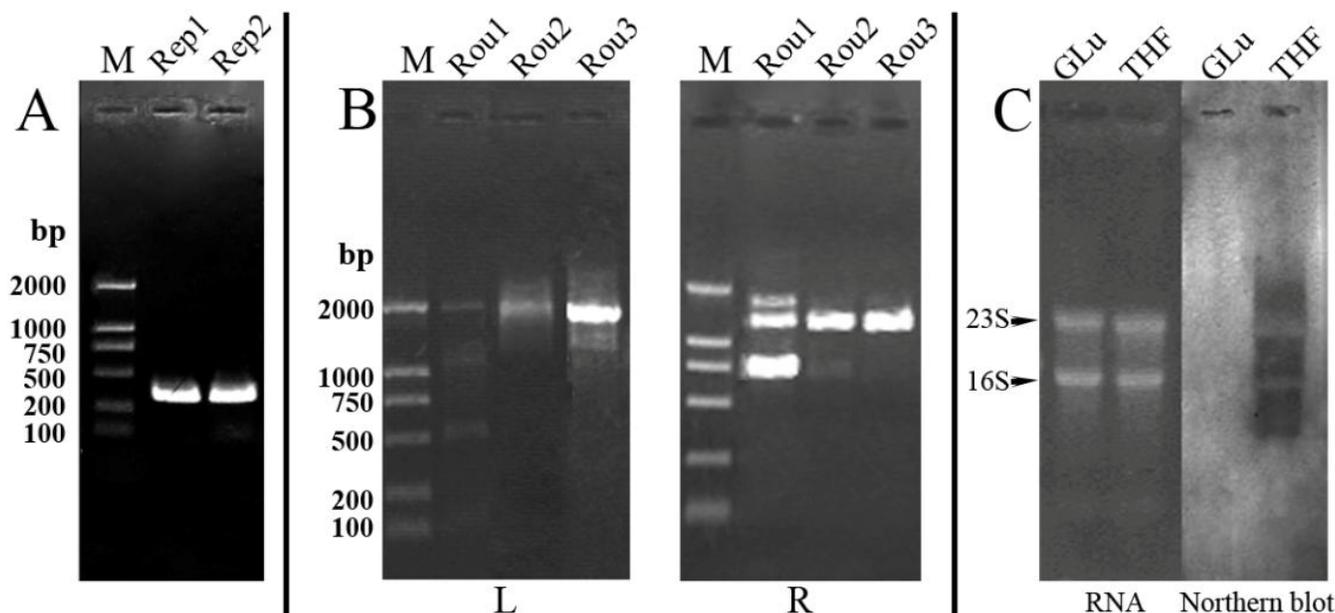


Figure 2. Cloning of a tetrahydrofuran monooxygenase gene cluster from *Rhodococcus* sp.YYL and expression validation by northern blot analysis. Five microlitres of PCR products from each pair of primers combination along with a DNA marker DL2000(M) were loaded into each well in the check of PCR product. **A)** Agarose gel check of the product of CODEHOP PCR. Lane Rep1 and Rep2 are the products of two repeats of CODEHOP PCR; **B)** Flanking sequence cloning of the fragment amplified by CODEHOP PCR using TAIL-PCR. L and R are the products of left and right flanking sequence cloning by TAIL-PCR. Lane Rou 1 to Rou 3 are the products of the first, second and third round of TAIL-PCR. Only one band is visible in third round reaction; **C)** Northern blotting analysis of the expression of *tdmA*. Lane 1 and 2 are the total RNA isolated from the cells cultured on glucose and THF as sole carbon source at the exponential phase, respectively. Northern blot analysis of the expression of *tdmA* showed that *tdmA* was specially induced when *Rhodococcus* sp.YYL was cultured with THF (lane 4) as the sole carbon source compared to culturing with glucose (lane 3).

(DQ388161) and *Pseudonocardia* sp. K1 (AJ296087), respectively. It also had similarities with the methane monooxygenase in *Frankia* sp. Ccl3 (ABD11887), hydrogenase subunit of propane monooxygenase in *Pseudonocardia* sp. TY-7 (AB250941) and α -subunit of toluene monooxygenase in *Bradyrhizobium* sp. ORS278 (CU234118). From the result of Blast searching in GenBank, the sequence was what we targeted at.

Amplification of the complete ORF of the tetrahydrofuran monooxygenase α -subunit by TAIL PCR

The fragment isolated by CODEHOP PCR was supplied as template for special nested primers designing to amplify the flanking sequence by TAIL PCR. A single band was visible on the agarose gel in amplifying the left and right flanking sequence after three round amplifications of TAIL PCR (Figure 2B). PCR products of the third round amplification were directly sent for sequencing. Fragments of 1,187 bp and 1,575 bp in length were obtained and blasted in the GenBank database. The result showed that both fragments had the highest similarity to *thm* gene cluster in *Pseudonocardia*

sp. K1 (AJ296087).

Special expression of the tetrahydrofuran monooxygenase α -subunit gene

A fragment of 3 262 bp in length was obtained after three sequencings above were assembled with the help of the ContigExpress program. A complete ORF with high similarity (96%) to the gene encoding multicomponent tetrahydrofuran monooxygenase α -subunit in *Pseudonocardia* sp.K1 was identified and designated as *tdmA*. Special expression of *tdmA* in THF degradation was examined by northern blot analysis under different growth conditions. Total RNA was isolated from *Rhodococcus* sp.YYL at the exponential phase when the strain grew on tetrahydrofuran or glucose (Figure 2C).

Northern blot analysis showed that transcription of this gene was only detectable in RNA isolated from tetrahydrofuran-growing cells (Figure 2C). The result demonstrated clearly that transcription of this gene was specifically induced by tetrahydrofuran and it was involved in the breakdown of this compound, which was consistent with *Pseudonocardia* sp. K1 (Thiemer et al., 2003).

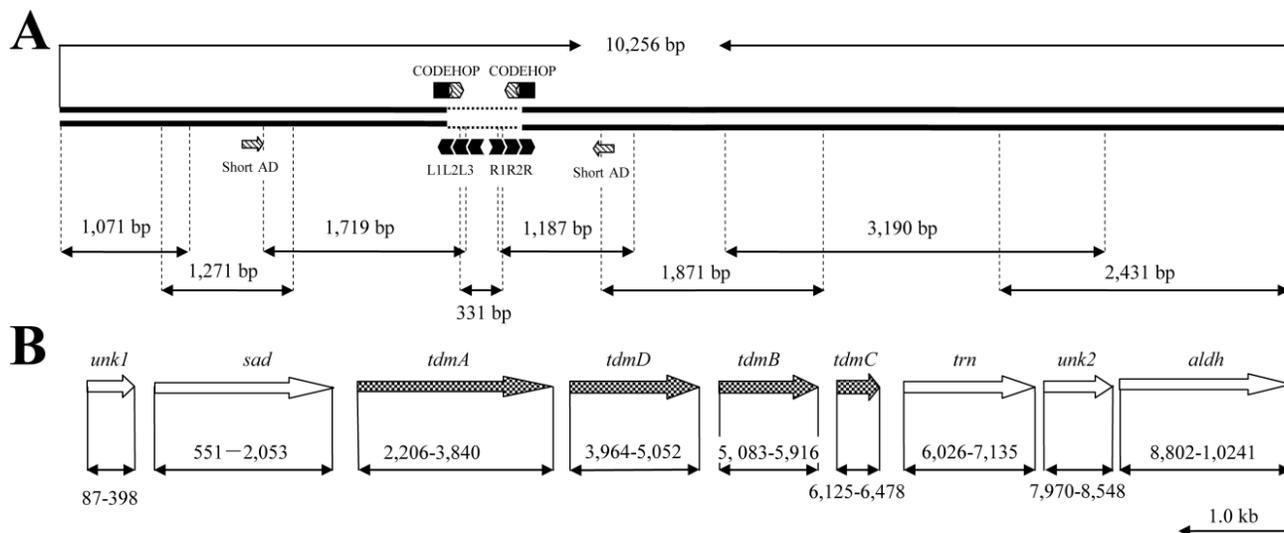


Figure 3. A 10,256-bp tetrahydrofuran monooxygenase gene cluster in *Rhodococcus* sp.YYL. **A)** Cloning of the tetrahydrofuran monooxygenase gene cluster. A pair of CODEHOPs was used to isolate partial fragment and TAIL PCR with special primers (L1L2L3 and R1R2R3) and degenerate primer (AD) was adopted to obtain the complete ORF of the α -subunit of the monooxygenase. Remaining genes were cloned directly using special primers; **B)** Genetic organization of the gene cluster. The position and orientation of the identified genes are indicated by arrows.

Table 2. Similarity of the THF degrading related genes between *Rhodococcus* sp.YYL and *Pseudonocardia* sp.K1.

Gene	Similarity (%)	Encoded protein
<i>sad:sad</i>	97	succinate semialdehyde dehydrogenase
<i>tdmA:thmA</i>	96	α -subunit of the tetrahydrofuran monooxygenase
<i>tdmD:thmD</i>	92	NADH:acceptor oxidoreductase
<i>tdmB:thmB</i>	94	β -subunit of the tetrahydrofuran monooxygenase
<i>tdmC:thmA</i>	96	coupling protein
<i>aldH:aldH</i>	95	4-hydroxybutyraldehyde dehydrogenase

Cloning and analysis of a 10, 256-bp tetrahydrofuran monooxygenase gene cluster

Genes encoding other THF monooxygenase subunits and other genes in the THF-degrading monooxygenase gene cluster were amplified with continuous gene walking. The final total assembled gene cluster was 10, 256 bp in length. Analysis of the gene cluster enucleated 9 ORFs with nucleotide sequence length more than 300 bp. Seven of them were designated as genes with definite function encoding four components of the tetrahydrofuran degrading monooxygenase (Tdm), two dehydrogenases and a transposase, respectively, by aligning with homologous sequences using BLAST programs available through the National Center for Biotechnology Information. The 9 ORFs were designated as *unk1*, *sad*, *tdmA*, *tdmD*, *tdmB*, *tdmC*, *trp*, *unk2* and *aldH*, respectively, in the order of expression orientation for differentiation with *Pseudonocardia* sp.K1 (Figure 3).

All the genes except *trp* had high similarities to these in *Pseudonocardia* sp. K1 has showed in Table 2. The expressing direction of these genes was also the same as that in *Pseudonocardia* sp.K1. A transposase gene was first identified between the *tdmC* and *aldH* in *Rhodococcus* sp.YYL, which was not found in *Pseudonocardia* sp. K1. Analysis of its nucleotide sequence revealed that it had the highest similarity of 75% to the one (Accession No. CP000850) in *Salinispora arenicola* CNS-205.

Sequence specificity of the tetrahydrofuran monooxygenase genes

The family of monooxygenases composed of a large number of members. They exist in a lot of organisms. Nucleotide sequence analysis of the tetrahydrofuran monooxygenase showed that this monooxygenase was

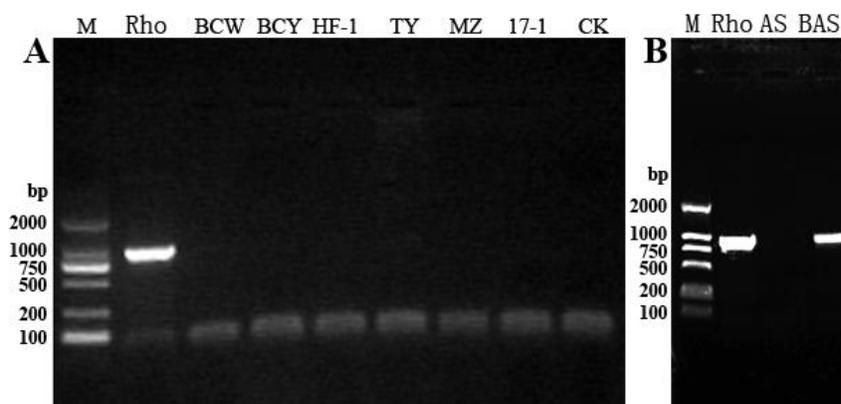


Figure 4. Sequence specificity test of the tetrahydrofuran monooxygenase genes by PCR amplification with gene-special primers. The mean of the labels is shown as following: M represents a DL2000 marker, Rho *Rhodococcus* sp.YYL, BCW *Bacillus cereus*, BCY *Bacillus aquimaris*, HF-1 *Pseudomonas* sp.HF-1, TY *Sphingomonas* sp.TY, MZ *Acinetobacter* sp.MZ, 17-1 *Pseudomonas aeruginosa*, CK black control, AS activated sludge, BAS bioaugmented activated sludge.

special to other monooxygenases in sequence similarity, though its encoding peptide contained the four highly conserved glutamate residues and two histidine residues which might be involved in coordination of the active-site binuclear-iron center (Westerheide et al., 2000). The specificity of these monooxygenase genes were further experimentally verified by PCR amplification using gene-special primers. Cells of six bacteria and two samples of activated sludge were collected for isolating each genomic DNA. The primers for amplifying northern blot probe were used to amplify partial monooxygenase α -subunit gene from the genomic DNA of the above samples. As shown in Figure 4A, nothing could be seen on the gel using other bacterial genomic DNA as templates except *Rhodococcus* sp.YYL. The activated sludge was once used for treating THF contamination and isolating THF degrading microorganisms, however non THF degrading event happened and no bacteria capable of degrading THF was obtained. In this study, it was exciting that after it was bioaugmented with *Rhodococcus* sp.YYL, the targeted fragment could be amplified from the bioaugmented activated sludge but not from the original one (Figure 4B). The experiments demonstrated that this monooxygenase was special to other one and the α -subunit gene of the tetrahydrofuran monooxygenase could be used for probe design for detecting THF degrading microorganisms, as it is difficult to isolate THF degrading microorganisms with traditional route.

DISCUSSION

Since no information about THF degrading genes in the genus *Rhodococcus* had been reported, the procedure involving in the use of two kinds of PCR (CODEHOP

PCR and TAIL PCR) was adopted in cloning of the THF-degrading monooxygenase gene cluster. CODEHOP PCR is efficient in designing degenerate primers and has been successfully utilized for cloning short homogenous genes from distant related organisms as diverse as fish, frog, protozoa, plants, viruses and bacteria (Rose et al., 2003, Liang et al., 2006; Luzhetskyy et al., 2007). TAIL PCR is an available way for genomic walking. However, the walking length is often shorter than 1.0kb (Tan et al., 2005). In this study, arbitrary degenerate primers were designed from homogenous region to improve the efficiency in amplification length and reduced the artificial product as recommended by Meng Haijun (Meng et al., 2005). Combination of CODEHOP PCR and improved TAIL PCR in this study was successfully applied to clone the tetrahydrofuran monooxygenase gene cluster.

The sequencing analysis of the gene cluster revealed all the subunits of the tetrahydrofuran monooxygenase between *Rhodococcus* sp.YYL and *Pseudonocardia* sp.K1 had high similarities and this made us to be interested in the occurrence of this monooxygenase. The ORFs were submitted for protein conserved domain searching and protein sequence similarity analysis in the GenBank. Sequences with relative high similarity to the subunits of Tdm in *Rhodococcus* sp.YYL and 16S rDNA sequences of the related strains were retrieved for phylogenetic analysis. As shown in Figure 5A, each subunit of the tetrahydrofuran monooxygenase in *Rhodococcus* sp.YYL and *Pseudonocardia* sp.K1 was clustered into a group along with evolutionary distance nearly close to 0 and a confidence level of 100%, but had a relative long distance with other monooxygenases. Whereas, there was a long evolutionary distance between *Rhodococcus* sp. YYL and *Pseudonocardia*.sp.K1 in the taxonomy position (Figure 5B). The high sequence similarity of the tetrahydrofuran

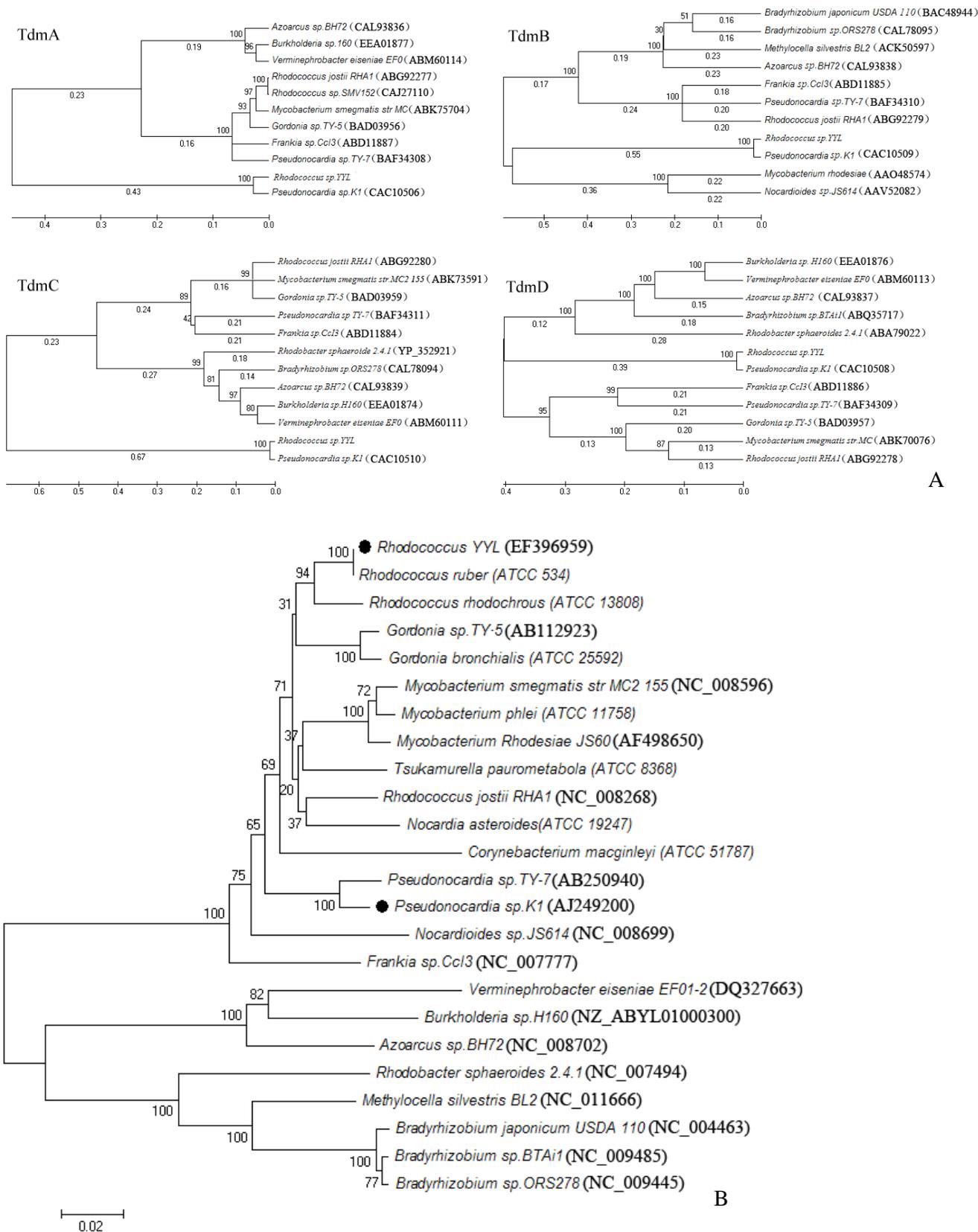


Figure 5. Phylogenetic analysis of each subunit of the tetrahydrofuran monooxygenase (A) and evolution analysis of strains related to monooxygenases mentioned in (A) based on the 16S rDNA sequences (B). The trees were constructed by the neighbor-joining method and bootstrapping analysis.

monooxygenase between different strains and relative distant phylogenetic relationship of two strains meant that this monooxygenase might originate from the same ancestor and occur between different species by horizontal transfer, which was suspected by Kohlweyer et al (Kohlweyer et al., 2000).

It was once suspected that the tetrahydrofuran monooxygenase might occur in *Pseudonocardia* sp.K1 by horizontal gene transfer (Kohlweyer et al., 2000). One of the reasons was that the GC content of these monooxygenase genes was differential with that of the genome. Genus *Pseudonocardia* was recognized as high G+C bacteria with GC content always high than 70% (Kampfer et al., 2006; Park et al., 2008; Chen et al., 2009; Duangmal et al., 2009). Base composition analysis clarified that GC content of the monooxygenase gene cluster was about 60% both in *Pseudonocardia* sp.K1 and *Rhodococcus* sp.YYL, which was mostly close to that of the genome of genus *Rhodococcus* usually between 60 to 70% (Goodfellow and Alderson, 1977). It shows that this gene cluster was mostly consistent with the genome of *Rhodococcus* sp.YYL. On the other hand, since no plasmid was extracted from the cells of *Rhodococcus* sp.YYL and the *tdm* gene cluster was amplified from the genomic DNA of *Rhodococcus* sp.YYL, the *tdm* gene cluster in *Rhodococcus* sp.YYL seems to be vertically inherited.

Furthermore, sequencing of *Rhodococcus* sp.RHA1 genomic indicated that strain RHA1 had acquired fewer genes by recent horizontal transfer than most bacteria characterized to date, which implies that the genus *Rhodococcus* rarely accepted foreign DNA by horizontal gene transfer and the THF monooxygenase in the genus *Rhodococcus* mostly occurred by vertical inheritance (McLeod et al., 2006). All the results above indicated that this gene cluster most likely originated from the genus *Rhodococcus*. The transfer of the gene cluster might be mediated by a transposon as the existence of a transposase next to the monooxygenase gene cluster.

ACKNOWLEDGEMENT

This work was financially supported by National Natural Science Foundation of China (No.21107092).

REFERENCES

- Baines JE, McGovern RM, Persing D, Gostout BS (2005). Consensus-degenerate hybrid oligonucleotide primers (CODEHOP) for the detection of novel papillomaviruses and their application to esophageal and tonsillar carcinomas. *J. Virol. Methods* 123(1):81-87.
- Bernhardt D, Diekmann H (1991). Degradation of dioxane, tetrahydrofuran and other cyclic ethers by an environmental *Rhodococcus* strain. *Appl. Microbiol. Biotechnol.* 36(1):120-123.
- Chen HH, Qin S, Li J, Zhang YQ, Xu LH, Jiang CL, Kim CJ, Li WJ (2009). *Pseudonocardia endophytica* sp. nov., isolated from the pharmaceutical plant *Lobelia clavata*. *Int. J. Syst. Evol. Microbiol.* 59(Pt 3):559-563.
- Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991). 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.* 19(14):4008.
- Duangmal K, Thamchaipenet A, Matsumoto A, Takahashi Y (2009). *Pseudonocardia acaciae* sp. nov., isolated from roots of *Acacia auriculiformis* A. Cunn. ex Benth. *Int. J. Syst. Evol. Microbiol.* 59(Pt6):1487-1491.
- Elisabeth C, Patrick S (2000). Isolating chromosomal DNA from bacteria. In: Rapley R (ed) *The Nucleic Acid Protocols Handbook* Humana Press Inc, Totowa, NJ, pp. 29-31.
- Goodfellow M, Alderson G (1977). The actinomycete-genus *Rhodococcus*: a home for the "rhodochrous" complex. *J. Gen. Microbiol.* 100(1):99-122.
- Kampfer P, Kohlweyer U, Thiemer B, Andreesen JR (2006). *Pseudonocardia tetrahydrofuranoxydans* sp. nov. *Int. J. Syst. Evol. Microbiol.* 56(Pt 7):1535-1538.
- Kohlweyer U, Thiemer B, Schrader T, Andreesen JR (2000). Tetrahydrofuran degradation by a newly isolated culture of *Pseudonocardia* sp. strain K1. *FEMS Microbiol. Lett.* 186(2):301-306.
- Li ZT, Gray DJ (2005). Isolation by improved thermal asymmetric interlaced PCR and characterization of a seed-specific 2S albumin gene and its promoter from grape (*Vitis vinifera* L.). *Genome* 48(2):312-320.
- Liang LK, Wang XK, Zhu KL, Chi ZM (2006). Trehalose accumulation in a high-trehalose-accumulating mutant of *Saccharomycopsis fibuligera* sdu does not respond to stress treatments. *Biochemistry (Mosc.)* 71(12):1291-1297.
- Liu Y, Mitsukawa N, Oosumi T, Whittier R (1995). Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* 8:457-463.
- Liu Y, Whittier R (1995). Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 25:674-681.
- Liu YG, Chen Y (2007). High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *Biotechniques* 43(5):649-650, 652, 654 passim.
- Lourenco PM, Almeida T, Mendonca D, Simoes F, Novo C (2004). Searching for nitrile hydratase using the Consensus-Degenerate Hybrid Oligonucleotide Primers strategy. *J. Basic Microbiol.* 44(3):203-214.
- Luzhetskyy A, Weiss H, Charge A, Welle E, Linnenbrink A, Vente A, Bechthold A (2007). A strategy for cloning glycosyltransferase genes involved in natural product biosynthesis. *Appl. Microbiol. Biotechnol.* 75(6):1367-1375.
- Mahendra S, Alvarez-Cohen L (2005). *Pseudonocardia dioxanivorans* sp. nov., a novel actinomycete that grows on 1,4-dioxane. *Int. J. Syst. Evol. Microbiol.* 55(Pt 2):593-598.
- McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJ, Holt R, Brinkman FS, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD (2006). The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. *Proc. Natl. Acad. Sci. USA.* 103(42):15582-15587.
- Meng H, Xu J, Guo W, Xu Q, Li D, Deng X (2005). A protocol for rapid isolation of flanking regions from short known sequences. *Plant Mol. Biol. Rep.* 23:75a-75g.
- Monier A, Claverie JM, Ogata H (2007). Horizontal gene transfer and nucleotide compositional anomaly in large DNA viruses. *BMC Genomics.* 8:456.
- Nakamiya K, Hashimoto S, Ito H, Edmonds JS, Morita M (2005). Degradation of 1,4-dioxane and cyclic ethers by an isolated fungus. *Appl. Environ. Microbiol.* 71(3):1254-1258.
- Nakamura Y, Itoh T, Matsuda H, Gojobori T (2004). Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat. Genet.* 36(7):760-766.
- Oppenheimer M, Pierce BS, Crawford JA, Ray K, Helm RF, Sobrado P (2010). Recombinant expression, purification, and characterization of ThmD, the oxidoreductase component of tetrahydrofuran monooxygenase. *Arch. Biochem. Biophys.* 496(2):123-131.
- Parales RE, Adamus JE, White N, May HD (1994). Degradation of

- 1,4-dioxane by an actinomycete in pure culture. *Appl. Environ. Microbiol.* 60(12):4527-4530.
- Park SW, Park ST, Lee JE, Kim YM (2008). *Pseudonocardia carboxydivorans* sp. nov., a carbon monoxide-oxidizing actinomycete, and an emended description of the genus *Pseudonocardia*. *Int. J. Syst. Evol. Microbiol.* 58(Pt 11):2475-2478.
- Rose TM, Henikoff JG, Henikoff S (2003). CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. *Nucleic Acids Res.* 31(13):3763-3766.
- Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM, Henikoff S (1998). Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res.* 26(7):1628-1635.
- Sambrook J, Russell DW (2001). *Molecular Cloning: A Laboratory Manual*. In. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Tajima T, Hayashida N, Matsumura R, Omura A, Nakashimada Y, Kato J (2012). Isolation and characterization of tetrahydrofuran-degrading *Rhodococcus aetherivorans* strain M8. *Process Biochem.* 47(11):1665-1669.
- Tan G, Gao Y, Shi M, Zhang X, He S, Chen Z, An C (2005). SiteFinding-PCR: a simple and efficient PCR method for chromosome walking. *Nucleic Acids Res.* 33(13):e122.
- Thiemer B, Andreesen JR, Schrader T (2003). Cloning and characterization of a gene cluster involved in tetrahydrofuran degradation in *Pseudonocardia* sp. strain K1. *Arch. Microbiol.* 179(4):266-277.
- Thomas CM, Nielsen KM (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3(9):711-721.
- Wang S, He J, Cui Z, Li S (2007). Self-formed adaptor PCR: a simple and efficient method for chromosome walking. *Appl. Environ. Microbiol.* 73(15):5048-5051.
- Westerheide L, Pascaly M, Krebs B (2000). Methane monooxygenase and its related biomimetic models. *Curr. Opin. Chem. Biol.* 4(2):235-241.
- Yao Y, Lv Z, Min H, Jiao H (2009). Isolation, identification and characterization of a novel *Rhodococcus* sp. strain in biodegradation of tetrahydrofuran and its medium optimization using sequential statistics-based experimental designs. *Bioresour. Technol.* 100(11):2762-2769.
- Zhou YY, Chen DZ, Jin XJ, Chen JM, He J (2011a). Characteristics of tetrahydrofuran degradation by *Pseudomonas oleovorans* DT4. *Huan Jing Ke Xue* 32(1):266-271.
- Zhou YY, Chen DZ, Zhu RY, Chen JM (2011b). Substrate interactions during the biodegradation of BTEX and THF mixtures by *Pseudomonas oleovorans* DT4. *Bioresour. Technol.* 102(12):6644-6649.