PHTHALATE BIODEGRADATION:
GENE ORGANIZATION, REGULATION AND DETECTION

By

Ruyang Han

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Microbiology and Molecular Genetics

is written under the direction of

Dr. Gerben J. Zylstra

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2008

©2008

Ruyang Han

PHTHALATE DEGRADATION:

GENE ORGANIZATION, REGULATION AND DETECTION

By RUYANG HAN

Dissertation Director: Prof. Gerben J. Zylstra

The three phthalate isomers are widely found in the environment due to their extensive use in the manufacture of plastics. Many microorganisms have been isolated for their ability to degrade phthalate isomers. In this study, we focused on nine different phthalate degrading bacterial strains (YZW-A, -B, -C, -D, -E, -F, -G, -H, and -I) which were isolated from Passaic River sediment and belong to different genera (*Comamonas*, *Pseudomonas*, *Acinetobacter*, and *Arthrobacter*).

Our work aims to identify the presence and divergence of the phthalate, isophthalate and terephthalate degradative genes in the nine strains isolated from the same sediment sample. The *oph, iph*, and/or *tph* genes in *Comamonas testosteroni* strains (YZW-B, -E, and -F) and *Pseudomonas* strains (YZW-A and -G) were determined by PCR and inverse PCR. Sequence analyses indicate that

phthalate, isophthalate and terephthalate degrading bacterial isolates at the same location are not simply clones of each other and that the genes identified are linked specifically to these bacterial strains.

In order to investigate whether each phthalate isomer would specifically induce the corresponding degradative gene cluster and how regulatory genes control phthalate isomers degradation, we used quantitative real time PCR to measure the expression of genes encoding phthalate (*ophA2*), isophthalate (*iphA2*), terephthalate (*tphA2*) dioxygenase in YZW-B*.* qPCR data showed that the *ophA2*, *iphA2*, and *tphA2* genes were specifically induced by phthalate, isophthalate, and terephthalate, respectively. The *tphA2* gene was slightly upregulated by isophthalate. Furthermore, we knocked out phthalate regulatory gene *ophR* and isophthalate regulatory gene *iphR* in YZW-B and analyzed the *ophA2*, *iphA2*, and *tphA2* gene expression patterns in the *ophR* or *iphR* knock out mutant using qPCR. Gene knockout and qPCR showed that the *ophR* gene and the *iphR* gene encoded repressors that negatively controlled the phthalate and the isophthalate gene expression, respectively.

Table of Contents

Acknowledgements

I'd like to express my deepest gratitude to my thesis advisor Prof. Gerben J. Zylstra. Thank you for giving me such a great opportunity to work with you, supporting my research, sharing your experiences with me, and teaching me things that I may benefit for the rest of my life. Your contribution to my microbiology world is invaluable.

Also I'd like to thank all of my labmates in the Zylstra lab who contributed their time, energy and talent to my research. Laurie Seliger, Michael Murillo, Hung-Kuang Chang, Jong-Chan Chae, Sinéad Ní Chadhain, Boris Wawrik, Marie Montes-Matías, María Cruz, Karen Pesce, Hisako Masuda, and Elyse Rodgers-Vieira, it's been a fun journey working with you all. You made this thesis possible.

Very special thanks to Dr. Boyd Woodruff and Mrs. Jeanette Woodruff who are the donors of the H. Boyd and Jeanette I. Woodruff Microbiology Fellowship. It's been a great honor to be awarded this fellowship during my first year at Rutgers. Your generosity changed my life dramatically ever since.

I'd like to thank Dr. Jerome Kukor, Dr. Costantino Vetriani, Dr. Elisabetta Bini, and

Dr. Donald Y. Kobayashi for serving as committee members and giving me excellent advice on my thesis and presentation. Also I'd like to thank all of the microbiologists on the Cook Campus, Dr. Theodore Chase, Dr. Douglas E. Eveleigh, Dr. Alan Antoine, Dr. Max Häggblom, Dr. Tamar Barkay, and Dr. Diane Davis, for those wonderful lectures and the most needed help on my TA work. You are the reasons for me to spend my lifetime exploring the beautiful and mysterious world of Microbiology.

I also want to thank my classmates in the Joint Program, my badminton and tennis friends around Rutgers for your belief in me. I really had a good time hanging out with you all.

Last but not least I want to thank my family. Thanks to my parents for bringing me strength and support. Thanks to Xin Cai, my lovely and beautiful wife and my best friend, for your love. Thank you so much for taking good care of me and sharing with me all my happiness and sadness.

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Tables

1. Introduction

1.1 Contamination and Toxicity of Phthalates

Phthalate esters, also called phthalates (Figure 1.1), have been widely used to increase flexibility, extensibility, and workability of plastic polymers as plasticizer (42), or to help make perfume fixatives, lubricants, adhesives, weather stripping, and safety glass (1). Many consumer products contain specific members of this family of chemicals, including building materials, household furnishings, clothing, cosmetics, pharmaceuticals, nutritional supplements, medical devices, dentures, children's toys, glow sticks, modeling clay, food packaging, automobiles, lubricants, waxes, cleaning materials and insecticides (103).

Because of their widespread use, phthalates have become one of the most abundant industrial pollutants in the environment. Phthalates may enter the environment through industrial waste water during manufacturing and processing, or by leaking or evaporation from container or landfill site (6). Consumer products containing phthalates can result in human exposures through direct contact and use, indirectly through leaching into other products, or general environmental contamination (103). Phthalates have been detected not only the

environment, but also in humans and animals (42). For example, di(2-ethylhexyl) phthalate DEHP has been found in at least 737 of the 1,613 current or former NPL (National Priorities List) sites, which are identified by EPA and targeted for long-term federal cleanup activities (7). Some phthalates, such as dibutyl phthalate, di(2-ethylhexyl) phthalate (DEHP), and dimethyl phthalate have been classified as toxic compounds in EPA toxic chemical list (36).

High levels of monoester metabolites of phthalates (monoethyl phthalate, monobutyl phthalate, and monobenzyl phthalate) have been detected in human urine samples, which indicates exposure to diethyl phthalate, dibutyl phthalate, and benzylbutyl phthalate (12). To address the impact of phthalates on human beings and natural environment, multiple biological systems have been used in toxicological studies. Standardized toxicity tests using the bacteria *Vibrio fischeri*, the green algae *Pseudokirchneriella subcapitata*, and the crustacean *Daphnia magna* showed that phthalate diesters are more toxic than the corresponding monoesters, including monomethyl, monoethyl, monobutyl, monobenzyl, mono(2-ethylhexyl), and monodecyl phthalate (52). In anaerobic reactors, high levels of DEHP or other recalcitrant phthalates are likely to affect the microbial populations and then compromise methanogenesis and removal of biodegradable phthalates in sludge digesters (3). The toxicity of phthalates in

animal models has been reviewed by different authors (49, 70, 71). Phthalates adversely affect the male, female, and infant reproductive systems in animals, inducing hypospadias, cryptorchidism, reduced testosterone production, decreased sperm counts, and sperm DNA damage (49, 70, 71). It has been documented that in utero phthalate exposure is associated with a shorter duration of pregnancy (64). Phthalate effects are much more severe after in utero than adult exposure (70).

However, although these findings raise concern about the safety of phthalate exposure for pregnant women and infants, the toxicity of phthalates on human beings is still being debated. An expert panel convened by the U.S. government claimed that there was insufficient evidence in humans that DEHP exposure during pregnancy, childhood, or adulthood caused any harm (53). They also found that the small number of subjects and possible confounding factors limited the usefulness of several other new human studies (53). Thus, large scale studies need to be done for a better understanding of the potential toxicity of phthalates on humans.
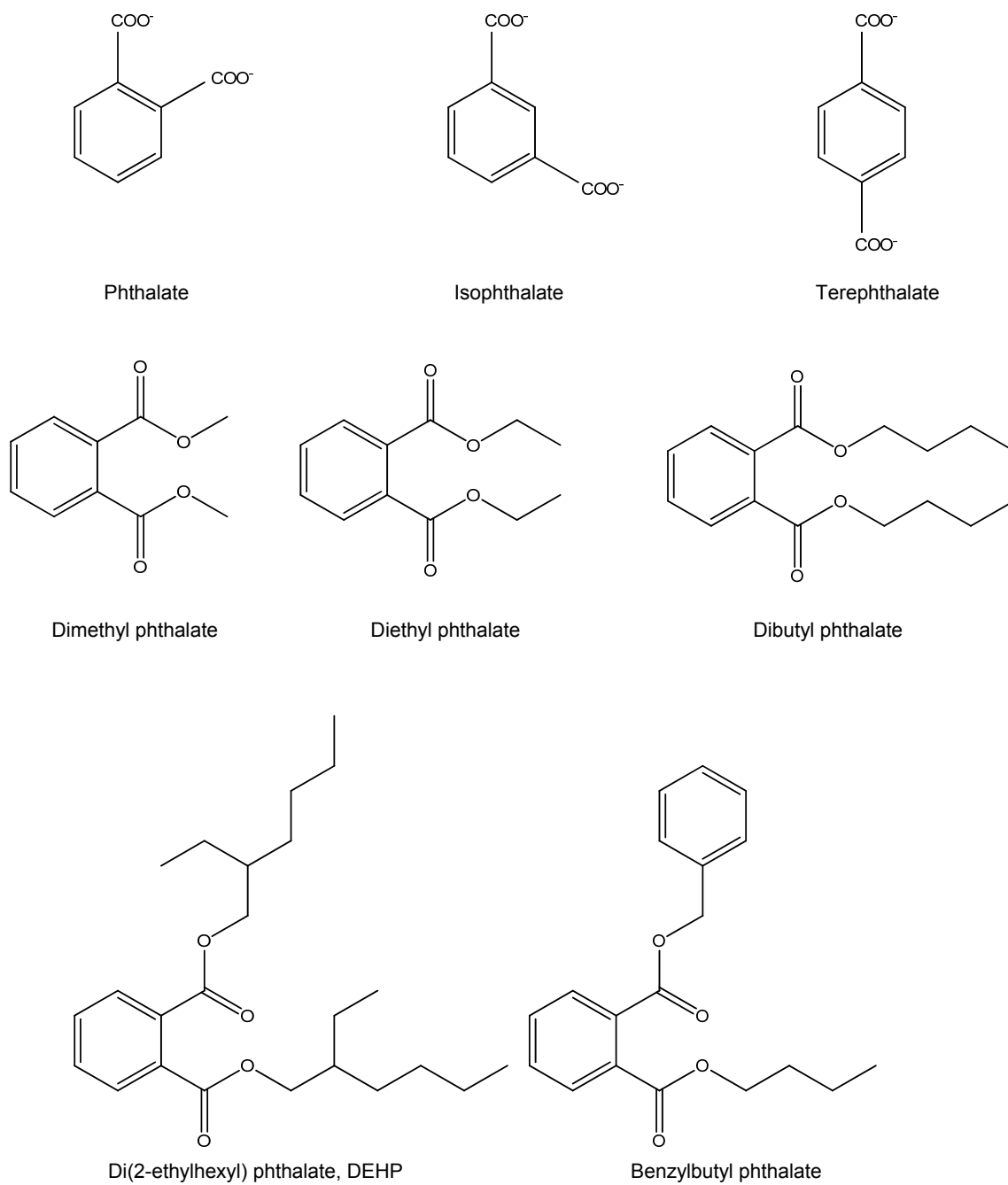
Figure 1.1 The molecular structures of phthalate isomers and some important phthalate esters.

1.2 Microbiology of Phthalate Degradation

1.2.1 Aerobic phthalate degradation

Biodegradation of phthalate esters in natural ecosystems, sewage, and laboratory cultures have been reviewed and bacteria are considered as major elements in phthalate biodegradation (57). Phthalate esters and intermediates of their degradation readily undergo ultimate degradation in different mixed microbial systems including unacclimated river water and acclimated activated sludge samples (99).

Phthalate esters could be hydrolyzed by an esterase to the corresponding monoesters and free phthalate, and then the latter could be metabolized via protocatechuate by ring cleavage (35). Recently, an esterase that specifically hydrolyzes medium-chain (C3-C5) monoalkyl phthalates has been purified from the phthalate degrading strain *Micrococcus* sp. strain YGJ1 (76). In addition, mono-2-ethylhexyl phthalate hydrolase has been purified from *Gordonia* sp. strain P8219 and it effectively hydrolyzed monophthalate esters, such as monoethyl, monobutyl, monohexyl, and mono-2-ethylhexyl phthalate (85).

Besides phthalate as the product of phthalate ester hydrolysis, phthalate has been identified as one of metabolic intermediates during polycyclic aromatic hydrocarbon (PAH) degradation in different bacteria, such as fluorene degradation in *Terrabacter* sp. strain DBF63(46, 47), phenanthrene degradation in *Comamonas testosteroni* GZ38A (41) and a *Bacillus* sp. strain (28), *Sinorhizobium* sp. strain C4 (56), and *Nocardioides* sp. strain KP7 (101), anthracene degradation by *Mycobacterium* sp. strain LB501T (120), 2-naphthoate degradation in *Burkholderia* sp. strain JT1500 (80), high molecular weight PAH pyrene degradation in *Mycobacterium vanbaalenii* PYR-1 (58), and pyrene and fluoranthene degradation in another *Mycobacterium* strain AP1 (69, 121). During pyrene degradation, *Mycobacterium* sp. strain AP1 attacks pyrene by either monooxygenation or dioxygenation at its C-4, C-5 positions to form *trans-* or *cis*-4,5-dihydroxy-4,5-dihydropyrene, respectively. In the dioxygenation pathway, *cis*-4,5-dihydroxy-4,5-dihydropyrene is degraded via phthalate by dehydrogenation, *ortho*-cleavage, decarboxylation (121). In strain AP1, fluoranthene degradation pathways involve dioxygenation at C-1 and C-2, or C-7 and C-8, which are followed by multiple steps to produce common intermediate benzene-1,2,3-tricarboxylic acid. Benzene-1,2,3-tricarboxylic finally forms phthalate by decarboxylation (69).

So far, two different phthalate degradation pathways have been identified in Gram negative and Gram positive bacteria (Figure 1.2). Gram negative and Gram positive bacteria transform phthalate via 4,5-dihydroxyphthalate and 3,4-dihydroxyphthalate respectively, and finally merge into the common protocatechuate catabolic pathway (15, 47). In Gram negative bacteria, phthalate is transformed into 4,5-dihydro-4,5-dihydroxyphthalate by phthalate dioxygenase, a two-component enzyme consisting of phthalate dioxygenase reductase and phthalate dioxygenase oxygenase. Phthalate dihydrodiol dehydrogenase transforms *cis*-phthalate dihydrodiol to 4,5-dihydroxyphthalate, which is followed by decarboxylation of 4,5-dihydroxyphthalate to form protocatechuate by 4,5-dihydroxyphthalate decarboxylase (15). While in Gram positive bacteria, phthalate is transformed to protocatechuate through *cis*-3,4-dihydroxy-3,4-dihydrophthalate and 3,4-dihydroxyphthalate (31) (Figure 1.2). The enzymes of the two pathways, such as reductive dioxygenases, dihydrodiol dehydrogenases, and decarboxylases, are not closely related, even though they catalyze similar reactions (31).
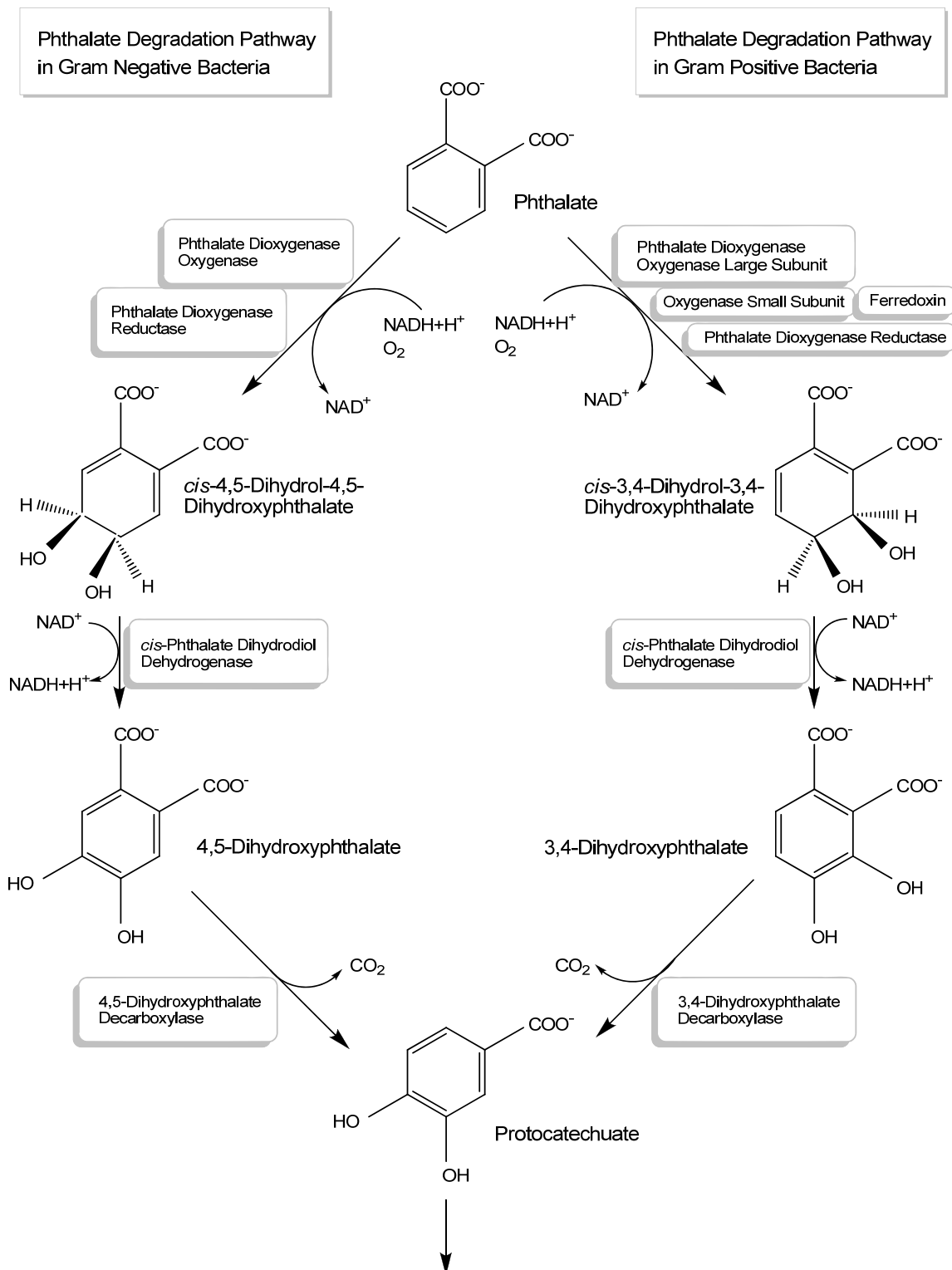
Figure 1.2 Phthalate degradation pathways in Gram negative bacteria (left branch) and Gram positive bacteria (right branch). The names of the intermediates and the enzymes in phthalate degradation pathways are shown in the figure.

Many microorganisms have been isolated from diverse sources for their ability to degrade phthalate aerobically (15) or anaerobically (60). Usually phthalate degrading bacteria belong to different genera including Gram negative bacteria *Burkholderia* (15), *Comamonas* (104, 123), *Delftia* (106), *Pseudomonas* (86), and Gram positve bacteria *Arthrobacter* (31), *Terrabacter* (47), *Mycobacterium* (110), *Rhodococcus* (21, 90) etc. For example, *Micrococcus* sp. strain 12B (renamed as *Arthrobacter keyseri* 12B(31)) was isolated by enriching for growth with dimethylphthalate and dibutylphthalate as the sole carbon and energy source (32, 33). Wang et al. isolated nine phthalate degrading bacterial strains from Passaic River sediments in New Jersey (123). They belonged to *Acinetobacter*, *Comamonas*, *Pseudomonas* or *Arthrobacter*. They have different capabilities to use phthalate, isophthalate, and/or terephthalate as the sole carbon and energy source for growth, even though they were isolated from the exact same location (123). A terephthalate-assimilating bacterium *Delftia tsuruhatensis* T7 was isolated from activated sludge collected from a domestic wastewater treatment plant in Japan. Besides terephthalate, T7 strain could utilize isophthalate and protocatechuate as well (107).

1.2.2 Anaerobic phthalate degradation

Although less understood than aerobic phthalate degradation, a few anaerobic phthalate degradation studies have been done under different anaerobic conditions. Liang et al. reported that under denitrifying conditions, dimethyl phthalate was first degraded to monomethyl phthalate (MMP), which was degraded to phthalate before complete mineralization (67). Microbial community in the denitrifying sludge was mainly composed of β-, α-*Proteobacteria* and some *Acidobacteria* (67). Aftring et al. showed that phthalate isomers were biodegradable under anaerobic conditions (2). Each enrichment culture used only one phthalate isomer under anaerobic conditions, but they were simultaneously adapted for the anaerobic catabolism of benzoate (2). In addition, Nozawa and Maruyama reported benzoate was detected as predominant metabolic intermediates and acyl-CoA synthetase activities for phthalates and benzoate were detected in a denitrifying strain *Pseudomonas* sp. strain P136 during phthalate isomer metabolism (87). Thus, the phthalate anaerobic degradation pathway involved phthalate isomer decarboxylation resulting in the formation of benzoyl CoA (87) (Figure 1.3).

Under sulfate-reducing conditions, dimethyl phthalate (DMP), dimethyl

isophthalate (DMI) and dimethyl terephthalate (DMT) could only be transformed to the monomethyl phthalate and/or phthalic acid, but could not be mineralized over 6 months (19). Some facultative anaerobes belonging to *Thauera* sp., *Xanthobacter* sp., and *Agrobacterium* sp. were identified as dominant bacterial species utilizing individual DMPE isomers (19).

Moreover, Battersby and Wilson studied the degradation potential of phthalates in an anaerobic digesting sludge under methanogenic conditions (10). This study showed that phthalic acid was completely degraded, while dimethyl phthalate was partially degraded under methanogenic conditions (10). Phthalate degradation under methanogenic condition has been proposed to proceed by syntrophic association between different physiological groups of anaerobes (60, 97) including fermentative bacteria, hydrogenotrophic and aceticlastic methanogens (97). Kleerebezem et al. recently reported three methanogenic consortia decomposing phthalate, isophthalate, or terephthalate (60). Cultures grown on one of the phthalate isomers were not capable of degrading the other phthalate isomers. All three cultures had the ability to degrade benzoate (60). Since the reaction performed by the phthalate degrading bacteria is energetically unfavorable under methanogenic condition, the presence of hydrogenotrophic and aceticlastic methanogens is thought to be essential to make the whole

reaction energetically feasible (97).

A few phthalate degrading anaerobes have been isolated and characterized under different anaerobic conditions. *Pelotomaculum terephthalicum* JT (96, 97) and *Pelotomaculum isophthalicum* JI (96) were isolated from anaerobic sludge treating wastewater from the manufacturing of terephthalic acid (97). Mixed cultures of *P. terephthalicum* JT or *P. isophthalicum* JI with *Methanospirillum hungatei* converted phthalate isomers or other aromatic compounds to acetate and methane, but they could not degrade these compounds in pure culture (96, 97). In coculture, the JT strain could degrade isophthalate and terephthalate, while JI strain could degrade all of three phthalate isomers (96).

A thermophilic terephthalate degrading methanogenic consortium was analyzed using terminal RFLPs, clone libraries and fluorescence in-situ hybridization with rRNA-targeted oligonucleotide probes (18). In this consortium, *Methanothrix thermophila*-related methanogens and *Desulfotomaculum*-related bacterial populations were the key members responsible for terephthalate degradation under thermophilic methanogenic conditions (18). Apparently, the dominant populations in the thermophilic hybrid reactor were different from those in mesophilic reactors (18, 97, 126), where Δ-*Proteobacteria* was the major

bacterial population, and *Methanosaeta* and *Methanospirillum*-related species were dominant anchaea populations (126). In another mesophilic reactor, the predominant bacteria were affiliated with a nonsulfate-reducing subcluster in the group '*Desulfotomaculum* lineage I' or a clone cluster in Δ-*Proteobacteria* (97).
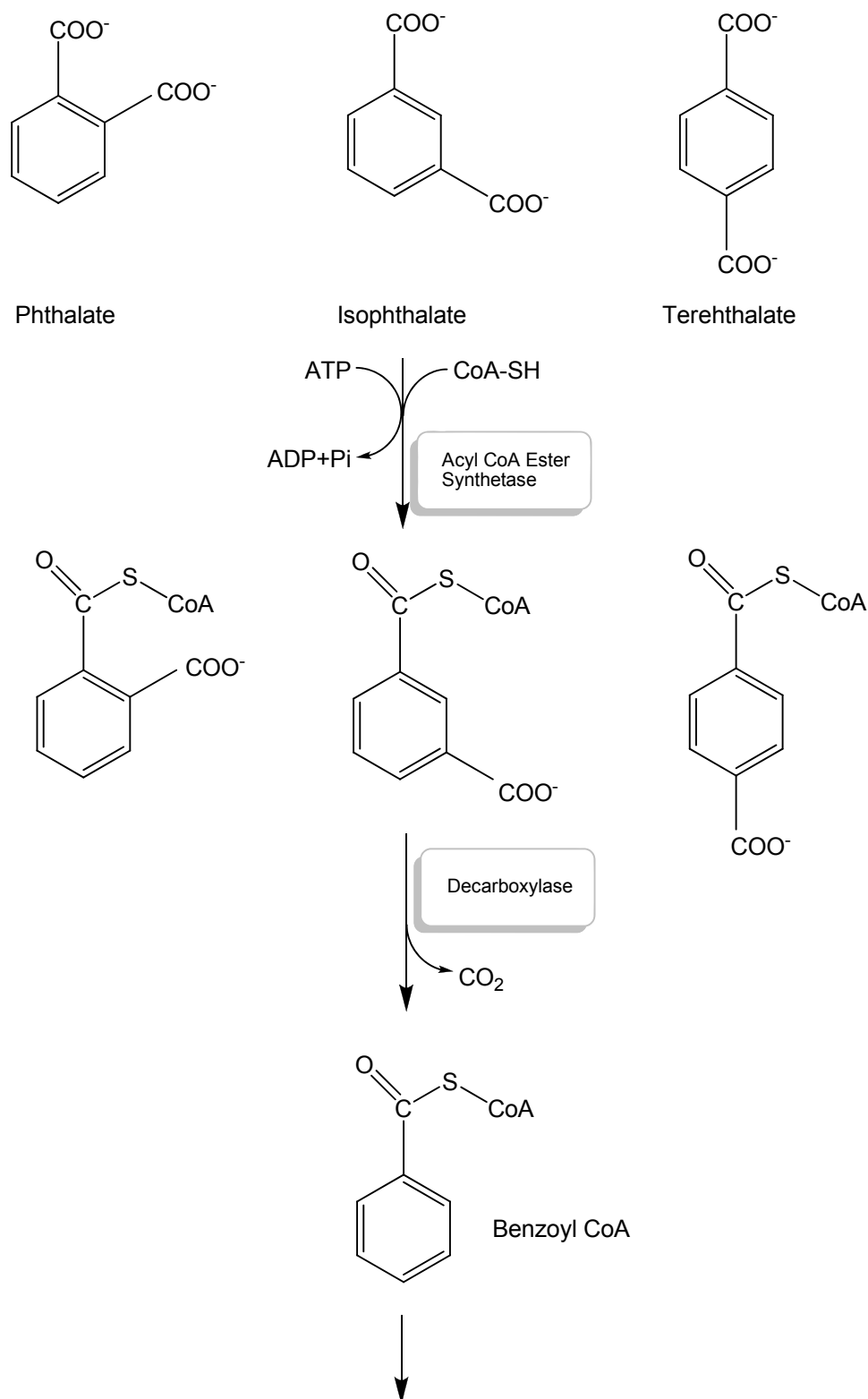
Figure 1.3 Phthalate, isophthalate and terephthalate degradation pathways in anaerobic bacteria. The names of the intermediates and the enzymes in the phthalate, isophthalate, and terephthalate degradation pathways are shown in the figure.

1.3 Molecular Basis of Phthalate Degradation

1.3.1 Phthalate degradation in Gram positive bacteria

The Gene organization for phthalate degradation has been revealed in several Gram positive bacteria, which include *Arthrobacter keyseri* 12B (31), *Terrabacter* sp. strain DBF63 (47), *Mycobacterium vanbaalenii* PYR-1 (110), *Rhodococcus* sp. strain DK17 (21), *Rhodococcus* sp. strain RHA1 (90) (Figure 1.4).

In *Arthrobacter keyseri* 12B, the *pcm* operon encoding protocatechuate catabolism, *pehA* encoding a possible phthalate ester hydrolase, and the *pht* operon (*phtBAaAbAcAdCR*) encoding the conversion of phthalate to protocatechuate are located in a 26,274 bp contiguous region (31). Activities of the eight enzymes involved in the catabolism of phthalate through protocatechuate to pyruvate and oxaloacetate were demonstrated in cells or cell extracts of recombinant *E. coli* strains (31). Phthalate catabolism gene expression is inducible by phthalate in *A. keyseri* 12B (31).

The phthalate degradative genes (*phtA1A2BA3A4CR*) of *Terrabacter* sp. strain

DBF63 encode the oxygenase large subunit of phthalate 3,4-dioxygenase (*phtA1*), oxygenase small subunit of phthalate 3,4-dioxygenase (*phtA2*), *cis*-3,4-dihydroxy-3,4-dihydrophthalate dehydrogenase (*phtB*), [3Fe-4S] or [4Fe-4S] type ferredoxin (*phtA3*), ferredoxin reductase (*phtA4*), 3,4-dihydroxyphthalate decarboxylase (*phtC*) and putative regulatory protein (*phtR*) (47). *pht* genes were found in the upstream region of the fluorine degradation gene (*flnRB-dbfA1A2-flnED1*-ORF16) (46, 47). Thus, the phthalate degradation pathway in DBF63 could be considered downstream of fluorene degradation pathway (46, 47).

The *Mycobacterium vanbaalenii* PYR-1 *pht* operon contains *phtR-phtAaAbBAcAd*. A putative regulatory protein (*phtR*) was encoded divergently with five tandem genes: phthalate dioxygenase large subunit (*phtAa*), small subunit (*phtAb*), phthalate dihydrodiol dehydrogenase (*phtB*), phthalate dioxygenase ferredoxin subunit (*phtAc*) and phthalate dioxygenase ferredoxin reductase (*phtAd*) (110). The operon differs from those of other Gram positive bacteria in both the placement and orientation of the regulatory gene (110). In addition, no decarboxylase gene has been identified in 37 kb sequenced region in *M. vanbaalenii* PYR-1 genome (110).

Rhodococcus sp. strain DK17 contains three plasmids (380-kb pDK1, 330-kb pDK2, 750-kb pDK3) (21). The phthalate and terephthalate operons are duplicated on both pDK2 and pDK3 (21). The putative phthalate operon contains the genes *ophA1A2-orf0-ophBA3A4C* encoding large and small subunits of a terminal oxygenase component of phthalate 3,4-dioxygenase, a protein of unknown function, a phthalate dihydrodiol dehydrogenase, a ferredoxin component of phthalate 3,4-dioxygenase, a ferredoxin reductase component of phthalate 3,4-dioxygenase, and a decarboxylase (21). RT-PCR and mutation analysis showed that two copies of the dihydroxyphthalate decarboxylase genes are simultaneously expressed during growth on phthalate, and both copies of the phthalate operon are equally functional in DK17 (20). The genes for the large and small subunits of terephthalate 1,2-dioxygenase (*tphA1-tphA2*) are located approximately 6.7 kb away from phthalate oxygenase gene *ophA1* and are transcribed in opposite directions (21). Both the phthalate and terephthalate operons are induced in response to terephthalate while expression of the terephthalate genes is undetectable in phthalate-grown cells (21). In addition, benzoate inhibits DK17 phthalate metabolism (22).

Rhodococcus sp. strain RHA1 genome sequencing has been completed (77) and identical copies of the phthalate degradative gene cluster *pad* were found on

plasmids pRHL1 and pRHL2 in RHA1 (90). The phthalate cluster contains seven genes encoding a regulatory protein (*padR*) and the enzymes that transform phthalate to protocatechuate: a 3,4-dioxygenase (*padAaAbAcAd*), a dehydrogenase (*padB*), and a decarboxylase (*padC*) (90). These ORFs share 99% sequence identity with phthalate degrading genes in *Rhodococcus* sp. strain DK17 (90). RHA1 strain contains not only phthalate gene but also terephthalate degradation genes (48, 77). Transcriptome analysis revealed that in RHA1 phthalate was degraded solely via the protocatechuate pathway, whereas terephthalate was degraded via a bifurcated pathway that additionally includes the catechol branch of the protocatechuate pathway (48).

1.3.2 Phthalate degradation in Gram negative bacteria

Most studies of Gram negative bacterial phthalate degradation have been done on *Burkholderia cepacia* DBO1 by Chang and Zylstra (13, 15, 16) (Figure 1.5). The phthalate degradative gene cluster (*oph*) of *Burkholderia cepacia* DBO1 has been identified in at least three operons *ophA1-ophDC-ophR-ophE-tnp-ophBA2*, which encode for phthalate dioxygenase reductase (*ophA1*), an inactive transporter (*ophD*), 4,5-dihydroxyphthalate decarboxylase (*ophC*), phthalate

dioxygenase oxygenase (*ophA2*) and *cis*-phthalate dihydrodiol dehydrogenase (*ophB*). *ophA1* and *ophDC* are adjacent to each other but are transcribed in opposite directions, while *ophBA2* is located 4 kb away. A transposase gene (*tnp*) was inserted between *ophE* and *ophB* (15).

The *ophD* gene of DBO1 strain contains a frameshift mutation in contrast to genes for other permeases (15). In addition, the *ophD* gene of another phthalate degrader *Burkholderia cepacia* ATCC 17616 has been cloned and expressed (13). *ophD* knockout mutants of DBO1 and 17616 were still able to take up phthalate at rates equivalent to that of the wild-type strain which suggested that they must have a second phthalate-inducible phthalate uptake system (13, 15).

The *ophE* gene in the DBO1 strain encodes for quinolinate phosphoribosyl transferase, which is involved in the biosynthesis of NAD+ (16). DBO1 and other phthalate degrading *B. cepacia* strains have two dissimilar genes for this enzyme, while non-phthalate degrading *B. cepacia* strains have only a single gene (16). Insertional knockout mutants lacking *ophE* grow noticeably slower on phthalate while exhibiting normal rates of growth on other substrates (16). The recruitment of the *ophE* gene thus gives *B. cepacia* an advantage for growing on phthalate (16).

Besides the *ophD* gene, another 4-methylphthalate permease gene *mopB* has been identified in *Burkholderia cepacia* Pc701 and its specific function for transporting 4-methylphthalate and phthalate has been proven by substrate uptake assay (100). Moreover, Northern hybridization experiments showed that *mopB* gene expression was only seen in cells grown on 4-methylphthalate and not in cells grown on phthalate (100).

Similar organization of phthalate degradative genes was also identified in *Burkholderia vietnamiensis* G4, *Burkholderia multivorans* ATCC 17616, and *Ralstonia pickettii* 12D. Comparing with strain DBO1, strains G4 and ATCC 17616 have a complete open reading frame for OphD. Phthalate gene cluster in *R. pickettii* 12D contains a porin-encoding gene *ophP*, but the transposase gene in *Burkholderia* strains is not contained in phthalate cluster in strain 12D.

In *Pseudomonas putida*, the phthalate genes (*pth12345*) are encoded by the plasmid pNMH102-2 and are organized in a single operon with the same orientation (86). In addition, the 4,5-dihydroxyphthalate decarboxylase gene *phtD* from *Pseudomonas testosteroni* M4-1 (renamed as *Comamonas testosteroni*) were isolated by transposon insertion mutagenesis (66). The putative amino acid sequence of *phtD* has 77.6% homology with that of the *pht*5

gene from *Pseudomonas putida* (66, 86) (Figure 1.5).

1.3.3 Isophthalate and terephthalate degradation

The genes responsible for isophthalate and terephthalate degradation have been cloned and sequenced from *Comamonas testosteroni* YZW-D, one of the nine bacterial strains isolated from river sediment in New Jersey (123) (Figure 1.6 and 1.7). The isophthalate degradative gene cluster (*iphA2CBA1R*) encodes genes for a dioxygenase oxygenase, transport protein, dehydrogenase, dioxygenase reductase and a regulatory protein. Isophthalate dioxygenase oxygnease and dioxygenase reductase transform isophthalate to *cis*-isophthalate dihydrodiol, which is further transformed to protocatechuate by *cis*-isophthalate dihydrodiol dehydrogenase (*iphB*) (123) (Figure 1.6 and 1.8).

Interestingly, the terephthalate gene operon (*tphRCA2A3BA1*) in *C. testosteroni* YZW-D was found to be adjacent to the *iph* genes. The terephthalate degradative gene operon In YZW-D encodes genes for a regulatory protein, transport protein, dioxygenase oxygenase large subunit, dioxygenase oxygenase small subunit, dehydrogenase, and reductase (123). Similar to isophthalate

degradation, *cis*-terephthalate dihydrodiol is the product of terephthalate oxygenation by terephthalate dioxygenase, a two-component dioxygenase system consisting of dioxygenase oxygenase large subunit, dioxygenase oxygenase small subunit, dioxygenase reductase. Finally, *cis*-terephthalate dihydrodiol is transformed to protocatechuate (123) (Figure 1.7 and 1.8).

Another terephthalate degrader *Comamonas* sp. strain E6 has two almost identical terephthalate gene clusters, *tphRICIA2IA3IBIA1I* and *tphRIICIIA2IIA3IIBIIA1II* (102). Mutagenesis analysis showed that both terephthalate dioxygenase genes are involved in terephthalate degradation(102). Terephthalate 1,2-dioxygenase systems have been purified from *Comamonas testosteroni* T-2 (104) and *Delftia tsuruhatensis* T7 (106).

In addition, the terephthalate gene operon has been identified in the Gram positive bacterium *Rhodococcus* sp. strains DK17 *tphRA2A3BA1C* (21)and RHA1 (*tpaAaAbBCK*) (90). Comparing to the YZW-D *tph* gene operon, the *Rhodococcus* transporter gene (*tphC* of DK17 or tpaK of RHA1) is located in different positions in the two operons and does not share any similarity with *tphC* of YZW-D.
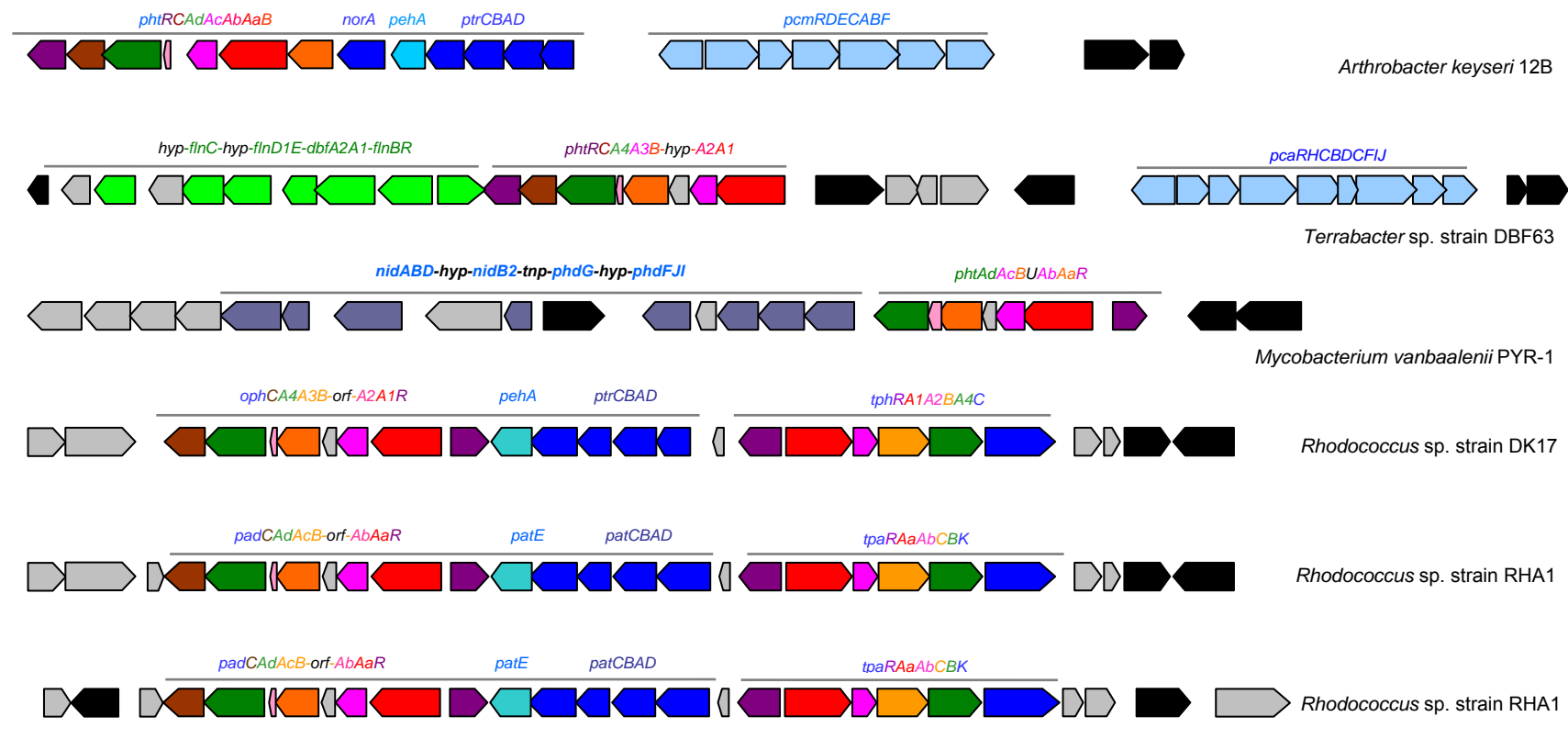
Figure 1.4 Gene organization of the phthalate degradation operon/cluster in Gram positive bacteria. The gene map for each bacterial strain is based on the Genbank sequences. In *Arthrobacter keyseri* 12B (AF331043), *phtRCAdAcAbAaB*, *pehA*, *ptrCBAD* and *pcmRDECABF* encode the phthalate degradation pathway, phthalate esterase, transporter proteins, and the protocatechuate degradation pathway. In *Terrabacter* sp. strain DBF63 (AP008980), *flnCD1EdbfA2A1flnBR*, *phtRCA4A3BA2A1* and *pcaRHCBDCFIJ* encode the fluorine, phthalate and protocatechuate degradation pathways. In *Mycobacterium vanbaalenii* PYR-1 (AY365117), *nidABDB2-phdGFJI* and *phtAdAcBUAbAaR* encode the phenanthrene and phthalate degradation pathways. In *Rhodococcus* sp. strain DK17 (AY502076), *ophCA4A3BA2A1R*, *pehA*, *ptrCBAD*, and *tphRA1A2BA4C* encode the phthalate degradation pathway, phthalate esterase, transporter proteins, and the terephthalate degradation pathway. In *Rhodococcus* sp. strain RHA1 (NC_008269 and NC_008270), there exist two identical copies of *padCAdAcB-orf-AbAaR*, *patE*, *patCBAD*, and *tpaRAaAbCBK* which encode the phthalate degradation pathway, phthalate esterase, transporter protein, and the terephthalate degradation pathway. Hypothetical proteins and transposase genes are shown in gray and black, respectively.

23
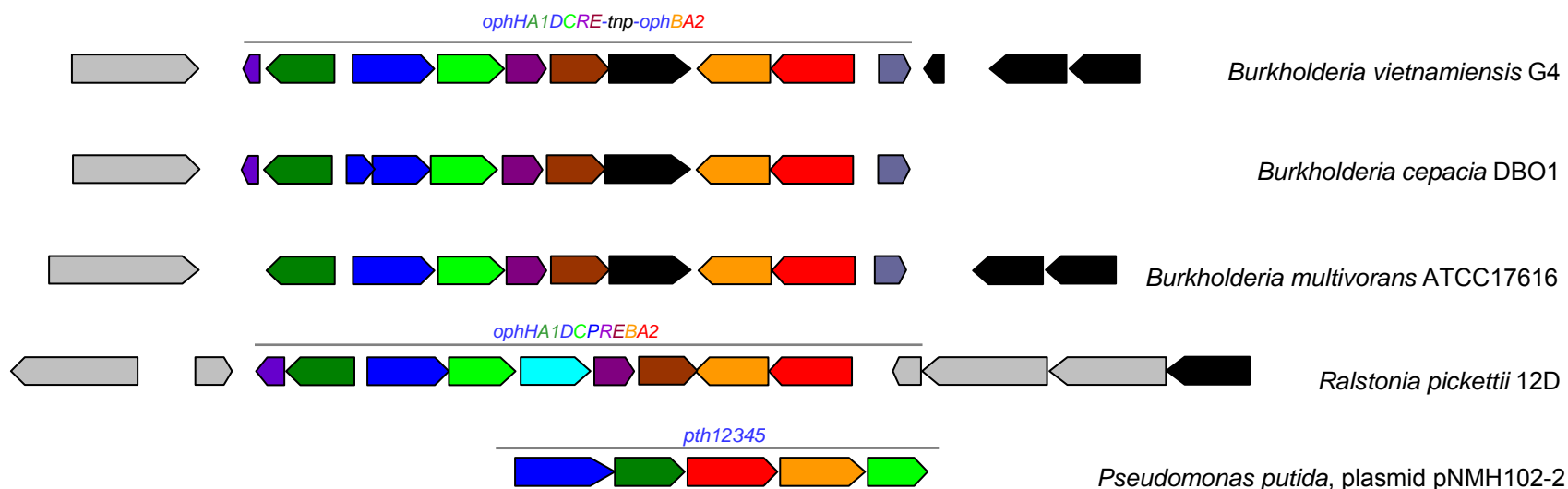
Figure 1.5 Gene organization of the phthalate degradation operon/cluster in Gram negative bacteria. The gene map for each bacterial strain is based on the Genbank sequences. In *Burkholderia vietnamiensis* G4 (NC_009255), *Burkholderia cepacia* DBO1 (AF095748), *Burkholderia multivorans* ATCC 17616 (NZ_AAVB01000002), *Ralstonia pickettii* 12D (NZ_ABDZ01000008), phthalate degradative genes *ophHA1-ophDCRE-ophBA2* encode for a hemerythrin, phthalate dioxygenase reductase, transporter, 4,5-dihydroxyphthalate decarboxylase, regulatory protein, quinolinate phosphoribosyltransferase, *cis*-phthalate dihydrodiol dehydrogenase and phthalate dioxygenase oxygenase. There exists a porin-encoding gene *ophP* in *Ralstonia pickettii* 12D. In *Pseudomonas putida* (D13229), the phthalate genes *pth12345* encode a transporter, phthalate dioxygenase reductase, phthalate dioxygenase oxygenase, *cis*-phthalate dihydrodiol dehydrogenase, and 4,5-dihydroxyphthalate decarboxylase. Hypothetical proteins and transposase genes are shown in gray and black, respectively.
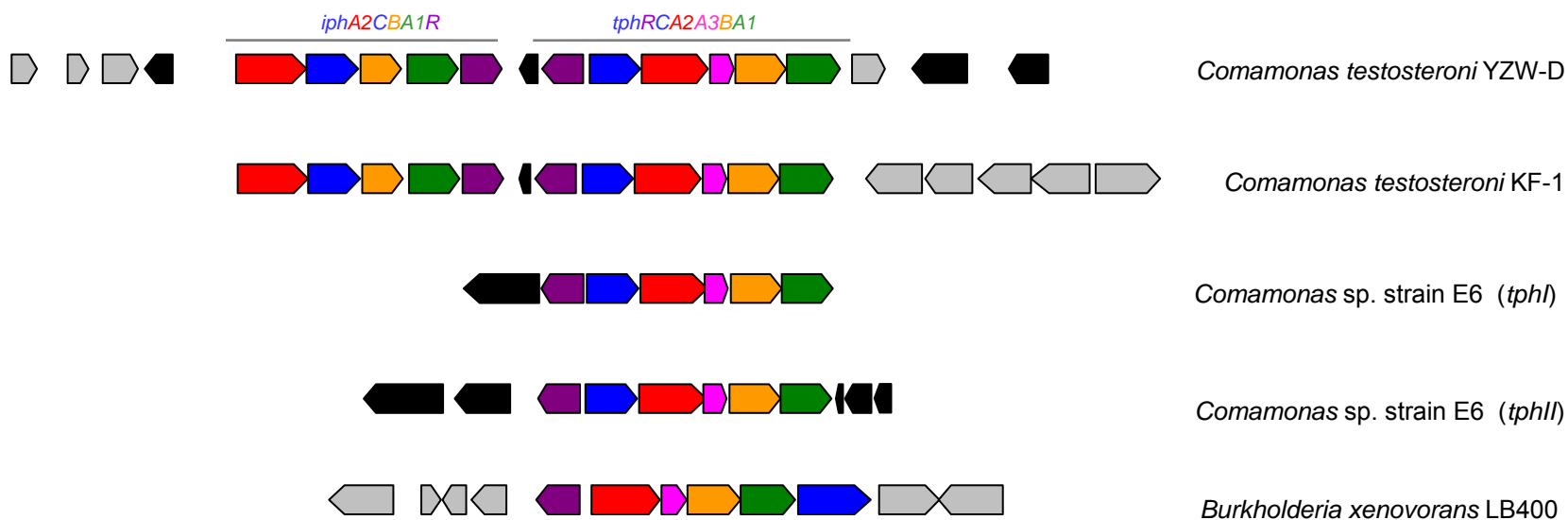
Figure 1.6 Gene organization of the isophthalate and terephthalate degradation operon/cluster in Gram negative bacteria. The gene map for each bacterial strain is based on the Genbank sequences. The isophthalate degradative genes *iphA2CBA1R* encode for an isophthalate dioxygenase oxygenase, transporter protein, dehydrogenase, isophthalate dioxygenase reductase, and regulatory protein. Terephthalate degradative genes encode for a regulatory protein, transporter protein, terephthalate dioxygenase oxygenase large subunit, dioxygenase oxygenase small subunit, dehydrogenase, and terephthalate dioxygenase reductase. In *Comamonas testosteroni* YZW-D (AY923836) and *Comamonas testosteroni* KF-1 (NZ_AAUJ01000006), the *iph* genes and *tph* genes are physically linked. *Comamonas* sp. strain E6 contains two copies of the *tph* gene operon, *tphI* (AB238678) and *tphII* (AB238679). A gene organization similar to the *tph* genes in Gram positive bacteria was identified in *Burkholderia xenovorans* LB400 (NC_007952). Hypothetical proteins and transposase genes are shown in gray and black, respectively.
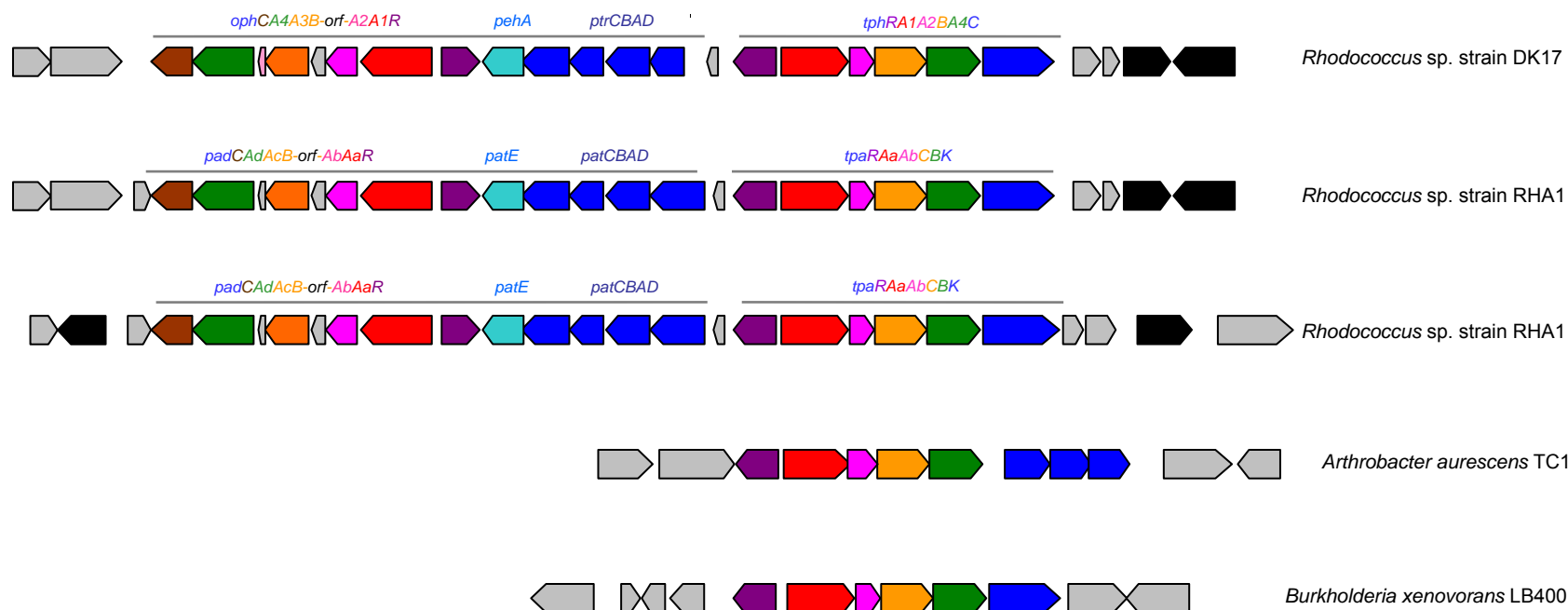
Figure 1.7 Gene organization of the terephthalate degradation operon/cluster in Gram positive bacteria. The gene map for each bacterial strain is based on the Genbank sequences. The terephthalate degradative genes in *Rhodococcus* sp. strains DK17 and RHA1 are linked with the phthalate and phthalate ester degradative genes. In *Rhodococcus* sp. DK17 (AY502076), *ophCA4A3BA2A1R*, *pehA*, *ptrCBAD*, and *tphRA1A2BA4C* encode the phthalate degradation, phthalate esterase, transporter proteins, and the terephthalate degradation pathway. In *Rhodococcus* sp. strain RHA1 (NC_008269 and NC_008270), there exist two identical copies of *padCAdAcB-orf-AbAaR*, *patE*, *patCBAD*, and *tpaRAaAbCBK* which encode the phthalate degradation, phthalate esterase, transporter protein, and the terephthalate degradation pathway. The terephthalate degradative genes (*tphRA1A2BA4C* in DK17 or *tpaRAaAbCBK* in RHA1) encode for a regulatory protein, transporter protein, terephthalate dioxygenase oxygenase large subunit, dioxygenase oxgenase small subunit, dehydrogenase, and terephthalate dioxygenase reductase. A similar gene organization was identified in *Arthrobacter aurescens* TC1 (NC_008712) and Gram negative *Burkholderia xenovorans* LB400 (NC_007952). Hypothetical proteins and transposase genes are shown in gray and black, respectively.

26
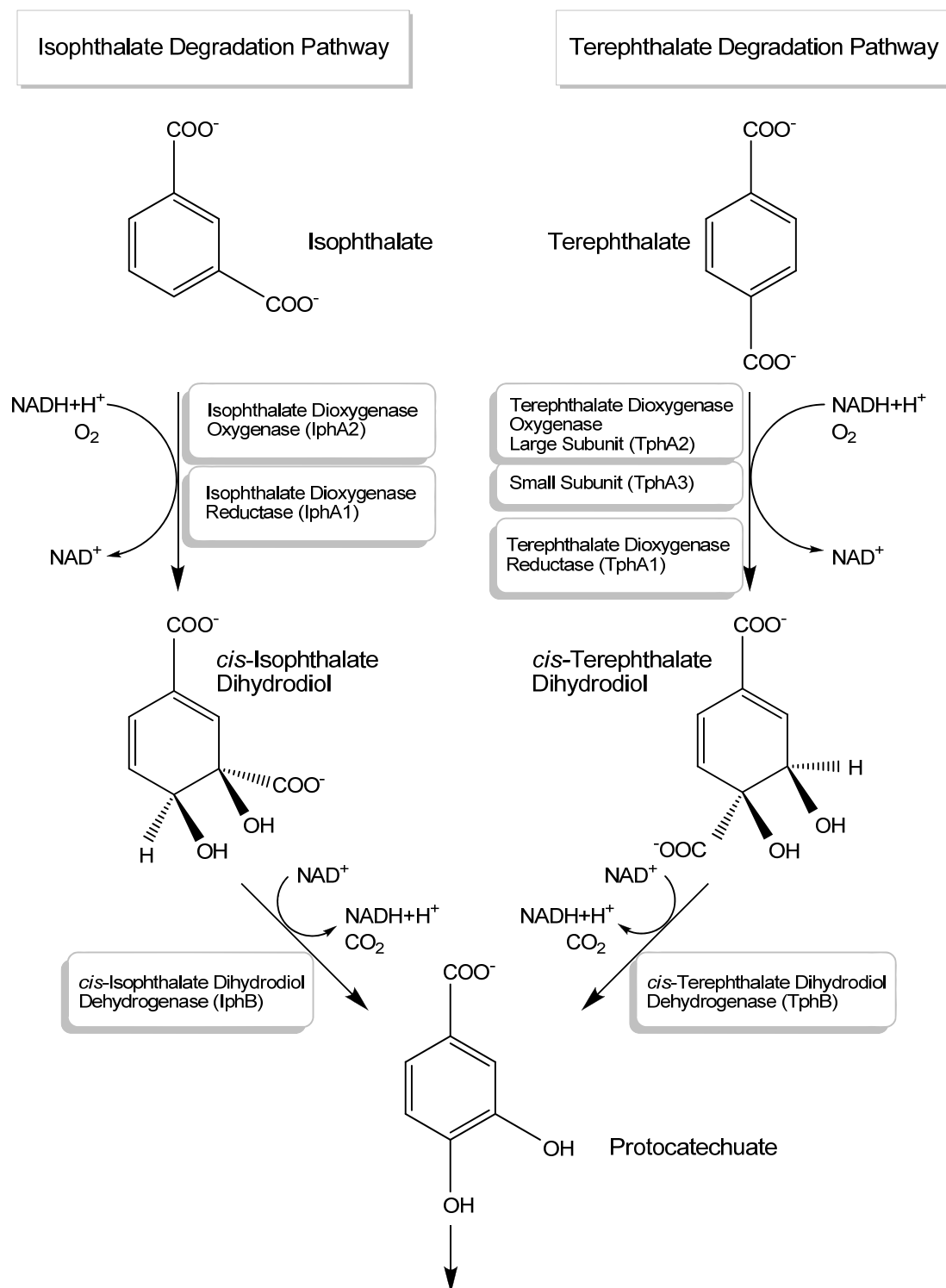
Figure 1.8 Isophthalate (left branch) and terephthalate (right branch) degradation pathways in bacteria. The names of the intermediates and the enzymes in the isophthalate and terephthalate degradation pathway are shown in the figure.

1.4 Biochemical Mechanisms of Phthalate Degradation

1.4.1 Phthalate dioxygenase

Phthalate dioxygenase and its reductase are parts of a two-component Rieske dioxygenase system that initiates the aerobic breakdown of phthalate by forming *cis*-4,5-dihydro-4,5-dihydroxyphthalate (113). Most of biochemical studies on the phthalate dioxygenase have been done by the Ballou group. Phthalate dioxygenase and phthalate dioxygenase reductase from *Pseudomonas cepacia* (renamed as *Burkholderia cepacia* DB01) was purified and characterized by Batie et al. in 1987 (8, 9). This system contains two proteins: phthalate dioxygenase reductase (PDR), a flavo-iron-sulfur protein with NADH-dependent oxidoreductase activity, and phthalate dioxygenase oxygenase (PDO), a nonheme iron protein with oxygenase activity (9). Detailed structure of phthalate dioxygenase has been produced using different techniques, such as Electron-nuclear double resonance spectroscopy (45), X-ray absorption spectroscopy (93, 118), Magnetic circular dichroism (38, 92), Electron paramagnetic resonance measurements (26), and NMR (116). The crystal structure of oxidized and reduced PDR from *B. cepacia* DB01 has been analyzed at 2.0 and 2.7 angstrom

resolution, respectively (25).

The PDO contains two iron-based sites, a Rieske-type [2Fe-2S] cluster serving as an electron-transferring cofactor, and a mononuclear iron site for substrate oxygenation (26). PDR is folded into three domains, the N-terminal FMN binding domain, central NAD(H) binding domains, and C-terminal [2Fe-2S] domain (39). Kinetic studies have identified sequential steps in the reaction of Phthalate dioxygenase reductase with NADH (39). PDR utilizes flavin mononucleotide (FMN) to mediate electron transfer from the two-electron donor NADH to the one-electron acceptor [2Fe-2S] (25). Product formation in steady-state reactions is tightly coupled to electron delivery, with 1 dihydrodiol (DHD) of phthalate formed for every 2 electrons delivered from NADH (112). The binding of phthalate or PDR to PDO each results in greater reactivity of PDO with O_2 (114). The presence of both the substrate and PDR was synergistic, making PDO fully catalytically active (114). Moreover, Aspartate D178 has been considered an important site in PDO (113). Mutation with D178N and D178A disrupted the interactions between the Rieske and mononuclear centers in PDO, and affected protonation of the Rieske center histidine and conformation of subunits within the PDO multimer to create a more open structure with more solvent-accessible Rieske centers (113).

1.4.2 Terephthalate dioxygenase

Terephthalate dioxygenase system from *Comamonas testosteroni* T-2 (104) and *Delftia tsuruhatensis* T7 (106) has been purified and studied. *Comamonas testosteroni* T-2 synthesizes inducible terephthalate dioxygenase system (TERDOS) to convert TER to (1R,2S)-dihydroxy-3,5-cyclohexadiene-1,4-dicarboxylic acid (DCD) and protocatechuate (104). TERDOS in T-2 strain contains fraction R showing an NADH-dependent reductase activity, and Z with UV-visible spectrum and electron paramagnetic resonance characteristics of a Rieske [2Fe-2S] protein. In the presence of O_2, NADH, and fraction R, component Z catalyzed the stoichiometric transformation of TER to PC, with the intermediate formation of DCD (104). In addition, the oxygenase component of terephthalate 1,2-dioxygenase system *Delftia tsuruhatensis* T7 has been purified and predicted to form α3β3 subunit structure (106).

1.5 Background of Current Project

The present project is focused on nine phthalate, isophthalate or terephthalate degrading bacteria designated YZW-A, -B, -C, -D, -E, -F, -G, -H, and -I, which were previously isolated from river sediments and have different capabilities to

use phthalate, isophthalate, and/or terephthalate as the sole carbon and energy

source for growth (123). The genes responsible for isophthalate and

terephthalate degradation have been cloned and sequenced from *Comamonas*

testosteroni YZW-D previously (123). As we already discussed before, the

isophthalate degradative gene cluster (*iphA2CBA1R*) and terephthalate

degradative cluster (*tphRCA2A3BA1*) in strain YZW-D have been identified (123).

For better understanding of the diversity of degradative genes, it is necessary to

determine and compare gene sequence and gene organization for phthalate,

isophthalate, terephthalate degradation in different bacterial species. In Chapter

1-3, we characterized nine bacterial isolates on the basis of 16S rDNA gene

sequences and aromatic substrate utilization, and identified the presence and

divergence of the phthalate, isophthalate, and terephthalate degradation genes in

different bacterial strains.

Additionally, based on Blast (4) and conserved domain searching (73, 74), the

phthalate regulatory protein OphR of *C. testosteroni* YZW-B, -E, -F, and

Pseudomonas sp. strain YZW-G was identified as belonging to the MarR family,

but these two OphR genes have different gene organization and transcription

direction. IphR and TphR belong to IclR family. IclR regulators are generally

transcriptional repressors; but so far, all IclR in catabolic pathway are described

as activators. The IclR gene is located upstream of its target gene cluster and is transcribed in the opposite direction (117), such as *tphR* of YZW-B, -D, -E, and -F. Apparently, *iphR* is an exception. *iphR* and other *iph* genes have the same transcription direction. The mechanism of gene regulation of phthalate degradation, in terms of inducer molecules, regulator-DNA binding properties, transcription activation and repression, still remain unknown. Thus, gene regulation of phthalate degradation is another interesting and important topic since bacteria should have evolved different mechanisms to respond to different phthalates and other aromatic compounds. In Chapter 4, we clarified the mechanism of gene regulation of phthalate, isophthalate, and terephthalate genes in *C. testosteroni* YZW-B using gene knockout and quantitative real-time PCR.

2. Materials and Methods

2.1 Bacterial Strains and Growth Media

Phthalate isomers (phthalate, isophthalate, or terephthalate) degrading bacteria YZW-A, B, C, D, E, F, G, H, and I were used in this study (123). These strains were originally isolated from Passaic River sediment and have been classified as *Pseudomonas* (YZW-A, YZW-G), *Acinetobacter* (YZW-C, YZW-H), *Comamonas* (YZW-B, YZW-D, YZW-E, YZW-F) and *Arthrobacter* (YZW-I) using the Biolog identification system (123). The degradation capacity of these strains for specific phthalate isomers are shown in Table 3.1. These bacterial strains grew on Luria-Bertani (LB) agar or minimal agar plates with certain carbon sources at 30°C, or in LB broth or minimal liquid media containing certain carbon sources at 30°C with 200 rpm shaking. *E. coli* DH5α was used as host strain for DNA transformation in the cloning experiments.

LB media containing different antibiotics such as ampicillin (25, 50, 100 μg/ml), kanamycin (20, 50, 100 μg/ml), tetracycline (15, 30, 60 μg/ml), streptomycin (10, 25, 50 μg/ml), gentamycin (10, 20, 40 μg/ml), or chloramphenical (5, 10, 20 μg/ml) were used for the YZW-B antibiotics test. Usually, LB media containing

ampicillin (100 µg/ml), kanamycin (50 or 100 µg/ml), or tetracycline (10 µg/ml) were used to grow bacterial strains harboring plasmids with corresponding antibiotic resistance markers or to screen mutants. SOC medium was used for growing *E.coli* DH5α after transformation in TA cloning and growing the YZW-B strain after electroporation in gene knockout. Minimal medium containing different phthalate isomers and other soluble chemical compounds as carbon sources (20 mM) was used to determine the substrate range and degradation capability of the bacteria. When using insoluble PAHs such as naphthalene and phenathrene for substrate testing, these compounds were dissolved in ether and sprayed on the surface of the media after the bacteria were inoculated. Disappearance of PAH, formation of clear zone, and growth of a bacterial lawn indicates the capability of bacteria to degrade the PAH.

Phthalate, isophthalate, terephthalate, naphthalene, and phenathrene were purchased from Aldrich Chemical Inc. Succinate was obtained from Sigma-Aldrich Inc.

2.2 PCR and DNA Cloning

Bacterial DNA was extracted using the UltraClean™ Microbial DNA Isolation kit (MO BIO Laboratories, Inc.). Plasmid DNA was isolated using the NucleoSpin® Plasmid kit (Clontech Laboratories, Inc.). PCR reactions were performed using the GeneAmp® PCR System 9700 (Applied Biosystems). For PCR, the ReadyMix™ Taq PCR Reaction Mix (SIGMA-ALDRICH, Inc.) was used for general purposes, while Pfu Turbo® DNA polymerase (Stratagene) was used to amplify long PCR products or PCR products with high fidelity. Primers were ordered from Sigma-Genosys and prepared as 2.5 µM stock solution. PCR products were purified using the NulceoTrap® PCR purification Kit (BD BioSciences) or QIAquick® PCR purification Kit (QIAGEN Sciences, Inc.). If necessary, PCR products were TA-cloned using the pGEM®-T (or pGEM®-T easy) Vector System (Promega Corporation). DNA fragments in agarose gels were gel-purified using QIAEX II Gel Extraction Kit (QIAGEN Sciences, Inc.) or Wizard® SV Gel and PCR Clean-Up System (Promega Corporation) for further sequencing or cloning work. For general two-piece cloning, both vector and insert (2 µg for each) were double digested from two plasmids in a 50 µl reaction. After gel electrophoresis, both DNA fragments were gel-purified, eluted in 50 µl solution, and religated at 17°C overnight in a 100 µl reaction containing in 10 µl

10X buffer (Invitrogen Corporation), 40 µl vector DNA, 40 µl insert DNA, 2 µl T4 ligase (Invitrogen Corporation) (5U/µl), and 8 µl ddH$_2$O. The ligation reaction solution was precipitated by ethanol and concentrated to 20 µl. 5 µl DNA was added in 25 µl competent *E. coli* DH5α cells for transformation. Positive clones were selected on LB media containing appropriate antibiotics.

2.3 Phylogenetic Analysis of Phthalate Degrading Strains

The nearly full length 16S rDNA sequences were amplified from the nine bacterial isolates (YZW-A, -B, -C, -D, -E, -F, -G, -H, and -I) using 16S universal primers, forward 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and reverse 1522R (5'-AAGGAGGTGATCCAICCGCA-3'). The PCR products were purified and sequenced for bacterial identification and phylogenetic analysis. 16S rDNA sequences of the nine bacterial isolates were compared with the related sequences in the Genbank database (11) and the ribosomal database (RDP-II) (24). Phylogenetic trees were generated by MEGA version 3.1 (63) using 16S rDNA sequences of the 9 bacterial isolates and 22 related strains which included *Acinetobacter calcoaceticus* DSM30006 (X81661) (98), *Acinetobacter calcoaceticus* ATCC23055T (Z93434) (50), *Acinetobacter haemolyticus* DSM6962 (X81662) (98), *Acinetobacter johnsonii* DSM6963 (X81663)(98),

Acinetobacter lwoffii DSM2403 (X81665) (98), *Acinetobacter radioresistens* DSM 6976 (X81666) (98), *Acinetobacter schindleri* LUH5832 (AJ278311) (83), *Arthrobacter gangotriensis* Lz1Y (AJ606061) (44), *Arthrobacter kerguelensis* KGN15 (AJ606062) (44), *Arthrobacter psychrophenolicus* DSM15454 (AJ616763) (75), *Arthrobacter sulfureus* DSM 20167 (X83409) (61), *Comamonas aquatica* LMG 2370 (AJ430344) (124), *Comamonas denitrificans* 123 (AF233877) (43), *Comamonas koreensis* KCTC12005 (AF275377) (17), *Comamonas terrigena* IMI 359870 (AF078772) (125), *Comamonas testosteroni* ATCC11996 (M11224) (17), *Delftia tsuruhatensis* T7 (AB075017) (107), *Pseudomonas oryzihabitans* IAM1568 (D84004) (5), *Pseudomonas monteilii* CIP104883 (AF064458) (34), *Pseudomonas plecoglossicida* FPC951 (AB009457) (84), *Pseudomonas putida* DSM291T (Z76667) (79), *Pseudomonas putida* KT2440 (AE015451) (82).

2.4 Inverse PCR

The theory and application of inverse PCR has been discussed by Ochman et al. (88). Genomic DNA is cleaved using restriction enzymes that do not cut the known sequence region. After restriction digestion, the DNA fragments are religated under self-ligation favorable conditions. The self-ligation products are

used as templates for PCR amplification. PCR primers are designed to locate at the both ends of the known region and amplify towards the outside of the known region. The resulting inverse PCR products include the unknown region which contains this particular restriction site, and two known region segments, one covering from the 5'-end to the left priming site and another from the 3'-end to the right priming site. PCR products are sequenced by primer walking using inverse PCR primers and newly designed primers until the entire PCR product sequence are completed (88).

The inverse PCR method used in this study was modified from Ochman et al. (88). Initially 2 μg or 4 μg of bacterial genomic DNA was used in restriction enzyme digestion using 10 units of different enzymes in 100 μl final volume. The restriction enzymes used for digesting the DNA are shown as follows: BamHI, ClaI, EcoRI, MluI, NcoI, NdeI, SalI, XhoI for YZW-A; ApaI, ClaI, EcoRI, HindIII, KpnI, MluI, NotI, SpeI, SphI, XhoI for YZW-B; AatII, EcoRI, EcoRV, NotI, ScaI, SpeI, XhoI for YZW-D; AatII, BamHI, CalI, EcoRI, EcoRV, KpnI, MluI, NotI, NcoI, NdeI, PstI, PvuI, SpeI, XbaI, XhoI for YZW-E; ApaI, ClaI, EcoRI, EcoRV, HindIII, KpnI, MluI, NcoI, NdeI, NotI, PstI, PvuI, NruI, SalI, ScaI, SphI, XbaI for YZW-F; ClaI, EcoRV, HindIII, MluI, NdeI, NotI, PstI, XbaI, XhoI, XmnI for YZW-G.

The restriction digests were incubated at 37°C for over 3 h and then the enzymes were inactivated at 75°C for 15 min. In order to favor self-ligation, the 100 µl digested DNA was diluted to 200 µl final solution containing 20 µl 10X ligation buffer and 2 µl T4 DNA ligase (Invitrogen Corporation) (5U/µl), and then incubated at 17°C overnight. Religated DNA was purified by ethanol precipitation and dissolved in 50 µl water.

Taq or Pfu polymerase was used in the PCR reaction. For expected amplicons shorter than 4 kb, the 50 µl PCR reaction contained: ReadyMix^{TM} Taq PCR Reaction Mix (SIGMA-ALDRICH, Inc.) 25 µl; 2.5 µM forward and reverse primer 5 µl each; DNA template 10 µl; ddH$_2$O 5 ul. PCR was performed using the following program: 3 min at 94°C; 30 cycles of 30 sec at 94°C, 30 sec at 50°C, and 4 min at 72°C; 10 min at 72°C; maintained at 4°C. For amplicons longer than 4 kb but shorter than 6 kb, PCR was performed using Pfu polymerase in a 50 µl reaction containing: 10X buffer 7.5 µl; 10 mM dNTP (2.5 mM dATP, dTTP, dCTP, and dGTP) 5 µl; 2.5 µM forward and reverse primer 5 µl each; Pfu polymerase (Stratagene) 0.5 µl (1U); DNA template 1 µl; ddH$_2$O 26 µl. PCR was performed using the following program: 3 min at 94°C; 30 cycles of 30 sec at 94°C, 30 sec at 50°C, and 12 min at 72°C; 15 min at 72°C, maintain at 4°C. To amplify PCR products longer than 6 kb, PCR was performed using Pfu

polymerase in a larger reaction system (100 µl) containing: 10X buffer 15 µl; 10

mM dNTP (2.5 mM dATP, dTTP, dCTP, and dGTP) 10 µl; 2.5 µM forward and

reverse primer 10 µl for each; Pfu polymerase (Stratagene) 1 µl (2U); DNA

template 1 µl; ddH₂O 53 µl. The following program was used for long PCR: 2 min

at 95°C; 30 cycles of 10 sec at 95°C, 30 sec at 60°C, and >12min at 68°C (for

Pfu PCR, elongation time will be estimated by amplicon length X 2 min/1 kb); 30

min at 68°C, maintain at 4°C. PCR products were purified after a single band

was amplified, and sequenced directly by the inverse PCR primers.

Using the DNA sequence analysis software Lasergene (DNASTAR, Inc.), the

particular restriction site was recognized and PCR product sequence was cut at

the position of restriction site to form two separated segments. The sequences

of each segment were reassembled with 5'-end and 3'-end of known region. In

some case, no appropriate restriction enzyme could be found, or only one side of

the flanking region needed to be amplified, we chose a restriction enzyme that

cut the known region once and considered the two digested halves two individual

known regions to design inverse PCR experiments to amplify each side of the

flanking region.

2.5 PCR of the Isophthalate and Terephthalate Genes of YZW-B, -E, and -F

YZW-B, -D, -E, and -F belong to the same species: *C. testosteroni*. The isophthalate genes (*iph*) and/or terephthalate genes (*tph*) from YZW-B, -E, and -F were determined by PCR amplification and sequencing using primers designed according to the YZW-D *iph* and *tph* gene (AY923836) (123). DNA sequences in the gap region between the *iph* and *tph* genes in YZW-B, YZW-D, and YZW-E were also obtained by PCR. Both sides of the flanking regions of *iph* and/or *tph* genes in YZW-B, YZW-D (AY923836) (123), YZW-E, and YZW-F were also determined by inverse PCR and sequencing.

2.6 PCR of the Phthalate genes of YZW-B, -E, -F and -G

The partial phthalate genes in YZW-F were originally PCR amplified and determined by inverse PCR after the flanking region of the terephthalate genes in YZW-F was sequenced. The phthalate gene clusters (*oph*) were found to be located upstream of the *tph* cluster in YZW-F. Then the entire *oph* gene cluster of YZW-F were obtained by inverse PCR and regular PCR using primers designed according to other known sequences, such as the 4,5-

dihydroxyphthalate decarboxylase gene from *C. testosteroni* M4-1 (Accession No.Q59727) (66). Similarly, the *oph* genes from YZW-B and YZW-E were determined by PCR amplification and sequencing using primers designed according to the YZW-F *oph* gene sequences. The *oph* genes of YZW-G were determined by regular PCR using primers designed according to sequences of the phthalate genes encoded on the plasmid pNMH102-2 from a *Pseudomonas putida* strain (D13229) (86) due to the high similarity of phthalate genes in these two strains. Inverse PCR was also used to amplify the flanking region of known sequences of the phthalate genes from all of these strains in this study.

2.7 Degenerate PCR of the Terephthalate Genes of YZW-A

Since the terephthalate genes (*tph*) in YZW-D (123) and *Rhodococcus* sp. strain DK17 (21) are known, degenerate PCR primers for the YZW-A *tph* genes were designed according to the conserved region of the *tph* genes of these two bacteria. The primers corresponding to the terephthalate dioxygenase gene *tphA2* are as follows: Forward primers A-tphA2-F1 5'-CTGGARCTACCTGTGCYTGG-3', A-tphA2-F2 5'-CTTCGAGAACCGSTGCGC-3', A-tphA2-F3 5'-CACGCSTGGAGCTACAACC-3', whose priming sites are located at positions 130-149, 231-248, and 313-331 of *tphA2* sequence of YZW-

D (123) and DK17 (21), respectively; reverse primers A-tphA2-R1 5'-AGCTCGAAGGTSGTGAAGAAC-3', A-tphA2-R2 5'-GAGGTTGGCCTGYTTGAGTC-3', A-tphA2-R3 5'-TTCCARAAGCCGCGKACCG-3', whose priming sites are located the position 668-648, 1063-1044, and 1302-1184 of *tphA2* sequence of YZW-D (123) and DK17 (21), respectively. Additionally, based on the *tphB* and *tphA3* gene sequences in YZW-D and DK17 strain, more degenerate PCR primers were designed to amplify the *tphB* and *tphA3* genes in YZW-A. These primers are as follows: *tphB* forward primers A-tphB-F1 5'-ATCGGCGACYCCMACGG-3', A-tphB-F2 5'-GAYGCCGTSATTGCCTG-3', A-tphB-F3 5'-TGATGYTGRTSGGTGC-3'; *tphB* reverse primers A-tphB-R1 5'-GGRTTGATSCCGAAKAC-3', A-tphB-R2 5'-GAATRTGGCCCTGATCGTG-3', A-tphB-R3 5'-GCCRTGGCCGACRCTGGA-3'; *tphA3* forward primers A-tphA3-F1 5'-GAAGCCAACWTCTACGAGC-3'; *tphA3* reverse primers A-tphA3-R1 5'-GCTCGTAGATTTGGCTTC-3', A-tphA3-R2 5'-TCATAGCGGGARYGCCA-3'. About a 2 kb region including the partial *tphA2*, *tphA3* and *tphB* genes was successfully amplified by degenerate PCR primers and sequenced. The rest of the *tph* gene operon in YZW-A and its flanking region was also determined by inverse PCR and sequencing.

2.8 DNA Sequence Analysis and Gene Annotation

ABI PRISM BigDye® Terminator cycle sequencing kit (Applied Biosystems) and an ABI 3100 automated DNA sequencer (Applied Biosystems) were used for cycle sequencing reactions as recommended by the manufacturer (Applied Biosystems). DNA sequences were assembled using the DNA sequence analysis software Lasergene (DNASTAR, Inc.). Open reading frames were analyzed using Blast against the GenBank database (4) and then assigned to gene function. Similarity and identity of two known DNA or protein sequences were analyzed using Lasergene (DNASTAR, Inc.) or Blast2 (115). Conserved domains of putative proteins were identified using Conserved Domain Database and Search Service V2.03 (73, 74). Insertion sequences and transposases were identified and classified using Blast (4) and IS finder (72). Phylogenetic trees were constructed by the Neighbor-Joining method and tested by 1000 bootstrap replications using MEGA 3.1 (63) or Mega 4 (111).

2.9 Southern Hybridization

Probes for Southern blots were generated using PCR DIG probe synthesis Kit (Roche Diagnostics GmbH). Positively charged nylon membrane, block reagent,

BCIP (5-Bromo-4-Chloro-3'-Indolyphosphate *p*-Toluidine Salt), and NBT (Nitro-Blue Tetrazolium Chloride) were purchased from Roche. The PCR reaction contains: 10X PCR buffer 5µl, PCR DIG labeling mix 5 µl, forward and reverse primer (2.5 µM) 2 µl each, enzyme mix 0.75 µl, DNA template 100 ng, add ddH$_2$O to 50µl. Gene specific primers were designed for amplification of full length *ophA2* (ophA2-N1, 5'-ATGCTGACTCATGAAGAAAAC-3'; ophA2-C1 5'-TTACTGCTTGACCTGGTAGTC-3'), *iphA2* (iphA2-N1, 5'-ATGAACAAGGAAATGTCCG-3'; iphA2-C1, 5'-TCAGAGGCTGAAGACCG-3') and *tphA2* (tphA2-N1, 5'-ATGCAAGAATCCATCATCC-3'; tphA2-C1, 5'-TCATGCTTGCATCTCCTG-3'). PCR was performed using the following program: 3 min at 94°C; 30 cycles of 1 min at 94°C, 1 min at 50°C, and 2 min at 72°C; 7 min at 72°C; maintain at 4°C. Probes were stored at -20°C.

Bacterial genomic DNA was single or double digested by the restriction enzymes BamHI, EcoRI, HindIII, PstI, or XbaI. In each reaction, 1 µg DNA was digested by 20 U restriction enzyme in single digestion or 10 U each enzyme in double digestion. DNA was separated in a 1% agarose gel by electrophoresis. After electrophoresis, DNA transfer and hybridization was followed as per the standard Southern blotting procedure. In this study, *ophA2*, *iphA2* and *tphA2* probes were used to determine the copy number of *ophA2*, *iphA2* and *tphA2* in the *C.*

testosteroni YZW-B genome. In addition, *iphA*2 probe was used to detect unknown isophthalate genes in YZW-C, YZW-H, and YZW-I by Southern blot.

2.10 Colony Hybridization

Once the positive bands detected by Southern blotting the *iphA2* gene with YZW-I DNA, the digested DNA was separated again using the same gel electrophoresis condition. Based on the first Southern blot result, DNA fragments in the same position as the positive Southern blot band were purified from the gel and cloned into the pGEM-3Z vector (Promega Corporation). Positive clones were screened by colony hybridization. Colonies were transferred into 96-well plates containing 200 μl media in each well and incubated overnight. Bacterial cultures were replicated from the 96-well plate onto a Nylon membrane (Roche Diagnostics Corporation) previously overlaid on an LB agar plate containing ampicillin (100 μg/ml) and incubated overnight until colonies grew up. Colonies were lysed on a cellulose paper saturated with Lysis solution containing 2X SSC and 5% SDS for 5 min at room temperature. The membrane was heated and dried using a microwave oven until the edges of the membrane were curled up. Then the membrane was transferred and incubated in 50 ml proteinase K solution (10 μl/ml) (containing 50 mM Tris pH 8.0, 50 mM

EDTA, 100 mM NaCl, and 1% N-laurylsarcosine) at 37°C for 2 h while rocking

the container occasionally. The membrane was washed using 100 ml 2X SSC,

air dried and UV crosslinked. The membrane was hybridized following the

standard Southern blotting procedure. Positive clones were verified by DNA

sequencing.

2.11 RNA Isolation

250 µl bacterial cells at OD$_{600nm}$ 0.4 (about 1×10^8 bacterial cells) were collected

for RNA isolation. RNA was isolated using RNeasy$^®$ Mini Kit (QIAGEN Sciences,

Inc.) under the manufacturer's instruction except that samples were treated by

lysozyme for 15 min and DNase on-column digestion using the RNase-Free

DNase set (QIAGEN Sciences, Inc.) for 30 min. RNA samples were finally eluted

in 50 µl RNase-free water for regular RT-PCR, quantified using a DU 800 UV

Spectrophotometer (Beckman Coulter, Inc.), and diluted to 0.5 ng/µl as initial

concentration for RT-qPCR.   RNA quality was checked using the

spectrophotometer, regular end-point RT-PCR, and real-time qPCR.   RNA

samples were stored at -70°C.

2.12 Reverse Transcription PCR

Reverse transcription PCR was performed using QIAGEN® One step RT-PCR Kit (QIAGEN Sciences, Inc.). The 25 µl RT-PCR reaction included: 5X QIAGEN One Step RT-PCR Buffer 5 µl; dNTP Mix (containing 10 mM of each dNTP) 2 µl; 2.5 µM forward and reverse primer 1 µl for each; QIAGEN One Step RT-PCR Enzyme Mix 1 µl; and RNA template 50 or 100 ng. PCR was performed with the following thermocycling program: 30 min at 60°C; 15 min at 95°C; 30 cycles of 1 min at 94°C, 1 min at 50°C, 2 min at 72°C; 7 min at 72°C; maintained at 4°C. In negative control, the same amount of RNA template was amplified in the same PCR condition except the RT step was omitted.

2.13 Operonic Structure of *oph*, *iph*, and *tph* Genes in YZW-B

In order to determine the operonic structure of the *oph* genes (*ophHA1DCRBA2*), *iph* genes (*iphA2CBA1R*), and *tph* genes (*tphRCA2A3BA1*) in YZW-B, we designed a series of primers to amplify each intergenic region between two adjacent genes. In each primer pair, the forward primer and reverse primer sites were located at upstream gene and downstream genes, respectively. Primers were confirmed by PCR using a YZW-B DNA template. RNA was isolated from

YZW-B cells grown on phthalate, isophthalate, or terephthalate. Positive RT-PCR amplification means two adjacent genes are transcribed on the same mRNA.

Primers for the *oph* genes are shown as follows: *ophH-ophA1*, forward 5'-CATTCATCTTCCTGCTCG-3', reverse 5'-TCTGGTCGACGACACATC-3'; *ophA1-ophD* forward 5'-ATTGCGCTCTTGCGAGTC-3', reverse 5'-TTGTCGGTCAGGTAGAAC-3'; *ophD-ophC* forward 5'-GCTCTGCCTTACTTGGTC-3', reverse 5'-AGCTCGCAGATGTCGTAGTC-3'; *ophC-ophR* forward 5'-ATGGTATGAGCGTCG-3', reverse 5'-ACTTCATACTGAGCCAGG-3'; *ophR-ophB* forward 5'-GCCTGGCTCAGTATG-3', reverse 5'-TGAATCCGGATGCCTACG-3'; *ophB-ophA2* forward 5'-CCGTCATCAGTTCTTCG-3', reverse 5'-AACCGCTTCTGGCAAGAC-3'.

In addition, another set of primers for amplifying each individual *oph* gene were used as positive control to show whether these genes actually are expressed in YZW-B grown on phthalate. These primers' target gene, direction, and sequences are shown as follows: *ophH* forward 5' CATTCATCTTCCTGCTCG-3', reverse 5'-ACCGATGCACGCTTACTG-3'; *ophA1* forward 5'-GTATCCATCAGCGACTGC-3', reverse 5'-

TGGCGCTTTGAACTTAC-3'; *ophD* forward 5'-TCGCAAGATCACTCTCAG-3', reverse 5'-GACAGGGCGAAATGC -3'; *ophC* forward 5'-CAAGCTTCAACTGTCCATCG-3', reverse 5'-AAACCATACGACCAGAAATC-3'; *ophR* forward 5'-GCCTGGCTCAGTATG-3', reverse 5'-GAAGGAATCCGCAGCAC-3'; *ophB* forward 5'-AGTTCGTCGATCACCTCG-3', reverse 5'-CAACGCAAGGACTTACACAG-3'; *ophA2* forward 5'-TTGCCGTCCTGCTTG-3', reverse 5'-GTCAAACACACTGCCTATCC-3'.

Similarly, primers used to show the operonic structure of the *iph* and *tph* genes were designed as follows: *iphA2-iphC* forward 5'-GGCTGTTCCCCTGGTTCAC-3', reverse 5'-GTGCCAGATCCTTGAGC-3'; *iphC-iphB* forward 5'-CAAGCTACAGGAGGCATTG-3', reverse 5'-TTCACCAGGGTATCGAGACC-3'; *iphB-iphA1* forward 5'-CAGCATCGTCAACGTGTC-3', reverse 5'-TCATACGAATCCAGAAAGG-3'; *iphA1-iphR* forward 5'-CCGTGAAGGTGTATG-3', reverse 5'-AGATCCTGCTTGTCCACG-3'; *iphR-tphR* forward 5'-GGCTATGCTGTGCTC-3', reverse 5'-TTACCTGTCGCCAGTCATTG-3'; *tphR* forward 5'-GAACTCGCTCATGCTGAG-3', reverse 5'-CTGTCAGAGGTAGCCAGG-3'; *tphR-tphC* forward 5'-GCCAAGCTGTACCAG-3', reverse 5'-TTTGAGAACGCTGCTCGTCG-3'; *tphC-tphA2* forward 5'-AATGGCCAGCCATGC-3', reverse 5'-CGTGAAGCGACCGATGACT-3'; *tphA2-*

tphA3 forward 5'-CCTTGGCCCTGAGATTTGC-3', reverse 5'-GAAATTCGGTCGGTGAG-3'; *tphA3-tphB* forward 5'-CTTCAATGCCGCCTACG-3', reverse 5'-CGAGCAAAGATGGGTAGC-3'; *tphB-tphA1* forward 5'-GTGCCTAAACCGCAAGTC-3', reverse 5'-GGTCAGTGATGGCTGCAATG-3'.

2.14 5'-RACE (Rapid Amplification of cDNA ends)

In order to determine the start site of each transcript in the *oph*, *iph* and *tph* gene cluster/operon, the rapid amplification of cDNA ends (5'-RACE) method was performed using Invitrogen 5'-RACE System (Invitrogen Corporation). RNA used in the 5'-RACE experiment was isolated from YZW-B grown on phthalate, isophthalate, or terephthalate. The procedure was modified from the Invitrogen 5'-RACE instruction manual (Figure 2.1). For each operon, four different gene specific primers (GSP1-4) were designed for cDNA synthesis, nested PCR, or DNA sequencing. In each 5'-RACE experiment, a 14.5 µl RNA sample (0.4 -1 µg) was initially added for cDNA synthesis using the GSP1 primer in the reverse transcription reaction. First strand cDNA was purified using a S.N.A.P. column. The 3'-end of purified cDNA (maximal amount, 16.5 µl) was tailed with dCTP by TdT (Terminal deoxynucleotidyl transferase). Tailed cDNA was then amplified by first nested PCR using second gene specific primer (GSP2) and abridged anchor

primer (AAP). PCR products were purified and amplified by second nested PCR using third gene specific primer (GSP3) and abridged universal amplification primer (AUAP). Then this PCR product was purified again and directly sequenced using third or fourth specific primer (GSP3 or GSP4). Finally, the start site of the transcripts was determined by looking for poly C at the end of PCR product sequence.
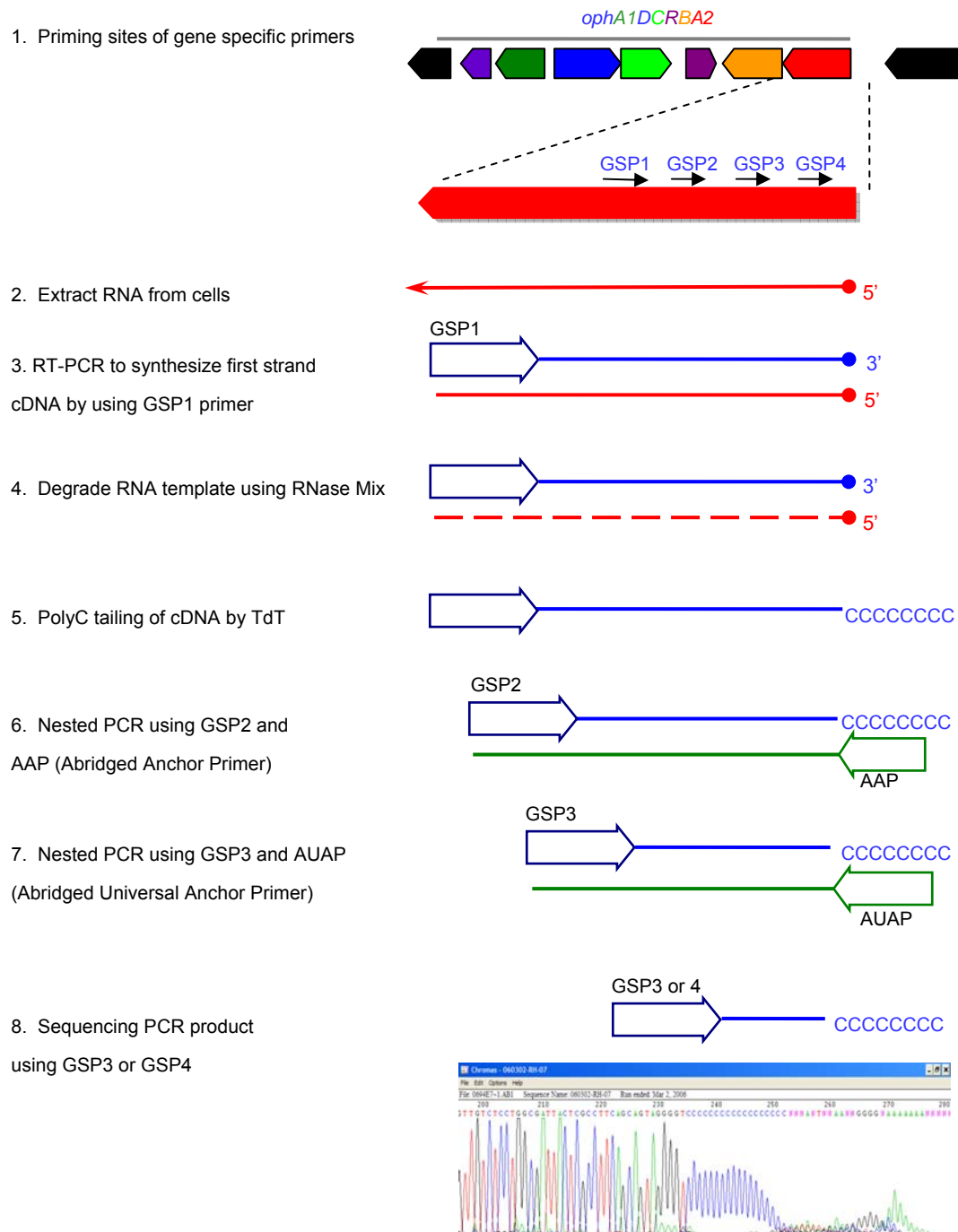
Figure 2.1 Overview of the 5'-RACE assay. In 5'-RACE assay, GSP1-4 are gene specific primers designed for cDNA synthesis, nested PCR and sequencing. AAP is abridged anchor primer (5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3') and AUAP is abridged universal anchor primer (5'-GGCCACGCGTCGACTAGTAC-3').

2.15 Quantitative Real Time PCR

Quantitative real time PCR was performed in an ABI 7300 real time system (Applied Biosystems) and data were analyzed using the ABI PRISM 7300 Sequence Detection System software (Applied Biosystems). qPCR primers were designed using Primer Express software (Applied Biosystems). All of these qPCR primers were checked by regular end-point PCR using DNA templates before qPCR was performed. TaqMan[®] Reverse Transcription Reagents (Applied Biosystems) and SYBR[®] Green PCR Master Mix (Applied Biosystems) were used for cDNA synthesis and gene detection in RT-qPCR assay.

An absolute quantification assay was carried out to check RNA sample quality, measure DNA background, check primers quality, and generate standard curves. For DNA templates, the qPCR reaction (25 µl) contained 2X SYBR Green PCR Master Mix 12.5 µl, forward and reverse Primer (2.5 µM) 2.5 µl for each, DNA template 5 µl, and ddH$_2$O 2.5 µl. The thermocycling parameters for qPCR were used as follows: 2 min at 50°C; 10 min at 95°C; 40 cycles of 15 sec at 95°C, 1 min at 60°C. For RNA samples, One Step RT-qPCR reaction (25 µl) included: 2X SYBR Green PCR Master Mix 12.5 µl, forward and reverse Primer (2.5 µM) 2.5 µl for each, RNase Inhibitor 0.5µl, MultiScribe Reverse transcriptase (50 U/µl)

0.125 µl, RNA template (0.5ng/µl) 2 µl (1 ng), RNase-free H_2O 4.875 µl. The

thermocycling parameters for One Step RT-qPCR were used as follows: 30 min

at 60°C; 15 min at 95°C; 40 cycles of 15 sec at 95°C and 1 min at 60°C.

Relative quantification assay was performed to measure gene expression of

housekeeping genes (16S, 23S, *gyrB,* or *rpoB*) and three phthalate isomer

dioxygenase genes (*ophA2*, *iphA2*, or *tphA2*) in YZW-B. Two-step PR-qPCR

was used for relative quantification assay. In the RT step, reaction mixtures (10

µl) contained: 10X RT Buffer 1 µl, $MgCl_2$ (25 mM) 2.2 µl, dNTP mixture (2.5 mM)

2 µl, random hexamers (50 µM) 0.5µl, RNase Inhibitor (20 U/µl) 0.2 µl,

MultiScribe reverse transcriptase (50 U/µl) 0.25 µl, RNA template (0.5 ng/µl) 2 µl

(1ng), RNase free H_2O 1.85 µl. The thermocycling parameters for the RT

reaction are as follows: Hexamer incubation at 25°C for 10 min, RT at 48°C for

30 min, reverse transcriptase inactivation at 95°C for 5 min. The qPCR reaction

(25 µl) contained 2X SYBR Green PCR Master Mix 12.5 µl, forward and reverse

Primer (2.5 µM) 2.5 µl for each, cDNA template 1 µl, RNase free H_2O 6.5 µl. The

thermocycling parameters for qPCR were used as follows: 10 min at 95°C; 40

cycles of 15 sec at 95°C; 1 min at 60°C.

2.16 RT-qPCR Analysis of *ophA2*, *iphA2*, and *tphA2* Genes Expression in YZW-B

Comamonas testosteroni YZW-B was chosen for real-time PCR analysis of *ophA2*, *iphA2*, and *tphA*2 gene expression because YZW-B is one of two *C. testosteroni* isolates, YZW-B and YZW-E, which could degrade phthalate, isophthalate, and terephthalate; and comparing among Comamonas isolates, YZW-B is phylogenetically closer to *C. testosteroni* YZW-D and YZW-F than to YZW-E. LB media and minimal media containing succinate or phthalate isomers as carbon sources were used to grow bacterial cells for RNA isolation. Eight different substrate treatments and their initial concentrations in media were used as followed: phthalate (20 mM); isophthalate (20 mM); terephthalate (20 mM); phthalate (10 mM) and isophthalate (10 mM); phthalate(10 mM) and terephthalate (10 mM); isophthalate (10 mM) and terephthalate (10 mM); phthalate (6.67 mM), isophthalate (6.67 mM) and terephthalate (6.67 mM).

YZW-B cells grown on LB plate overnight were transferred on minimal plate containing certain carbon sources and incubated for 24 hours. A loop of bacterial cells were inoculated in a 500 ml flask containing 100 ml minimal liquid medium with the same carbon sources and incubated in a 30°C water bath with 200 rpm

shaking for about 12 hours. This culture was transferred to 100 ml fresh minimal

liquid medium at OD_{600nm} of 0.05, and incubated in a 30°C water bath with 200

rpm shaking until an OD_{600nm} of 0.4. 250 µl cells (about $1.0X10^8$ cells) were

collected for RNA isolation.

Four housekeeping genes (16S, 23S, *gyrB*, *rpoB*) and three phthalate isomers

dioxygenase genes (*ophA2*, *iphA2*, and *tphA2*) expression in YZW-B were

analyzed by RT-qPCR. First, nearly full length 16S, 23S, *gyrB*, and *rpoB* genes

and full length *ophA2*, *iphA2*, and *tphA2* genes were amplified by PCR using

corresponding specific primers. The features of these primers are shown as

follows: *16S*, forward 5'-AGAGTTTGATCCTGGCTCAG-3', reverse 5'-AAGGAGGTGATCCAICCGCA-3'; *23S* forward 5'-CCGATAGTGAACCAGTACC-3', reverse 5'-CGCTTAGATGCTTTCAGC-3'; *gyrB* forward 5'-TTCGAGGTCGTGGACAACTC-3', reverse 5'-CGAGTCTTCGCTGACCTTGC-3'; *rpoB* forward 5'-TGCCATCTACCGCATGATG-3', reverse 5'-CCGTGGTCAGCTCCTTG-3'; *ophA2* forward 5'-ATGCTGACTCATGAAGAAAAC-3', reverse 5'-TTACTGCTTGACCTGGTAGTC-3'; *iphA2* forward 5'-ATGAACAAGGAAATGTCCG-3', reverse 5'-TCAGAGGCTGAAGACCG-3', *tphA2* forward 5'-ATGCAAGAATCCATCATCC-3', reverse 5'-TCATGCTTGCATCTCCTG-3'. All of these PCR products were

purified and quantified by spectrophotometry. The copy number concentration of each PCR product was calculated on the basis of the average mass of base pair (1.096×10^{-21} g/bp). Each PCR product was serially diluted to 2×10^4, 2×10^3, 2×10^2, 2×10^1, and 2 copy/µl, which were used as DNA templates in generating a standard curve.

qPCR primers were designed using Primer Express software (Applied Biosystems). The features of these Primers are shown as follows (target gene, primer name, direction, and primer sequence): 16S 16S-1146F forward 5'-TCAGTTGAGCACTCTAATGGGAC-3', 16S-1299R reverse 5'-TTATGGGATTAGCTCCCCCTC-3'; 23S 23S-1300F forward 5'-CCGTAAGCGCAAGGTTTCC-3', 23S-1502R reverse 5'-TCCCCACATCGCACTATTGA-3'; *gyrB* gyrB-1057F forward 5'-CGCCGAGTTCCTCGAAGA-3', gyrB-1124R reverse 5'-GGGCGGCTTCCACAATCT-3'; *rpoB* rpoB-534F forward 5'-GCGTGGGCCGTATGAAGTT-3', rpoB-624R reverse 5'-ACCACGGCCAGGATGTCTT-3'; *ophA2* ophA2-273F forward 5'-CCACGGCTGGAAGTTCGA-3', ophA2-391R reverse 5'-TGCCGCCCCATTCCTT-3'; *iphA2* iphA2-713F forward 5'-GGCTGTTCCCCTGGTTCAC-3'; iphA2-797R reverse 5'-

TTGTGGTCGTCGATAGGCAC-3'; *tphA2* tphA2-501F forward 5'-CCTTGGCCCTGAGATTTGC-3', tphA2-573R reverse 5'-CGTGAAGCGACCGATGACT-3'.

In order to determine the qPCR efficiency and primer limitation, serially diluted PCR products were used as DNA templates in generating a standard curve, and final copy numbers of the PCR product in the qPCR reaction were $1X10^5$, $1X10^4$, $1X10^3$, $1X10^2$, 10 copies/reaction correspondingly after adding 5 µl PCR product from each dilution ($2X10^4$, $2X10^3$, $2X10^2$, $2X10^1$, and 2 copy/µl). Standard curves were plotted using C_T as Y-axis and Log10 input gene copy numbers as X-axis, and the linear relationship between C_T and Log10 gene copy numbers was expressed as Y=AX+B, in which Y represents C_T and X represents Log10 gene copy numbers.   In addition, after qPCR, a dissociation assay was carried out for each sample to check PCR specificity.

Another validation experiment was carried out to check whether the qPCR amplification efficiencies for targets and endogenous controls from YZW-B cDNA samples were approximately equal. cDNA was synthesized from 1 ng RNA isolated from YZW-B cells grown on phthalate, and then 10 fold diluted serially from 1 ng -10^{-4} ng. A series of qPCRs were performed using different diluted

cDNA as templates. Similarly, cDNA standard curves were plotted using C_T as Y-axis and Log10 input amount of RNA as X-axis, and the linear relationship between C_T and Log10 input amount of RNA was expressed by Y=AX+B, in which Y represents C_T and X represents Log10 input amount of RNA. Also, ΔC_T ($C_{T,\ reference}$ - $C_{T,\ target}$) was calculated and data were plotted using ΔC_T as Y-axis and Log10 input amount of cDNA as X-axis.

The comparative C_T ($\Delta\Delta C_T$) method was used to calculate relative quantification of *ophA2*, *iphA2*, and *tphA2* gene expression in YZW-B. Using this method, the amount of target, normalized to an endogenous control and relative to a calibrator, is given by $2^{-\Delta\Delta C_T}$ (68). In this study, *ophA2, iphA2,* and *tphA2* gene were measured as target genes. One of the control genes (16S, 23S, *gyrB*, and *rpoB*) were used as the endogenous control, while other genes were treated as target genes. RNA samples isolated from cells grown on succinate were used as the calibrator. RNA samples of triplicates in the same treatment were isolated from three independent cultures. Two Step RT-qPCR was performed in this study. In each assay, cDNA was synthesized from 1 ng RNA in 10 μl reaction. 1μl cDNA was used as the qPCR template for each gene. RT-qPCR assay of each RNA sample was triplicated. Relative quantification was calculated using

ABI PRISM 7300 Sequence Detection System (SDS) (Applied Biosystems) by the following formula:

$$- \frac{[(C_{T,\ Target} - C_{T,\ endo})_{Pht} - (C_{T,\ Target} - C_{T,\ endo})_{Suc}]}{2}$$

In this formula, $C_{T,\ Target}$ represents target gene C_T value, while $C_{T,\ endo}$ represents endogenous control C_T value. Pht or Suc means RNA sample was extracted from cells grown on phthalate isomers or succinate. For data analysis, one of the four housekeeping genes (16S, 23S, *gyrB*, and *rpoB* gene) was chosen as endogenous control, while other three housekeeping genes, *ophA2*, *iphA2*, and *tphA2* gene were treated as targets to calculate fold change of gene expression.

2.17 Knock Out of *ophR*, *iphR*, and *tphR* Gene in YZW-B

An *ophR* knockout mutant of YZW-B was generated using gene replacement with the tetracycline resistance gene cassette *tetRA* (Figure 2.2). About a 1.2kb fragment containing *ophC* gene on the left of *ophR* was amplified by PCR using primers 5'-<u>CATATG</u>ACATCTGCGAGCTG-3' (underlined is NdeI site) and 5'-<u>GGTACC</u>GCTTTCTCCTTCACG-3' (underlined is KpnI site) so that the resulting fragment, designed fragment C, contained an extra NdeI and KpnI site. Similarly,

about 1.2 kb fragment containing the *ophB* gene on the right of *ophR* was amplified by PCR using the primers 5'-<u>GGTACC</u>GAGGAGTGCGAAAACTG-3' (underlined is KpnI site) and 5'-<u>CCATGG</u>AACAGCGCCGTTTAC-3' (underlined is NcoI site) so that the resulting fragment, designed fragment B, contained an extra KpnI and NcoI site. PCR was performed with Pfu polymerase using the following thermocycling program: 3.0 min at 94°C; 30 cycles of 30 s at 94°C, 1 min at 50°C, and 2.0 min at 72°C; 10 min at 72°C; maintained at 4°C. Both fragments were purified, A-tailed by Taq polymerase, and TA cloned into pGEM-T-Easy vector (Promega Corporation) to generate two new constructs, pC and pB, respectively. Both pC and pB were double digested by KpnI and NdeI. 1.2 kb fragment C and 4.2 kp linearized pB with KpnI and NdeI site in ends were released from pC and pB after digestion. A new plasmid pCB containing fragment C and B joined at the KpnI site was constructed by ligation of these two fragments.

Tetracycline resistance gene cassette *tetRA* was also amplified from plasmid pRK414 by PCR using primers 5'-<u>GGTACC</u>TCAATCGTCACCC-3' and 5'-<u>GGTACC</u>TCAGCGATCGGCTC-3' (underlined are KpnI site) which were designed according to *tetRA* genes sequences of pRK404 (Accession No. AY204475) (105). PCR was performed with Pfu polymerase using the following

thermocycler program: 3.0 min at 94°C; 30 cycles of 30 s at 94°C, 1 min at 50°C, and 4.0 min at 72°C; 10 min at 72°C; maintained at 4°C. This PCR product was purified, A-tailed by Taq polymerase, and TA cloned into pGEMT-Easy vector to construct plasmid pTetRA. Both pTetRA and pCB were digested by KpnI, and the *tetRA* cassette was released from pTetRA while pCB was linearized with KpnI site in both ends. Finally, the *tetRA* cassette and linearized pCB were ligated to generate the final construct pCTB for electroporation.

Electrocompetent YZW-B cells were originally prepared from cells grown on LB Broth until OD_{600nm} equaled 0.6. Cells (1.5 ml) were centrifuged, washed with 500 μl of 10% glycerol twice and resuspended in 50 μl 10% glycerol. 1 μg plasmid pCTB DNA was electroporated into 50 μl electrocompetent YZW-B cells using the following conditions: 25 μFD, 200 Ω, 1.25 kV, 0.1 cm cuvette. Electroporated cells were incubated in 500 μl SOC media at 30°C overnight, and then spread on LB plate containing tetracycline (10 μg/ml). After tetracycline resistant mutants grew on the plate, positive mutants, whose *ophR* gene was replaced by tetracycline resistance cassette *tetRA* via a double crossover, were screened and confirmed by PCR and DNA sequencing.

Primers ophF4 (5′-ATGGTATGAGCGTCG-3') and ophR5 (5′-CCAAGCTCAAGGCTGC-3') were used to distinguish the wild type cells, double crossover knockout mutants, or single crossover cells, from which PCR produced a 1.6 kb amplicon (containing partial *ophC*, *ophR*, and partial *ophB* in YZW-B genome), 3.0 kb amplicon (containing partial *ophC*, *tetRA*, and partial *ophB*), or both 1.6 and 3.0 kb amplicon, respectively. In addition, another primer pair ophF6 (5′-GCCTGGCTCAGTATG-3') and ophR1-1 (5′-GAAGGAATCCGCAGCAC-3') were used to amplify the internal region of *ophR* yielding about 500 bp amplicon from wild type YZW-B cells and no product from double crossover knockouts since the *ophR* gene was supposed to be replaced by *tetRA* in the knockout mutant. *ophA2* gene primes ophF7 (5′-TTGCCGTCCTGCTTG-3') and ophA2-R1 (5′-GTCAAACACACTGCCTATCC-3') were used as a positive control to show that the knockout mutants shared the same genetic background with the wild type strain except that the *ophR* gene was replaced by *tetRA*.

For the *iphR* gene knockout (Figure 2.3), primer pair RH7 (forward, 5′-TGCTGATGCTGGACG-3') and iphR-XhoI (reverse, 5′-<u>CTCGAG</u>GCCGATACTCTACCC-3', with underlined XhoI site) was used to amplify the 1.8 kb left fragment, and primer pair iphF-XhoI (forward, 5′-

CTCGAGTGTCCAGTCTCATAC-3') and tphC-R4 (reverse, 5'-TGCTCGGAACACTAGATCC-3') was used to amplify the 1.9 kb right fragment. Then the left and right fragments were cloned into the pGEMT vector to generate plasmids pIphL and pIphR. Both pIphL and pIphR were double digested by XhoI, whose cutting site was introduced by the primer, and NsiI, whose cutting site originally exists in pGEMT. pIphL released about a 4.8 kb large fragment containing the 1.8 kb left fragment and pGEMT vector, and about a 60 bp short fragment, while pIphR was cut into 1.9 kb right fragment and 3.0 kb vector fragment. The 4.8 kb pIphL and 1.9 kb right fragment from pIphR were religated to construct a new plasmid pIphLR which contained both left and right fragments joined at the XhoI site. The tetracycline resistance gene cassette *tetRA* was also amplified from plasmid pRK414 by PCR using primers tetR-XhoI (forward, 5'-CTCGAGTCAATCGTCACCC-3', with underlined XhoI site) and tphR-XhoI (5'-CTCGAGTCAGCGATCGGCTC-3', with underlined XhoI site). The *tetRA* fragment with XhoI site in both ends was cloned in pGEMT, released by XhoI, and cloned into pIphLR as a new construct pIphLTR. Both intact plasmid pIphLTR and linearized plasmid were used in electroporation for *iphR* gene knockout. For intact plasmid electroporation, the same method was used as for *ophR* gene knockout. Since we only got single crossover mutant instead of double crossover *iphR* gene knockout mutants, linearized plasmid was used in

another electroporation. While using linearized plasmid, 2 µg plasmid DNA was digested by 2 µl NsiI (20 U) in 50 reaction and concentrated by ethanol precipitation into a 10 µl final solution. All of the 10 µl DNA was added to 50 µl of electrocompetent YZW-B cells for electroporation.

Primers iphR-L (5′-CTGGATTGCTGAGCGACATC-3') and iphR-R (5′-TTTGCAGCACCTAGAGCGTG-3') were used to distinguish the wild type cells, double crossover knockout mutants, or single crossover mutants, from which PCR produced 1.0 kb amplicon (containing upstream flanking region, *iphR*, and downstream flanking region in YZW-B genome), 2.0 kb amplicon (containing upstream flanking region of *iphR*, *tetRA*, and downstream flanking region of *iphR*), or both 1.0 and 2.0 kb amplicon, respectively. In addition, another primer pair iphR-F1 (5'-GCTCTTGACGTGCTTCTCG-3') and iphR-R2 (5'-CTGAGGAGCACAGCATAGC-3') were used to amplify the internal region of *iphR* yielding about a 650 bp amplicon from wild type cells and no product from double crossover knockout mutant cells. *ophA2* gene primes ophF7 (5'-TTGCCGTCCTGCTTG-3') and ophA2-R1 (5'-GTCAAACACACTGCCTATCC-3') were used as positive control.

Similar methods were used in *tphR* gene knockout (Figure 2.3) except that RH10 (forward, 5'-CCGTGAAGGTGTATG-3') and tphR-XhoI (reverse, 5'-<u>CTCGAG</u>TTTACACGCAAAGCTG-3', with underlined XhoI site) were used to amplify the 1.8 kb left fragment, and tphF-XhoI (forward, 5'-<u>CTCGAG</u>AAAGTTCTTGTCCTGC-3', with underlined XhoI site) and RH-tph-R (reverse, 5'-GTCCCATATGCTTGCGGTAG-3') were used to amplify the 2.4 kb right fragment. The construct pTphLTR contained 1.8 kb left fragment, *tetRA*, and 2.4 kb right fragment. pTphLTR was digested by SphI and about 5.3 kb insert containing 1.8 kb left fragment, *tetRA*, and 1.5 kb remaining right fragment was released and cloned into another plasmid pK18 with the kanamycin resistance marker (95) as pK18-tphLTR for electroporation. Positive double crossover *tphR* gene knockout was screened by colony phenotype tetracycline resistant but kanamycin sensitive. In addition, primers tphR-L (5'-ACCAGCCCGTCAAAGCAATC-3') and tphR-R (5'-GTGTGGGGAACTGCAGTATG-3') were used to distinguish the wild type cells, double crossover knockout mutant, or single crossover cells, from which PCR produced 1.0 kb amplicon (containing upstream flanking region, *tphR*, and downstream flanking region in YZW-B genome), 2.0 kb amplicon (containing upstream flanking region of *tphR*, *tetRA*, and downstream flanking region of *tphR*), or both 1.0 and 2.0 kb amplicon, respectively.
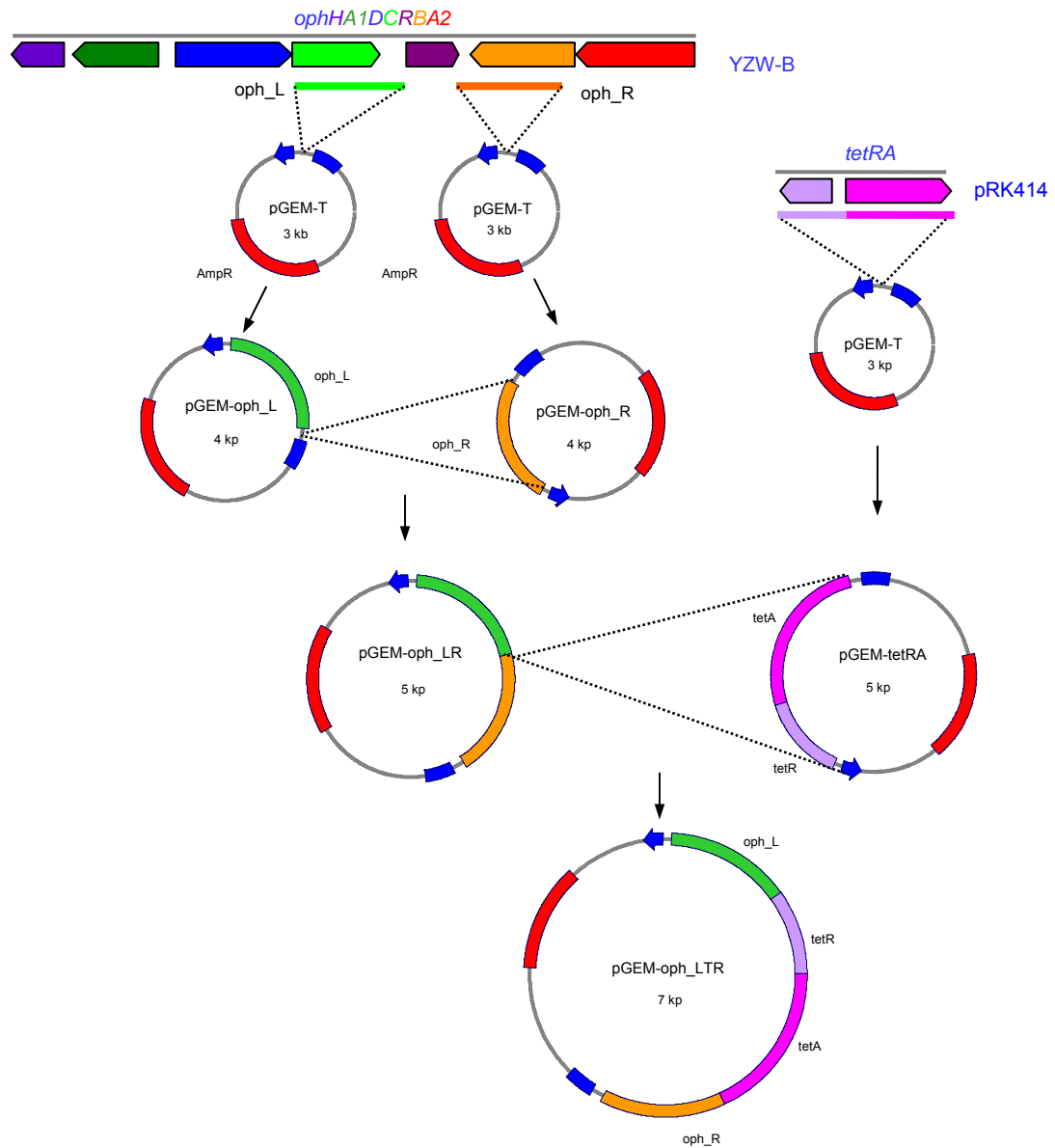
Figure 2.2 Generation of the plasmid construct for the *ophA2* gene knock out. First, PCR products of *ophR* left flanking region (oph_L), the *ophR* right flanking region (oph_R), and the tetracycline resistant gene *tetRA* were amplified from YZW-B and plasmid pRK414, respectively, and cloned into pGEM-T to generate pGEM-oph_L, pGEM-oph_R, and pGEM-tetRA. Then, oph_R was released and cloned into pGEM-oph_L to form pGEM-ophLR. Finally, *tetRA* fragment was released and inserted into pGEM-ophLR to produce the final construct pGEM-oph_LTR for the *ophR* gene knock out.
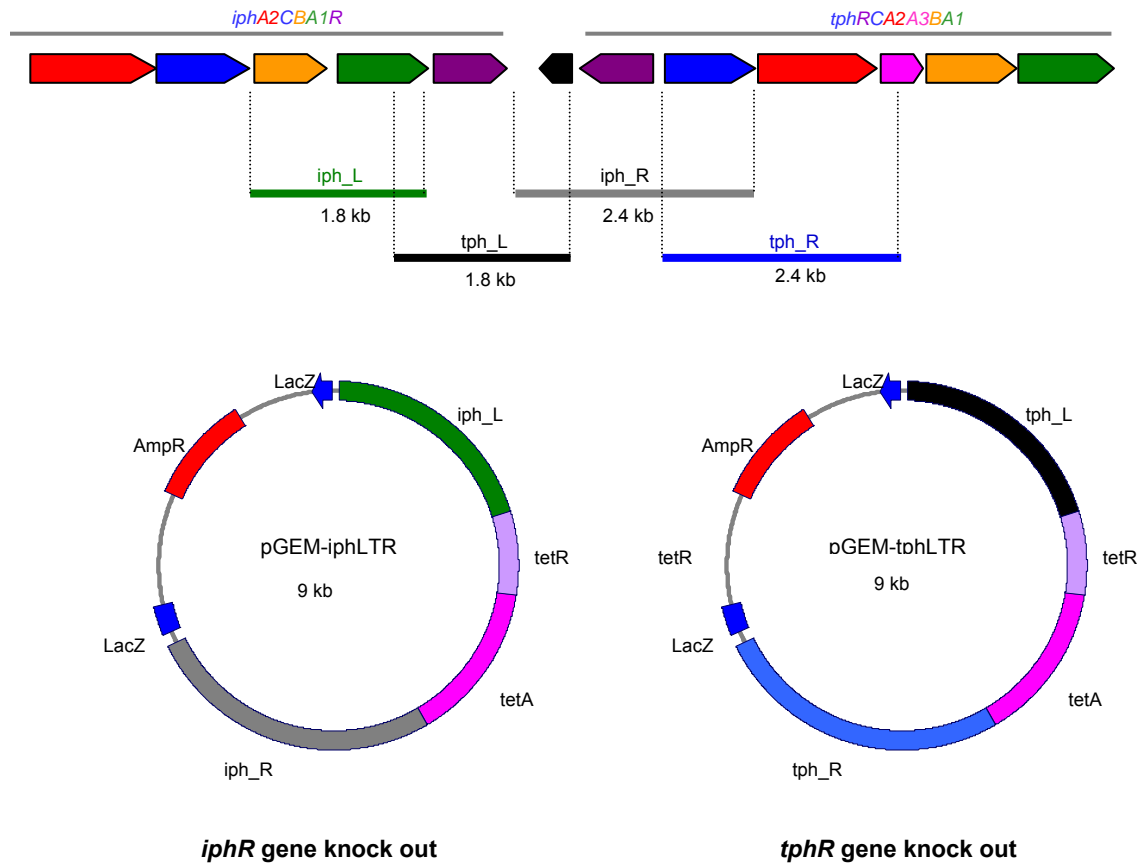
Figure 2.3 Generation of plasmid constructs for the *iphA2* and *tphA2* gene knock outs.
The *iph* (*iphA2CBA1R*) and *tph* (*tphRCA2A3BA1*) gene organizations are shown in the figure.
The 1.8 kb *iphR* left flanking region (iph_L), the 2.4 kb *iphR* right flanking region (iph_R), the 1.8
kb *tphR* left flanking region, and the 2.4 kb *tphR* right flanking region were amplified to generate
the constructs for the *iphA2* and *tphA2* gene knock out.   Plasmid pGEM-iphLTR (containing
iph_L, tetRA, iph_R) and pGEM-tphLTR (containing tph_L, tetRA, tph_R) were generated for the
iphR and *tphR* gene knock out.

2.18 *ophA2*, *iphA2*, and *tphA2* Gene Expression in Knockout Mutants

In order to determine the *ophR* and *iphR* gene function for phthalate isomer degradation in YZW-B, relative quantification of RT-qPCR was carried out to measure *ophA2*, *iphA2*, and *tphA2* gene expression in the *ophR* and *iphR* gene knockout mutants YZW-BΔ*ophR* and YZW-BΔ*iphR*. For the *ophR* gene, RNA samples were isolated from YZW-BΔ*ophR* cells grown in minimal liquid media containing 20 mM succinate or phthalate. For each substrate, RNA was sampled in triplicate independently. For each RNA sample, real time qPCR was carried out in triplicate to measure the C_T value for *rpoB*, *ophA2*, *iphA2*, and *tphA2* using the same method as was discussed before. Finally, the relative expression of the *ophA2*, *iphA2*, and *tphA2* genes in YZW-BΔ*ophR* grown on phthalate was given by $2^{-\Delta\Delta C_T}$ (68) after normalization to the endogenous control *rpoB* gene expression and relative to cells grown on succinate as calibrator.

For *iphR*, a similar RT-qPCR method was used to measure *rpoB*, *ophA2*, *iphA2*, and *tphA2* except that RNA samples of YZW-BΔ*iphR* were isolated from cells grown on isophthalate or succinate.

3. Results and Discussion

3.1 Phylogenetic Analysis of Phthalate Degrading Bacteria

These nine phthalate isomers degradative strains were originally isolated from Passaic river sediment and have been classified as *Pseudomonas* (YZW-A, YZW-G), *Acinetobacter* (YZW-C, YZW-H), *Comamonas* (YZW-B, YZW-D, YZW-E, YZW-F) and *Arthrobacter* (YZW-I) using the Biolog identification system (123). In order to further classify each isolate down to species level, nearly full length 16S rDNAs from these nine strains were PCR amplified and sequenced.

16S rDNA sequences analysis show that the nine phthalate degrading bacterial isolates belong to different genera such as *Pseudomonas*, *Acinetobacter*, *Comamonas*, and *Arthrobacter* (Figure 3.1), which is consistent with the results from the carbon source utilization-based Biolog identification system (123). Blastn results show that the YZW-A 16S rDNA shares 99% identity with 16S rDNA of *Pseudomonas putida* KT2440, which actually contains three identical copies of 16S rDNA in its genome (82), and *Pseudomonas monteilii* CIP104883 (AF064458) (34). Also, YZW-G 16S rDNA shares 99% identity with 16S rDNA from *P. putida* KT2440 (82) and *P. monteilii* CIP104883 (34). YZW-A and YZW-

G 16S rDNA share 99% identity and are grouped with other *Pseudomonas* species, such as *P. putida* DSM 291T (Z76667) (79), *P. plecoglossicida* FPC951 (AB009457) (84), and *P. oryzihabitans* IAM1568 (D84004) (5).

YZW-C and YZW-H are closest to *Acinetobacter lwoffii* DSM2403 (X81665) (98) and *A. haemolyticus* DSM6962 (X81662) (98), with 99% and 98% 16S rDNA identity, respectively. Certainly, YZW-C and YZW-H are clustered with other *Acinetobacter* species, such as *A. calcoaceticus* ATCC 23055T (DNA group one) (Z93434) (50), *A. radioresistens* DSM6976 (X81666) (98), *A. johnsonii* DSM6963 (X81663) (98), *A. schindleri* LUH5832 (AJ278311) (83), and *A. calcoaceticus* DSM30006 (X81661) (98).

YZW-B, YZW-D, YZW-E, and YZW-F were identified as *Comamonas testosteroni* on the basis of 16S rDNA sequences, which exhibit 99% identical to *C. testosteroni* ATCC11996 (M11224) (17). These four strains share 100% 16S rDNA identity with each other. They are clustered with other related *Comamonas* species, such as *C. koreensis* KCTC12005 (AF275377) (17), *C. aquatica* LMG2370 (AJ430344) (124), *C. denitrificans* 123 (AF233877) (43), and *C. terrigena* IMI359870 (AF078772) (125). In addition, the 16S rDNA sequences

from *C. testosteroni* isolates YZW-B,-D,-E, and -F share 95% identity to another

terephthalate degradative bacterium, *Delftia tsuruhatensis* T7 (AB075017) (107).

YZW-I belongs to the Gram positive bacteria *Arthrobacter*, which is

phylogenetically distant from the other isolates. YZW-I shows 98% identical to *A.*

kerguelensis KGN15 (AJ606062) (44) and *A. psychrophenolicus* DSM15454

(AJ616763) (75), and groups with *Arthrobacter sulfureus* DSM20167 (X83409)

(61) and *Arthrobacter gangotriensis* Lz1Y (AJ606061) in a single cluster (44).

The ability of these strains to degrade specific phthalate isomers are as follows:

YZW-A: terephthalate; YZW-B: phthalate, isophthalate, and terephthalate; YZW-

C: isophthalate; YZW-D: isophthalate and terephthalate; YZW-E: phthalate,

isophthalate, and terephthalate; YZW-F: phthalate and terephthalate; YZW-G:

phthalate; YZW-H: isophthalate; and YZW-I: isophthalate. Thus, *Pseudomonas*,

Acinetobacter, and *Arthrobacter* isolates could utilize only one of the three

phthalate isomers, while the *Comamonas* isolates could degrade two or three of

the phthalate isomers. For instance, YZW-B and YZW-E have the capacity to

degrade all of three isomers. In addition, YZW-B could grow on naphthalene as

a sole carbon source (Table 3.1).
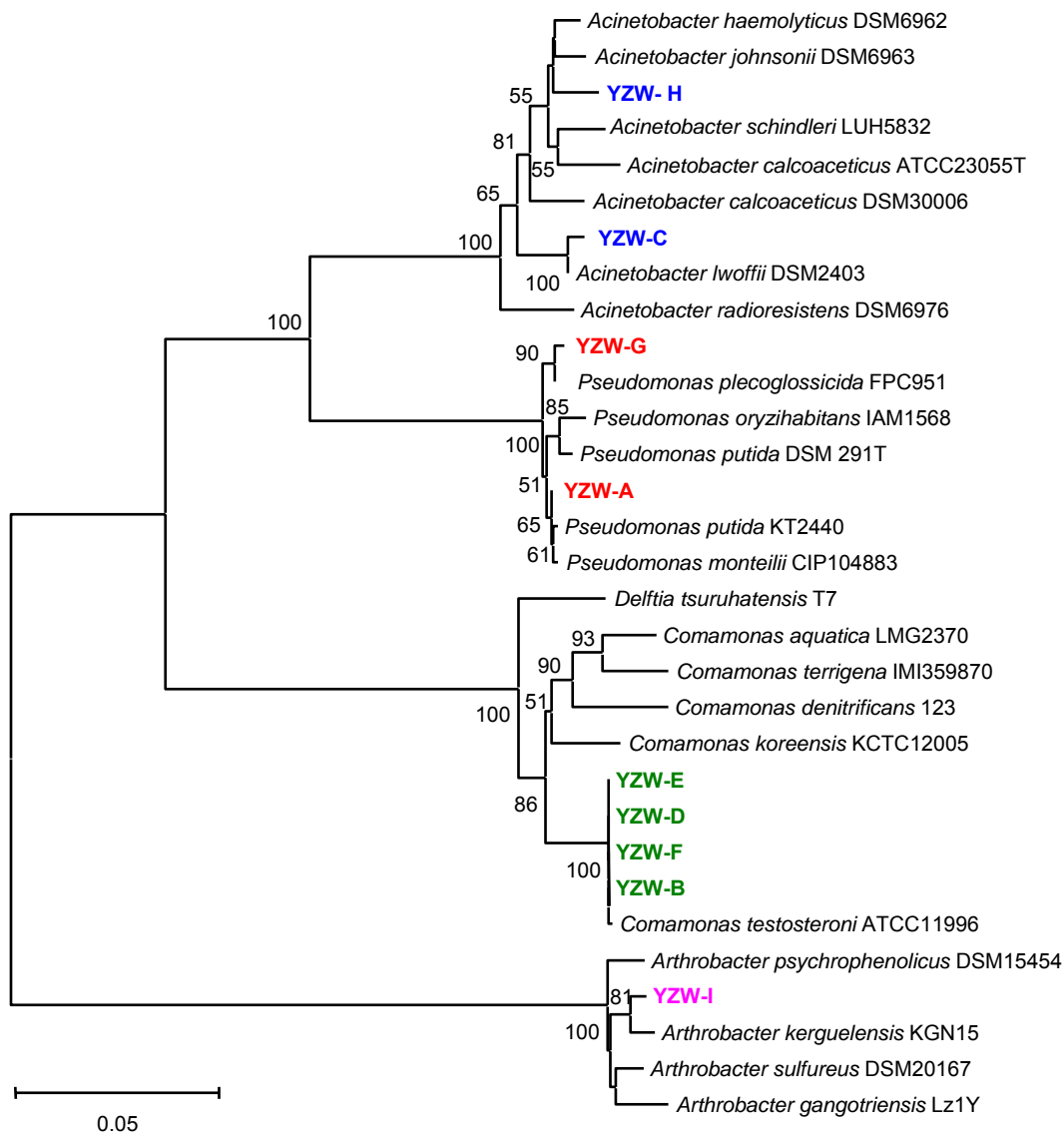
Figure 3.1 Phylogenetic tree of 16S rDNA of YZW-A, -B, -C, -D, -E, -F, -G, -H, -I and their related bacterial species.

Table 3.1 Phthalate, isophthalate, terephthalate and naphthalene as substrates for bacterial growth

Bacterial strain name	Phthalate	Isophthalate	Terephthalate	Naphthalene
Pseudomonas sp. strain YZW-A	-	-	+	-
Comamonas testosteroni YZW-B	+	+	+	+
Acinetobacter sp. strain YZW-C	-	+	-	-
Comamonas testosteroni YZW-D	-	+	+	-
Comamonas testosteroni YZW-E	+	+	+	-
Comamonas testosteroni YZW-F	+	-	+	-
Pseudomonas sp. strain YZW-G	+	-	-	-
Acinetobacter sp. strain YZW-H	-	+	-	-
Arthrobacter sp. strain YZW-I	-	+	-	-

3.2 Comparative Analysis of Genes for Isophthalate and Terephthalate Degradation in Different Bacterial Species

3.2.1 Gene organization of the *iph* and *tph* genes in YZW-A, B, D, E, and F

The isophthalate (*iph*) and/or terephthalate (*tph*) genes from YZW-B, -E, and -F were determined by PCR amplification and sequencing using primers designed according to the YZW-D *iph* or *tph* genes (123). So far, the *iph* and *tph* gene clusters (*iphA2CBA1R*) have only been examined in *C. testosteroni* YZW-D (123) and *C. testosteroni* KF-1. In addition, *Comamonas* sp. strain E6 has two almost identical terephthalate gene clusters, $tphR_IC_IA2_IA3_IB_IA1_I$ and $tphR_{II}C_{II}A2_{II}A3_{II}B_{II}A1_{II}$ (102). Sequence analysis shows that the isophthalate or terephthalate genes from YZW-B, -E, and/or -F have the same gene organization as those from YZW-D and KF-1 (Figure 3.2). The *iph* gene cluster in YWZ-B, -D, or -E contain an isophthalate gene cluster (*iphA2CBA1R*) encoding for a dioxygenase oxygenase (*iphA2*), permease (*iphC*), dehydrogenase (*iphB*), and dioxygenase reductase (*iphA1*) and regulatory protein (*iphR*) (Table 3.2-3.4). The YZW-B, -D, -E, or -F terephthalate gene operon (*tphRCA2A3BA1*) encodes for a regulatory protein (*tphR*), permease (*tphC*), dioxygenase large subunit (*tphA2*), dioxygenase small subunit (*tphA3*), dehydrogenase (*tphB*), and dioxygenase reductase (*tphA1*) (Table 3.2-3.5). These two gene clusters in YZW-B, -D and -E are linked by a gap region with differing lengths.

The terephthalate gene operon in YZW-A was determined by PCR using degenerate primers designed on the basis of YZW-D and *Rhodococcus* sp. strain DK17 *tph* gene sequences. Different from the *C. testosteroni* YZW-B, -D, -E, -F, KF-1, and E6 *tph* gene operon (*tphRCA2A3BA1*), the *tph* genes in *Pseudomonas* sp. strain YZW-A (*tphRA2A3BA1C*) exhibit the similar gene organization to the Gram positive *Rhodococcus* sp. strains DK17 (21) and RHA1 (77, 90). The YZW-A *tphC* is located downstream of *tphA1* in YZW-A, while the *Comamonas* *tphC* gene is inserted between *tphR* and *tphA2* in YZW-B, D, E, or F. Interestingly, the TphC in YZW-A and *Comamonas* strains belong to different protein families. A conserved domain search shows that YZW-A TphC contains the sugar transportor domain, while YZW-B TphC contains the Bug domain (*Bordetella* uptake gene product).
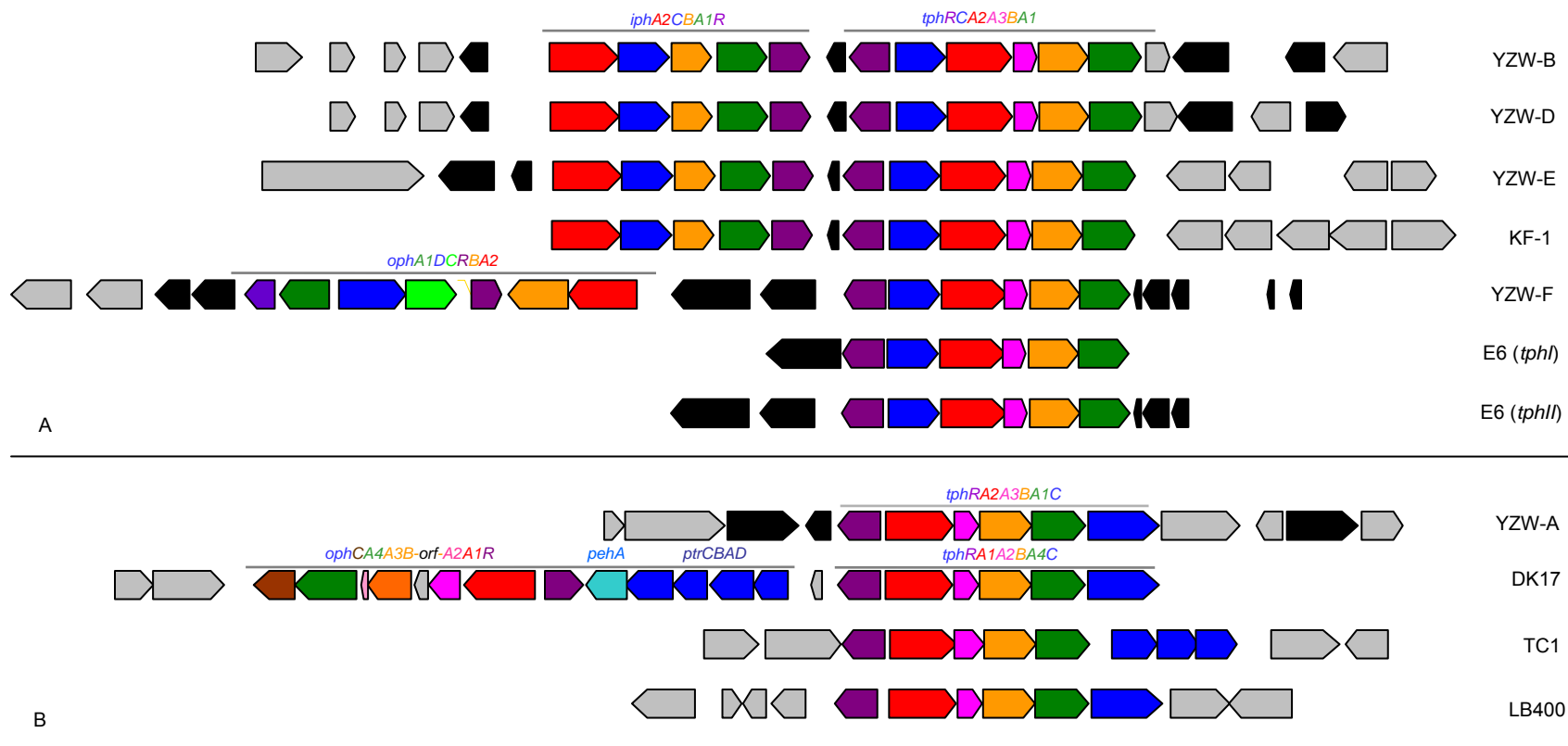
Figure 3.2 Gene organization of the isophthalate and/or terephthalate degradation in YZW-A, -B, -D, -E, and -F. The sequences of the isophthalate (*iph*) and/or terephthalate (*tph*) degradative genes and their flanking regions in YZW-A, -B, -E, and -F were determined in this study. The sequences of strain YZW-B and other reference bacterial strains were retrieved from Genbank. Hypothetical protein and transposase genes are shown in gray and black, respectively. A) In *Comamonas testosteroni* YZW-B, YZW-D (AY923836), YZW-E, and KF-1 (NZ_AAUJ01000006), *iph* genes and *tph* genes are physically linked. YZW-F *tph* genes are linked with the phthalate degradative gene (*ophA1DCRBA2*). *Comamonas* sp. strain E6 contains two copies of *tph* gene operon, *tphI* (AB238678) and *tphII* (AB238679). B) Gene organization of the terephthalate degradative genes (*tphRA2A3BA1C*) in YZW-A is similar to the *tph* operons in *Rhodococcus* sp. strains DK17 and RHA1. The terephthalate degradative genes in DK17 and RHA1 are linked with the phthalate and phthalate ester degradative genes (only DK17 shown). In *Rhodococcus* sp. strain DK17 (AY502076), *ophCA4A3BA2A1R*, *pehA*, *ptrCBAD*, and *tphRA1A2BA4C* encode the phthalate degradation pathway, phthalate esterase, transporter proteins, and the terephthalate degradation pathway. A similar gene organization was identified in *Arthrobacter aurescens* TC1 (NC_008712) and *Burkholderia xenovorans* LB400 (NC_007952).

78

3.2.2 Comparative analysis of the *iph* genes

The phylogenetic trees based on the putative protein sequences of *iphA2*, *iphB*, *iphC*, *iphA1*, and *iphR* from YZW-B, YZW-D, and YZW-E are clustered in a single group (Figure 3.3-3.7), which indicate they are divergent from other reference sequences, but share high similarity among *C. testosteroni* isolates. Comparative analysis of gene sequences and putative protein sequences show that the YZW-B and YZW-D *iph* genes are closer to each other, but they are relatively distant from YZW-E *iph* genes. For instance, the isophthalate dioxygenase gene *iphA2* from YZW-B shares 99% and 92% DNA identity with *iphA2* form YZW-D and YZW-E, and 100% and 98% protein identity with YZW-D and YZW-E IphA2.

To compare the *Comamonas iph* genes with other related sequences in the database, we used YZW-B as representative of the *Comamonas* group for Blast searching (Table 3.2). Blastp results show that YZW-B isophthalate dioxygenase IphA2 shares 98% identity and 99% similarity with a potential Rieske non-heme iron oxygenase in *C. testosteroni* KF-1 (ZP_01521495). The IphA2 in YZW-B, YZW-D, YZW-E and KF-1 (ZP_01521495) are grouped in a single cluster. In addition, YZW-B IphA2 shared 37% identity and 53% similarity to *cbaA* in *C. testosteroni* BR60, which encodes for 3-chlorobenzoate -3,4/4,5- dioxygenase (Q44256) (81). Since the 3-chlorobenzoate molecular structure is similar to isophthalate, it is not surprising that dioxygenases for 3-chlorobenzoate and

isophthalate somehow share certain similarity, which indicates that they could catalyze their substrate using similar mechanisms in the initial step. Conserved domain searching shows that the putative isophthalate dioxygenase of YZW-B contains a conserved Rieske [2Fe-2S] domain (C69, H71, C88, H91) in the N-terminus and an iron-bind domain (E170, D174, H177, H182) in the central region.

The YZW-B isophthalate dioxygenase reductase IphA1 exhibits 99% identical to IphA1 in YZW-D, and shares 95% identity with a potential ferrodoxin sequence in *C. testosteroni* KF-1. IphB belongs to the short chain dehydrogenase family. YZW-B isophthalate dihydrodiol dehydrogenase IphB shares 100% identity with the YZW-D IphB protein sequence, and 96% identity with a potential dehydrogenase in KF-1. A conserved domain search shows that the isophthalate permease IphC sequence contains the Bug domain (*Bordetella* uptake gene). YZW-B IphC shares 98% and 93% identity with YZW-D IphC and an uncharacterized protein in KF-1, respectively. IphR belongs to the IclR family of transcriptional regulators. YZW-B IphR shares 99% and 98% identity with the YZW-D isophthalate regulator and a regulator in KF-1. In order to investigate the biological function of IphR in isophthalate degradation, we knocked out the *iphR* gene of YZW-B and analyzed the expression of isophthalate dioxygenase gene *iphA2* using quantitative real time PCR. RT-qPCR analysis revealed that IphR played the role of a repressor in isophthalate degradation, as discussed in detail in Chapter 4.

Figure 3.3 Phylogenetic tree of the isophthalate dioxygenase oxygenase and related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.

Figure 3.4 Phylogenetic tree of *cis*-isophthalate dihydrodiol dehydrogenase and related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.

Figure 3.5 Phylogenetic tree of the isophthalate transporter and related proteins.

The Genbank accession number and the bacterial strain are shown on the dendrogram.
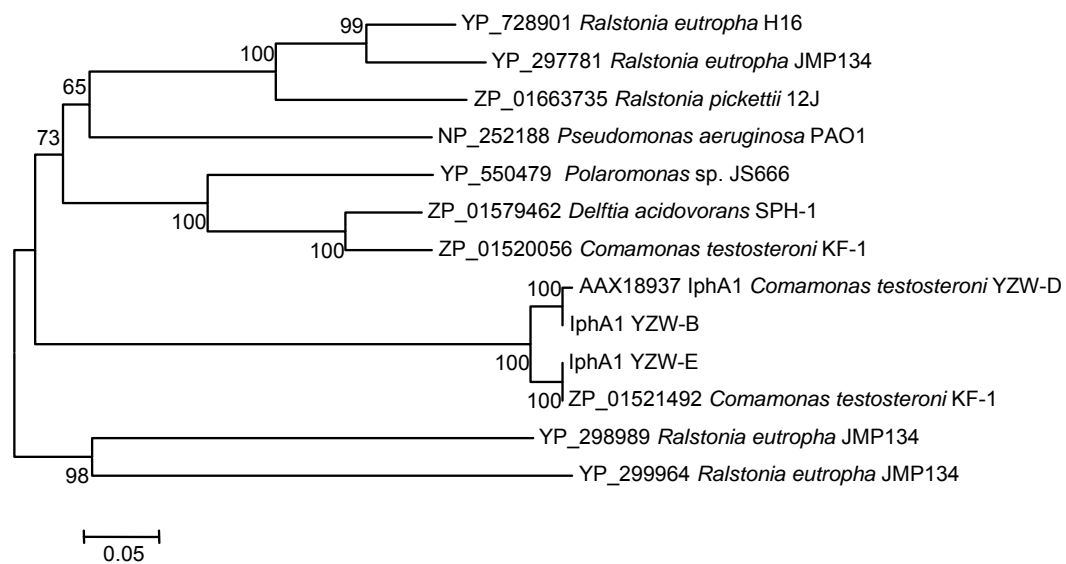
Figure 3.6 Phylogenetic tree of the isophthalate dioxygenase reductase and related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.
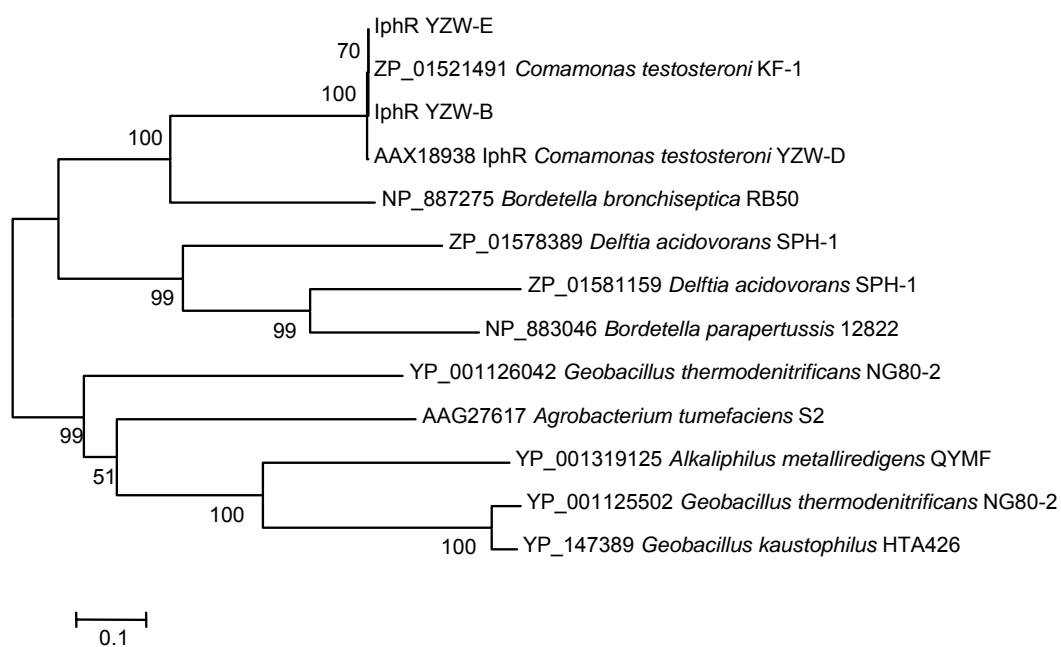
Figure 3.7 Phylogenetic tree of the isophthalate regulator and related proteins. The

Genbank accession number and the bacterial strain are shown on the dendrogram.

3.2.3 Comparative analysis of *tph* genes

The *tphR* gene of YZW-A, -B, -D, -E, and -F encodes for a putative transcriptional regulatory protein for terephthalate degradation, which is a member of the IclR family. TphR from YZW-B, -D, -E, and -F clusters with related sequences from *Comamonas* strains in a phylogenetic tree, while TphR of YZW-A groups with potential regulators from *Rhodococcus* sp. strain DK17 (21) and *Rhodococcus* sp. strain RHA1 (YP_707378) (77) (Figure 3.8). This observation is also consistent with the low identity and similarity (41% identity, 60% similarity) between TphR of YZW-A and *C. testosteroni* isolates, and high identity and similarity (98%-100%) among TphR from *Comamonas* strains. TphR of YZW-A shares 65% and 56% identity with the putative terephthalate regulator from *Rhodococcus* sp. strains DK17 (AAR90186) (21) and RHA1 (YP_707378 and YP_708551) (77), respectively. In addition, TphR of YZW-B shares 42% identity with PcaR, a putative regulator in the protocatechuate catabolic gene cluster from *Rhodococcus opacus* (AAC38247) (37).

Terephthalate dioxygenase TphA2 of YZW-B shares 99% protein sequence identity with TphA2 in YZW-D (123), E6 (BAE47085, BAE47077) (102), and KF-1, and share 98% protein sequence identity with the terephthalate 1,2-dioxygenase large subunit of *Delftia tsuruhatensis* T7 (BAC15591) (106). But TphA2 of YZW-B only shares 68% identity with the putative terephthalate 1,2-dioxygenase in *Rhodococcus* sp. strains DK17 (AAR90187) (21) and RHA1

(YP_707379 and YP_708551) (77). A conserved Rieske [2Fe-2S] domain (C82, H84, C102, H105) and iron-binding domain (E203, D207, H210, H215) are found in the N-terminus and C-terminus of terephthalate dioxygenase. The TphA2 protein sequence of *Pseudomonas* sp. strain YZW-A only shares 68% identity with that of YZW-B, but exhibits a much closer relationship with *Rhodococcus* sp. strains DK17 (AAR90187) (21) and RHA1 (YP_707379) (77) (77% identity). As shown in the phylogenetic tree (Figure 3.9), TphA2 from the *Comamonas* strains (YZW-B, -D (123), -E, -F, E6 (102), KF-1) and *D. tsuruhatensis* T7 (BAC15591) (106) clusters in a single group with above 98% identity, while TphA2 from YZW-A and *Rhodococcus* sp. strains DK17 (AAR90187) (21) and RHA1 (YP_707379 and YP_708550) (77) form another group.

The small subunit of terephthalate 1,2 dioxygenase TphA3 from YZW-B shares 100% identity with YZW-D TphA3 (123), and shares 97% and 95% identity with two copies of TphA3 in *Comamonas* sp. strain E6 (BAE47086 and BAE47078) (102). TphA3 from YZW-A shares 61% identity and 60% identity with the putative terephthalate 1,2 dioxygenase small subunit from *Rhodococcus* sp. strains DK17 (AAR90188) (21) and RHA1 (YP_707380 and YP_708549) (77). TphA3 of YZW-B only shares 50% identity with TphA3 from YZW-A.

The identities of TphA1 among *Comamonas* strains are higher than 90%. For instance, TphA1 of YZW-B shares 96%, 95%, 93%, 92% identity with YZW-D TphA1 (AAX18944) (123), KF-1 reductase (ZP_01521485), and two reductases

(BAE47088 and BAE47080) in the E6 strain (102), respectively. While *Pseudomonas* sp. strain YZW-A TphA1 shares 42% identity with those from *Comamonas* isolates, YZW-A TphA1 is phylogenetically closer to the terephthalate 1,2-dioxygenase ferredoxin reductase subunit of *Rhodococcus* sp. strains RHA1 (YP_707382 and YP_708547) (77) and DK17 (AAR90190) (21) with 66% and 65% identity, respectively (Figure 3.10).

Similarly, TphB genes from *Comamonas* strains groups in a single cluster in the phylogenetic tree (Figure 3.11), indicating high identity with each other (above 95% in the protein sequence). YZW-A TphB clusters closely to a putative protein from *Rhodococcus* sp. strains DK17 (AAR90189) (21) and RHA1 (YP_707381 and YP_708548) (77) (65% identity).

Interestingly, the terephthalate permease gene *tphC* of YZW-A and *C. testosteroni* YZW-B, -D, -E, or -F do not share any similarity at the DNA and protein sequence level. A conserved domain search shows that YZW-A TphC contains the sugar transportor domain, while YZW-B TphC contains the Bug domain (*Bordetella* uptake gene product). TphC of YZW-A is related to putative permease in *Rhodococcus* sp. strains DK17 (AAR90191) (21) and RHA1 (YP_708546 and YP_707383) (77) with 69% identity. TphC from *Pseudomonas* strain YZW-A and *Comamonas* strains including YZW-B, -D (AAX18940) (123), -E, -F, E6 (BAE47076 and BAE47084) (102), and KF-1 (ZP_01521569) are obviously clustered into two different groups in phylogenetic tree (Figure 3.12).

The different *tphC* genes were probably recruited from different ancestors, and inserted into the genes of YZW-A and other *Comamonas* strains during bacterial evolution.
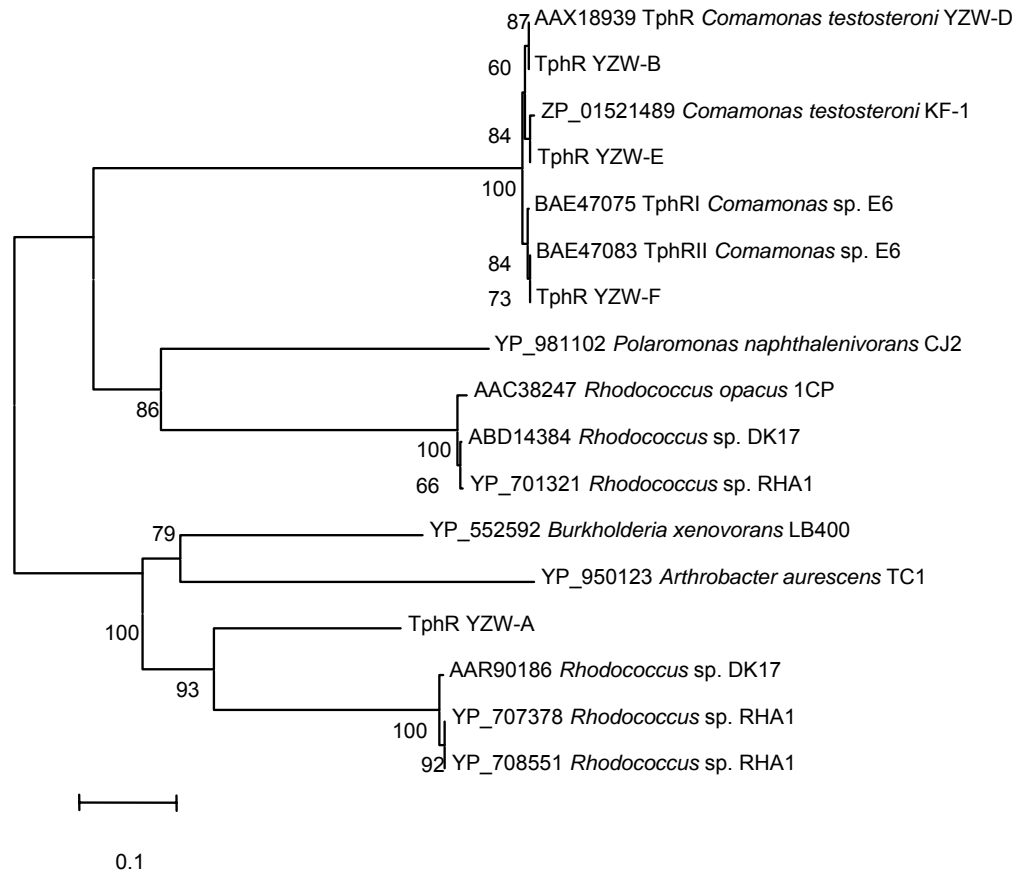
Figure 3.8 Phylogenetic tree of the terephthalate regulator and related proteins. The

Genbank accession number and the bacterial strain are shown on the dendrogram.

BAE47085 TphA2II*Comamonas* sp. E6

TphA2 YZW-B

AAX18941 TphA2 *Comamonas testosteroni* YZW-D

54

BAE47077 TphA2I *Comamonas* sp. E6

83

TphA2 YZW-E

ZP_01521488 *Comamonas testosteroni* KF-1

100

62

TphA2 YZW-F

BAC15591 TerZa *Delftia tsuruhatensis* T7

YP_950122 *Arthrobacter aurescens* TC1

YP_552593 *Burkholderia xenovorans* LB400

TphA2 YZW-A

50

YP_707379 *Rhodococcus* sp. RHA1

61

100

AAR90187 *Rhodococcus* sp. DK17

YP_708550 *Rhodococcus* sp. RHA1

0.05
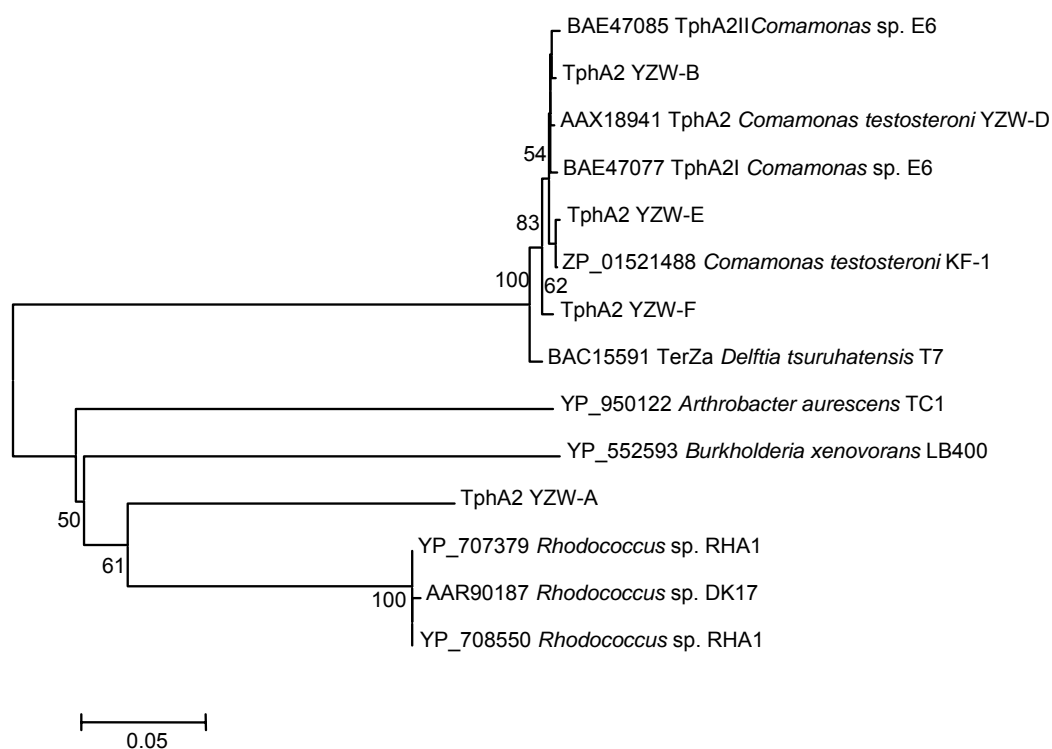
Figure 3.9 Phylogenetic tree of the terephthalate dioxygenase oxygenase large subunit and related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.
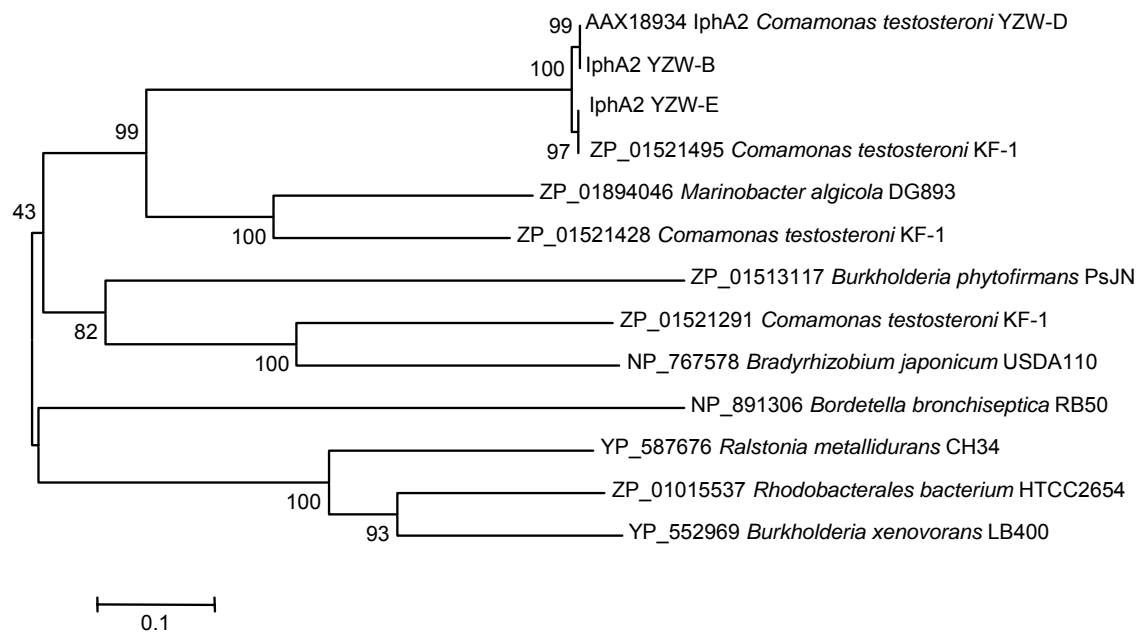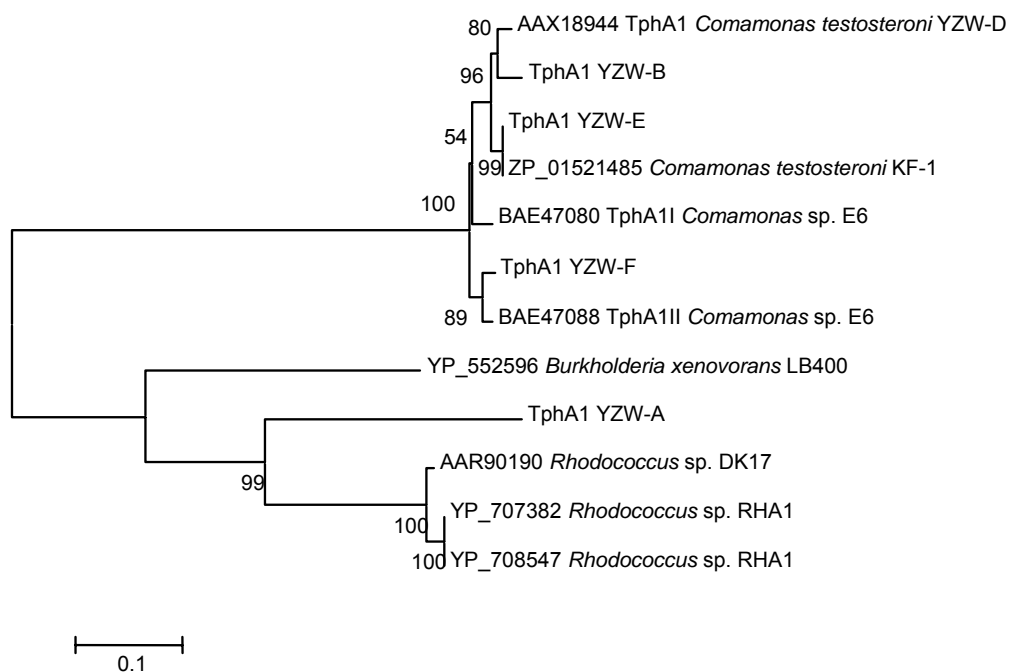
Figure 3.10 Phylogenetic tree of the terephthalate dioxygenase reductase and related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.
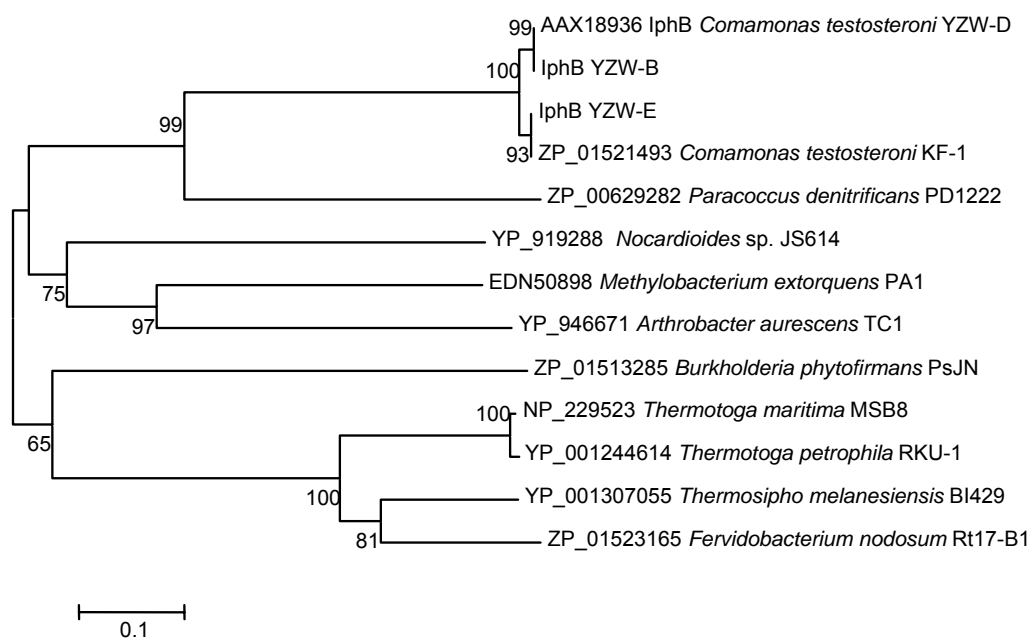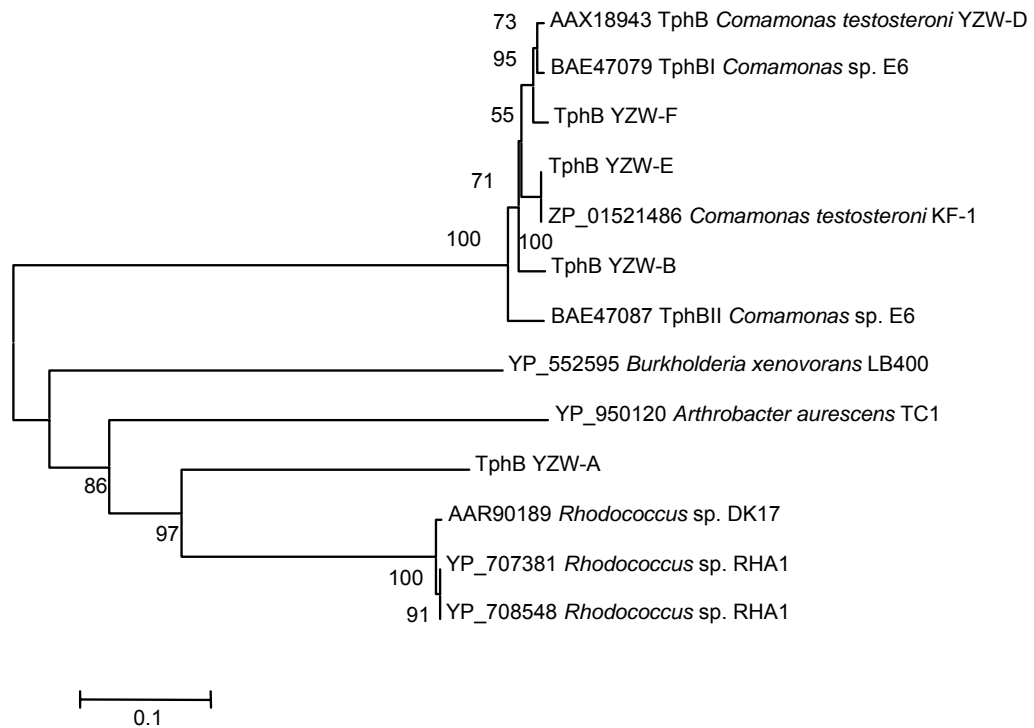
Figure 3.11 Phylogenetic tree of the *cis*-terephthalate dihydrodiol dehydrogenase and related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.
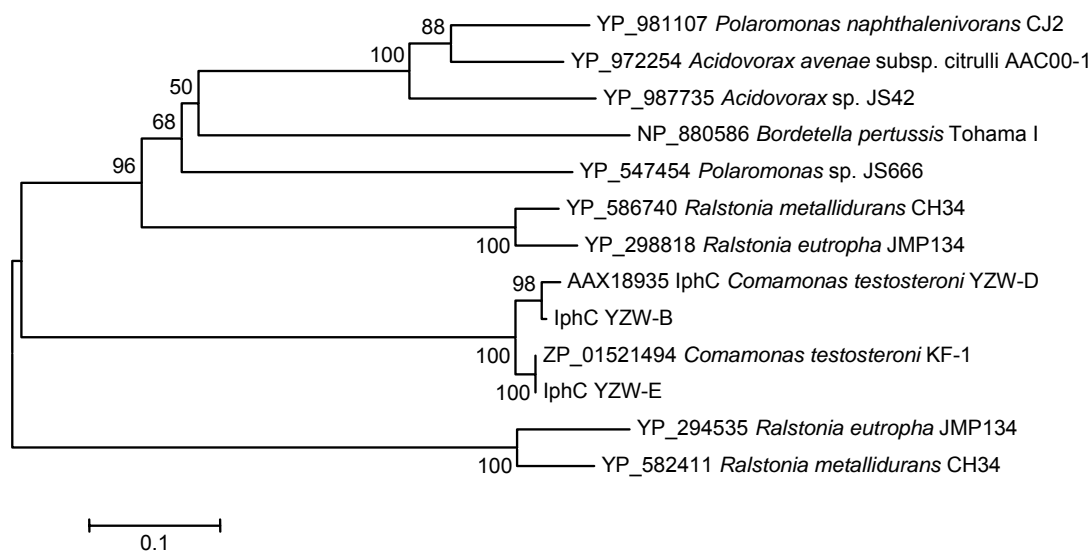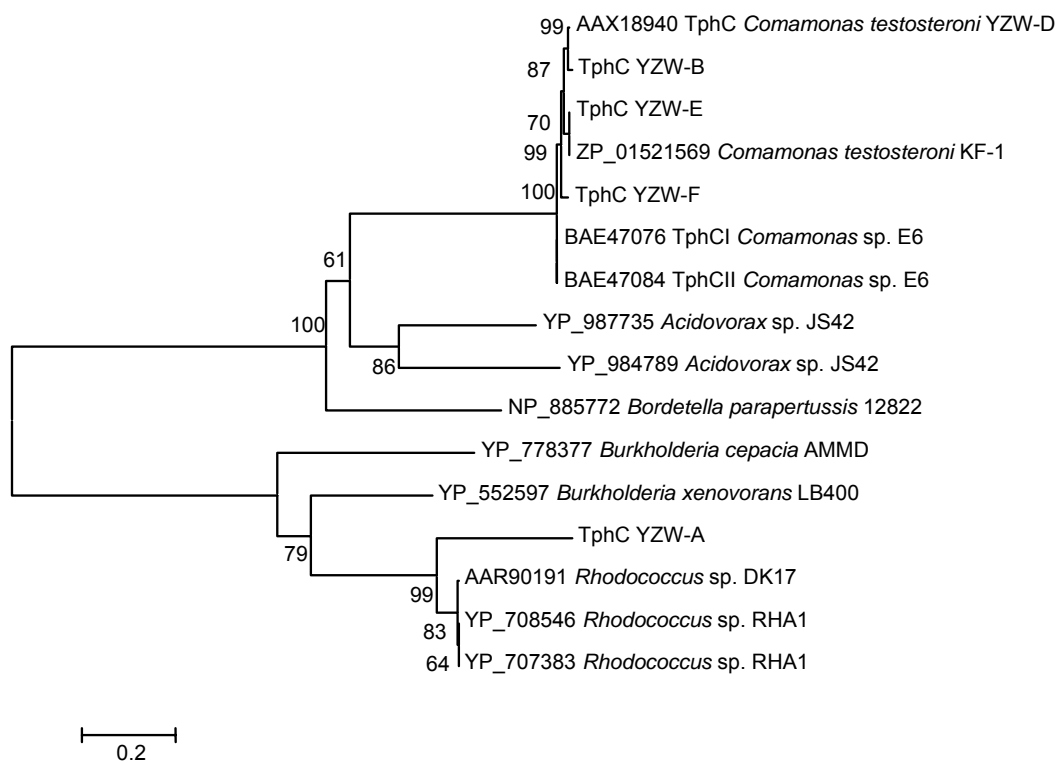
Figure 3.12 Phylogenetic tree of the terephthalate permease and related proteins.

The Genbank accession number and the bacterial strain are shown on the dendrogram.

3.2.4 Gap and flanking regions of the *iph* and *tph* genes

In *C. testosteroni* YZW-B, -D, and -E, there are gap regions located between the *iph* and *tph* gene clusters. Comparing among YZW-B, D, and E, both YZW-B and YZW-D gap regions are 823 bp long and share 99% identity with each other. Gap regions of YZW-B and YZW-E share 97% and 100% identity in two separated aligned regions. The length of the gap regions in the two strains are different, 823 bp in YZW-B and 628 bp in YZW-E due to a 199 bp sequences loss in YZW-E. YZW-B, -D and -E gap regions contain partial transposase genes which share 92% identity with *Comamonas* sp. strain E6 transposase sequence (AB238679) upstream of the *tph* gene operon (102), and 89% identity with an integrase gene from *Ralstonia metallidurans* CH34 (CP000352), which is not physically associated with any degradative genes.

The 11,239 bp sequence upstream of the *iph* gene cluster of YZW-B contains several ORFs encoding putative integrase, reductase, isochorismatase hydrolase, ABC transporter related protein. These contiguous ORFs share high identity with the corresponding ORFs in KF-1 encoding isochorismatase hydrolase (ZP_01520132), an outer membrane protein (ZP_01517644), and an ABC transporter related protein (ZP_01517645). The upstream flanking regions of the *iph* gene cluster in YZW-D (4,713 bp) and YZW-B (11,239 bp) share 97% identity.

Similarly, comparing the flanking sequences downstream of the *tph* gene cluster of YZW-B (5,601 bp) and YZW-D (4,148 bp), an 3,404 bp aligned sequence in that regions shares 99% identity and contains a putative metalloprotease and transposase gene. The region further downstream from this consensus region in YZW-B and YZW-D does not show any similarity at the nucleotide or amino acid sequence level. The downstream region of YZW-B contains two ORFs, of which the first one shares 82% identity with a resolvase-like protein in *Delftia acidovorans* SPH-1 and the second one shares 98% identity with a hypothetical protein (ZP_01522430) in *C. testosteroni* KF-1. However, the flanking regions of this hypothetical protein in KF-1 do not contain any ORFs associated with the *iph* and *tph* genes. The downstream sequence of this region in YZW-D shares 47% identity with the phage integrase family protein in *Methylobacterium* sp. strain 4-46 (ZP_01848567).

Although the *iph* and *tph* gene operons of YZW-B share 93% and 95% identity with those of YZW-E, their flanking regions do not show any similarity. Instead, ORFs upstream of the YZW-E *iph* genes share high identity with transposase (ZP_01522922) (100%), integrase (ZP_01521471) (96%), SMC protein-like protein (ZP_01522156) (100%), and a hypothetical protein (ZP_01522157) (100%) in KF-1. So far, since the *C. testosteroni* KF-1 genome sequence has not been completed, we could not determine whether this region in KF-1 is located upstream of the *iph* and *tph* genes or not. Downstream sequences of the *tph* operon in YZW-E share 100% identity with ORFs in KF-1 which encode

uncharacterized protein (ZP_01521484), enoyl-CoA hydratase/isomerase (ZP_01521483), putative acyl-CoA dehydrogenase (ZP_01521482), putative acyl-CoA dehydrogenase (ZP_01521481), and L-carnitine dehydratase/bile acid-inducible protein F (ZP_01521480). Moreover, the 18 kb sequence region including the *iph* genes, *tph* genes, and downstream region of the *tph* operon in YZW-E and KF-1 share 99% identity at the DNA level, which indicates that YZW-E is phylogenetically closer to KF-1 than to YZW-B or YZW-D.

YZW-F lacks isophthalate degradation capability, but does grow on both phthalate and terephthalate. Surprisingly, the phthalate gene cluster (*oph*) was found upstream of the *tph* gene cluster in YZW-F. The *oph* and *tph* gene cluster in YZW-F are separated by about 4 kb gap region containing two different transposase genes, which belong to the IS3 and IS110 insertion sequence family, respectively. Similarly, the two transposase genes (BAE47081 and BAE47082) associated with the second copy of the terephthalate degradation operon (*tph$_{II}$*) in *Comamonas* sp. strain E6 are identical to the transposase genes in YZW-F. In addition, these two transposase genes were also found in YZW-B. Interestingly, they are located downstream of the *oph* genes in YZW-B instead of upstream of the *tph* genes. Downstream of the *tph* genes in YZW-F are partial transposase gene and insertion element sequences.

3.2.5 Transposon carrying the *tph* gene operon in *Pseudomonas* sp. strain YZW-A

Two identical copies of a 1,302 bp transposase gene (*tnp*) were found in the upstream and downstream flanking region of the *tph* gene operon in YZW-A (Figure 3.13; Table 3.6). These two *tnp* genes are 9,452 bp away from each other. 24 bp inverted repeats are located upstream and downstream of each copy of the *tnp* gene (GGCTCTTCGCATTTAAGGGTGTAG in the 5'-end of each *tnp* gene and CTACACCCTTAAATGCGAAGAGCC in the 3'-end of each *tnp* gene). 8 bp direct repeats (TCTAATCA) are situated immediately upstream of first *tnp* gene and immediately downstream of the second *tnp* gene.

Blastp and IS finder search show that these two transposases belong to the ISL3 family. The *tnp* gene shares 99% DNA identity with insertion sequence IS1411 (M57500) in plasmid EST1226 of *Pseudomonas putida* EST1001, which was located upstream of a catechol 1, 2-dioxygenase gene (*pheB*) and a phenol monooxygenase genes (*pheA*) (54). In *Pseudomonas* sp. strain TW3, there exists a 161 bp region identical to a partial YZW-A *tnp* gene sequence. Interestingly, this region was found to be in the middle of a benzyl alcohol dehydrogenase gene (*ntnB*) in the nitrotoluene degradative gene cluster (AF043544). Thus, the *ntnB* gene actually is a pseudogene without *ntnB* original function (51). Moreover, both insertion sequences in *P. putida* EST1001 and *Pseudomonas* sp. strain TW3 contain the exact same 24 bp inverted repeats as

those in YZW-A. In addition, 13 identical copies of this *tnp* gene with inverted repeats at both ends of each gene were found in the *P. mendocina* ymp (CP000680) genome. Other similar inverted repeats (GGCTCTTCGCAGTTGAGGGTGTAG) were also found in the *Mycobacterium smegmatis* mc^{2}155 genome (CP000480, M76495) (23).

Besides the terephthalate gene operon (*tphRA2A3BA1C*), this transposon contains several ORFs encoding hypothetical proteins with unknown function (Table 3.6). The upstream flanking region of the transposon shares 94% nucleotide and 97% amino acid sequence identity with N-terminal and central region of a hypothetical protein PP3616 (1-2,111bp of total 2,991bp) from *P. putida* KT2440 (NP_745752) (82). The downstream flanking region of transposon shares 56% protein identity with the C-terminus of PP3616 (2,137-2,991bp of 2,991bp), but it does not show any DNA identity with C-terminus of PP3616. In addition, in *P. putida* F1 (YP_001268425) and *P. putida* GB-1 (ZP_01712902), PP3616-like hypothetical proteins and their flanking genes are organized in the same order as in *P. putida* KT2440 (Figure 3.13). This result suggests that YZW-A probably shares a genome backbone with *P. putida* KT2440, F1, and GB-1. This transposon inserted into the PP3616 gene in a common ancestor of *P. putida* so that YZW-A could have been evolved by recruiting the *tph* gene cluster to gain capability of terephthalate degradation, while KT2440 still contains the intact PP3616 gene. Based on the fact that YZW-A *tph* genes are closer to *Rhodococcus tph* genes and they exhibit the same

gene organization, a horizontal gene transfer event could have occurred between *Pseudomonas* and *Rhodococcus* bacterial species. Interestingly, another PP3616-like ORF was found about 2 kb away from the downstream of transposon in YZW-A, which shares 97% amino acid sequence identity and 96% nucleotide sequence identity with that upstream copy. The second copy of a gene encoding a PP3616-like protein probably maintains the normal biological function to replace the first copy which has been inactivated by the transposon insertion.
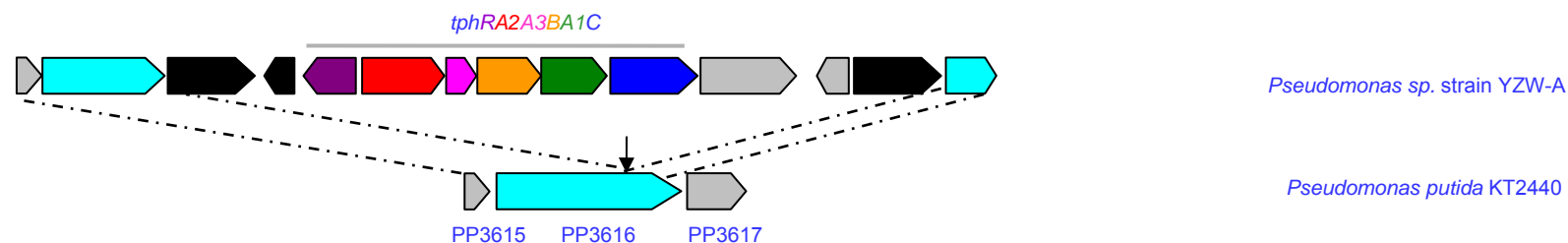
Figure 3.11 A

Figure 3.11 B
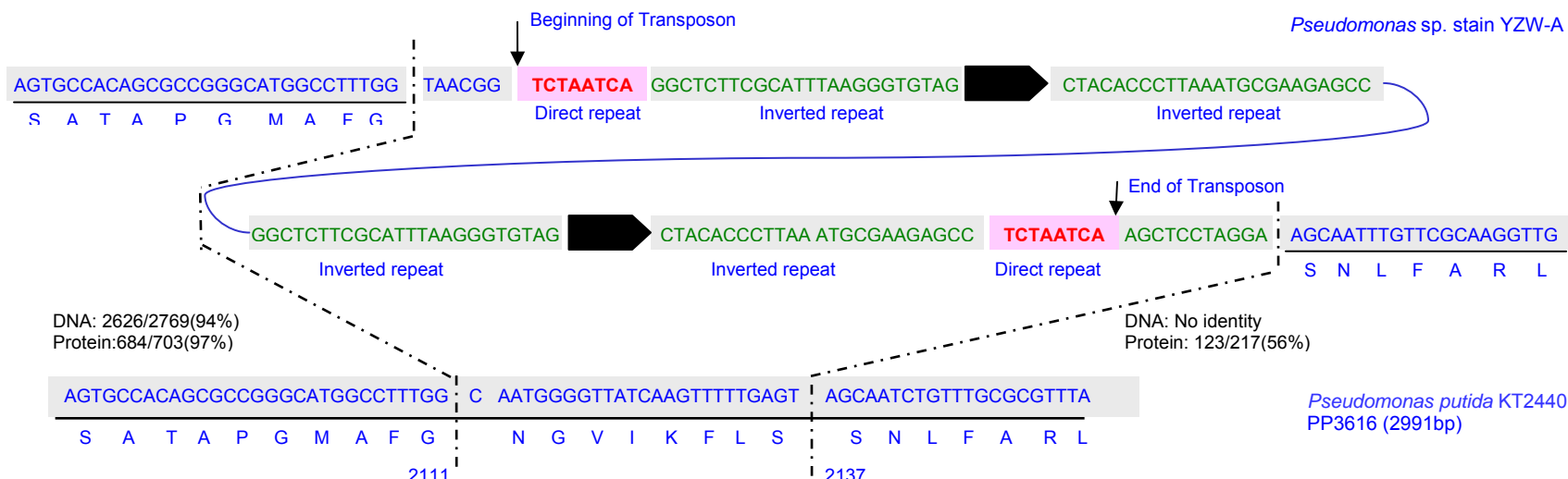
Figure 3.13 Organziation of the terephthalate catabolic transposon in *Pseudomonas* sp. strain YZW-A. A. The terephthalate degradation genes are associated with two identical transposase genes in YZW-A. This transposon is flanked by partial gene (light blue) encoding hypothetical protein which shares similarity with PP3616 protein in *P. putida* KT2440. Black arrow indicates the approximate location of the putative insertion site of the terephthalate catabolic transposon in *P. putida* genome. B. Detailed structure of transposon and its putative insertion site in PP3616 encoding gene. DNA sequences in the boundary region of transposases and its putative insertion site in PP3616 encoding gene are shown in Figure B. Direct repeats (pink) and inverted repeats (green) flanking transposase genes were identified. Black arrows indicate the beginning and end of the transposon. DNA and amino acid sequences of putative PP3616 protein are shown above and under the black lines. Dash lines indicate the insertion region of transposon in PP3616 gene. The upstream flanking region of the transposon shared 94% DNA and 97% protein identity with N-terminus and central region of a hypothetical protein PP3616 (1-2111 bp of total 2991 bp) from *P. putida* KT2440 (NP_745752) (82). The downstream flanking region of the transposon shares 56% protein identity with C-terminus of PP3616 (2137-2991 bp of 2991 bp), but it doesn't show any DNA identity with C-terminus of PP3616.

101

Table 3.2 BlastX of YZW-B *iph* and *tph* genes and flanking regions (28,168 bp)

Gene	Region	Length	Direction	Function	Source	Identity	Similarity	Gap	Accession No.	Reference
orf	1-1563	1563	→	ABC transporter related	*Comamonas testosteroni* KF-1	496/520 (95%)	507/520 (97%)	0/520 (0%)	ZP_01517645	
orf	1566-2954	1389	→	RND efflux system, outer membrane lipoprotein, NodT family	*Comamonas testosteroni* KF-1	391/461 (84%)	423/461 (91%)	0/461 (0%)	ZP_01517644	
orf	3408-4319	912	→	transcriptional regulator, LysR family	*Polaromonas* sp. JS666	204/297 (68%)	241/297 (81%)	2/297 (0%)	YP_551536	
orf	4482-5360	879	→	Pirin domain protein domain protein	*Burkholderia multivorans* ATCC 17616	197/285 (69%)	228/285 (80%)	0/285 (0%)	ZP_01570252	
orf	5487-6374	888	→	Predicted PecM protein	*Ralstonia eutropha* H16	118/266 (44%)	156/266 (58%)	2/266 (0%)	YP_724597	(94)
orf	6926-7555	630	→	isochorismatase hydrolase	*Comamonas testosteroni* KF-1	196/209 (93%)	201/209 (96%)	0/209 (0%)	ZP_01520132	
orf	7841-8401	561	→	Rieske (2Fe-2S) domain protein	*Candidatus Methanoregula boonei* 6A8	47/138 (34%)	82/138 (59%)	13/138 (9%)	YP_001405472	
orf	8781-9362	582	→	NADPH-dependent FMN reductase	*Ralstonia eutropha* H16	102/180 (56%)	131/180 (72%)	2/180 (1%)	YP_726173	(94)
orf	9414-10046	633	←	Integrase, catalytic region	*Burkholderia cenocepacia* MC0-3	221/310 (71%)	252/310 (81%)	1/310 (0%)	ZP_01567141	
iphA2	11240-12514	1275	→	isophthalate dioxygenase oxygenase component	*Comamonas testosteroni* YZW-D	424/424 (100%)	424/424 (100%)	0/424 (0%)	AAX18934	(123)
iphC	12556-13545	990	→	isophthalate permease	*Comamonas testosteroni* YZW-D	323/329 (98%)	325/329 (98%)	0/329 (0%)	AAX18935	(123)
iphB	13549-14328	780	→	isophthalate dihydrodiol dehydrogenase	*Comamonas testosteroni* YZW-D	259/259 (100%)	259/259 (100%)	0/259 (0%)	AAX18936	(123)
iphA1	14309-15310	1002	→	isophthalate dioxygenase reductase	*Comamonas testosteroni* YZW-D	323/325 (99%)	324/325 (99%)	0/325 (0%)	AAX18937	(123)
iphR	15432-16220	789	→	isophthalate regulator	*Comamonas testosteroni* YZW-D	261/262 (99%)	262/262 (100%)	0/262 (0%)	AAX18938	(123)
orf	16723-17061	339	→	Integrase, catalytic region	*Comamonas testosteroni* KF-1	84/97 (86%)	85/97 (87%)	0/97 (0%)	ZP_01518983	
tphR	17044-17790	747	←	terephthalate regulator	*Comamonas testosteroni* YZW-D	254/254 (100%)	254/254 (100%)	0/254 (0%)	AAX18939	(123)
					Comamonas sp. E6	249/252 (98%)	249/252 (98%)	0/252 (0%)	BAE47075	(102)
					Comamonas sp. E6	248/252 (98%)	249/252 (98%)	0/252 (0%)	BAE47083	(102)
tphC	17918-18886	969	→	terephthalate permease	*Comamonas testosteroni* YZW-D	319/321 (99%)	321/321 (100%)	0/321 (0%)	AAX18940	(123)
					Comamonas sp. E6	311/321 (96%)	317/321 (98%)	0/321 (0%)	BAE47076	(102)
									BAE47084	
tphA2	18901-20142	1242	→	terephthalate dioxygenase oxygenase component large subunit	*Comamonas testosteroni* YZW-D	411/413 (99%)	411/413 (99%)	0/413 (0%)	AAX18941	(123)
					Comamonas sp. E6				BAE47077	(102)
									BAE47085	
tphA3	20139-20603	465	→	terephthalate dioxygenase oxygenase component small subunit	*Comamonas testosteroni* YZW-D	154/154 (100%)	154/154 (100%)	0/154 (0%)	AAX18942	(123)
					Comamonas sp. E6	150/154 (97%)	151/154 (98%)	0/154 (0%)	BAE47086	(102)
					Comamonas sp. E6	147/154 (95%)	149/154 (96%)	0/154 (0%)	BAE47078	(102)
tphB	20600-21547	948	→	terephthalate dihydrodiol dehydrogenase	*Comamonas testosteroni* YZW-D	302/314 (96%)	308/314 (98%)	0/314 (0%)	AAX18943	(123)
					Comamonas sp. E6	302/315 (95%)	309/315 (98%)	0/315 (0%)	BAE47087	(102)
					Comamonas sp. E6	299/314 (95%)	306/314 (97%)	0/314 (0%)	BAE47079	(102)
tphA1	21557-22567	1011	→	terephthalate dioxygenase reductase	*Comamonas testosteroni* YZW-D	323/336 (96%)	324/336 (96%)	0/336 (0%)	AAX18944	(123)
					Comamonas sp. E6	314/336 (93%)	323/336 (96%)	0/336 (0%)	BAE47088	(102)
					Comamonas sp. E6	310/336 (92%)	319/336 (94%)	0/336 (0%)	BAE47080	(102)
orf	22807-23352	546	→	zinc metalloprotease	*Enterobacter sakazakii*	54/125 (43%)	77/125 (61%)	5/125 (4%)	ABK56826	(62)
orf	23475-24518	1044	←	transposase	*Mesorhizobium loti* MAFF303099	140/194 (72%)	164/194 (84%)	0/194 (0%)	NP_106702	(55)
orf	25997-26569	573	←	Resolvase-like	*Delftia acidovorans* SPH-1	156/190 (82%)	173/190 (91%)	0/190 (0%)	ZP_01580932	
orf	26920-27873	954	←	hypothetical protein	*Comamonas testosteroni* KF-1	312/317 (98%)	314/317 (99%)	0/317 (0%)	ZP_01522430	

Table 3.3 BlastX of YZW-D *iph* and *tph* genes and flanking regions (20,190 bp)

Gene	Region	Length	Direction	Function	Source	Identity	Similarity	Gap	Accession No.	Reference
orf	393-1022	630	→	isochorismatase hydrolase	*Comamonas testosteroni* KF-1	197/209 (94%)	201/209 (96%)	0/209 (0%)	ZP_01520132	
orf	1472-1867	396	→	Rieske (2Fe-2S) domain protein	*Candidatus Methanoregula boonei* 6A8	46/138 (33%)	81/138 (59%)	13/138 (9%)	YP_001405472	
orf	2247-2828	582	→	NADPH-dependent FMN reductase	*Ralstonia eutropha* H16	100/180 (55%)	130/180 (72%),	2/180 (1%)	YP_726173	(94)
orf	2880-3512	633	←	Integrase, catalytic region	*Burkholderia cenocepacia* MC0-3	160/211 (75%)	179/211 (84%)	1/211 (0%)	ZP_01567141	
iphA2	4714-5988	1275	→	isophthalate dioxygenase oxygenase component	*Comamonas testosteroni* YZW-D				AAX18934	(123)
iphC	6030-7019	990	→	isophthalate permease	*Comamonas testosteroni* YZW-D				AAX18935	(123)
iphB	7023-7802	780	→	isophthalate dihydrodiol dehydrogenase	*Comamonas testosteroni* YZW-D				AAX18936	(123)
iphA1	7807-8784	978	→	isophthalate dioxygenase reductase	*Comamonas testosteroni* YZW-D				AAX18937	(123)
iphR	8907-9695	789	→	isophthalate regulator	*Comamonas testosteroni* YZW-D				AAX18938	(123)
orf	10198-10536	339	→	Integrase, catalytic region	*Comamonas testosteroni* KF-1	49/57 (85%)	50/57 (87%)	0/57 (0%)	ZP_01518983	
tphR	10519-11265	747	←	terephthalate regulator	*Comamonas testosteroni* YZW-D				AAX18939	(123)
					Comamonas sp. E6	243/246 (98%)	243/246 (98%)	0/246 (0%)	BAE47075	(102)
					Comamonas sp. E6	242/246 (98%)	243/246 (98%)	0/246 (0%)	BAE47083	(102)
tphC	11393-12361	969	→	terephthalate permease	*Comamonas testosteroni* YZW-D				AAX18940	(123)
					Comamonas sp. E6	310/322 (96%)	318/322 (98%)	0/322 (0%)	BAE47076 BAE47084	(102)
tphA2	12376-13617	1242	→	terephthalate dioxygenase oxygenase component large subunit	*Comamonas testosteroni* YZW-D				AAX18941	(123)
					Comamonas sp. E6	411/413 (99%)	411/413 (99%)	0/413 (0%)	BAE47077 BAE47085	(102)
tphA3	13614-14078	465	→	terephthalate dioxygenase oxygenase component small subunit	*Comamonas testosteroni* YZW-D				AAX18942	(123)
					Comamonas sp. E6	150/154 (97%)	151/154 (98%)	0/154 (0%)	BAE47086	(102)
					Comamonas sp. E6	147/154 (95%)	149/154 (96%)	0/154 (0%)	BAE47078	(102)
tphB	14075-15022	948	→	terephthalate dihydrodiol dehydrogenase	*Comamonas testosteroni* YZW-D				AAX18943	(123)
					Comamonas sp. E6	296/314 (94%)	306/314 (97%)	0/314 (0%)	BAE47087	(102)
					Comamonas sp. E6	310/315 (98%)	313/315 (99%)	0/315 (0%)	BAE47079	(102)
tphA1	15032-16042	1011	→	terephthalate dioxygenase reductase	*Comamonas testosteroni* YZW-D				AAX18944	(123)
					Comamonas sp. E6	314/336 (93%)	323/336 (96%)	0/336 (0%)	BAE47088	(102)
					Comamonas sp. E6	319/336 (94%)	327/336 (97%)	0/336 (0%)	BAE47080	(102)
orf	16281-16828	546	→	zinc metalloprotease	*Enterobacter sakazakii*	53/125 (42%)	76/125 (60%)	5/125 (4%)	ABK56826	(62)
orf	16949-17992	1044	←	transposase	*Mesorhizobium loti* MAFF303099	140/194 (72%)	164/194 (84%)	0/194 (0%)	NP_106702	(55)
orf	18591-19175	585	←	hypothetical protein	*Psychrobacter arcticus* 273-4	51/196 (26%)	89/196 (45%)	32/196 (16%)	YP_263750	
orf	19646-20188	543	→	phage integrase family protein	*Methylobacterium* sp. 4-46	83/175 (47%)	113/175 (64%)	4/175 (2%)	ZP_01848567	

Table 3.4 BlastX of YZW-E *iph* and *tph* genes and flanking regions (24,387 bp)

Gene	Region	Length	Direction	Function	Source	Identity	Similarity	Gap	Accession No.	Reference
orf	105-701	597	→	hypothetical protein	*Comamonas testosteroni* KF-1	198/198 (100%)	198/198 (100%)	0/198 (0%)	ZP_01522157	
orf	721-1257	537	→	SMC protein-like	*Comamonas testosteroni* KF-1	55/55 (100%)	55/55 (100%)	0/55 (0%)	ZP_01522156	
orf	1170-4301	3132	→	SMC protein-like	*Comamonas testosteroni* KF-1	1043/1043 (100%)	1043/1043 (100%)	0/1043 (0%)	ZP_01522156	
orf	4447-5619	1173	←	Integrase, catalytic region	*Comamonas testosteroni* KF-1	268/277 (96%)	272/277 (98%)	0/277 (0%)	ZP_01521471	
orf	5731-6501	771	→	transposase IS116/IS110/IS902	*Comamonas testosteroni* KF-1	138/138 (100%)	138/138 (100%)	0/138 (0%)	ZP_01522922	
iphA2	6667-7941	1275	→	isophthalate dioxygenase oxygenase component	*Comamonas testosteroni* YZW-D	424/424 (100%)	424/424 (100%)	0/424 (0%)	AAX18934	(123)
iphC	7983-8969	987	→	isophthalate permease	*Comamonas testosteroni* YZW-D	328/328 (100%)	328/328 (100%)	0/328 (0%)	AAX18935	(123)
iphB	8973-9752	780	→	isophthalate dihydrodiol dehydrogenase	*Comamonas testosteroni* YZW-D	251/259 (96%)	254/259 (98%)	0/259 (0%)	AAX18936	(123)
iphA1	9757-10734	978	→	isophthalate dioxygenase reductase	*Comamonas testosteroni* YZW-D	309/325 (95%)	314/325 (96%)	0/325 (0%)	AAX18937	(123)
iphR	10856-11644	789	→	isophthalate regulator	*Comamonas testosteroni* YZW-D	257/262 (98%)	260/262 (99%)	0/262 (0%)	AAX18938	(123)
orf	11906-12187	282	←	Integrase, catalytic region	*Comamonas testosteroni* KF-1	93/93 (100%)	93/93 (100%)	0/93 (0%)	ZP_01521490	
tphR	12273-13019	747	←	terephthalate regulator	*Comamonas testosteroni* YZW-D	245/248 (98%)	245/248 (98%)	0/248 (0%)	AAX18939	(123)
					Comamonas sp. E6	243/246 (98%)	243/246 (98%)	0/246 (0%)	BAE47075	(102)
					Comamonas sp. E6	242/246 (98%)	243/246 (98%)	0/246 (0%)	BAE47083	(102)
tphC	13147-14115	969	→	terephthalate permease	*Comamonas testosteroni* YZW-D	313/322 (97%)	317/322 (98%)	0/322 (0%)	AAX18940	(123)
					Comamonas sp. E6	314/322 (97%)	317/322 (98%)	0/322 (0%)	BAE47076 BAE47084	(102)
tphA2	14130-15371	1242	→	terephthalate dioxygenase oxygenase component large subunit	*Comamonas testosteroni* YZW-D	410/413 (99%)	410/413 (99%)	0/413 (0%)	AAX18941	(123)
					Comamonas sp. E6				BAE47077 BAE47085	(102)
tphA3	15368-15832	465	→	terephthalate dioxygenase oxygenase component small subunit	*Comamonas testosteroni* YZW-D	146/154 (94%)	149/154 (96%)	0/154 (0%)	AAX18942	(123)
					Comamonas sp. E6	145/154 (94%)	149/154 (96%)	0/154 (0%)	BAE47086	(102)
					Comamonas sp. E6	146/154 (94%)	151/154 (98%)	0/154 (0%)	BAE47078	(102)
tphB	15829-16776	948	→	terephthalate dihydrodiol dehydrogenase	*Comamonas testosteroni* YZW-D	305/314 (97%)	309/314 (98%)	0/314 (0%)	AAX18943	(123)
					Comamonas sp. E6	298/315 (94%)	308/315 (97%)	0/315 (0%)	BAE47087	(102)
					Comamonas sp. E6	302/314 (96%)	307/314 (97%)	0/314 (0%)	BAE47079	(102)
tphA1	16786-17796	1011	→	terephthalate dioxygenase reductase	*Comamonas testosteroni* YZW-D	328/336 (97%)	330/336 (98%)	0/336 (0%)	AAX18944	(123)
					Comamonas sp. E6	319/336 (94%)	327/336 (97%)	0/336 (0%)	BAE47088	(102)
					Comamonas sp. E6	321/336 (95%)	329/336 (97%)	0/336 (0%)	BAE47080	(102)
orf	18483-19502	1020	←	Uncharacterized protein	*Comamonas testosteroni* KF-1	326/326 (100%)	326/326 (100%)	0/326 (0%)	ZP_01521484	
orf	19555-20328	774	←	Enoyl-CoA hydratase /isomerase	*Comamonas testosteroni* KF-1	257/257 (100%)	257/257 (100%)	0/257 (0%)	ZP_01521483	
orf	20336-21415	1080	←	acyl-CoA dehydrogenase-like	*Comamonas testosteroni* KF-1	359/359 (100%)	359/359 (100%)	0/359 (0%)	ZP_01521482	
orf	21412-21756	345	←	acyl-CoA dehydrogenase-like	*Comamonas testosteroni* KF-1	114/114 (100%)	114/114 (100%)	0/114 (0%)	ZP_01521481	
orf	21758-22552	795	←	acyl-CoA dehydrogenase-like	*Comamonas testosteroni* KF-1	246/254 (96%)	250/254 (98%)	2/254 (0%)	ZP_01521481	
orf	22651-23478	828	→	L-carnitine dehydratase/ bile acid-inducible protein F	*Comamonas testosteroni* KF-1	144/144 (100%)	144/144 (100%)	0/144 (0%)	ZP_01521480	
orf	23831-24253	423	→	transcriptional regulator, TetR family	*Comamonas testosteroni* KF-1	138/138 (100%)	138/138 (100%)	0/138 (0%)	ZP_01521479	

Table 3.5 BlastX of YZW-F *oph* and *tph* genes and flanking regions (26,695 bp)

Gene	Region	Length	Direction	Function	Source	Identity	Similarity	Gap	Accession No.	Reference
orf	790-1092	303	→	isophthalate dioxygenase reductase	*Comamonas testosteroni* YZW-D	74/97 (76%)	79/97 (81%)	2/97 (2%)	AAX18937	(123)
orf	1140-2279	1140	←	porin, Gram-negative type	*Comamonas testosteroni* KF-1	344/374 (91%)	356/374 (95%)	0/374 (0%)	ZP_01521416	
orf	2715-3689	975	←	uncharacterized protein	*Polaromonas naphthalenivorans* CJ2	173/327 (52%)	225/327 (68%)	3/327 (0%)	YP_973788	
tnp	4032-4586	555	←	hypothetical protein	*Sinorhizobium medicae* WSM419	92/186 (49%)	125/186 (67%)	5/186 (2%)	YP_001314563	
tnp	4829-5431	603	←	putative integrase protein	*Rhizobium etli* CFN 42	124/151 (82%)	132/151 (87%)	0/151 (0%)	NP_659803	(40)
ophH	5894-6388	495	←	hemerythrin-like metal-binding protein	*Ralstonia pickettii* 12D	126/164 (76%)	141/164 (85%)	0/164 (0%)	ZP_02008301	
ophA1	6454-7416	963	←	ferredoxin	*Ralstonia pickettii* 12D	277/320 (86%)	293/320 (91%)	0/320 (0%)	ZP_02008300	
ophD	7693-9036	1344	→	d-galactonate transporter	*Ralstonia pickettii* 12D	380/441 (86%)	404/441 (91%)	0/441 (0%)	ZP_02008299	
ophC	9080-10072	993	→	4,5-dihydroxyphthalate decarboxylase	*Comamonas testosteroni* M4-1	328/330 (99%)	329/330 (99%)	0/330 (0%)	Q59727	(66)
					Ralstonia pickettii 12D	307/330 (93%)	318/330 (96%)	0/330 (0%)	ZP_02008298	
ophR	10363-10956	594	→	transcriptional regulator, MarR family	*Ralstonia pickettii* 12D	168/193 (87%)	176/193 (91%)	1/193 (0%)	ZP_02008296	
*Gap										
ophB	18-1193	1176	←	oxidoreductase domain protein	*Ralstonia pickettii* 12D	345/391 (88%)	361/391 (92%)	0/391 (0%)	ZP_02008294	
ophA2	1213-2544	1332	←	Rieske (2Fe-2S) domain protein	*Ralstonia pickettii* 12D	391/443 (88%)	413/443 (93%)	0/443 (0%)	ZP_02008293	
tnp	3487-5025	1539	←	hypothetical transposase	*Comamonas* sp. E6	419/419 (100%)	419/419 (100%)	0/419 (0%)	BAE47081	
tnp	5324-6400	1077	←	hypothetical transposase	*Comamonas* sp. E6	358/358 (100%)	358/358 (100%)	0/358 (0%)	BAE47082	
tphR	6817-7566	750	←	terephthalate regulator	*Comamonas* sp. E6	249/249 (100%)	249/249 (100%)	0/249 (0%)	BAE47083	(102)
					Comamonas sp. E6	247/249 (99%)	249/249 (100%)	0/249 (0%)	BAE47075	(102)
					Comamonas testosteroni YZW-D	242/246 (98%)	243/246 (98%)	0/246 (0%)	AAX18939	(123)
tphC	7694-8662	969	→	terephthalate permease	*Comamonas* sp. E6	316/322 (98%)	319/322 (99%)	0/322 (0%)	BAE47076 BAE47084	(102)
					Comamonas testosteroni YZW-D	309/322 (95%)	316/322 (98%)	0/322 (0%)	AAX18940	(123)
tphA2	8678-9919	1242	→	terephthalate dioxygenase oxygenase component large subunit	*Comamonas* sp. E6	410/413 (99%)	410/413 (99%)	0/413 (0%)	BAE47077	(102)
					Comamonas sp. E6	408/413 (98%)	408/413 (98%)	0/413 (0%)	BAE47085	(102)
					Comamonas testosteroni YZW-D	410/413 (99%)	410/413 (99%)	0/413 (0%)	AAX18941	(123)
tphA3	9916-10380	465	→	terephthalate dioxygenase oxygenase component small subunit	*Comamonas* sp. E6	153/154 (99%)	154/154 (100%)	0/154 (0%)	BAE47078	(102)
					Comamonas sp. E6	152/154 (98%)	152/154 (98%)	0/154 (0%)	BAE47086	(102)
					Comamonas testosteroni YZW-D	148/154 (96%)	149/154 (96%)	0/154 (0%)	AAX18942	(123)
tphB	10344-11324	948	→	terephthalate dihydrodiol dehydrogenase	*Comamonas testosteroni* YZW-D	306/315 (97%)	311/315 (98%)	0/315 (0%)	AAX18943	(123)
					Comamonas sp. E6	307/315 (97%)	311/315 (98%)	0/315 (0%)	BAE47079	(102)
					Comamonas sp. E6	297/314 (94%)	308/314 (98%)	0/314 (0%)	BAE47087	(102)
tphA1	11334-12344	1011	→	terephthalate dioxygenase reductase	*Comamonas* sp. E6	328/336 (97%)	330/336 (98%)	0/336 (0%)	BAE47088	(102)
					Comamonas sp. E6	323/336 (96%)	329/336 (97%)	0/336 (0%)	BAE47080	(102)
					Comamonas testosteroni YZW-D	314/336 (93%)	324/336 (96%)	0/336 (0%)	AAX18944	(123)

* Gap is about 100 bp unsequenced regaion between the *ophR* and *ophB* gene.

Table 3.6 BlastX of YZW-A *tph* genes and flanking region (19,872 bp)

Gene	Region	Length	Direction	Function	Source	Identity	Similarity	Gap	Accession No.	Reference
orf	1-371	371	←	succinate dehydrogenase cytochrome b subunit family protein	*Pseudomonas putida* KT2440	119/123 (96%)	120/123 (97%)	0/123 (0%)	NP_745750	(82)
orf	434-823	390	→	hypothetical protein	*Pseudomonas putida* KT2440	115/129 (89%)	121/129 (93%)	1/129 (0%)	NP_745751	(82)
orf	1156-3266	2111	→	hypothetical protein	*Pseudomonas putida* F1	687/712 (96%)	701/712 (98%)	0/712 (0%)	YP_001267438	
					Pseudomonas putida KT2440	688/712 (96%)	696/712 (97%),	0/712 (0%)	NP_745752	(82)
tnp	3386-4687	1302	→	transposase, IS204/IS1001/ IS1096/IS1165 family protein	*Pseudomonas mendocina* ymp	432/433 (99%)	433/433 (100%)	0/433 (0%)	YP_001185932	
				putative transposase	*Pseudomonas* sp. Strain EST1001, Plasmid pEST1226	431/433 (99%)	432/433 (99%)	0/433 (0%)	AAC64902	(59)
orf	4837-5334	498	←	hypothetical protein	*Polaromonas* sp. JS666	108/149 (72%)	126/149 (84%)	0/149 (0%)	YP_551931	
tphR	5524-6312	789	←	putative transcriptional regulator	*Rhodococcus* sp. DK17	164/249 (65%)	196/249 (78%)	1/249 (0%)	AAR90186	(21)
tphA2	6448-7702	1260	→	terephthalate 1,2-dioxygenase alpha subunit	*Rhodococcus* sp. RHA1	314/405 (77%)	356/405 (87%)	4/405 (0%)	YP_707379 YP_708550	(77)
					Rhodococcus sp. DK17				AAR90187	(21)
tphA3	7720-8190	471	→	terephthalate 1,2-dioxygenase beta subunit	*Rhodococcus* sp. RHA1	94/153 (61%)	121/153 (79%)	0/153 (0%)	YP_707380 YP_708549	(77)
					Rhodococcus sp. DK17				AAR90188	(21)
tphB	8187-9185	999	→	terephthalate dihydrodiol dehydrogenase	*Rhodococcus* sp. RHA1	212/322 (65%)	251/322 (77%)	1/322 (0%)	YP_707381 YP_708548	(77)
				putative pyridoxal phosphate biosynthesis protein	*Rhodococcus* sp. DK17	192/294 (65%)	228/294 (77%)	1/294 (0%)	AAR90189	(21)
tphA1	9182-10192	1011	→	terephthalate 1,2-dioxygenase ferredoxin reductase subunit	*Rhodococcus* sp. RHA1	221/331 (66%)	262/331 (79%)	0/331 (0%)	YP_707382 YP_708547	(77)
				putative flavodoxin oxidoreductase	*Rhodococcus* sp. DK17	218/331 (65%)	262/331 (79%)	0/331 (0%)	AAR90190	(21)
tphC	10258-11598	1341	→	terephthalate transporter, MFS superfamily protein	*Rhodococcus* sp. RHA1	302/435 (69%)	359/435 (82%)	0/435 (0%)	YP_707383 YP_708546	(77)
				putative permease	*Rhodococcus* sp. DK17	302/435 (69%)	359/435 (82%)	0/435 (0%)	AAR90191	(21)
orf	11820-13252	1431	→	hypothetical protein	*Pseudomonas putida* KT2440	264/440 (60%)	336/440 (76%),	3/440 (0%)	NP_745490	(82)
orf	13487-14005	519	←	hypothetical protein	*Pseudomonas fluorescens* PfO-1	132/172 (76%)	150/172 (87%)	0/172 (0%)	YP_350472	
tnp	14140-15441	1302	→	transposase, IS204/IS1001/ IS1096/IS1165 family protein	*Pseudomonas mendocina* ymp	432/433 (99%)	433/433 (100%)	0/433 (0%)	YP_001185932	
				putative transposase	*Pseudomonas* sp. strain EST1001, Plasmid pEST1226	431/433 (99%)	432/433 (99%)	0/433 (0%)	AAC64902	(59)
orf	15557-16330	774	→	hypothetical protein	*Pseudomonas putida* KT2440	161/288 (55%)	204/288 (70%)	0/288 (0%)	NP_745752	(82)
orf	16330-17223	894	→	hypothetical protein	*Pseudomonas putida* KT2440	218/297 (73%)	243/297 (81%)	0/297 (0%)	NP_745754	(82)
orf	17220-19379	2160	→	hypothetical protein	*Pseudomonas putida* KT2440	691/712 (97%)	702/712 (98%)	0/712 (0%)	NP_745752	(82)

3.3 Comparative Analysis of Gene Organization for Phthalate Degradation in Different Bacterial Species

3.3.1 Gene organization of the *oph* gene cluster in different bacterial species

Since the sequences of the *tph* genes (*tphRA2CA3BA1*) in *C. testosteroni* YZW-F are known, we used inverse PCR to amplify the flanking regions of the *tph* genes. Initially, a DNA region encoding N-terminus of phthalate dioxygenase was found upstream of the *tph* operon. After further inverse PCR and sequencing, a phthalate degradation cluster (*ophHA1DCRBA2*) in YZW-F was identified. Sequence analysis shows that the *oph* gene cluster and the *tph* gene operon in YZW-F are linked by a 6 kb gap region containing two transposase genes (*tnp1* and *tnp2*). The phthalate gene clusters in YZW-B and YZW-E were also amplified using primers designed according to the YZW-F *oph* genes sequence. Sequence analysis shows that the phthalate genes from YZW-B, -E, and -F have the same gene organization and exhibit high identity (Figure 3.14; Table 3.5, 3.7, 3.8).

The phthalate gene cluster (*ophHA1-DCR-BA2*) encodes for a hemerythrin-like protein (*ophH*), dioxygenase reductase (*ophA1*), permease (*ophD*), decarboxylase (*ophC*), regulatory protein (*ophR*), dehydrogenase (*ophB*), and dioxygenase oxygenase (*ophA2*). Similar to the phthalate gene clusters of

Burkholderia cepacia DBO1 (15) and *Burkholderia vietnamiensis* G4, the three operons (*ophHA1, ophDCR,* and *ophBA2*) in this cluster are transcribed in different directions. The difference between the *oph* gene cluster in *Comamonas* and *Burkholderia* strains is that there are a quinolinate phosphoribosyl transferase (QAPRTase) gene *ophE* and a transposase gene *tnp* inserted between *ophR* and *ophB* in DBO1 (15) and G4, while both *ophE* and *tnp* are missing in YZW-B, -E, and -F (Table 3.5, 3.7, 3.8).

The phthalate gene operon (*ophRDA1A2BCFE*) in *Pseudomonas* sp. strain YZW-G was also revealed by PCR and sequencing. Compared to the *oph* gene cluster in *C. testosteroni* YZW-B, -E, or -F, the *oph* gene operon in YZW-G shows a different gene organization (Figure 3.14). The genes for a permease (*ophD*), dioxygenase reductase (*ophA1*), dioxygenase oxygenase (*ophA2*), dehydrogenase *(ophB*), decarboxylase (*ophC*), porin (*ophF*), quinolinate phosphoribosyl transferase (*ophE*) are transcribed in the same direction, while the regulatory protein (*ophR*) is transcribed in the opposite direction. The *oph* gene organization in YZW-G is similar to the plasmid encoded phthalate gene operon (*pth12345*) in *P. putida* plasmid pNMH102-2 (86) (Table 3.9).
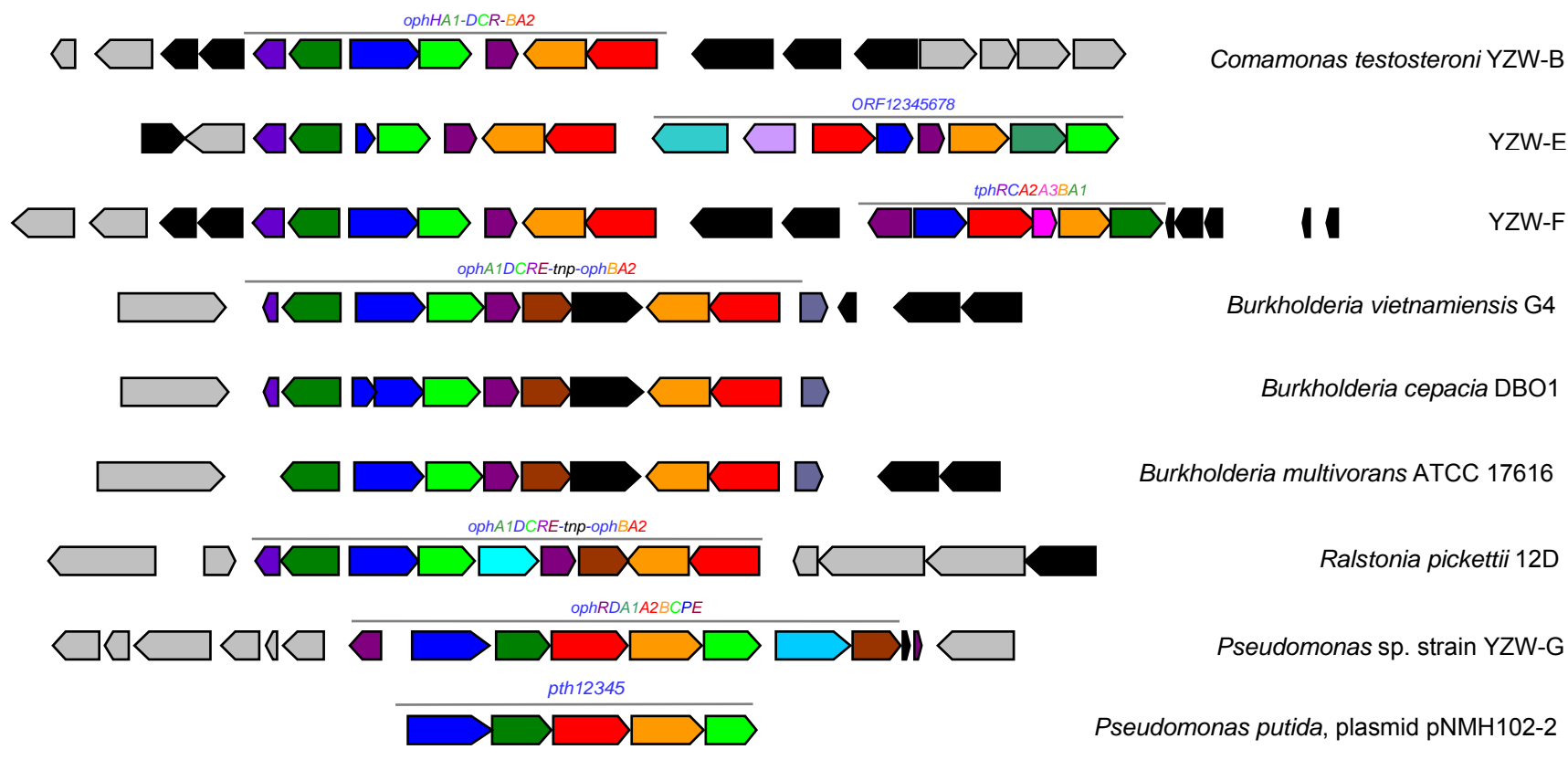
Figure 3.14 Gene organization of the phthalate degradation operon/cluster in YZW-B, -E, -F and -G. The sequences of the phthalate degradative genes (*oph*) and their flanking regions in YZW-B, -E, -F and -G were determined in this study. The sequences of other reference bacterial strains were retrieved from Genbank. Hypothetical proteins and transposase genes are shown in gray and black, respectively. In *Comamonas testosteroni* YZW-B, -E, and -F, the phthalate degradative genes *ophHA1-DCR-ophBA2* encode for a hemerythrin, phthalate dioxygenase reductase, transporter, 4,5-dihydroxyphthalate decarboxylase, regulatory protein, *cis*-phthalate dihydrodiol dehydrogenase and phthalate dioxygenase oxygenase. YZW-E *oph* gene clusters are associated with another unknown degradative gene cluster. YZW-F *oph* gene clusters are linked with the terephthalate degradative genes. In *Pseudomonas* sp. strain YZW-G, the phthalate genes *ophRDA1A2BCPE* encode a regulatory protein, transporter protein, phthalate dioxygenase reductase, phthalate dioxygenase oxygenase, *cis*-phthalate dihydrodiol dehydrogenase, 4,5-dihydroxyphthalate decarboxylase, porin, and quinolinate phosphoribosyltransferase.

3.3.2 Comparative analysis of *oph* genes

The first ORF in the *oph* gene cluster encodes a putative hemerythrin-like metal-binding protein. Hemerythrin is a non-heme diiron oxygen transport protein found in four marine invertebrate phyla. So we designated this ORF *ophH*. OphH in YZW-B shares 80% protein identity to hemerythrin-like protein in *Ralstonia pickettii* 12D (ZP_02008301) and 77% identity to hemerythrin-like protein in *Burkholderia multivorans* ATCC 17616 (ZP_01568748) and *Burkholderia vietnamiensis* G4 (YP_001116719). In *C. testosteroni* YZW-B, -E, and -F, a 495 bp *ophH* is located upstream of *ophA1*; while in *Burkholderia* stains, only about 100 bp sequence with 78% similarity to YZW-B *ophH* was found upstream of *ophA1*. Interestingly, another *ophH*-similar 477 bp ORF was found downstream of the phthalate dioxygenase gene *ophA2* instead of next to *ophA1*. No similar ORF was found in the flanking region of the phthalate gene operon in *Pseudomonas* sp. strain YZW-G.

The *ophA1* gene encodes for phthalate dioxygenase reductase in YZW-B, -E, -F and -G. OphA1 of YZW-B shares 86% identity with a ferredoxin gene in *R. pickettii* 12D (ZP_02008300), 81% identity with phthalate dioxygenase reductase in *B. cepacia* DBO1 (AAD03550) and *B. vietnamiensis* G4 (ZP_00427200), 59% identity with phthalate 4,5-dioxygenase reductase of *P. putida* (Q05182) (86). OphA1 in YZW-G shares 90% identity with OphA1 from *P. putida* (Q05182) (86).

The *ophA1* genes from YZW-G and YZW-B share no similarity at the DNA level but are 59% identity at the protein level (Figure 3.15).

The phthalate permease gene *ophD* is located downstream of *ophA1*. OphD of YZW-B shares 85% identity with a D-galactonate transporter in *R. pickettii* 12D (ZP_02008299) and 82% identity with the same gene from *B. vietnamiensis* G4 (YP_001116712), *B. multivorans* ATCC 17616 (ZP_01568741), *B. cepacia* ATCC 17616 (AAD41517) (13), and *B. cepacia* DBO1 (AAD03552) (13, 15). In YZW-E, only a 258 bp *ophD* gene sequence encoding the N-terminus of phthalate permease shares 97% identity with that in YZW-B and YZW-F, but the rest of the *ophD* gene in YZW-E is missing in genome. The permease gene in *Pseudomonas* sp. strain YZW-G shares only 64% identity with that from YZW-B (Figure 3.16).

Comparing with *ophD* in *B. cepacia* ATCC17616, *ophD* in DBO1 contains a frameshift mutation which produces a nonfunctional truncated protein. Previously, Chang and Zylstra showed that *ophD* knockout strain was able to transport phthalate into the cell at rates equivalent to that of the wild-type organism (13, 15). Thus, this gene actually is not required for phthalate utilization in DBO1.

In *B. vietnamiensis* G4, there are two copies of the D-galactonate transporter gene (YP_001116712 and YP_001116750) closely related to the phthalate

permease gene in YZW-B with 82% and 77% identity at the amino acid sequence level, respectively. The two copies of this gene are located in different positions in G4 genome. The first copy (YP_001116712) is associated with phthalate gene cluster with the same organization as the gene cluster in DBO1, but the second copy (YP_001116750) is between the upstream 4,5-dihydroxyphthalate decarboxylase gene (YP_001116749) and the downstream chemotaxis sensory transducer gene (YP_001116751). The two copies of the permease genes share 82% protein identity. In addition, the second copy of the *ophC* gene (YP_001116749) was also identified in strain G4 and is closely related to *ophC* in the phthalate gene cluster of *B. vietnamiensis* G4 (YP_001116713) and *B. cepacia* DBO1 (AAD03553) (82% identity). But other ORFs in the flanking region of the second permease copy (ZP_00423961) do not show any similarity to the ORFs in the phthalate gene cluster. These two genes encoding for decarboxylase and transporter could have been co-transferred from a common source during evolution so that one copy of the *ophD* gene was recruited for phthalate degradation and another copy could be used for other unknown functions, or both of these two copies of *ophD* genes could be involved in phthalate transport. Previous studies have shown that an *ophD* knockout mutant of *B. cepacia* ATCC17616 grew slightly more slowly on phthalate but was still able to take up phthalate at rates equivalent to that of the wild-type strain (13). Similarly, *B. multivorans* ATCC 17616 also contains a second copy of *ophD* (ZP_01568768). So far, we still don't know whether there exists a second copy of the *ophD* gene in DBO1 and YZW-E. The second copy of *ophD* gene could be

expressed when DBO1 and YZW-E grow on phthalate since the truncated copy of *ophD* in DBO1 has been proved as nonfunctional (13), and YZW-E has been missing most part of *ophD* gene in *oph* gene cluster.   Recently, Chang and Zylstra found that an ABC transporter system OphFGH in *B. cepacia* 249 (ATCC 17616) was also involved in phthalate transport (14).   The transport assays showed that The OphFGH system had higher phthalate transport efficiency than OphD when cells grown on phthalate (5 mM-50 mM) (14).

The 4,5-dihydroxyphthalate decarboxylase gene *ophC* is located downstream of *ophD* in YZW-B, -E and -F.   This gene exhibits high identity to the 4,5-dihydroxyphthalate decarboxylase gene *phtD* (Q59727) in the other phthalate degrading strains *C. testosteroni* M4-1 (99% identity) (66) and *R. pickettii* 12D (ZP_02008298) (93% identity).   The *ophC* gene in YZW-B shares 91%, 81%, 77% identity with two copies of the 4,5 -dihydroxyphthalate decarboxylase gene in *B. vietnamiensis* G4 (YP_001116713 and YP_001116749), *B. multivorans* ATCC 17616 (ZP_01568742 and ZP_01568767), and *pht5* (Q05185) in *P. putida* plasmid pNMH102-2.   In *Burkholderia* strains, each of the two *ophC* genes are physically linked with one of the *ophD* genes and the second copy of *ophC* is closely related to *ophC* in the phthalate gene cluster of *B. vietnamiensis* G4, *B. multivorans* ATCC 17616, and *B. cepacia* DBO1 (82% identity) (15).   Comparing among YZW-B, -E, -F, and -G, *ophC* in *C. testosteroni* stains shows the highest identity to each other (99% or 100% at the amino acid sequence level) and shares only 79% identical to that from YZW-G (Figure 3.17).

Based on sequence similarity, *ophR* in YZW-B, -E, or -F encodes a MarR family phthalate regulator. *ophR* in YZW-B shares 87% identity to *ophR* in the phthalate operon of *R. pickettii* 12D (ZP_02008296), *B. multivorans* ATCC 17616 (ZP_01568743), *B. vietnamiensis* G4 (YP_001116714), and *B. cepacia* DBO1 (AAD03554). Interestingly, another MarR-encoding ORF (381 bp) in G4 (YP_001116738) was found to be located downstream of *ophR* and shares 74% DNA identity to *ophR* of G4. Similarly in ATCC 17616 (ZP_01568760), there exists a MarR-encoding ORF (204 bp) exhibiting 75% DNA identity to its *ophR* gene. The second MarR-encoding ORF in G4 and ATCC17616 seems like truncated versions of *ophR*. *ophR* in YZW-B is 61% identical to *ophR* in YZW-G, which also belongs to the MarR family (Figure 3.18). Similar to the function analysis of *iphR*, we knocked out the *ophR* gene of YZW-B and analyzed the expression of the three phthalate isomer dioxygenase genes *ophA2*, *iphA2*, and *tphA2* using quantitative real time PCR. RT-qPCR analysis showed that OphR also played the role of repressor in phthalate degradation, as discussed in detail in Chapter 4.

The 4,5-dihydroxyphthalate dehydrogenase gene *ophB* in YZW-B shares 88% protein identity with an oxidoreductase gene in *R. pickettii* 12D (ZP_02008294), 86% identity with *ophB* in the phthalate gene cluster of *B. vietnamiensis* G4 (YP_001116717) and *B. cepacia* DBO1 (AAD03557) (15), and 68% identity with putative 4,5-dihydroxyphthalate dehydrogenase gene *pht4* (Q05184) in *P. putida* plasmid pNMH102-2 (86). The *ophB* gene in *Pseudomonas* sp. strain YZW-G

shares 90% identity with *pht4* (Q05184) in *P. putida* plasmid pNMH102-2, but only 67% identity with the *ophB* in YZW-B (Figure 3.19).

The phthalate dioxygenase gene *ophA2* in YZW-B shares 88% identity with a Rieske (2Fe-2S) domain protein in *R. pickettii* 12D (ZP_02008293), 84% identity with *ophA2* in phthalate gene cluster of *Burkholderia* strains including *B. multivorans* ATCC 17616 (ZP_01568747), *B. vietnamiensis* G4 (YP_001116718) and *B. cepacia* DBO1 (AAD03558). *ophA2* in YZW-B shares 75% identity with the phthalate dioxygenase gene *pht3* (Q05183) encoded on plasmid pNMH102-2 of *P. putida*, which is closely related to *ophA2* in YZW-G (92% identity) (Figure 3.20).

3.3.3 Other genes associated with *oph* gene cluster

Besides the genes for phthalate degradation in the different bacterial species, there are other conserved ORFs associated with the phthalate gene cluster in specific species, but the exact function of those ORFs for phthalate degradation still remains unknown.

ophE is a second copy of *nadC* gene encoding quinolinate phosphoribosyl transferase (QAPRTase) which is key enzyme involved in NAD biosynthesis. Phthalate degrading bacteria, such as *B. cepacia* DBO1, *B. vietnamiensis* G4, *B. multivorans* ATCC17616, *R. pickettii* 12D, have two different copies of the *nadC*

gene (*nadC* and *ophE*), while non-phthalate-degrading strains, such as *Burkholderia dolosa* AUO158, *R. pickettii* 12J and most of *Pseudomonas* strains, only contain a single housekeeping *nadC* gene. A previous study showed that *ophE* in DBO1 was inducible by phthalate and an *ophE* knockout resulted in slower growth on phthalate, which suggested that *ophE* expression on phthalate enhanced QAPRTase level and thus overcame the toxicity of phthalate on QAPRTase as a competitive inhibitor (16). *ophE* and its adjacent transposase gene *tnp* in *Burkholderia* strains are located between *ophR* and *ophB*. However, this *ophE* was found neither in this corresponding location nor somewhere else in *oph* gene cluster in the phthalate degrading *Comamonas* strains. In *C. testosteroni* YZW-B, -E, and -F, *ophR* and *ophB* are adjacent to each other, which suggests that an insertion or deletion event occur in this intergenic space. *ophE* in *Pseudomonas* sp. strain YZW-G is suited downstream of *oph* gene operon. To clarify the phylogenetic relationship of *ophE* and *nadC*, we amplified the housekeeping *nadC* gene in strain YZW-G and DBO1. Phthalate-inducible OphE of YZW-G shares 72% identity with housekeeping NadC. NadC in YZW-G shares 97% and 96% identity with nicotinate-nucleotide pyrophosphorylase of *P. putida* F1 (YP_001266157) and *P. putida* KT2440 (NP_742948) (82), while YZW-G OphE exhibits 74% identity to nicotinate-nucleotide pyrophosphorylase of *P. fluorescens* PfO-1 (YP_346517). On a dendrogram (Figure 3.21), NadC from *Pseudomonas* strains including YZW-G NadC clustered in their own branch, while YZW-G OphE formed a single family.

Similarly, OphE in DBO1 is 100% identical to that in *B. vietnamiensis* G4 (YP_001116715) and *B. multivorans* ATCC 17616 (ZP_01568744), and NadC in DBO1 is 99% identical to NadC in *B. vietnamiensis* G4 (YP_001120413) and *B. multivorans* ATCC 17616 (ZP_01571590); however, OphE in DBO1 shares only 61% identity with its housekeeping NadC. OphE and NadC in *Burkholderia* strains are grouped in their own subclusters. Interestingly, *R. pickettii* 12D also contains both *nadC* and *ophE* genes. Thus, the *Ralstonia oph* gene operon is relatively closer to that from *Burkholderia* strains than from *Comamonas* strains.

As clearly shown in the phylogenetic tree (Figure 3.21), NadC and OphE from *Burkholderia* and *Ralstonia* strains are clustered in the same group, and NadC and OphE from *Pseudomonas* strains cluster in their own group. NadC and OphE from different strains including phthalate degrading and nonphthalate degrading strains group in their own subclusters. These results suggest that *ophE* could have evolved from the housekeeping *nadC* gene in the same strain or have been recruited from closely related species by gene transfer. Thus, the *ophE* gene could be considered a parolog of the *nadC* gene whose function is especially involved in phthalate detoxification, instead of NAD biosynthesis as quinolinate phosphoribosyl transferase.

The transposase gene (*tnp*) (AAD03556), which is located between *ophE* and *ophB* in DBO1, encodes an IS256 family transposase and exhibits 98% identity to IS1356 (AAC44170) in an infectious strain *B. cepacia* ET12 (119). This *tnp*

gene is also conserved in *Burkholderia* strains including *B. multivorans* ATCC 17616 and *B. vietnamiensis* G4. There are six highly identical copies (98%-100% identity) existing in ATCC17616 (ZP_01568745, ZP_01568759, ZP_01572719, ZP_01573299, ZP_01569485, ZP_01571700), while only two copies of this *tnp* gene (YP_001116716, YP_001116787) are found in the G4 genome. Among these multiple copies of *tnp* genes, one *tnp* gene in ATCC17616 (ZP_01568745) and G4 (YP_001116716) is inserted between *ophE* and *ophB* in their genomes and formed the same gene organization as the *oph* gene cluster in DBO1. Compared to the *Burkholderia* phthalate degrading strains, *R. pickettii* 12D lacks the *tnp* gene in the intergenic region between *ophE* and *ophB*.
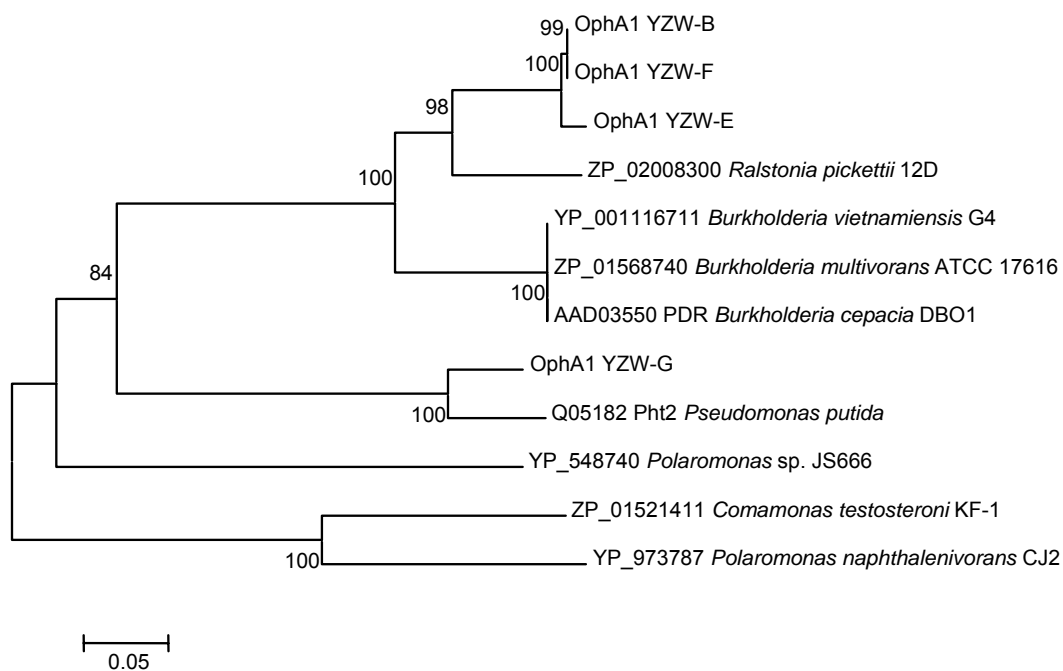
Figure 3.15 Phylogenetic tree of the phthalate dioxygenase reductase and related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.

ZP_01568741 *Burkholderia multivorans* ATCC 17616

69

AAD03552 OphD (C-terminal) *Burkholderia cepacia* DBO1

100

AAD41517 OphD *Burkholderia cepacia* ATCC 17616

99

YP_001116712 *Burkholderia vietnamiensis* G4

ZP_02008299 *Ralstonia pickettii* 12D

OphD YZW-B

87

100 OphD YZW-F

100

YP_001116750 *Burkholderia vietnamiensis* G4

100 ZP_01568768 *Burkholderia multivorans* ATCC 17616

OphD YZW-G

100 Q05181 Pht1 *Pseudomonas putida*

YP_530958 *Rhodopseudomonas palustris* BisB18

ZP_01556272 *Burkholderia ambifaria* MC40-6

0.1

Figure 3.16 Phylogenetic tree of the phthalate permease and related proteins. The

Genbank accession number and the bacterial strain are shown on the dendrogram.
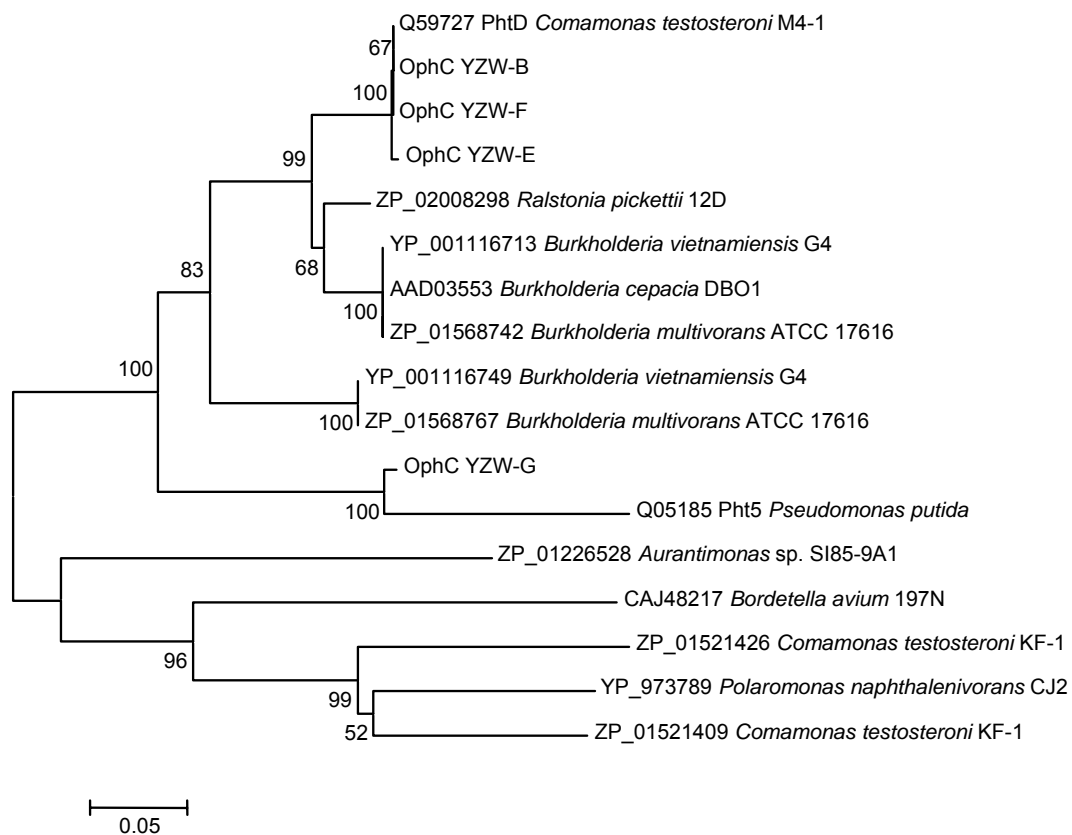
Figure 3.17 Phylogenetic tree of the dihydroxyphthalate decarboxylase and related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.
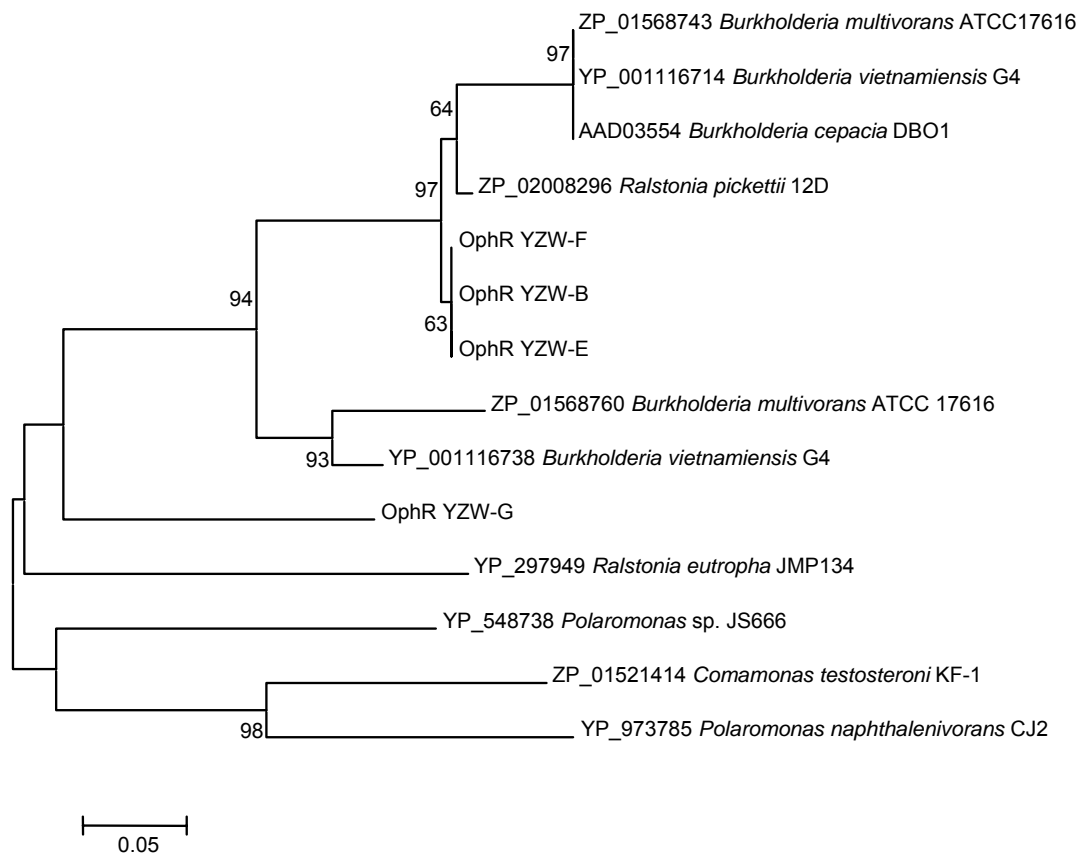
Figure 3.18 Phylogenetic tree of the phthalate regulator and related proteins. The

Genbank accession number and the bacterial strain are shown on the dendrogram.
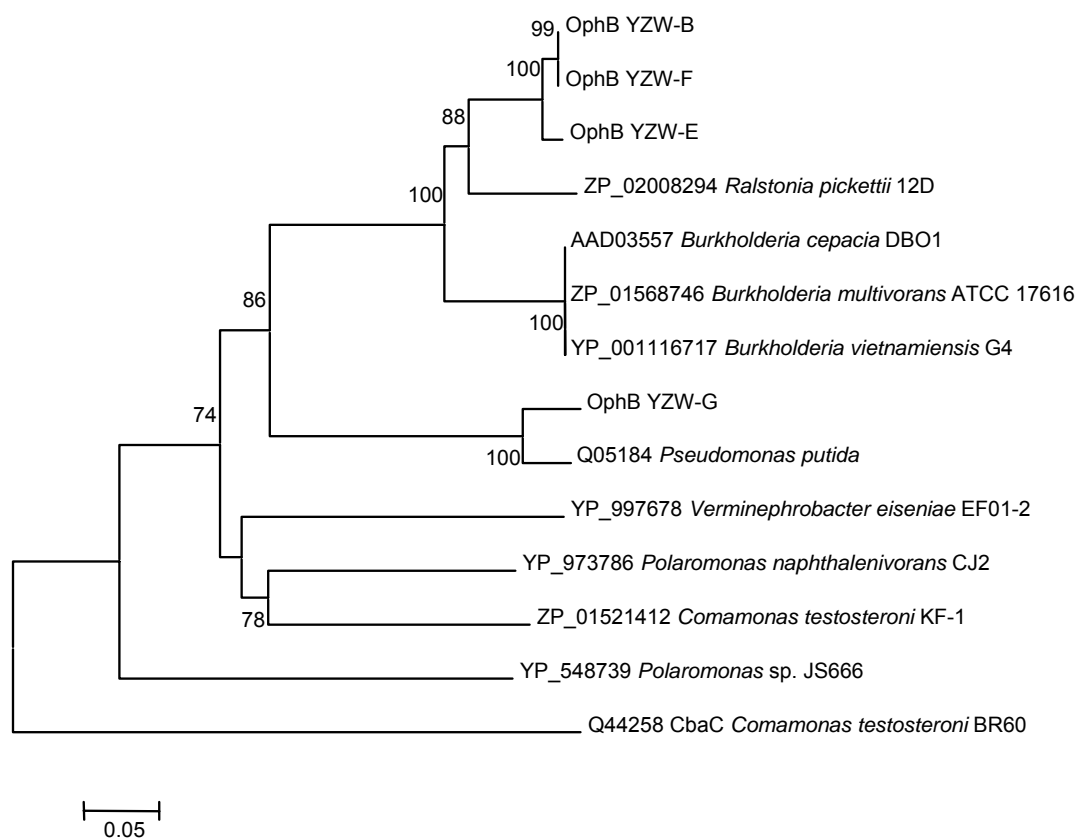
Figure 3.19 Phylogenetic tree of the phthalate dihydrodiol dehydrogenase and related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.
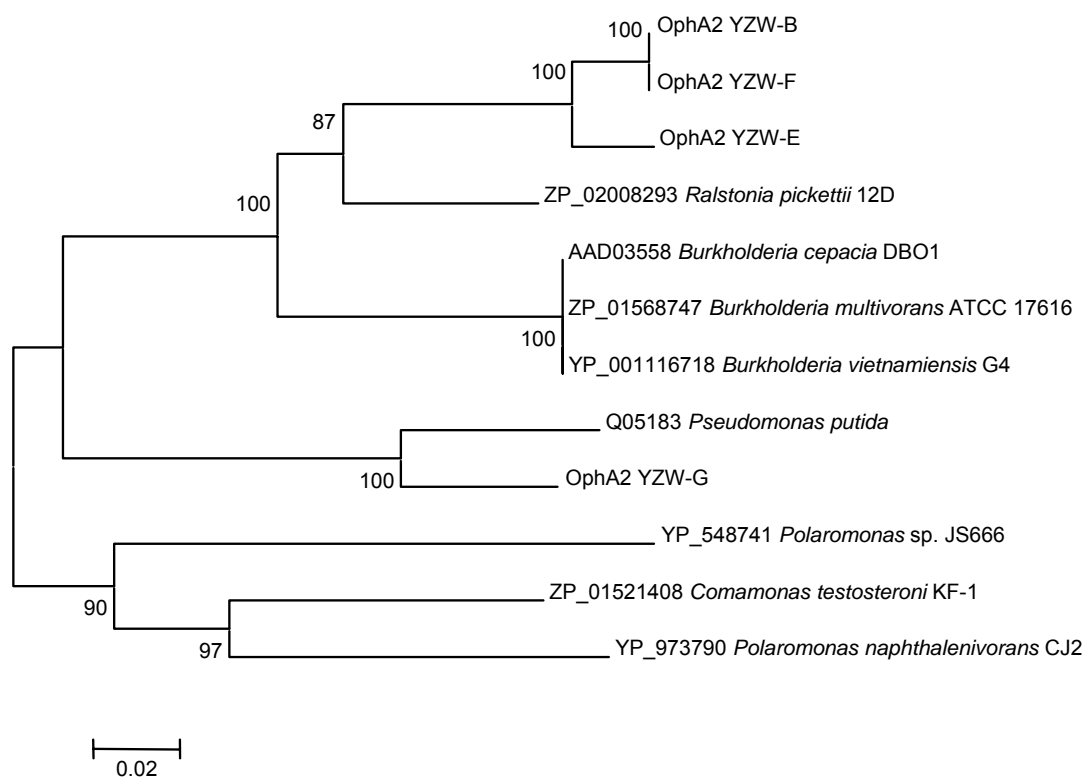
Figure 3.20 Phylogenetic tree of the phthalate dioxygenase and related proteins. The

Genbank accession number and the bacterial strain are shown on the dendrogram.
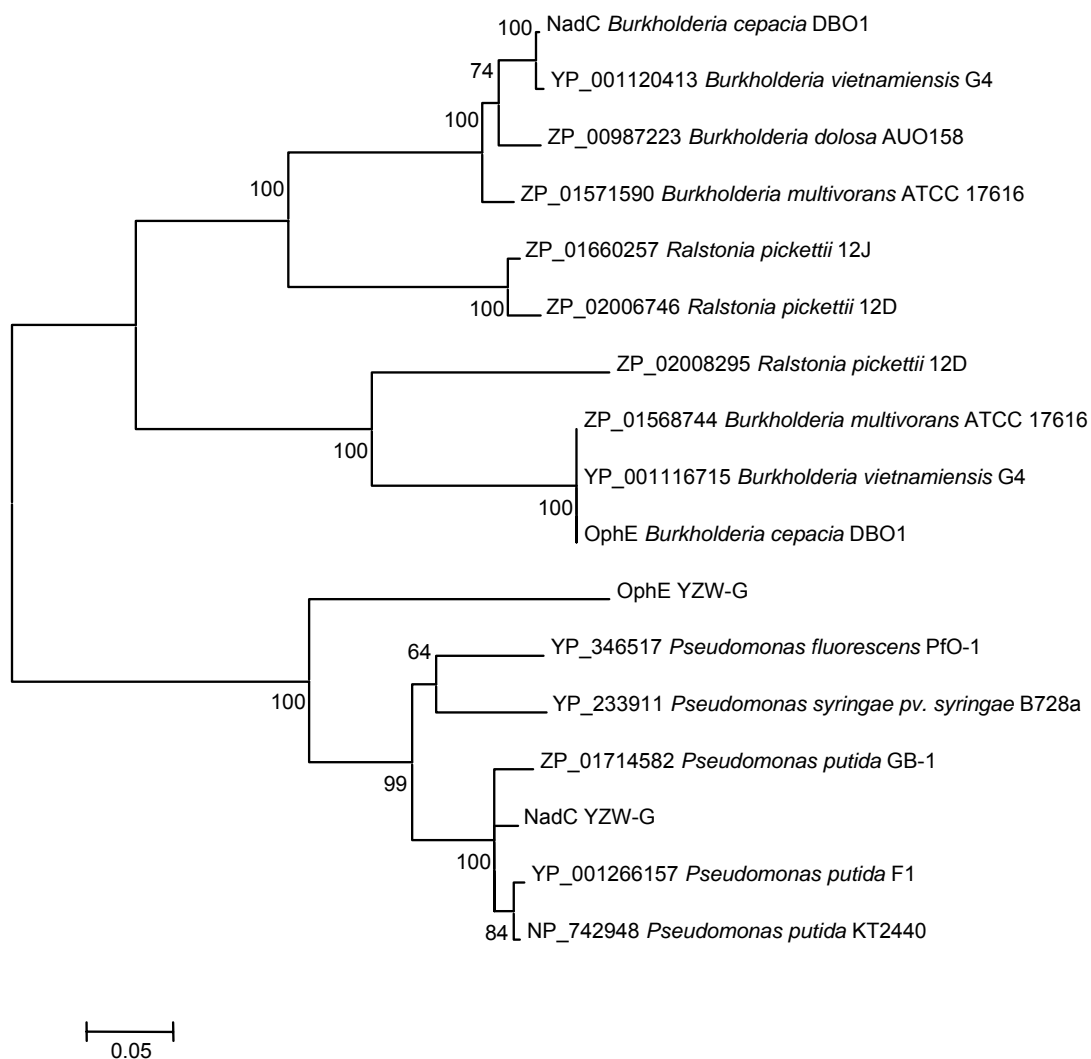
Figure 3.21 Phylogenetic tree of the quinolinate phosphoribosyl transferases (NadC and OphE) and their related proteins. The Genbank accession number and the bacterial strain are shown on the dendrogram.

3.3.4 Flanking regions of *oph* gene clusters

Besides the *oph* gene cluster, the left flanking region of the *oph* cluster in YZW-B (Table 3.7) and YZW-F (Table 3.5) are highly conserved with 99% DNA identity. This flanking region contains two transposase genes encoding IS481 and IS5 family transposase, respectively. However, these two transposase genes don't exist in the corresponding position in the YZW-E sequence. Instead, another IS5 family transposase gene is present in YZW-E, while it is not contained in YZW-B and YZW-F. An ORF encoding an uncharacterized protein located downstream of the transposase genes is also conserved in YZW-B, -E, and -F sequence.

The downstream flanking region of the *oph* gene clusters in YZW-B contains three consecutive transposase genes encoding IS3, IS110, and IS4 family transposases, respectively (Talbe 3.7). The first and second *tnp* genes in YZW-B are 100% identical to two transposase genes associated with terephthalate gene operon in YZW-F and *Comamonas* sp. strain E6 (BAE47081 and BAE47082) (102). The first *tnp* gene is highly conserved and shares 95% protein identity with transposases in *R. metallidurans* CH34 (YP_585133), *C. testosteroni* KF-1 (ZP_01518983), and *D. acidovorans* SPH-1 (ZP_01581429). No transposase genes were found in the above bacterial strains that shares high similarity with the second transposase. In addition, the third *tnp* gene shows 100% identity to transposase genes in other bacterial species, such as *R. metallidurans* CH34 (YP_585021 and YP_586290) and *B. vietnamiensis* G4

(YP_001115308, YP_001115958, YP_001120252, YP_001120487, and YP_001116699). However, none of these *tnp* genes is associated with phthalate or other degradation pathway.

Unlike YZW-F (Table 3.5) in which the *tph* gene operon and *oph* gene cluster are linked by two transposase genes, the four ORFs further downstream of the *oph* gene cluster and the *tnp* genes in YZW-B (Table 3.7) encode putative L-carnitine dehydratase/bile acid-inducible protein F, naphthoate synthase, acyl-CoA dehydrogenase, and acyl-CoA transferase/carnitine dehydratase. The functions of these ORFs still remain unknown, but it doesn't seem likely that they are associated with phthalate degradation.

In YZW-E, no transposase gene linked with the *oph* gene cluster was identified (Table 3.6). Instead, several novel genes were found in the flanking region of the *oph* gene cluster in YZW-E. These genes encode a Gram negative type porin, hypothetical protein, putative oxygenase, aldolase, regulatory protein, dehydrogenase, reductase, decarboxylase, and transporter, respectively. The porin protein exhibits 90% identity to a Gram negative type porin in *C. testosteroni* KF-1 (ZP_01521416). The oxygenase shares 47% identity with Rieske (2Fe-2S) domain protein *Alkalilimnicola ehrlichei* MLHE-1 (YP_743189), 44% identity with phthalate dioxygenase-like protein in *Sphingomonas wittichii* RW1 (YP_001263550), 40% identity with LigX in *Sphingomonas paucimobilis* SYK-6 (BAA36168), which is involved in oxygenative O-demethylation of 5,5'-

dehydrodivanillic acid (DDVA) as a terminal enzyme (109). The dehydrogenase shares 42% identity with putative phthalate dihydrodiol dehydrogenase in *Sphingomonas* sp. strain KA1 (YP_718168) (108). KA1 contains the plasmid pCAR3 encoding the complete mineralization of carbazole. Several ORFs in pCAR3 share low identity (40-50% identity) with phthalate degradation genes in *B. cepacia* DBO1 (108). The downstream reductase shares 52% identity with the ferredoxin in *C. testosteroni* KF-1 (ZP_01521411), and 49% identity with phthalate dioxygenase reductase in *Burkholderia* strains such as *B. cepacia* DBO1 (AAD03550), *B. vietnamiensis* G4 (YP_001116711), and *B. multivorans* ATCC 17616 (ZP_01568740). The decarboxylase shares 76% identity with 4,5-dihydroxyphthalate decarboxylase in *Polaromonas naphthalenivorans* CJ2 (YP_973789). In addition, there are two decarboxylase genes in *C. testosteroni* KF-1 (ZP_01521409 and ZP_01521426) sharing 74% and 71% identity with this decarboxylase gene in YZW-E. Based on the above comparative analysis, we could predict that the degradation operon could be involved in phthalate related but more complicated aromatic compound degradation. In addition, three downstream transporter genes in YZW-E are highly conserved in *C. testosteroni* KF-1 and exhibit 86% 96%, and 99% identical to major facilitator superfamily protein (ZP_01521358) efflux transporter (ZP_01521359), and hydrophobe/amphiphile efflux-1 family transporter (ZP_01521360).

In *Pseudomonas* sp. strain YZW-G, the gene organization upstream of the *oph* gene cluster is highly conserved in *P. putida* F1 and *P. putida* W619 (Figure 3.22;

Table 3.9). The F1 strain contains several ORFs encoding hypothetical proteins (YP_001268225-YP_001268230) in this region. However, ORFs downstream of *ophE* encode a putative permease, monooxygenase, regulatory protein, dehydrogenase, etc, which do not exist in the F1 strain genome. Unlike the *oph* operon and its downstream flanking region in YZW-G, a different degradation genetic island was found in this corresponding region of the F1 strain genome which contains two degradative gene clusters, *p*-cumate (*cmtAaAbAcCBAdDIEFHG*) (U24215) (29) and toluene degradation operon (*todRXFC1C2BADEGIH*) (U18304, J04996, and U09250) (65, 78, 122, 127). Thus, we consider this region a hot spot where foreign DNA could relatively easily insert and integrate into common genome backbone so that different strains could have gained different metabolic capacities during evolution.
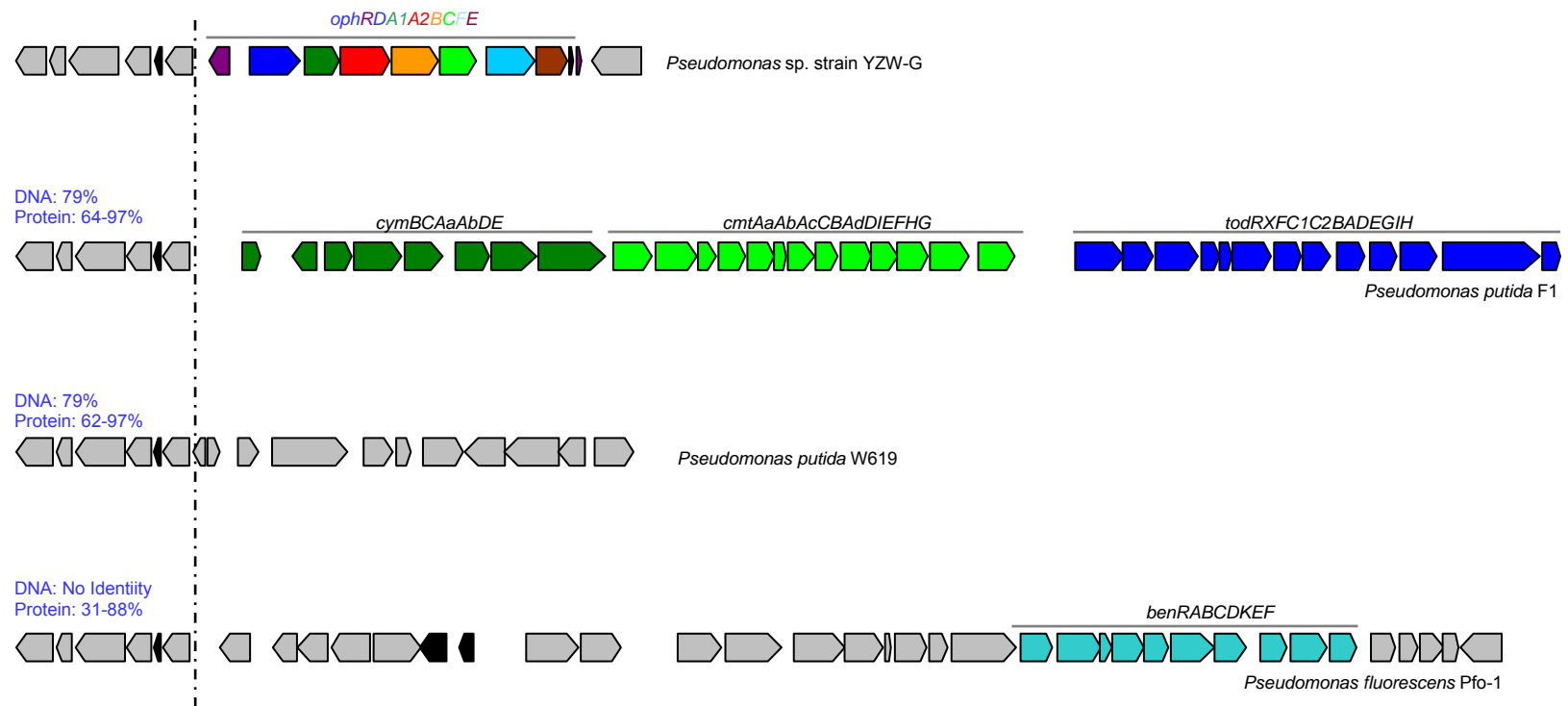
Figure 3.22 Homologous comparison of flanking regions of certain degradative genes in *Pseudomonas* strains. The sequences of YZW-G were obtained in this study, and others were retrieved from the Genbank database. In this figure, the ORFs to the left of the dashed line are conserved regions in different *Pseudomonas* strains (protein ID from left to right: YP_001268230- YP_001268225 in *P. putida* F1 (NC_009512); ZP_01639294-ZP_01639289 in *P. putida* W619 (NZ_AAVY01000010); YP_348726- YP_348721 in *P. fluorescens* Pfo-1 (NC_007492)). DNA and protein identities above the gene map indicate the identity of the conserved region in each reference strain compared with YZW-G. Degradative genes are located to the right of the dashed line. *ophRDA1A2BCFE* encode the phthalate degradation pathway in YZW-G. *cymBCAaAbDE*, *cmtAaAbAcCBAdDIEFHG*, *todRXFC1C2BADEGIH* encode *p*-cymene (30) (U24215), *p*-cumate (29) (U24215), and toluene degradation (U18304, J04996, U09250) (65, 78, 122, 127) in *P. putida* F1 (NC_009512), respectively. Putative benzoate degradative genes (*benRABCDKEF*) in *P. fluorescens* Pfo-1 (NC_007492) were annotated according to homology with *ben* genes in *P. putida* PRS2000 (AAF63447- AAF63455 in AF218267) (27). Hypothetical proteins and transposase genes are shown in gray and black, respectively.

Table 3.7 BlastX of YZW-B *oph* genes and flanking regions

Gene	Region	Length	Direction	Function	Source	Identity	Similarity	Gap	Accession No.	Reference
orf	1-308	308	←	porin, Gram-negative type	*Comamonas testosteroni* KF-1	84/91 (92%)	87/91 (95%)	0/91 (0%)	ZP_01521416	
orf	744-1718	975	←	uncharacterized protein	*Polaromonas naphthalenivorans* CJ2	173/327 (52%)	225/327 (68%)	3/327 (0%)	YP_973788	
tnp	2061-2615	555	←	hypothetical protein	*Sinorhizobium medicae* WSM419	92/186 (49%)	125/186 (67%)	5/186 (2%)	YP_001314563	
tnp	2858-3460	603	←	putative integrase protein	*Rhizobium etli* CFN 42	124/151 (82%)	132/151 (87%)	0/151 (0%)	NP_659803	(40)
ophH	3923-4417	495	←	hemerythrin-like metal-binding protein	*Ralstonia pickettii* 12D	126/164 (76%)	141/164 (85%)	0/164 (0%)	ZP_02008301	
ophA1	4483-5445	963	←	ferredoxin	*Ralstonia pickettii* 12D	277/320 (86%)	293/320 (91%)	0/320 (0%)	ZP_02008300	
ophD	5723-7066	1344	→	d-galactonate transporter	*Ralstonia pickettii* 12D	380/441 (86%)	404/441 (91%)	0/441 (0%)	ZP_02008299	
ophC	7110-8102	993	→	4,5-dihydroxyphthalate decarboxylase	*Comamonas testosteroni* M4-1	328/330 (99%)	329/330 (99%)	0/330 (0%)	Q59727	(66)
					Ralstonia pickettii 12D	307/330 (93%)	318/330 (96%)	0/330 (0%)	ZP_02008298	
ophR	8393-8986	594	→	transcriptional regulator, MarR family	*Ralstonia pickettii* 12D	168/193 (87%)	176/193 (91%)	1/193 (0%)	ZP_02008296	
* Gap										
ophB	21-1196	1176	←	oxidoreductase domain protein	*Ralstonia pickettii* 12D	345/391 (88%)	361/391 (92%)	0/391 (0%)	ZP_02008294	
ophA2	1216-2547	1332	←	Rieske (2Fe-2S) domain protein	*Ralstonia pickettii* 12D	391/443 (88%)	413/443 (93%)	0/443 (0%)	ZP_02008293	
tnp	3491-5029	1539	←	hypothetical transposase	*Comamonas* sp. E6	419/419 (100%)	419/419 (100%)	0/419 (0%)	BAE47081	
tnp	5328-6404	1077	←	hypothetical transposase	*Comamonas* sp. E6	358/358 (100%)	358/358 (100%)	0/358 (0%)	BAE47082	
tnp	6871-8073	1203	←	transposase, IS4	*Ralstonia metallidurans* CH34	400/400 (100%)	400/400 (100%)	0/400 (0%)	YP_585021	
orf	8107-9153	1047	→	L-carnitine dehydratase /bile acid-inducible protein F	*Ralstonia metallidurans* CH34	124/333 (37%)	181/333 (54%)	10/333 (3%)	YP_587753	
orf	9183-9941	759	→	naphthoate synthase	marine actinobacterium PHSC20C1	137/248 (55%)	174/248 (70%)	2/248 (0%)	ZP_01131603	
orf	10019-11176	1158	→	acyl-CoA dehydrogenase-like	*Delftia acidovorans* SPH-1	309/384 (80%)	340/384 (88%)	0/384 (0%)	ZP_01579884	
orf	11206-12363	1158	→	acyl-CoA transferase /carnitine dehydratase	*Ralstonia eutropha* H16	250/390 (64%)	304/390 (77%)	9/390 (2%)	YP_726883	(94)

* Gap is about 100 bp unsequenced regaion between the *ophR* and *ophB* gene.

Table 3.8 BlastX of YZW-E *oph* genes and flanking regions

Gene	Region	Length	Direction	Function	Source	Identity	Similarity	Gap	Accession No.	Reference
tnp	61-378	318	←	Integrase, catalytic region	*Comamonas testosteroni* KF-1	79/83 (95%)	80/83 (96%)	0/83 (0%)	ZP_01521471	
tnp	458-739	282	→	transposase IS3/IS911	*Burkholderia cenocepacia* MC0-3	43/71 (60%)	53/71 (74%)	1/71 (1%)	ZP_01566900	
tnp	900-1637	738	→	transposase, IS4 family	*Acidovorax* sp. JS42	196/237 (82%)	209/237 (88%)	5/237 (2%)	YP_986017	
orf	1618-2655	1038	←	uncharacterized protein	*Polaromonas naphthalenivorans* CJ2	176/327 (53%)	226/327 (69%)	3/327 (0%)	YP_973788	
ophH	3002-3496	495	←	hemerythrin-like metal-binding protein	*Ralstonia pickettii* 12D	127/164 (77%)	142/164 (86%)	0/164 (0%)	ZP_02008301	
ophA1	3600-4562	963	←	ferredoxin	*Ralstonia pickettii* 12D	274/320 (85%)	290/320 (90%)	0/320 (0%)	ZP_02008300	
ophD*	4842-5098	257	→	d-galactonate transporter (truncated)	*Ralstonia pickettii* 12D	66/85 (77%)	70/85 (82%)	0/85 (0%)	ZP_02008299	
	5099-5129	31								
ophC	5173-6165	993	→	4,5-dihydroxyphthalate decarboxylase	*Comamonas testosteroni* M4-1	327/330 (99%)	328/330 (99%)	0/330 (0%)	Q59727	(66)
					Ralstonia pickettii 12D	306/330 (92%)	317/330 (96%)	0/330 (0%)	ZP_02008298	
ophR *Gap	6456-7049	594	→	transcriptional regulator, MarR family	*Ralstonia pickettii* 12D	166/193 (86%)	174/193 (90%)	1/193 (0%)	ZP_02008296	
ophB	21-1196	1176	←	oxidoreductase domain protein	*Ralstonia pickettii* 12D	343/391 (87%)	359/391 (91%)	0/391 (0%)	ZP_02008294	
ophA2	1216-2547	1332	←	Rieske (2Fe-2S) domain protein	*Ralstonia pickettii* 12D	394/443 (88%)	416/443 (93%)	0/443 (0%)	ZP_02008293	
orf	3665-4861	1197	→	porin, Gram-negative type	*Comamonas testosteroni* KF-1	338/374 (90%)	352/374 (94%)	0/374 (0%)	ZP_01521416	
orf	4914-5969	1059	←	hypothetical protein	*Bordetella pertussis* Tohama I	122/322 (37%)	184/322 (57%)	10/322 (3%)	NP_881739	(89)
orf	6390-7667	1278	→	Rieske (2Fe-2S) domain protein	*Alkalilimnicola ehrlichei* MLHE-1	199/415 (47%)	258/415 (62%)	11/415 (2%)	YP_743189	
orf	7664-8362	699	→	Ribulose-5-phosphate 4-epimerase and related epimerases and aldolases	*Nostoc punctiforme* PCC73102	72/191 (37%)	109/191 (57%)	1/191 (0%)	ZP_00110197	
orf	8413-9000	588	→	transcriptional regulator, MarR family	*Polaromonas* sp. JS666	76/147 (51%)	111/147 (75%)	0/147 (0%)	YP_548738	
orf	8997-10178	1182	→	putative phthalate dihydrodiol dehydrogenase	*Sphingomonas* sp. KA1	158/372 (42%)	210/372 (56%)	6/372 (1%)	YP_718168	(108)
orf	10175-11125	951	→	phthalate dioxygenase reductase	*Burkholderia cepacia* DBO1	154/312 (49%)	204/312 (65%)	2/312 (0%)	AAD03550	(15)
orf	11138-12127	990	→	4,5-dihydroxyphthalate decarboxylase	*Polaromonas naphthalenivorans* CJ2	254/329 (77%)	287/329 (87%)	0/329 (0%)	YP_973789	
orf	12136-13677	1542	→	major facilitator superfamily MFS_1	*Comamonas testosteroni* KF-1	324/374 (86%)	338/374 (90%)	13/374 (3%)	ZP_01521358	
orf	13930-15264	1335	→	efflux transporter, RND family, MFP subunit	*Comamonas testosteroni* KF-1	429/444 (96%)	434/444 (97%)	0/444 (0%)	ZP_01521359	
orf	15283-16182	900	→	transporter, hydrophobe/amphiphile efflux-1 (HAE1) family	*Comamonas testosteroni* KF-1	297/300 (99%)	299/300 (99%)	0/300 (0%)	ZP_01521360	

ophD* : truncated phthalate transporter gene, the first region (257bp, 4842-5098) is homologous to YZW-B *ophD* gene (1-257) encoding for N-terminal of transporter protein and the second region (31 bp, 5099-5129) is homologous to YZW-B *ophD* gene (1314-1344) encoding for C-terminal of transporter protein.

* Gap is about 100 bp unsequenced regaion between the *ophR* and *ophB* gene.

Table 3.9 BlastX of YZW-G *oph* genes and flanking region (25,863 bp)

Gene	Region	Length	Direction	Function	Source	Identity	Similarity	Gap	Accession No.	Reference
orf	1-838	838	←	hypothetical protein	*Pseudomonas putida* F1	257/304 (84%)	268/304 (88%)	0/304 (0%)	YP_001268230	
					Pseudomonas putida W619	250/304 (82%)	263/304 (86%)	0/304 (0%)	ZP_01639294	
orf	906-1457	552	←	hypothetical protein	*Pseudomonas putida* W619	179/183 (97%)	180/183 (98%)	0/183 (0%)	ZP_01639293	
					Pseudomonas putida F1	178/183 (97%)	179/183 (97%)	0/183 (0%)	YP_001268229	
orf	1473-2984	1512	←	hypothetical protein	*Pseudomonas putida* F1	371/514 (72%)	415/514 (80%)	14/514 (2%)	YP_001268228	
					Pseudomonas putida W619	368/505 (72%)	416/505 (82%)	9/505 (1%)	ZP_01639292	
orf	3072-3875	804	←	Cobyrinic acid a,c-diamide synthase	*Pseudomonas putida* W619	231/249 (92%)	235/249 (94%)	0/249 (0%)	ZP_01639291	
					Pseudomonas putida F1	229/259 (88%)	233/259 (89%)	2/259 (0%)	YP_001268227	
orf	3880-4095	216	←	phage transcriptional regulator, AlpA	*Pseudomonas putida* W619	63/71 (88%)	65/71 (91%)	0/71 (0%)	ZP_01639290	
					Pseudomonas putida F1	61/68 (89%)	62/68 (91%)	0/68 (0%)	YP_001268226	
orf	4209-4946	738	←	hypothetical protein	*Pseudomonas putida* F1	157/245 (64%)	195/245 (79%)	0/245 (0%)	YP_001268225	
					Pseudomonas putida W619	154/245 (62%)	193/245 (78%)	0/245 (0%)	ZP_01639289	
ophR	5353-5961	609	←	transcriptional regulator, MarR family	*Ralstonia pickettii* 12D	120/191 (62%)	148/191 (77%)	1/191 (0%)	ZP_02008296	
ophD	6502-7857	1356	→	Phthalate transporter	*Pseudomonas putida*, plasmid pNMH102-2	388/451 (86%)	407/451 (90%)	0/451 (0%)	BAA02509	(86)
ophA1	7967-8944	978	→	Phthalate 4,5-dioxygenase reductase subunit	*Pseudomonas putida* plasmid pNMH102-2	293/324 (90%)	307/324 (94%)	0/324 (0%)	BAA02510	(86)
ophA2	8993-10312	1320	→	Phthalate 4,5-dioxygenase oxygenase subunit	*Pseudomonas putida* plasmid pNMH102-2	404/438 (92%)	420/438 (95%)	0/438 (0%)	BAA02511	(86)
ophB	10309-11538	1230	→	Putative 4,5,-dihydroxyphthalate dehydrogenase	*Pseudomonas putida* plasmid pNMH102-2	371/409 (90%)	387/409 (94%)	0/409 (0%)	BAA02512	(86)
ophC	11535-12527	993	→	4,5-dihydroxyphthalate decarboxylase	*Pseudomonas putida* plasmid pNMH102-2	213/219 (97%),	216/219 (98%)	1/219 (0%)	BAA02513	(86)
ophP	12845-14131	1287	→	Putative porin	*Pseudomonas putida* KT2440	244/418 (58%)	308/418 (73%)	1/418 (0%)	NP_745530	(82)
ophE	14238-15086	849	→	nicotinate-nucleotide pyrophosphorylase	*Pseudomonas fluorescens* Pf-5	211/282 (74%)	252/282 (89%)	0/282 (0%)	YP_257983	(91)
orf	15717-17087	1371	←	major facilitator superfamily MFS_1	*Rubrobacter xylanophilus* DSM 9941	161/410 (39%)	235/410 (57%)	10/410 (2%)	YP_645215	
orf	17310-18845	1536	→	flavoprotein involved in K+ transport-like protein	*Sphingomonas wittichii* RW1	225/491 (45%)	295/491 (60%)	18/491 (3%)	YP_001261397	
orf	19996-20949	954	→	transcriptional regulator, LysR family	*Burkholderia* sp. 383	161/410 (39%)	235/410 (57%)	10/410 (2%)	YP_373250	
orf	21579-23180	1602	→	Aldehyde dehydrogenase	*Roseovarius* sp. HTCC2601	399/489 (81%)	435/489 (88%)	2/489 (0%)	ZP_01441917	

3.4 Gene Regulation for Phthalate, Isophthalate, Terephthalate Degradation in *C. testosteroni* YZW-B

3.4.1 Operonic structure of the *oph*, *iph* and *tph* genes

The operonic structure of *oph*, *iph*, and *tph* genes were revealed by RT-PCR using RNA extracted from YZW-B cells grown on phthalate, isophthalate, or terephthalate, respectively. *oph* genes are transcribed in three transcripts *ophHA1*, *ophDCR*, and *ophBA2* with opposite transcription directions, which are consistent with the ORF prediction (Figure 3.23). The *iph* genes are transcribed in single transcript *iphA2CBA1R* since all of the *iph* genes have the same direction for transcription. The *tph* genes are organized as a typical operon structure in which two transcripts *tphR* and *tphCA2A3BA1* are transcribed in opposite directions (Figure 3.24).

3.4.2 *oph*, *iph*, or *tph* genes are inducible by phthalate isomers

In order to determine whether the *oph*, *iph*, and *tph* genes are inducible by phthalate isomers, RT-PCR was performed to amplify *ophA2*, *iphA2*, and *tphA2* gene from RNA templates extracted from YZW-B cells grown on phthalate, isophthalate, or terephthalate. RT-PCR shows that the *ophA2*, *iphA2* and *tphA2* genes are induced by phthalate, isophthalate and terephthalate, respectively (Figure 3.25).
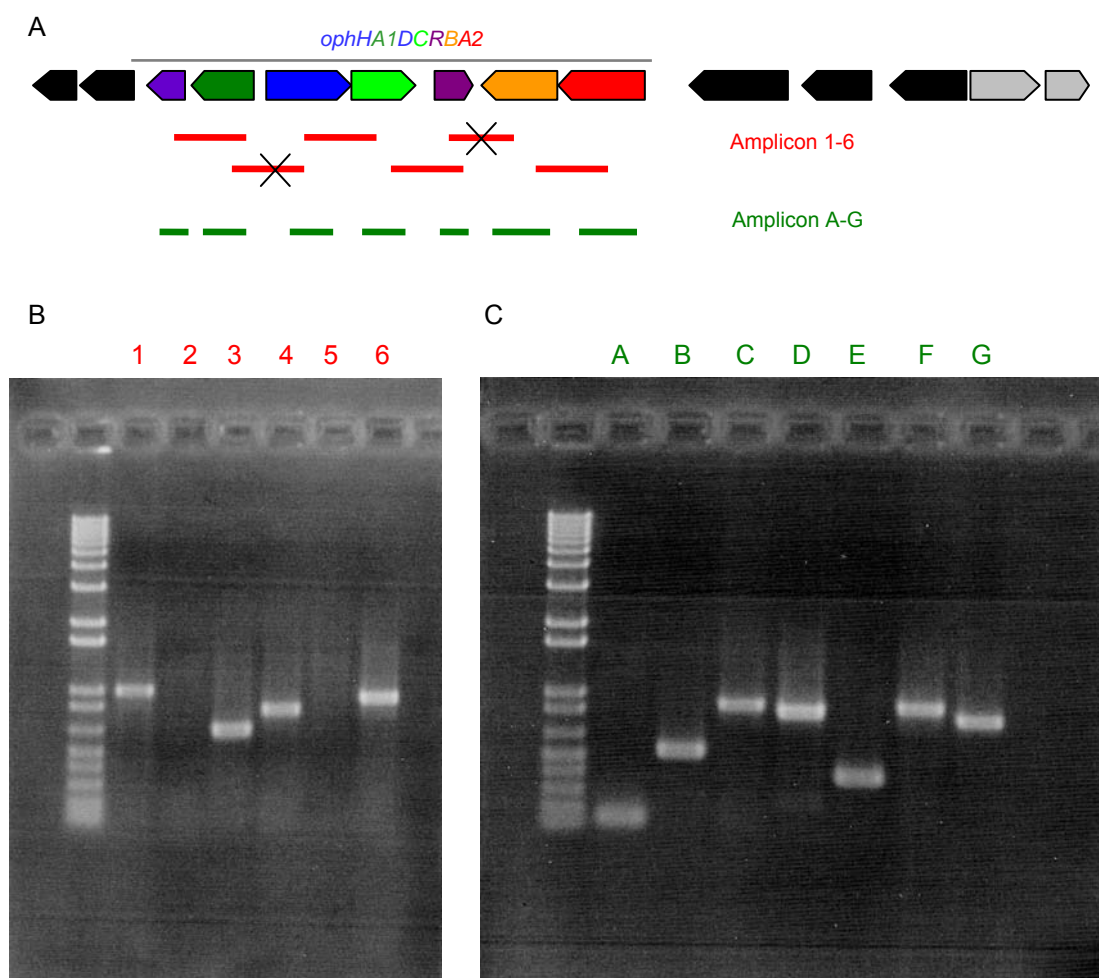
Figure 3.23 A. Location of RT-PCR products according to the *oph* gene organization. Amplicons 1-6 are RT-PCR products containing intergenic regions between two adjacent genes. Amplicons A-G are RT-PCR products from internal regions in each individual gene.

Figure 3.23 B. RT-PCR products containing intergenic regions between two adjacent *oph* genes. Amplicons 1, 3, 4 and 6 are positive and amplicons 2 and 5 are negative, which indicate that *ophHA1*, *ophDCR*, and *ophBA2* are transcribed in different directions.

Figure 3.23 C. RT-PCR products from each *oph* gene. Amplicon A-G are positive, which means that each *oph* gene is expressed in the presence of phthalate.
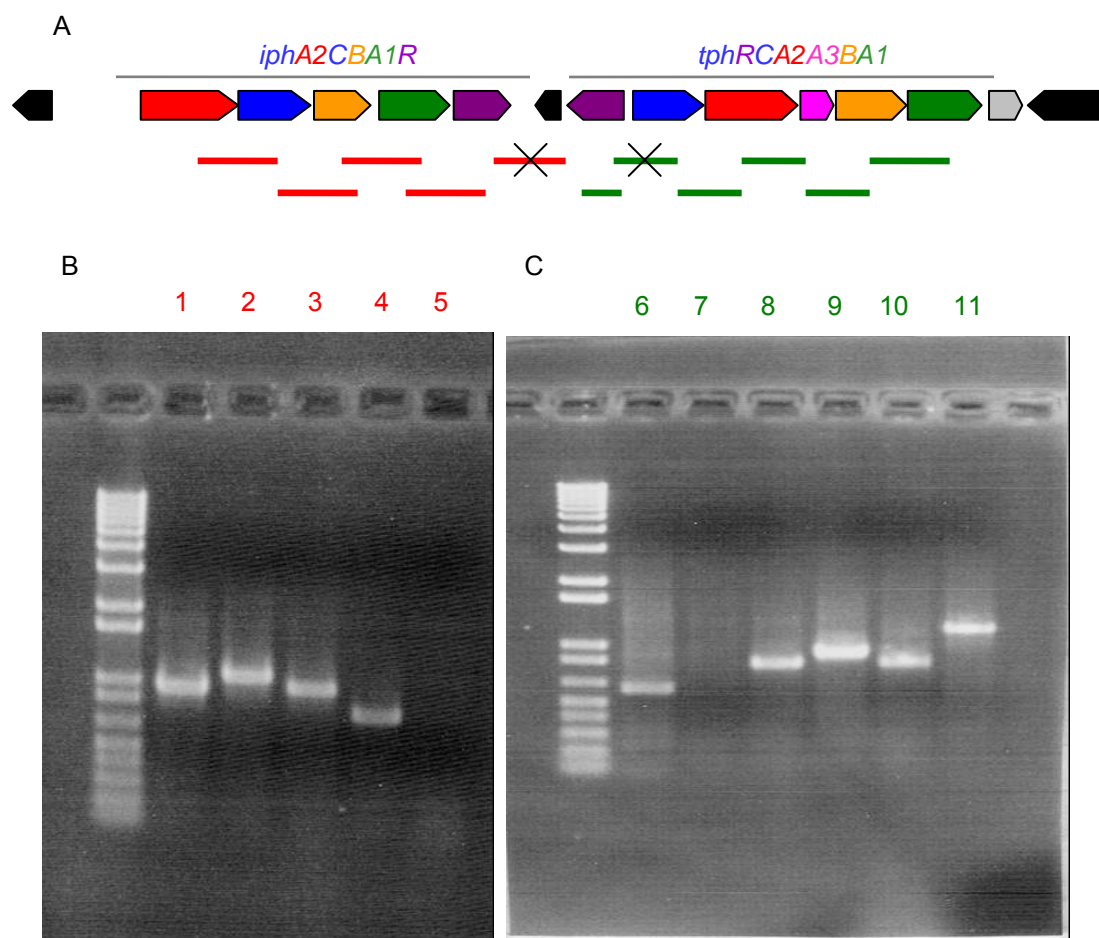
Figure 3.24 A. Location of RT-PCR products according to the *iph* and *tph* gene organization. Amplicons 1-5 (red) and 6-11 (green) are RT-PCR products containing intergenic region between two adjacent *iph* and *tph* genes.

Figure 3.24 B. RT-PCR products containing intergenic regions between two adjacent *iph* genes. Amplicons 1-4 are positive and amplicon 5 is negative, which indicate that *iphA2CBA1R* are transcribed in the same direction.

Figure 3.24 C. RT-PCR products containing intergenic regions between two adjacent *tph* genes. Amplicons 6 and 8-11 are positive and amplicon 7 is negative, which indicates that *tphR* and *tphCA2A3BA1* are transcribed in opposite directions.
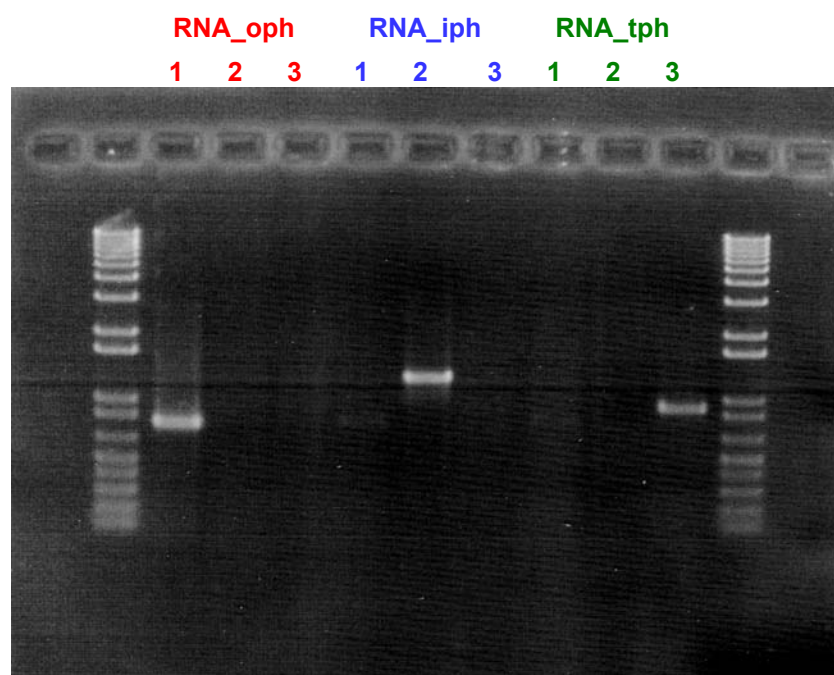
Figure 3.25 RT-PCR showing inducibility fo the *oph*, *iph*, and *tph* genes by phthalate isomers. RT-PCR was performed to the amplify *ophA2*, *iphA2*, and *tphA2* genes from RNA templates extracted from YZW-B cells grown on phthalate (red 1,2,3), isophthalate (blue 1,2,3), or terephthalate (green 1,2,3). Primers for RT-PCR: 1. *ophA2* gene primers; 2. *iphA2* gene primers; 3, *tphA2* gene primers.

3.4.3 5'-RACE (Rapid Amplification of cDNA ends)

In order to define the initiation site of each transcript in the *oph*, *iph* and *tph* gene cluster/operon, the rapid amplification of cDNA ends (5'-RACE) method was performed using RNA templates isolated from YZW-B grown on the corresponding substrate phthalate, isophthalate, or terephthalate, respectively (Figure 3.26). For the *iph* gene cluster, the start site of the transcript is located at 52 or 53 bp (A or T) upstream of the predicted translation initiation site of *iphA2*. Comparing the size of the peak representing G and T immediate upstream of A(52), we could conclude that most of the transcripts from *iphA2* start from A(52) and others start from T(53).

For the *tph* gene operon, since *tphR* and *tphCA2A3BA1* are transcribed in opposite directions, two 5'-RACE assays were performed separately for each transcript.   The start site of the *tphCA2A3BA1* transcript is located at 22 bp upstream of the translation initiation codon in *tphCA2A3BA1*. For *tphR*, it wasn't successful to define the exact transcription start site of *tphR* since the readings from the sequencer gave mixed peaks in the region around 60 or 61(C or G) bp upstream of the *tphR* gene. Since the 3'-end of purified cDNAs have been tailed with dCTP by terminal deoxynucleotidyl transferase (TdT), we couldn't determine whether C(61) in the 5'-RACE assay originally came from the mRNA template or was C-tailed by TdT.   In addition, there is a palindrome sequence (TTTTTGCGCTATGCGCAAAAA) located 44 bp-66 bp upstream of the *tphR*

gene, which probably allows hairpin formation at the end of *tphR* transcript to interrupt cDNA synthesis, polyC tailing, or final sequencing step.

Similarly, three 5'-RACE assays were performed for the three transcripts produced from the *oph* gene clusters. For most transcripts of *ophA1*, the start sites are T, which is located 29 bp upstream of the *ophA1* translation initiation site, while others are A(28) downstream next to T(29), as shown in sequence profile. Similarly, T(130) or A(129) were determined as the transcription start sites of *ophDCR*, and A(40) was determined as the start site for *ophA2B*.

Since the start sites for each transcript from *iph*, *tph*, and *oph* gene operon/clusters have been revealed by 5'-RACE, it allows us to align the promoter sequences of each transcript to find out the consensus promoter region such as the -10 and -35 region. Sequence alignment shows that there are three consensus regions located upstream of the *iphA2* and *tphC* transcripts, which are -15 region (5'-NATCCN-3'), -30 region (5'-TTTNACACA-3'), and -45 region (5'-CAAAACA-3'), respectively.   And two consensus regions, -10 region (-5'-TATNNAAT-3') and -35 region (5'-GCTTGTC-3'), were identified at the upstream region of *ophHA1*, *ophDCR*, and *ophBA2* transcripts. Moreover, the -10 region and the space between -10 and -35 region are AT rich. But we didn't find any of the above consensus regions in the upstream region of the *tphR* transcript.

ophHA1-ophDCR-ophBA2

iphA2CBA1R

tphR-tphCA2A3BA1

Initiation site of transcription

```
TGTCCACAGGGTGGACAAAAACAATAAATTTGACACAGATCAGATAAATCCATAAGATGGACAT  T(53)/A(52)    iphA2CBA1R
TTTGCGCATAGCG--CAAAAAC---AGGTTTAACACAAAGTACGACATCCCCATACTGCAGTT   C(22)          tphCA2A3BA1

GTTTCGCGCCAGGGGAAAACCCTGCTTGTCGCTACTCTCGAAGCCTCGTAATCTAAT---AAA   T(130)/A(129)  ophDCR
ACCGTGCGCAAGTCAACTACTCCGCTTG-CATAATTA--GATCTCTAATAATAGAATTGCAAA   T(29)/A(28)    ophA1H
GCCACACA-ATCTTGTCAAACAGGCTTGTCAATATAACTAGA-TCTC-TAATATAATTA-AGT   A(40)          ophA2B
```

Figure 3.26 A) Operonic structure of *oph*, *iph*, and *tph* genes revealed by RT-PCR. Arrows under the gene organization map indicate the direction of transcription. Dots at the end of each arrow indicate the approximate location of the transcription initiation site.
B) Alignment of upstream sequences of *oph*, *iph*, and *tph* gene transcription initiation sites in YZW-B. The black arrow indicates the initiation sites for each transcript. The numbers show the distance from the transcription initiation site to the ORF start site for each transcript. Red boxes indicate the consensus sequences in the putative promoter regions.

140

3.4.4 Quantitative real time PCR analysis of *ophA2*, *iphA2*, and *tphA2* gene expression in YZW-B

3.4.4.1 Development and validation of quantitative RT-PCR

In this study, four housekeeping genes (16S, 23S, *gyrB* and *rpoB*) were used as endogenous controls for relative quantification. The phthalate dioxygenase gene *ophA2*, isophthalate dioxygenase gene *iphA2*, and terephthalate dioxygenase gene *tphA2* were measured as target genes representative of the *oph*, *iph*, and *tph* gene expression, respectively. Since the sequences of the 16S, *ophA2*, *iphA2*, and *tphA2* genes are known, the exact lengths of the PCR products of these genes also will be known after amplification with gene specific primers. For those genes with unknown sequences such as 23S, *gyrB*, and *rpoB*, PCR products were amplified using universal primers or degenerate primers, then sequenced by primer walking, and finally re-PCR amplified using newly designed gene specific primers. Generally, 1 kb to 2 kb long PCR products of these genes were amplified from YZW-B genomic DNA. The qPCR primers were designed according to the amplicon sequences. To illustrate the detection limitation, primer efficiency, and amplification specificity of qPCR, absolute quantification followed by a dissociation curve assay were carried out using specific primer pairs and purified PCR products as templates. The linear relationship of the threshold cycle C_T (Y) and log10 of DNA copy number of each gene (X) is given

by the following formulas: For 16S, Y=-3.487X+36.675 (R^2=0.9983); for 23S, Y=-3.7X+37.472 (R^2=0.9994); for *gyrB*, Y=-3.2834X+37.189 (R^2=0.9991); for *rpoB*, Y=-3.792X+37.85 (R2=0.9957); for *ophA2*, Y=-3.838X+39.108 (R^2=0.9983); for *iphA2*, Y=-3.316X+35.146 (R^2=0.9954); for *tphA2*, Y=-3.3689X+34.926 (R^2=0.9997). Good linear relationship between C_T and log10 of input PCR product copy number was revealed by regression analysis (R^2 > 0.99). The following dissociation curve assays for each qPCR reaction clearly show that neither primer dimer nor nonspecific PCR product was detected in each primer pair and different template concentrations. Only one Tm peak displayed in dissociation curve for each PCR reaction and peaks in dissociation curves from same template in different dilutions had same Tm value, indicative of the specificity of qPCR reaction. Moreover, RT-qPCR system theoretically should be able to detect a single copy of the target genes under the current PCR conditions including the thermocycling setting and reaction component concentration. So the primer pairs and PCR conditions could be used for further qPCR experiments.

The $2^{-\Delta\Delta C_T}$ method was used for calculation of relative quantification, in which the amount of target, normalized to an endogenous reference and relative to a calibrator, is given by $2^{-\Delta\Delta C_T}$. For validation of application of the $2^{-\Delta\Delta C_T}$ method, the amplification efficiencies of both target and endogenous gene should be approximately equal. So we performed a series of qPCR using serially diluted RNA as template to amplify four housekeeping genes (16S, 23S, *gyrB*, and *rpoB*)

and the target gene *ophA2*. Linear relationship of threshold cycle C_T (Y) and log10 of input amount of RNA (X) is given by the following formulas: for 16S, Y=-3.242X+13.564 (R^2=0.9996); for 23S, Y=-3.188X+13.734 (R^2=0.9985); for *gyrB*, Y=-2.983X+26.683 (R^2=0.9986); for *rpoB*, Y=-3.385X+24.24 (R^2=0.9986); for *ophA2*, Y=-3.295X+21.28 (R^2=0.9994). Comparing to the standard curves generated by PCR products, the slopes from RNA as templates are slightly lower than for those from PCR products, which is probably because other RNA species in the total RNA somehow interferes with the efficiency and specificity of reverse transcription and the subsequent qPCR steps, or the introduction of RT into qPCR in one step affects the qPCR reaction. The above formulas (R^2>0.99) also show that there is a good relationship established between C_T and log10 of the input amount of RNA.

In addition, validation curves were generated by plotting ΔC_T ($C_{T,target}$ - $C_{T,endogenous}$) versus log10 of the input amount of RNA, which were shown by the following formulas: for *ophA2*-16S, Y=-0.102X+7.667; for *ophA2*-23S, Y=0.016X+7.669; for *ophA2-gyrB*, Y=-0.312X-5.403; for *ophA2-rpoB*, Y=0.09X-2.96. Theoretically, if the slopes from standard curves for target gene and endogenous gene are equal, the slope of the validation curve will be zero, which means the efficiency of qPCR for the target gene and the endogenous gene are equal and that the endogenous gene will be the perfect one for calculating the relative quantification of target gene expression. Practically, we consider the slopes of validation curves 0.1 a cutting off value, i.e. when the slope is smaller

than or around 0.1, the amplification efficiency of target gene and endogenous gene are thought to be approximately equal. Thus, 16S, 23S, and *rpoB* gene are qualified as endogenous control for RT-qPCR assay. Additionally, because the gene expression level of *ophA2* is relatively closer to *rpoB* expression level than 16S, 23S, or *gyrB* expression levels, *rpoB* was considered as the best endogenous control for relative quantification using the $2^{-\Delta\Delta C_T}$ method, even though similar gene expression profiles were seen no matter which housekeeping gene was used as an endogenous control.

3.4.4.2 Housekeeping gene expression in YZW-B

Relative quantification of housekeeping genes 16S, 23S, *gryB*, *rpoB*, and target genes *ophA2*, *iphA2* and *tphA2* expression was measured in *C. testosteroni* YZW-B grown on succinate, phthalate, isophthalate, terephthalate, and two or three phthalate isomers (phthalate and isophthalate; phthalate and terephthalate; isophthalate and terephthalate; phthalate, isophthalate and terephthalate) (Figure 3.27).

The relative quantification (RQ) for different housekeeping genes was calculated using one housekeeping gene as an endogenous control and RNA from succinate as calibrator (Figure 3.27). In order to determine the stability of housekeeping gene expression, the average of each housekeeping gene expression in YZW-B grown on different substrates was calculated. When 16S

was used as an endogenous control to calculate the relative quantification of the other three housekeeping genes, the 23S gene showed the most constant expression with an average RQ of 1.02025, which means that these two genes are always expressed at same level. Both *gyrB* (RQ 0.815625) and *rpoB* (RQ 0.67) are expressed at a relatively lower level than 16S and 23S. When the 23S gene was used as an endogenous control, the 16S was more constantly expressed (RQ 1.009), while both *gyrB* (RQ 0.824375) and *rpoB* (RQ 0.6665) were slightly downregulated. Similarly, when *gyrB* or *rpoB* was used as an endogenous control, *rpoB* or *gyrB* was expressed more constantly than 16S and 23S. These results show that 16S and 23S share the same gene expression pattern on the different substrates and so do *rpoB* and *gyrB*. It is probably due to the different features of rRNA (16S and 23S) and mRNA (*rpoB* and *gyrB*), which include that rRNAs (16S and 23S) are always maintained in much higher levels than mRNA, and rRNAs tend to form secondary structures. Both of these factors could affect the performance of the RT-qPCR.

3.4.4.3 ophA2, iphA2 and tphA2 gene expression on phthalate isomers

Similarly, relative quantification of *ophA2*, *iphA2*, and *tphA2* were calculated using each of the four housekeeping genes 16S, 23S, *gyrB* and *rpoB* as the endogenous control. The qPCR results show that the target gene expression patterns are similar no matter which housekeeping gene was used as the endogenous control to calculate relative quantification (Figure 3.28). For

instance, when bacteria grow on phthalate, the *ophA2*, *iphA2* and *tphA2* genes are upregulated by about 236, 1040, and 958 fold with 16S as the endogenous control; while the fold change for *ophA2*, *iphA2* and *tphA2* gene expression turned to be 332, 1597, and 1117 using *rpoB* as endogenous control. Therefore, in this section we discuss *ophA2*, *iphA2* and *tphA2* gene expression in detail according to RQ normalized by *rpoB*.

The relative quantification results show that *ophA2*, *iphA2*, *tphA2* gene expression is highly induced by the corresponding substrate, i.e. phthalate ($RQ_{ophA2/rpoB}$ 322), isophthalate ($RQ_{iphA2/rpoB}$ 1597), or terephthalate ($RQ_{tphA2/rpoB}$ 1117), respectively. The level of relative quantification of gene expression of *ophA2*, *iphA2* or *tphA2* is probably correlated with the initial concentration of phthalate, isophthalate, or terephthalate in liquid media. The highest RQ of *ophA2*, *iphA2* or *tphA2* gene expression was measured in RNA samples from phthalate, isophthalate, or terephthalate (20 mM) as the solo carbon source. For *iphA2* or *tphA2*, relative quantification in bacteria grown on media containing two isomers (10 mM each) are larger than that in three isomers (6.67 mM each). The only exception is the *ophA2* gene, which was expressed at a higher level in the three isomers as mixed substrates ($RQ_{ophA2/rpoB}$ 279) than that in phthalate and isophthalate ($RQ_{ophA2/rpoB}$ 238), or phthalate and terephthalate mixed substrates ($RQ_{ophA2/rpoB}$ 181). Since RNA was extracted from mixed carbon sources in the above RQ assay, we couldn't rule out the possible crossregulation or inhibition on *ophA2* gene expression by isophthalate and terephthalate. To elucidate the

relationship between *ophA2*, *iphA2*, or *tphA2* gene expression and the corresponding substrate concentration, another relative quantification or absolute quantification assay should be done to measure the *ophA2*, *iphA2*, and *tphA2* gene expression in RNA samples exacted from YZW-B cells grown on phthalate, isophthalate, or terephthalate with different concentrations.

Relative quantification values of *ophA2* gene expression are always much lower than *iphA2* and *tphA2* gene expression in YZW-B, regardless of which concentration of phthalate, isophthalate, or terephthalate were used as solo substrate or mixed substrates. For instance, *iphA2* gene expressed in isophthalate culture (RQ$_{iphA2/rpoB}$ 1597) and *tphA2* gene expressed in terephthalate culture (RQ$_{tphA2/rpoB}$ 1117) are about 5 and 4 fold higher than *ophA2* gene expression in phthalate culture (RQ$_{ophA2/rpoB}$ 322), respectively. In phthalate and terephthalate cultures, the RQ of *ophA2* and *tphA2* is 181 and 489, respectively. The lower *ophA2* fold change values are due to relatively higher *ophA2* gene expression background in YZW-B cells grown on succinate since the succinate culture is used as the calibrator to calculate RQ. Actually, the absolute copy number of the *ophA2* transcripts is much larger than that of *iphA2* or *tphA2*. According to the standard curves generated by purified *ophA2*, *iphA2*, or *tphA2* PCR product, we calculated the copy number of *ophA2*, *iphA2* and *tphA2* genes from succinate, phthalate, isophthalate and terephthalate culture. In succinate cultures, the copy numbers of the *ophA2*, *iphA2*, and *tphA2* genes are 7110, 180, and 320 copy/ng (input RNA), respectively. For YZW-B cells (OD 0.4)

grown on phthalate, isophthalate, or terephthalate as the solo carbon source, the copy number of the *ophA2* gene in the phthalate culture reached as high as 358,130 copy/ng (input RNA), which is much higher than *iphA2* expression on isophthalate (91,510 copy/ng (input RNA)) and *tphA2* expression on terephthalate (150,170 copy/ng (input RNA)), respectively.

Additionally, RT qPCR analysis show that in bacteria grown on two or three phthalate isomers, the three phthalate isomers dioxygenase genes *ophA2*, *iphA2*, and *tphA2* are highly induced by the corresponding isomers phthalate, isophthalate, and terephthalate, respectively. Moreover, *ophA2* and *tphA2* were slightly upregulated by isophthalate, which has been measured in cells grown on isophthalate ($RQ_{ophA2/rpoB}$ 3 and $RQ_{tphA2/rpoB}$ 13) or phthalate/isophthalate mixed substrate ($RQ_{tphA2/rpoB}$ 6); while phthalate or terephthalate didn't show the ability to induce *tphA2* or *ophA2* gene expression. Instead, qPCR results show that *ophA2* gene expression on phthalate and terephthalate mixed substrates ($RQ_{ophA2/rpoB}$ 181) is lower than that on phthalate and isophthalate mixed substrates ($RQ_{ophA2/rpoB}$ 238) and three isomers mixed substrates ($RQ_{ophA2/rpoB}$ 279). Thus, *ophA2* gene expression could be inhibited by terephthalate.

Figure 3.27 Relative quantification of housekeeping gene expression in YZW-B using 16S, 23S, *gyrB*, or *rpoB* as the endogenous control and RNA samples from cells grown on succinate as the calibrator. In the same substrate treatment, three RNA samples were isolated from three independent cultures from bacteria grown until the OD_{600nm}=0.4. Substrates 1-7 represent different substrates as follows: 1. phthalate, 2. isophthalate, 3. terephthalate, 4. phthalate+isophthalate, 5. phthalate+terephthalate, 6. isophthalate+terephthalate, 7. phthalate+isophthalate+terephthalate. For each RNA sample, RT-qPCR assays of 16S, 23S, *gyrB,* and *rpoB* were performed in triplicate. The averages of relative quantification of gene expression were calculated from nine assays using the $2^{-\Delta\Delta CT}$ method and the log10 values of the relative quantification were plotted. Error bars represent upper and lower limit which were calculated from Log10 RQmax and Log10 RQmin.
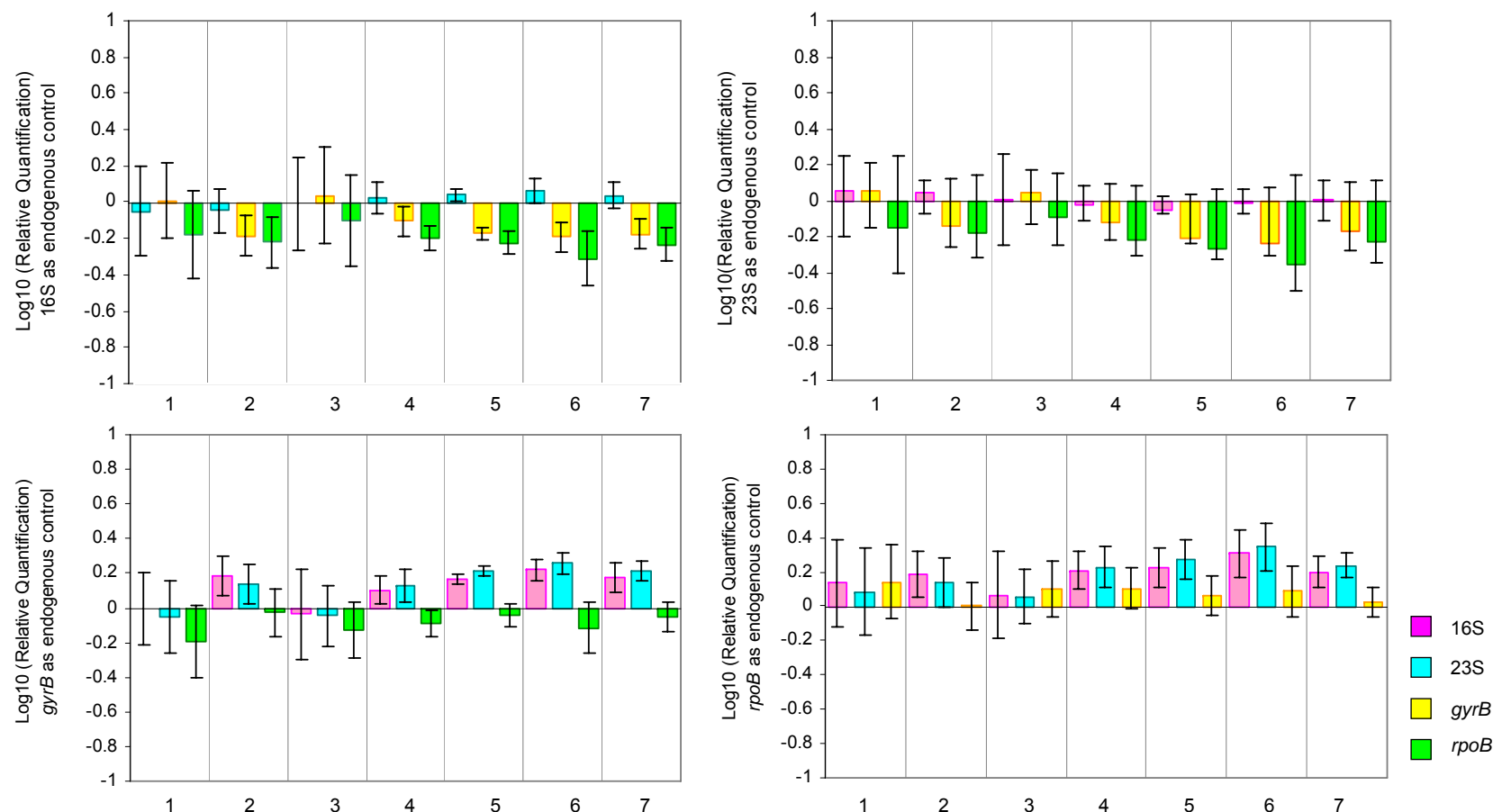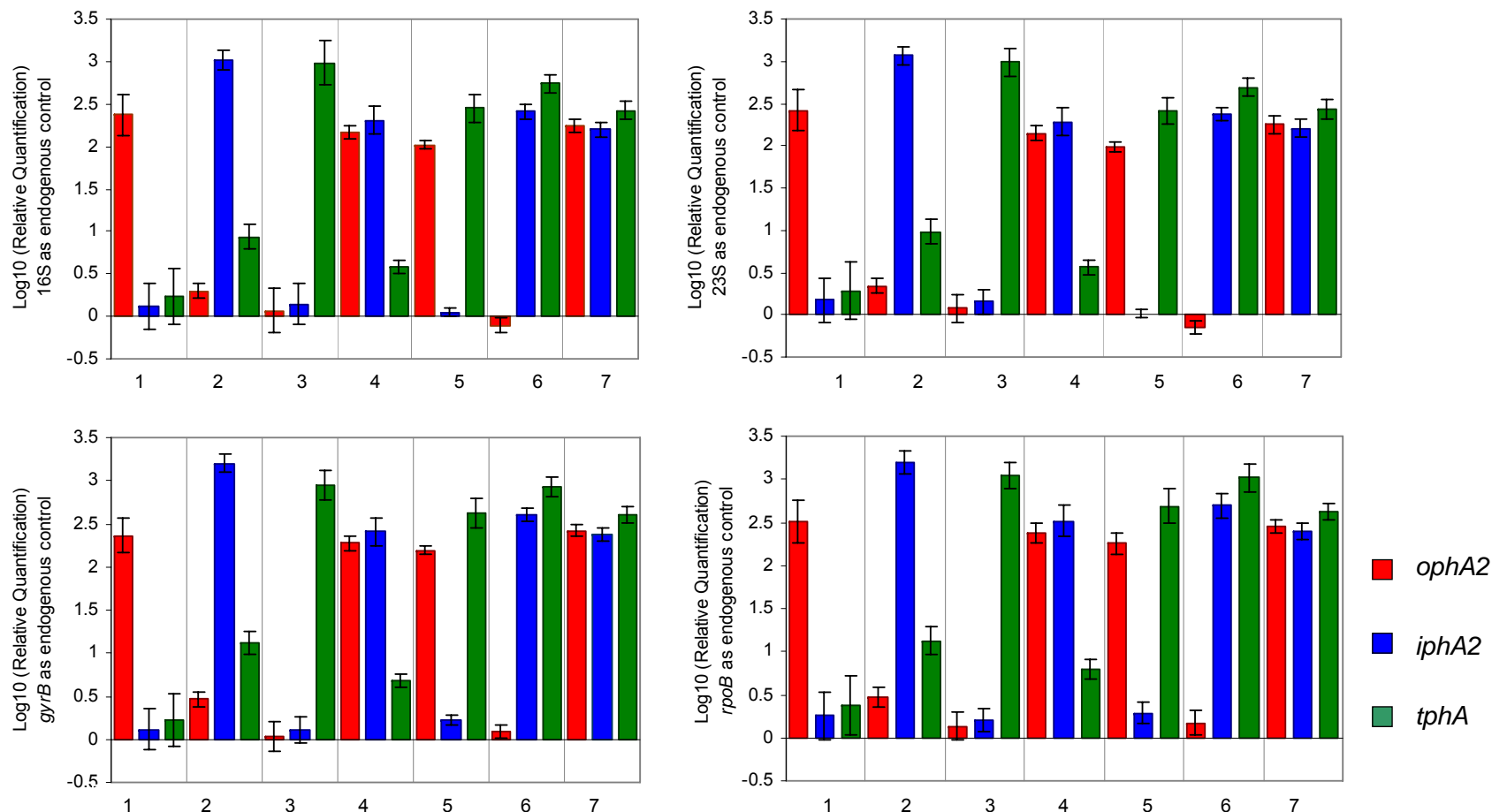
Figure 3.28 Relative quantification of phthalate gene expression in YZW-B using 16S, 23S, *gyrB*, or *rpoB* as endogenous control and RNA samples from cells grown on succinate as the calibrator. In the same substrate treatment, three RNA samples were isolated from three independent cultures from bacteria grown until the OD_{600nm}=0.4. Substrates 1-7 represent different substrates as follows: 1. phthalate, 2. isophthalate, 3. terephthalate, 4. phthalate+isophthalate, 5. phthalate+terephthalate, 6. isophthalate+terephthalate, 7. phthalate +isophthalate+terephthalate. For each RNA sample, RT-qPCR assays of 16S, 23S, *gyrB, rpoB, ophA2, iphA2, tphA2* were performed in triplicate. The averages of relative quantification of gene expression were calculated from nine assays using the $2^{-\Delta\Delta CT}$ method and the log10 values of relative quantification were plotted. Error bars represent upper and lower limit which were calculated from Log10 RQmax and Log10 RQmin.

3.4.5 Functional Analysis of *ophR*, *iphR*, and *tphR* using Knockout Mutagenesis and Quantitative Real Time PCR

3.4.5.1 Gene knockout of ophR, iphR and tphR in YZW-B

The sequence and location of *ophR*, *iphR*, and *tphR* have been revealed in the *oph*, *iph*, and *tph* gene clusters in YZW-B. Based on sequence similarity, the regulator OphR, IphR and TphR have been predicated as repressors controlling the genes for phthalate, isophthalate, and terephthalate degradation, respectively. To confirm the function of *ophR*, *iphR*, and *tphR* gene experimentally, we knocked out these regulatory genes in YZW-B and analyzed the *ophA2*, *iphA2*, and *tphA2* gene expression patterns in the knock out mutants using quantitative RT-PCR. Our hypotheses are that since *ophR*, *iphR* and *tphR* genes are repressors, deletion of these regulatory genes would not affect the phthalate isomers degradation capability of YZW-B and the phthalate, isophthalate, and terephthalate degradative genes would be expressed in an unregulated fashion.

Basically, knock out mutants were generated by gene replacement with tetracycline resistance gene cassette. For *ophR* gene knockout, a plasmid pCTB was constructed containing tetracycline resistance gene cassette *tetRA* flanking with two 1.2 kb fragments homologous to the flanking regions of *ophR* in YZW-B

genome. Intact plasmids were used for electroporation and after 24-48 h incubation. Nine colonies grew on an LB plate with tetracycline. Colony PCR using primers ophF4 and ophR5 distinguished effectively the wild type cells, double crossover knockout mutants, or single crossover cells, from which PCR produced a 1.6 kb amplicon (containing partial *ophC*, *ophR*, and partial *ophB* in YZW-B genome), 3.0 kb amplicon (containing partial *ophC*, *tetRA*, and partial *ophB*), or both 1.6 and 3.0 kb amplicon, respectively (Figure 3.29). One positive mutant out of nine colonies was screened out in the first round colony PCR. The other eight colonies turned out to be single crossover mutants. There were no spontaneous tetracycline resistant mutants growing on the plate. To ensure the positive mutant colony didn't mix with the wild type YZW-B cells, the knockout mutant was purified and confirmed by multiple PCR. Besides the first primer pair for colony PCR, another primer pair ophF6 and ophR1-1 were used to amplify the internal region of *ophR* yielding about a 500 bp amplicon from the wild type YZW-B cells and no product from double crossover knockout mutant cells since the *ophR* gene has been replaced by *tetRA* in the knockout mutant. In addition, the partial *ophA2* gene was amplified by specific primers as a positive control, which showed that the knockout mutants shared the same genetic background with the wild type strain except that the *ophR* gene was replaced by *tetRA*. We designated this *ophR* knockout mutant YZW-BΔ*ophR*.

A similar approach was used to make *iphR* gene knockout. Considering that longer flanking DNA fragments would be more favorable for homologous

recombination and gene replacement, we constructed a plasmid containing the tetracycline resistance gene *tetRA* with 1.8 kb and 2.4 kb flanking region on either end. Unfortunately, no positive double crossover mutant colonies were detected after electroporation using the intact plasmid. All of those colonies detected were single crossover mutants, which were confirmed by PCR. Next, linearized plasmid was used for electroporation, even though linearized DNA could dramatically reduce the efficiency of electrotransformation. Two positive double crossover colonies out of six colonies were screened out by colony PCR, and the purified mutant cells were confirmed by multiple PCR (Figure 3.30). This *iphR* knockout mutant was designated YZW-BΔ*iphR*.

For the *tphR* gene knockout, initially both intact and linearized pTphLTR containing a 1.8 kb left fragment, *tetRA*, and a 2.4 kb right fragment were used for electroporation. However, neither intact nor linearized plasmid generated double crossover mutant. Maybe very few double crossover mutants were generated so we were not able to screen positive mutants by PCR. In order to screen mutants easily, we released about the 5.3 kb insert containing the 1.8 kb left fragment, *tetRA*, and the 1.5 kb remaining right fragment from pTphLTR and cloned it into another vector pK18 with kanamycin resistance marker (95) as pK18-tphLTR for electroporation. Since YZW-B is both Tet^s and Km^s, it is relatively easy to screen single crossover mutants by antibiotics testing, i.e. the double crossover mutant is Tet^r and Km^s, while single crossover mutants are Tet^r and Km^r. After repeating electrotransformation with pK18-tphLTR, mutants with

Tet[r] and Km[s] phenotype were screened out and checked by PCR. Interestingly, the band representing the *tphR* gene was amplified from these Tet[r] and Km[s] mutants, which suggested that there exists another copy of *tphR* in YZW-B so that *tphR* could be amplified from one of the copies when the other copy is knocked out. The subsequent southern hybridization analysis confirmed that both strains YZW-B and YZW-E contained two copies of the *tphR* gene, while strains YZW-D and YZW-F contained only one copy. Currently, we are working on constructing the cosmid cloning library of YZW-B to reveal the sequence and location of the second copy of *tph* gene operon.

Figure 3.29 YZW-B *ophR* gene knock out and mutant screen. A) Gene organization of the *oph* genes in YZW-B, double crossover YZW-BΔ*ophR*, and single crossover YZW-B-pGEM-ophLTR. Primer pair 1 (ophF4+ophR5), 2 (ophF6+ophR1-1), 3 (ophF7+ophA2-R1) were used in colony PCR for mutant confirmation. Putative amplicons (solid line means positive and dash line means negative) from different primer pairs (on the left of the lines) are shown under the gene map. Primer pair 3 was used to amplify the *ophA2* gene as the positive control. B) YZW-BΔ*ophR* mutant screen by colony PCR using primer pair 1. In the left picture on the bottom, lane 2 shows the positive 3.0 kb band amplified from the double crossover mutant containing *tetRA*; other lanes show the 1.6 kb band amplified from spontaneous tetracycline resistant mutants containing the original *ophR*. C) Positive YZW-BΔ*ophR* mutant confirmed by PCR using mutiple primer pairs. For DNA template, Lane 1-3: wild type YZW-B DNA; Lane 4-6: YZW-B Δ*ophR* mutant DNA. For primers, Lane 1 and 4: primer pair 1; Lane 2 and 5: primer pair 2; Lane 3 and 6: primer pair 3.
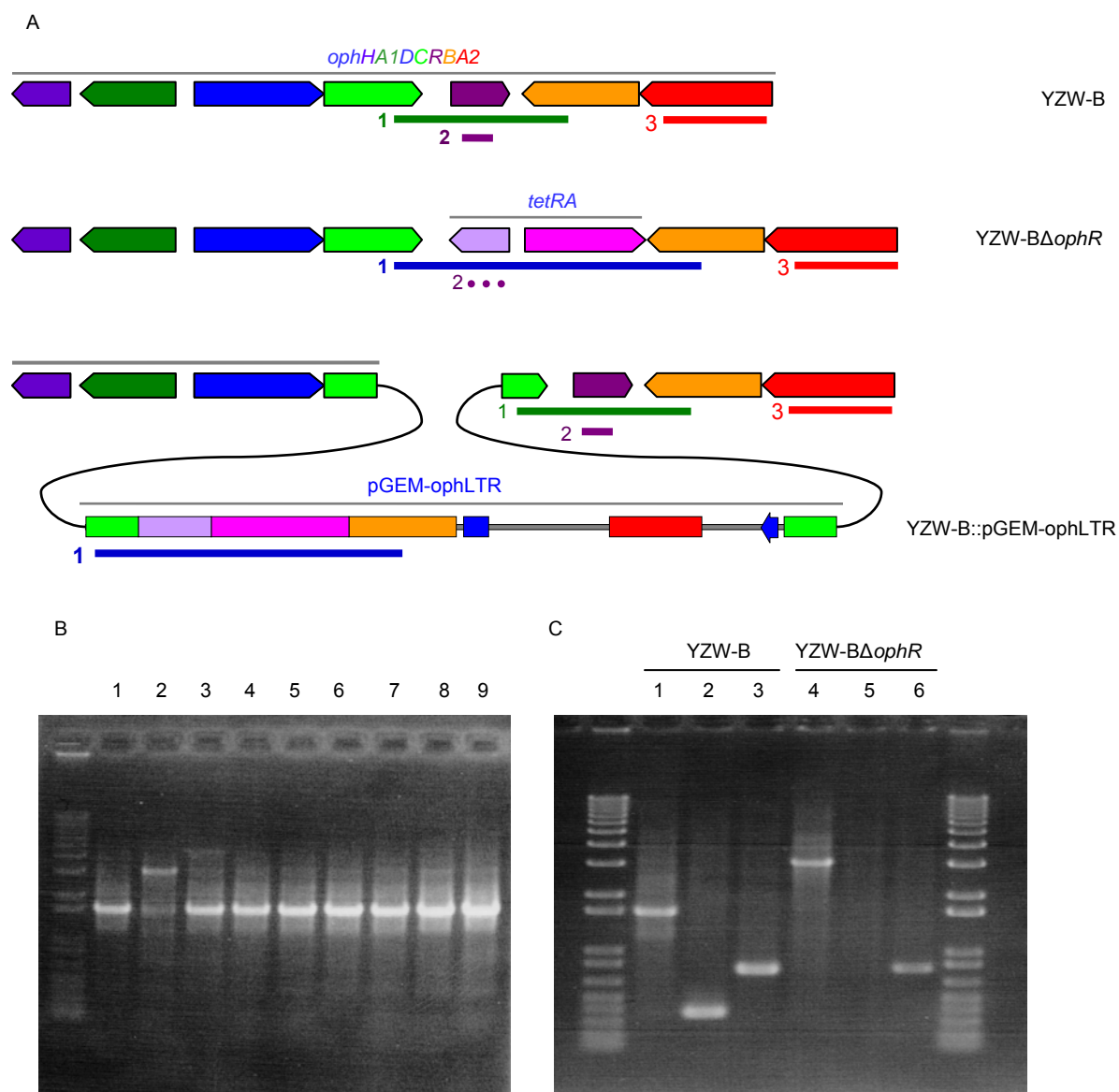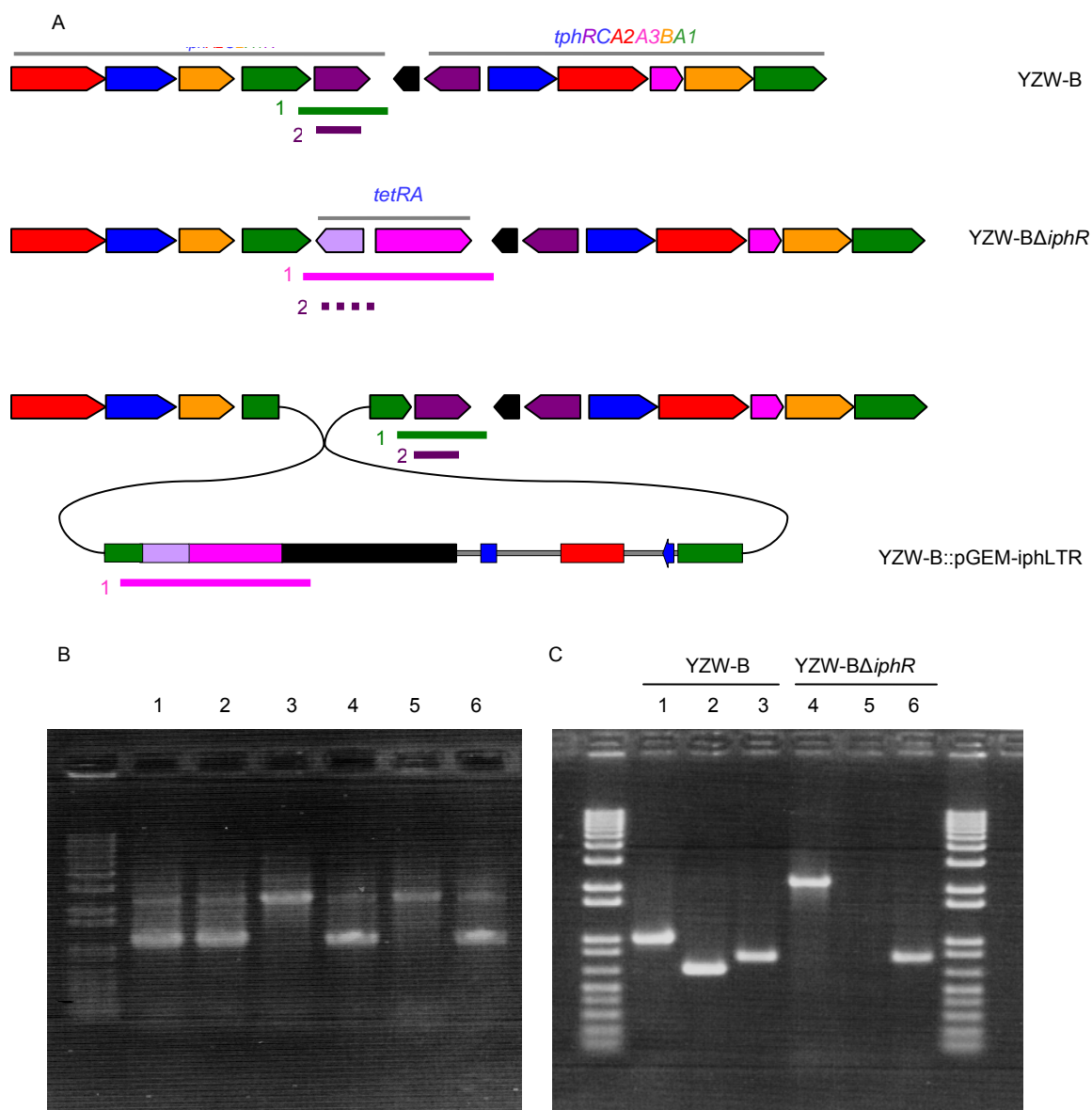
Figure 3.30 YZW-B *iphR* gene knock out and mutant screen. A) Gene organization of the *iph* genes in YZW-B, double crossover YZW-BΔ*iphR*, and single crossover YZW-B-pGEM-iphLTR. Primer pair 1 (iphR-L+iphR-R), 2 (iphR-F1+iphR-R2), 3 (ophF7+ophA2-R1) were used in colony PCR for mutants confirmation. Putative amplicons (solid line means positive and dash line means negative) from primer pair 1 and 2 (on the left of the lines) are shown under the gene map. Primer pair 3 was used to amplify the *ophA2* gene as the positive control. B) YZW-BΔ*iphR* mutant screen by colony PCR using primer pair 1. In the left picture on the bottom, lane 3 and 5 show the positive 2 kb band amplified from the double crossover mutant containing *tetRA*; Lane 1,2, 4, and 6 show a 1 kb band amplified from the double crossover mutant. C) Positive YZW-BΔ*iphR* mutant confirmed by PCR using multiple primer pairs. For DNA template, Lane 1-3: wild type YZW-B DNA; Lane 4-6: YZW-BΔ*iphR* mutant DNA. For primers, Lane 1 and 4: primer pair 1; Lane 2 and 5: primer pair 2; Lane 3 and 6: primer pair 3.

3.4.5.2 RT-qPCR analysis of ophA2, iphA2, tphA2 genes expression in YZW-BΔophR and YZW-BΔiphR

In order to determine the *ophR* gene function for phthalate isomer degradation in YZW-B, relative quantification of RT-qPCR assays were carried out to measure *ophA2*, *iphA2*, and *tphA2* gene expression in *ophR* gene knockout mutant YZW-BΔ*ophR*. Quantification of each target gene was performed with RNA samples isolated from cells of the wild type YZW-B and the *ophR* knock out mutant grown on succinate or phthalate. The level of gene expression under each growth condition was normalized by *rpoB* and calibrated by gene expression of wild type YZW-B cells grown on succinate.

RT-qPCR analysis showed that the expression level of *ophA2* in wild type cells grown on phthalate was 322 fold higher than that in cells grown on succinate. In contrast, the expression level of *ophA2* in YZW-BΔ*ophR* cells grown on phthalate or succinate is 226 and 117 fold higher than wild type cells grown on succinate, indicating constitutive expression of the *ophA2* gene in the YZW-BΔ*ophR* strain. These results suggest that the *ophR* gene encodes a repressor that negatively regulates phthalate gene expression. In addition, deletion of *ophR* did not affect *iphA2* and *tphA2* gene expression, which demonstrated that *ophR* is specifically responsible for controlling phthalate degradation (Figure 3.31; Table 3.10).

For *iphR* gene knockout, similar RT-qPCR method was used to measure *rpoB*, *ophA2*, *iphA2*, and *tphA2* except that RNA samples of YZW-BΔ*iphR* were isolated from cells grown on isophthalate or succinate. RT-qPCR analysis showed that the expression level of *iphA2* in wild type cells grown on isophthalate was 1597 fold higher than that in cells grown on succinate. In contrast, the expression level of *ophA2* in YZW-BΔ*iphR* cells grown on phthalate or succinate is 1095 and 428 fold higher than wild type cells grown on succinate, indicative of constitutive expression of the *iphA2* gene in the YZW-BΔ*iphR* strain. These results suggested that the *iphR* gene encodes a repressor that negatively regulates isophthalate gene expression. In addition, deletion of *iphR* did not affect *ophA2*, but slightly affect *tphA2* gene expression. *tphA2* gene expression in YZW-BΔ*iphR* grown on succinate is 2.8 fold higher than that in wild type, which demonstrated that *iphR* somehow is involved in terephthalate gene regulation (Figure 3.32; Table 3.11).

Another interesting phenomena is that in the present of phthalate or isophthalate, the target gene *ophA2* ($RQ_{ophA2/rpoB}$ 322) or *iphA2* ($RQ_{iphA2/rpoB}$ 1587) expression level in wild type cells are always higher than *ophA2* ($RQ_{ophA2/rpoB}$ 226) or *iphA2* ($RQ_{iphA2/rpoB}$ 1095) in *ophR* or *iphR* knockout mutants. The ratio of $RQ_{ophA2/rpoB}$ or $RQ_{iphA2/rpoB}$ from wild type cells versus mutant coincidently is about 3:2. The foreign tetracycline resistance genes introduced by gene replacement changed the gene organization and it could interfere and decrease the phthalate or isophthalate gene expression level in the *ophR* or *iphR* knockout mutants grow

on phthalate or isophthalate. Moreover, for knock out mutants, the ratio of $RQ_{ophA2/rpoB}$ in the present of phthalate versus succinate is also similar to the ratio of $RQ_{iphA2/rpoB}$ in the presence of isophthalate versus succinate, which is about 2:1. When mutant cells grow on succinate, *ophA2* or *iphA2* transcripts actually are not necessarily synthesized for bacterial growth. So there exists another secondary regulation pathway such as catabolite repression or simply negative feedback to control *ophA2* or *iphA2* constitutive expression in a relatively lower level even without regulatory protein negative control.

Figure 3.31 Relative quantification of *ophA2*, *iphA2*, and *tphA2* gene expression in YZW-B and YZW-BΔ*ophR* grown on succinate and phthalate. In the same substrate treatment, three RNA samples were isolated from three independent cultures from bacteria grown until an OD_{600nm} of 0.4. For each RNA sample, RT-qPCR assays of *rpoB, ophA2, iphA2,* and *tphA2* were performed in triplicate. The average of the relative quantification of gene expression was calculated from nine assays using the $2^{-\Delta\Delta C_T}$ method. Data were plotted with Log10 of relative quantification on the Y axis and target genes (1-4 for *ophA2*, 5-8 for *iphA2*, and 9-12 for *tphA2*) on the X axis. Red, gray, pink and blue columns also indicate YZW-B or YZW-BΔ*ophR* grown on succinate or phthalate. Error bars represent the upper and lower limits which were calculated from Log10 RQmax and Log10 RQmin.

Table 3.10 Relative quantification of *ophA2*, *iphA2*, and *tphA2* gene expression in YZW-B and YZW-BΔ*ophR* grown on succinate and phthalate.

Gene	YZW-B		YZW-BΔ*ophR*	
	phthalate	succinate	phthalate	succinate
ophA2	**322.339**	1	**226.482**	**115.712**
iphA2	1.807	1	1.496	0.859
tphA2	2.348	1	2.119	1.365

* In the same substrate treatment, three RNA samples were isolated from three independent cultures from bacteria grown until an OD$_{600nm}$ of 0.4. For each RNA sample, RT-qPCR assays of *rpoB, ophA2, iphA2,* and *tphA2* were performed in triplicates. The averages of relative quantification of gene expression were calculated from nine assays using the $2^{-\Delta\Delta CT}$ method.
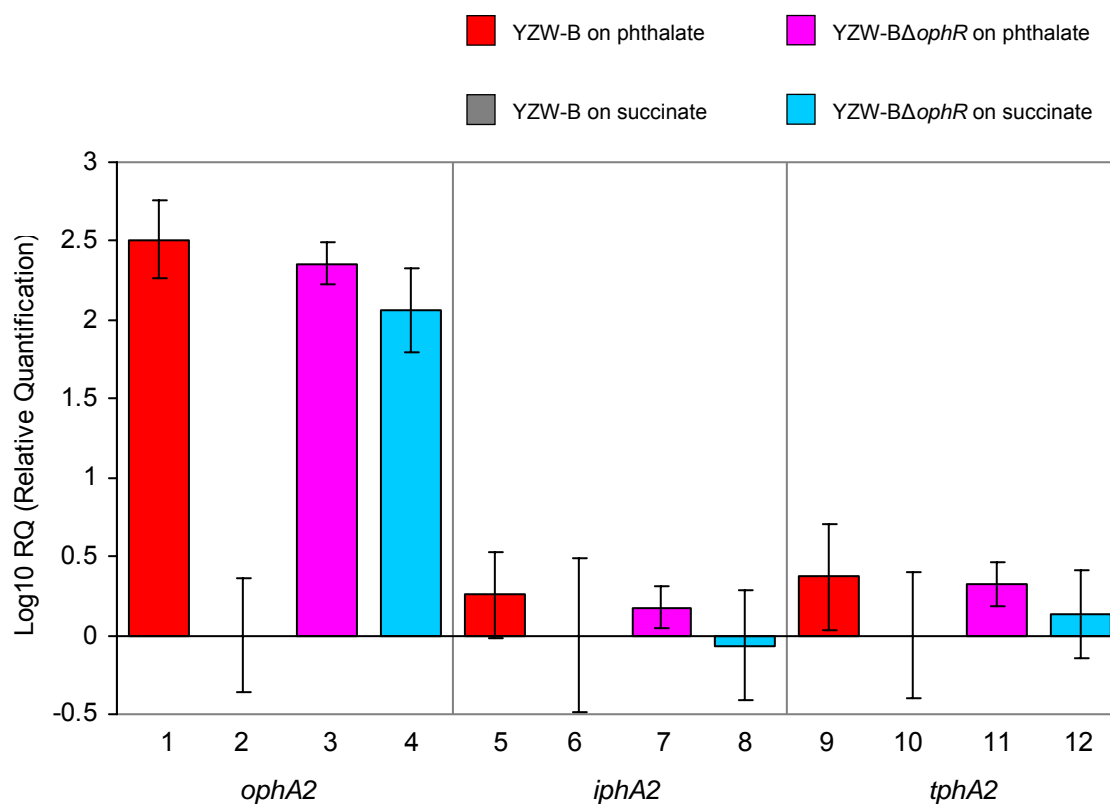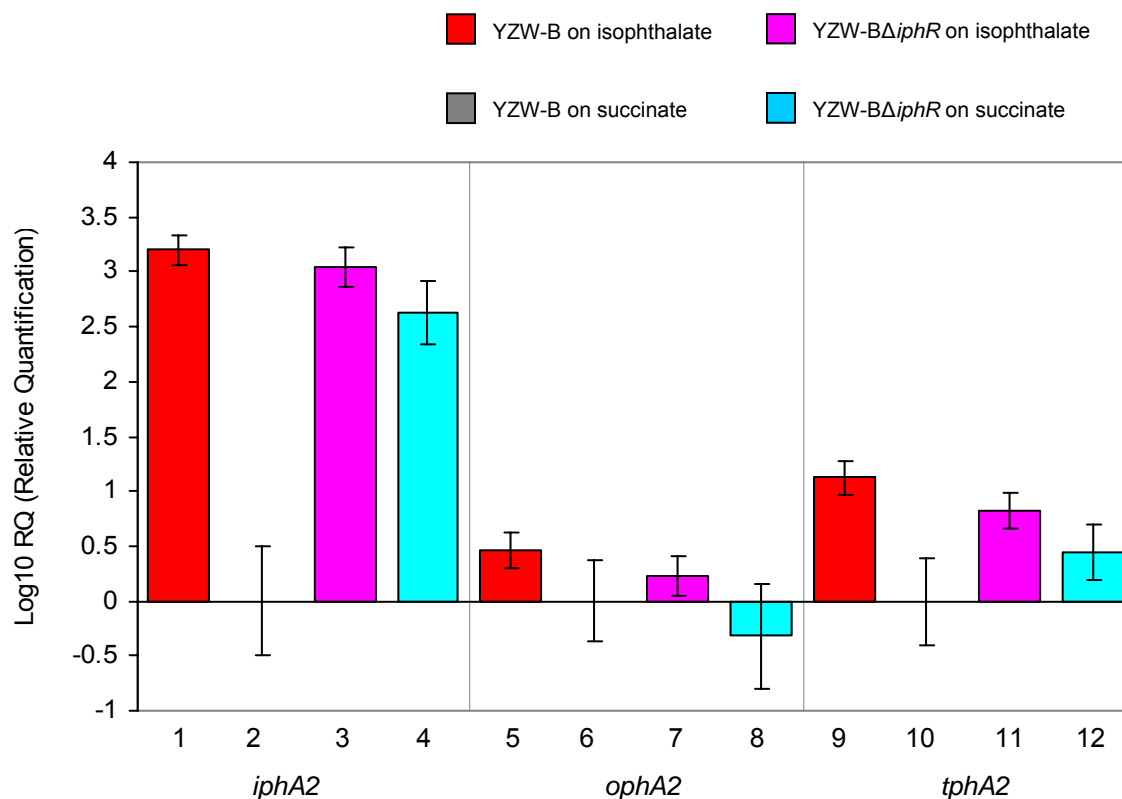
Figure 3.32 Relative quantification of *iphA2*, *ophA2*, and *tphA2* gene expression in YZW-B and YZW-BΔ*iphR* grown on succinate and isophthalate. For the same substrate treatment, three RNA samples were isolated from three independent cultures when bacteria grown until an OD_{600nm} of 0.4. For each RNA sample, RT-qPCR assays of *rpoB, ophA2, iphA2,* and *tphA2* were performed in triplicate. The average of the relative quantification of gene expression was calculated from nine assays using the $2^{-\Delta\Delta C_T}$ method. Data were plotted with Log10 of relative quantification on the Y axis and target genes (1-4 for *iphA2*, 5-8 for *ophA2*, and 9-12 for *tphA2*) on the X axis. Red, gray, pink and blue columns also indicate YZW-B or YZW-BΔ*iphR* grown on succinate or isophthalate. Error bars represent the upper and lower limits which were calculated from Log10 RQmax and Log10 RQmin.

Table 3.11 Relative quantification of *iphA2*, *ophA2*, and *tphA2* gene expression in YZW-B and YZW-BΔ*iphR* grown on succinate and isophthalate.

Gene	YZW-B		YZW-BΔ*iphR*	
	isophthalate	succinate	isophthalate	succinate
iphA2	**1597.381**	1	**1094.758**	**428.438**
ophA2	2.895	1	1.718	0.478
tphA2	13.278	1	6.707	2.793

* In the same substrate treatment, three RNA samples were isolated from three independent cultures when bacteria grown until an OD_{600nm} of 0.4. For each RNA sample, RT-qPCR assays of *rpoB, ophA2, iphA2,* and *tphA2* were performed in triplicate. The averages of relative quantification of gene expression were calculated from nine assays using the $2^{-\Delta\Delta CT}$ method.

References

1.   2006. http://www.phthalates.org/whatare/index.asp. Phthalates information center, American Chemistry Council.

2.   **Aftring, R. P., B. E. Chalker, and B. F. Taylor.** 1981. Degradation of phthalic acids by denitrifying, mixed cultures of bacteria. Appl. Environ. Microbiol. **41:**1177-1183.

3.   **Alatriste-Mondragon, F., R. Iranpour, and B. K. Ahring.** 2003. Toxicity of di-(2-ethylhexyl) phthalate on the anaerobic digestion of wastewater sludge. Water Research **37:**1260-1269.

4.   **Altschul, S., T. Madden, A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. **25:**3389-3402.

5.   **Anzai, Y., Y. Kudo, and H. Oyaizu.** 1997. The phylogeny of the genera *Chryseomonas*, *Flavimonas*, and *Pseudomonas* supports synonymy of these three genera. Int. J. Syst. Bacteriol. **47:**249-251.

6.   **ATSDR.** 1997. Toxicological profile for di-*n*-octylphthalate (DNOP). Agency for Toxic Substances and Disease Registry, Department of Health and Human Services.

7.   **ATSDR.** 2002. Toxicological profile for di(2-ethylhexyl)phthalate (DEHP). Agency for Toxic Substances and Disease Registry, Department of Health and Human Services.

8.   **Ballou, D., and C. Batie.** 1988. Phthalate oxygenase, a Rieske iron-sulfur protein from *Pseudomonas cepacia*. Prog. Clin. Biol. Res. **274:**211-226.

9.   **Batie, C. J., E. LaHaie, and D. P. Ballou.** 1987. Purification and characterization of phthalate oxygenase and phthalate oxygenase reductase from *Pseudomonas cepacia*. J. Biol. Chem. **262:**1510-1518.

10. **Battersby, N. S., and V. Wilson.** 1989. Survey of the anaerobic biodegradation potential of organic chemicals in digesting sludge. Appl. Environ. Microbiol. **55:**433-439.

11. **Benson, D. A., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and D. L. Wheeler.** 2007. GenBank. Nucl. Acids Res. **35:**D21-25.

12. **Blount, B. C., M. J. Silva, S. P. Caudill, L. L. Needham, J. L. Pirkle, E. J. Sampson, G. W. Lucier, R. J. Jackson, and J. W. Brock.** 2000. Levels of seven urinary phthalate metabolites in a human reference population. Environ. Health. Perspect. **108:**979-982.

13. **Chang, H.-K., and G. J. Zylstra.** 1999. Characterization of the phthalate permease OphD from *Burkholderia cepacia* ATCC 17616. J. Bacteriol. **181:**6197-6199.

14. **Chang, H.-K., and G. J. Zylstra.** 2008. Dual phthalate transport systems in *Burkholderia cepacia.* Unpublished.

15. **Chang, H.-K., and G. J. Zylstra.** 1998. Novel organization of the genes for phthalate degradation from *Burkholderia cepacia* DBO1. J. Bacteriol. **180:**6529-6537.

16. **Chang, H.-K., and G. J. Zylstra.** 1999. Role of quinolinate phosphoribosyl transferase in degradation of phthalate by *Burkholderia cepacia* DBO1. J. Bacteriol. **181:**3069-3075.

17. **Chang, Y. H., J. Han, J. Chun, K. C. Lee, M. S. Rhee, Y. B. Kim, and K. S. Bae.** 2002. *Comamonas koreensis* sp. nov., a non-motile species from wetland in Woopo, Korea. Int. J. Syst. Evol. Microbiol **52:**377-381.

18. **Chen, C. L., H. Macarie, I. Ramirez, A. Olmos, S. L. Ong, O. Monroy, and W. T. Liu.** 2004. Microbial community structure in a thermophilic anaerobic hybrid reactor degrading terephthalate. Microbiology **150:**3429-3440.

19. **Cheung, J. K., R. K. Lam, M. Y. Shi, and J. D. Gu.** 2007. Environmental fate of endocrine-disrupting dimethyl phthalate esters (DMPE) under sulfate-reducing condition. Sci. Total. Environ. **381:**126-133.

20. **Choi, K. Y., D. Kim, J.-C. Chae, G. J. Zylstra, and E. Kim.** 2007. Requirement of duplicated operons for maximal metabolism of phthalate by *Rhodococcus* sp. strain DK17. Biochem. Biophys. Res. Commun. **357:**766-771.

21. **Choi, K. Y., D. Kim, W. J. Sul, J.-C. Chae, G. J. Zylstra, Y. M. Kim, and E. Kim.** 2005. Molecular and biochemical analysis of phthalate and terephthalate degradation by *Rhodococcus* sp. strain DK17. FEMS Microbiol. Lett. **252:**207-213.

22. **Choi, K. Y., G. J. Zylstra, and E. Kim.** 2007. Benzoate catabolite repression of the phthalate degradation pathway in *Rhodococcus* sp. strain DK17. Appl. Environ. Microbiol. **73:**1370-1374.

23. **Cirillo, J. D., R. G. Barletta, B. R. Bloom, and W. R. Jacobs, Jr.** 1991. A novel transposon trap for mycobacteria: isolation and characterization of IS1096. J. Bacteriol. **173:**7772-7780.

24. **Cole, J. R., B. Chai, R. J. Farris, Q. Wang, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, A. M. Bandela, E. Cardenas, G. M. Garrity, and J. M. Tiedje.** 2006. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucl. Acids Res.**:**gkl889.

25. **Correll, C. C., C. J. Batie, D. P. Ballou, and M. L. Ludwig.** 1992. Phthalate dioxygenase reductase: a modular structure for electron transfer from pyridine nucleotides to [2Fe-2S]. Science **258:**1604-1610.

26. **Coulter, E. D., N. Moon, C. J. Batie, W. R. Dunham, and D. P. Ballou.** 1999. Electron paramagnetic resonance measurements of the ferrous mononuclear site of phthalate dioxygenase substituted with alternate divalent metal ions: direct evidence for ligation of two histidines in the copper(II)-reconstituted protein. Biochemistry **38:**11062-11072.

27.   **Cowles, C. E., N. N. Nichols, and C. S. Harwood.** 2000. BenR, a XylS homologue, regulates three different pathways of aromatic acid degradation in *Pseudomonas putida*. J. Bacteriol. **182:**6339-6346.

28.   **Doddamani, H. P., and H. Z. Ninnekar.** 2000. Biodegradation of phenanthrene by a *Bacillus* species. Curr. Microbiol. **41:**11-14.

29.   **Eaton, R.** 1996. *p*-Cumate catabolic pathway in *Pseudomonas putida* FI: cloning and characterization of DNA carrying the *cmt* operon. J. Bacteriol. **178:**1351-1362.

30.   **Eaton, R. W.** 1997. *p*-Cymene catabolic pathway in Pseudomonas putida F1: cloning and characterization of DNA encoding conversion of *p*-cymene to *p*-cumate. J. Bacteriol. **179:**3171-3180.

31.   **Eaton, R. W.** 2001. Plasmid-encoded phthalate catabolic pathway in *Arthrobacter keyseri* 12B. J. Bacteriol. **183:**3689-3703.

32.   **Eaton, R. W., and D. W. Ribbons.** 1982. Metabolism of dibutylphthalate and phthalate by *Micrococcus* sp. strain 12B. J. Bacteriol. **151:**48-57.

33.   **Eaton, R. W., and D. W. Ribbons.** 1982. Metabolism of dimethylphthalate by *Micrococcus* sp. strain 12B. J. Bacteriol. **151:**465-467.

34.   **Elomari, M., L. Coroler, S. Verhille, D. Izard, and H. Leclerc.** 1997. *Pseudomonas monteilii* sp. nov., isolated from clinical specimens. Int. J. Syst. Bacteriol. **47:**846-852.

35.   **Engelhardt, G., and P. R. Wallnofer.** 1978. Metabolism of di- and mono-*n*-butyl phthalate by soil bacteria. Appl. Environ. Microbiol. **35:**243-246.

36.   **EPA.** 2004. 2004 Reporting year list of TRI chemicals. U.S. Environmental protection agency.

37.   **Eulberg, D., S. Lakner, L. A. Golovleva, and M. Schlomann.** 1998. Characterization of a protocatechuate catabolic gene cluster from *Rhodococcus opacus* 1CP: Evidence for a merged enzyme with 4-

carboxymuconolactone-decarboxylating and 3-oxoadipate enol-Lactone-hydrolyzing activity. J. Bacteriol. **180:**1072-1081.

38. **Gassner, G. T., D. P. Ballou, G. A. Landrum, and J. W. Whittaker.** 1993. Magnetic circular dichroism studies on the mononuclear ferrous active site of phthalate dioxygenase from *Pseudomonas cepacia* show a change of ligation state on substrate binding. Biochemistry **32:**4820-4825.

39. **Gassner, G. T., M. L. Ludwig, D. L. Gatti, C. C. Correll, and D. P. Ballou.** 1995. Structure and mechanism of the iron-sulfur flavoprotein phthalate dioxygenase reductase. Faseb J. **9:**1411-1418.

40. **Gonzalez, V., P. Bustos, M. A. Ramirez-Romero, A. Medrano-Soto, H. Salgado, I. Hernandez-Gonzalez, J. C. Hernandez-Celis, V. Quintero, G. Moreno-Hagelsieb, L. Girard, O. Rodriguez, M. Flores, M. A. Cevallos, J. Collado-Vides, D. Romero, and G. Davila.** 2003. The mosaic structure of the symbiotic plasmid of *Rhizobium etli* CFN42 and its relation to other symbiotic genome compartments. Genome Biol. **4:**R36.

41. **Goyal, A. K., and G. J. Zylstra.** 1996. Molecular cloning of novel genes for polycyclic aromatic hydrocarbon degradation from *Comamonas testosteroni* GZ39. Appl. Environ. Microbiol. **62:**230-236.

42. **Graham, P. R.** 1973. Phthalate ester plasticizers-why and how they are used. Environ. Health. Perspect. **3:**3–12.

43. **Gumaelius, L., G. Magnusson, B. Pettersson, and G. Dalhammar.** 2001. *Comamonas denitrificans* sp. nov., an efficient denitrifying bacterium isolated from activated sludge. Int. J. Syst. Evol. Microbiol. **51:**999-1006.

44. **Gupta, P., G. S. N. Reddy, D. Delille, and S. Shivaji.** 2004. *Arthrobacter gangotriensis* sp. nov. and *Arthrobacter kerguelensis* sp. nov. from Antarctica. Int. J. Syst. Evol. Microbiol. **54:**2375-2378.

45. **Gurbiel, R. J., C. J. Batie, M. Sivaraja, A. E. True, J. A. Fee, B. M. Hoffman, and D. P. Ballou.** 1989. Electron-nuclear double resonance spectroscopy of 15N-enriched phthalate dioxygenase from *Pseudomonas*

cepacia proves that two histidines are coordinated to the [2Fe-2S] Rieske-type clusters. Biochemistry **28:**4861-4871.

46. **Habe, H., J.-S. Chung, H. Kato, Y. Ayabe, K. Kasuga, T. Yoshida, H. Nojiri, H. Yamane, and T. Omori.** 2004. Characterization of the upper pathway genes for fluorene metabolism in *Terrabacter* sp. Strain DBF63. J. Bacteriol. **186:**5938-5944.

47. **Habe, H., M. Miyakoshi, J. Chung, K. Kasuga, T. Yoshida, H. Nojiri, and T. Omori.** 2003. Phthalate catabolic gene cluster is linked to the angular dioxygenase gene in *Terrabacter* sp. strain DBF63. Appl. Microbiol. Biotechnol. **61:**44-54.

48. **Hara, H., L. D. Eltis, J. E. Davies, and W. W. Mohn.** 2007. Transcriptomic analysis reveals a bifurcated terephthalate degradation pathway in *Rhodococcus* sp. strain RHA1. J. Bacteriol. **189:**1641-1647.

49. **Hauser, R., J. D. Meeker, N. P. Singh, M. J. Silva, L. Ryan, S. Duty, and A. M. Calafat.** 2006. DNA damage in human sperm is related to urinary levels of phthalate monoester and oxidative metabolites. Hum. Reprod.

50. **Ibrahim, A., P. Gerner-Smidt, and W. Liesack.** 1997. Phylogenetic relationship of the twenty-one DNA groups of the genus Acinetobacter as revealed by 16S ribosomal DNA sequence analysis. Int. J. Syst. Bacteriol. **47:**837-841.

51. **James, K. D., and P. A. Williams.** 1998. *ntn* genes determining the early steps in the divergent catabolism of 4-nitrotoluene and toluene in *Pseudomonas* sp. strain TW3. J. Bacteriol. **180:**2043-2049.

52. **Jonsson, S., and A. Baun.** 2003. Toxicity of mono- and diesters of *o*-phthalic esters to a crustacean, a green alga, and a bacterium. Environ. Toxicol. Chem. **22:**3037-3043.

53. **Kaiser, J.** 2005. TOXICOLOGY: Panel finds no proof that phthalates harm infant reproductive systems. Science **310:**422a.

54. **Kallastu, A., R. Horak, and M. Kivisaar.** 1998. Identification and characterization of IS1411, a new insertion sequence which causes transcriptional activation of the phenol degradation genes in *Pseudomonas putida.* J. Bacteriol. **180:**5306-5312.

55. **Kaneko, T., Y. Nakamura, S. Sato, K. Minamisawa, T. Uchiumi, S. Sasamoto, A. Watanabe, K. Idesawa, M. Iriguchi, K. Kawashima, M. Kohara, M. Matsumoto, S. Shimpo, H. Tsuruoka, T. Wada, M. Yamada, and S. Tabata.** 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. DNA Res. **9:**189-197.

56. **Keum, Y.-S., J.-S. Seo, Y. Hu, and Q. Li.** 2006. Degradation pathways of phenanthrene by *Sinorhizobium* sp. C4. Appl. Microbiol. Biotech. **71:**935-941.

57. **Keyser, P., B. G. Pujar, R. W. Eaton, and D. W. Ribbons.** 1976. Biodegradation of the phthalates and their esters by bacteria. Environ. Health Perspect. **18:**159-166.

58. **Kim, S.-J., O. Kweon, R. C. Jones, J. P. Freeman, R. D. Edmondson, and C. E. Cerniglia.** 2007. Complete and integrated pyrene degradation pathway in *Mycobacterium vanbaalenii* PYR-1 based on systems biology. J. Bacteriol. **189:**464-472.

59. **Kivisaar, M., L. Kasak, and A. Nurk.** 1991. Sequence of the plasmid-encoded catechol 1,2-dioxygenase-expressing gene, *pheB,* of phenol-degrading *Pseudomonas* sp. strain EST1001. Gene **98:**15-20.

60. **Kleerebezem, R., L. W. H. Pol, and G. Lettinga.** 1999. Anaerobic degradation of phthalate isomers by methanogenic consortia. Appl. Environ. Microbiol. **65:**1152-1160.

61. **Koch, C., P. Schumann, and E. Stackebrandt.** 1995. Reclassification of *Micrococcus agilis* (Ali-Cohen 1889) to the genus *Arthrobacter* as *Arthrobacter agilis* comb. nov. and emendation of the genus *Arthrobacter.* Int. J. Syst. Bacteriol. **45:**837-839.

62. **Kothary, M. H., B. A. McCardell, C. D. Frazar, D. Deer, and B. D. Tall.** 2007. Characterization of the zinc-containing metalloprotease encoded by *zpx* and development of a species-specific detection method for *Enterobacter sakazakii*. Appl. Environ. Microbiol. **73:**4142-4151.

63. **Kumar, S., K. Tamura, and M. Nei.** 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. **5:**150-163.

64. **Latini, G., A. Del Vecchio, M. Massaro, A. Verrotti, and D. E. F. C.** 2006. In utero exposure to phthalates and fetal development. Curr. Med. Chem. **13:**2527-2534.

65. **Lau, P. C. K., H. Bergeron, D. Labbe, Y. Wang, R. Brousseau, and D. T. Gibson.** 1994. Sequence and expression of the *todGIH* genes involved in the last three steps of toluene degradation by *Pseudomonas putida* F1. Gene **146:**7-13.

66. **Lee, J.-H., T. Omori, and T. Kodama.** 1994. Identification of the metabolic intermediates of phthalate by Tn5 mutants of *Pseudomonas testosteroni* and analysis of the 4,5-dihydroxyphthalate decarboxylase gene. J. Ferment. Bioeng. **77:**583-590.

67. **Liang, D.-W., T. Zhang, and H. H. P. Fang.** 2006. Denitrifying degradation of dimethyl phthalate. Appl. Microbiol. Biotechnol. **74:**221-229.

68. **Livak, K. J., and T. D. Schmittgen.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. Methods **25:**402-408.

69. **López, Z., J. Vila, C. Minguillón, and M. Grifoll.** 2006. Metabolism of fluoranthene by *Mycobacterium* sp. strain AP1. Appl. Microbiol. Biotechnol. **70:**747-756.

70. **Lottrup, G., A. M. Andersson, H. Leffers, G. K. Mortensen, J. Toppari, N. E. Skakkebaek, and K. M. Main.** 2006. Possible impact of phthalates on infant reproductive health. Int. J. Androl. **29:**172-180.

71. **Lovekamp-Swan, T., and B. J. Davis.** 2003. Mechanisms of phthalate ester toxicity in the female reproductive system. Environ. Health Perspect. **111:**139-145.

72. **Mahillon, J., and M. Chandler.** 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. **62:**725-774.

73. **Marchler-Bauer, A., J. B. Anderson, P. F. Cherukuri, C. DeWeese-Scott, L. Y. Geer, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, G. H. Marchler, M. Mullokandov, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, R. A. Yamashita, J. J. Yin, D. Zhang, and S. H. Bryant.** 2005. CDD: a conserved domain database for protein classification. Nucl. Acids Res. **33:**D192-196.

74. **Marchler-Bauer, A., and S. H. Bryant.** 2004. CD-Search: protein domain annotations on the fly. Nucl. Acids Res. **32:**W327-331.

75. **Margesin, R., P. Schumann, C. Sproer, and A.-M. Gounot.** 2004. *Arthrobacter psychrophenolicus* sp. nov., isolated from an alpine ice cave. Int. J. Syst. Evol. Microbiol. **54:**2067-2072.

76. **Maruyama, K., K. Akita, C. Naitou, M. Yoshida, and T. Kitamura.** 2005. Purification and characterization of an esterase hydrolyzing monoalkyl phthalates from *Micrococcus* sp. YGJ1. J. Biochem. (Tokyo) **137:**27-32.

77. **McLeod, M. P., R. L. Warren, W. W. L. Hsiao, N. Araki, M. Myhre, C. Fernandes, D. Miyazawa, W. Wong, A. L. Lillquist, D. Wang, M. Dosanjh, H. Hara, A. Petrescu, R. D. Morin, G. Yang, J. M. Stott, J. E. Schein, H. Shin, D. Smailus, A. S. Siddiqui, M. A. Marra, S. J. M. Jones, R. Holt, F. S. L. Brinkman, K. Miyauchi, M. Fukuda, J. E. Davies, W. W. Mohn, and L. D. Eltis.** 2006. The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. PNAS **103:**15582-15587.

78. **Menn, F.-M., G. J. Zylstra, and D. T. Gibson.** 1991. Location and sequence of the *todF* gene encoding 2-hydroxy-6-oxohepta-2,4-dienoate hydrolase in *Pseudomonas putida* F1. Gene **104:**91-94.

79. **Moore, E. R. B., M. Mau, A. Arnscheidt, E. C. Boettger, R. A. Hutson, M. D. Collins, Y. Van de Peer, R. De Wachter, and K. N. Timmis.** 1996. The determination and comparison of the 16S rRNA gene sequences of species of the genus *Pseudomonas* (sensu stricto) and estimation of the natural intrageneric relationships. Syst. Appl. Microbiol. **19:**478-492.

80. **Morawski, B., R. W. Eaton, J. T. Rossiter, S. Guoping, H. Griengl, and D. W. Ribbons.** 1997. 2-Naphthoate catabolic pathway in *Burkholderia* strain JT 1500. J. Bacteriol. **179:**115-121.

81. **Nakatsu, C., N. Straus, and R. Wyndham.** 1995. The nucleotide sequence of the Tn5271 3-chlorobenzoate 3,4-dioxygenase genes (*cbaAB*) unites the class IA oxygenases in a single lineage. Microbiology **141:**485-495.

82. **Nelson, K. E., C. Weinel, I. T. Paulsen, R. J. Dodson, H. Hilbert, V. A. P. Martins dos Santos, D. E. Fouts, S. R. Gill, M. Pop, M. Holmes, L. Brinkac, M. Beanan, R. T. DeBoy, S. Daugherty, J. Kolonay, R. Madupu, W. Nelson, O. White, J. Peterson, H. Khouri, I. Hance, P. C. Lee, E. Holtzapple, D. Scanlan, K. Tran, A. Moazzez, T. Utterback, M. Rizzo, K. Lee, D. Kosack, D. Moestl, H. Wedler, J. Lauber, D. Stjepandic, J. Hoheisel, M. Straetz, S. Heim, C. Kiewitz, J. Eisen, and K. N. Timmis.** 2002. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. Environ. Microbiol. **4:**799-808.

83. **Nemec, A., T. De Baere, I. Tjernberg, M. Vaneechoutte, T. J. van der Reijden, and L. Dijkshoorn.** 2001. *Acinetobacter ursingii* sp. nov. and *Acinetobacter schindleri* sp. nov., isolated from human clinical specimens. Int. J. Syst. Evol. Microbiol. **51:**1891-1899.

84. **Nishimori, E., K. Kita-Tsukamoto, and H. Wakabayashi.** 2000. *Pseudomonas plecoglossicida* sp. nov., the causative agent of bacterial haemorrhagic ascites of ayu, *Plecoglossus altivelis.* Int. J. Syst. Evol. Microbiol. **50:**83-89.

85. **Nishioka, T., M. Iwata, T. Imaoka, M. Mutoh, Y. Egashira, T. Nishiyama, T. Shin, and T. Fujii.** 2006. A mono-2-ethylhexyl phthalate hydrolase from a *Gordonia* sp. that is able to dissimilate di-2-ethylhexyl phthalate. Appl. Environ. Microbiol. **72:**2394-2399.

86. **Nomura, Y., M. Nakagawa, N. Ogawa, S. Harashima, and Y. Oshima.** 1992. Genes in PHT plasmid encoding the initial degradation pathway of phthalate in *Pseudomonas putida*. J. Ferment. Bioeng. **74:**333-344.

87. **Nozawa, T., and Y. Maruyama.** 1988. Anaerobic metabolism of phthalate and other aromatic compounds by a denitrifying bacterium. J. Bacteriol. **170:**5778-5784.

88. **Ochman, H., A. S. Gerber, and D. L. Hartl.** 1988. Genetic applications of an inverse polymerase chain reaction. Genetics **120:**621-623.

89. **Parkhill, J., M. Sebaihia, A. Preston, L. D. Murphy, N. Thomson, D. E. Harris, M. T. Holden, C. M. Churcher, S. D. Bentley, K. L. Mungall, A. M. Cerdeno-Tarraga, L. Temple, K. James, B. Harris, M. A. Quail, M. Achtman, R. Atkin, S. Baker, D. Basham, N. Bason, I. Cherevach, T. Chillingworth, M. Collins, A. Cronin, P. Davis, J. Doggett, T. Feltwell, A. Goble, N. Hamlin, H. Hauser, S. Holroyd, K. Jagels, S. Leather, S. Moule, H. Norberczak, S. O'Neil, D. Ormond, C. Price, E. Rabbinowitsch, S. Rutter, M. Sanders, D. Saunders, K. Seeger, S. Sharp, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, L. Unwin, S. Whitehead, B. G. Barrell, and D. J. Maskell.** 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. Nat. Genet. **35:**32-40.

90. **Patrauchan, M. A., C. Florizone, M. Dosanjh, W. W. Mohn, J. Davies, and L. D. Eltis.** 2005. Catabolism of benzoate and phthalate in *Rhodococcus* sp. strain RHA1: redundancies and convergence. J. Bacteriol. **187:**4050-4063.

91. **Paulsen, I. T., C. M. Press, J. Ravel, D. Y. Kobayashi, G. S. Myers, D. V. Mavrodi, R. T. DeBoy, R. Seshadri, Q. Ren, R. Madupu, R. J. Dodson, A. S. Durkin, L. M. Brinkac, S. C. Daugherty, S. A. Sullivan, M. J. Rosovitz, M. L. Gwinn, L. Zhou, D. J. Schneider, S. W. Cartinhour, W. C. Nelson, J. Weidman, K. Watkins, K. Tran, H. Khouri, E. A. Pierson, L. S. Pierson, 3rd, L. S. Thomashow, and J. E. Loper.** 2005. Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. Nat. Biotechnol. **23:**873-878.

92. **Pavel, E. G., L. J. Martins, W. R. Ellis, Jr., and E. I. Solomon.** 1994. Magnetic circular dichroism studies of exogenous ligand and substrate

binding to the non-heme ferrous active site in phthalate dioxygenase. Chem. Biol. **1:**173-183.

93. **Penner-Hahn, J. E.** 1989. X-ray absorption spectroscopy of *Pseudomonas cepacia* phthalate dioxygenase. Basic Life Sci. **51:**177-186.

94. **Pohlmann, A., W. F. Fricke, F. Reinecke, B. Kusian, H. Liesegang, R. Cramm, T. Eitinger, C. Ewering, M. Potter, E. Schwartz, A. Strittmatter, I. Voss, G. Gottschalk, A. Steinbuchel, B. Friedrich, and B. Bowien.** 2006. Genome sequence of the bioplastic-producing "Knallgas" bacterium *Ralstonia eutropha* H16. Nat. Biotechnol. **24:**1257-1262.

95. **Pridmore, R. D.** 1987. New and versatile cloning vectors with kanamycin-resistance marker. Gene **56:**309-312.

96. **Qiu, Y. L., Y. Sekiguchi, S. Hanada, H. Imachi, I. C. Tseng, S. S. Cheng, A. Ohashi, H. Harada, and Y. Kamagata.** 2006. *Pelotomaculum terephthalicum* sp. nov. and *Pelotomaculum isophthalicum* sp. nov.: two anaerobic bacteria that degrade phthalate isomers in syntrophic association with hydrogenotrophic methanogens. Arch. Microbiol. **185:**172-182.

97. **Qiu, Y. L., Y. Sekiguchi, H. Imachi, Y. Kamagata, I. C. Tseng, S. S. Cheng, A. Ohashi, and H. Harada.** 2004. Identification and isolation of anaerobic, syntrophic phthalate isomer-degrading microbes from methanogenic sludges treating wastewater from terephthalate manufacturing. Appl. Environ. Microbiol. **70:**1617-1626.

98. **Rainey, F. A., E. Lang, and E. Stackebrandt.** 1994. The phylogenetic structure of the genus *Acinetobacter*. FEMS Microbiol. Lett. **124:**349-353.

99. **Saeger, V. W., and E. S. Tucker.** 1976. Biodegradation of phthalic acid esters in river water and activated sludge. Appl. Environ. Microbiol. **31:**29-34.

100. **Saint, C. P., and P. Romas.** 1996. 4-Methylphthalate catabolism in *Burkholderia* (*Pseudomonas*) *cepacia* Pc701: a gene encoding a phthalate-specific permease forms part of a novel gene cluster. Microbiology **142 (Pt 9):**2407-2418.

101. **Saito, A., T. Iwabuchi, and S. Harayama.** 2000. A novel phenanthrene dioxygenase from *Nocardioides* sp. Strain KP7: expression in *Escherichia coli.* J. Bacteriol. **182:**2134-2141.

102. **Sasoh, M., E. Masai, S. Ishibashi, H. Hara, N. Kamimura, K. Miyauchi, and M. Fukuda.** 2006. Characterization of the terephthalate degradation genes of *Comamonas* sp. strain E6. Appl. Environ. Microbiol. **72:**1825-1832.

103. **Schettler, T.** 2006. Human exposure to phthalates via consumer products. Int. J. Androl. **29:**134-139; discussion 181-185.

104. **Schlafli, H. R., M. A. Weiss, T. Leisinger, and A. M. Cook.** 1994. Terephthalate 1,2-dioxygenase system from *Comamonas testosteroni* T-2: purification and some properties of the oxygenase component. J. Bacteriol. **176:**6644-6652.

105. **Scott, H. N., P. D. Laible, and D. K. Hanson.** 2003. Sequences of versatile broad-host-range vectors of the RK2 family. Plasmid **50:**74-79.

106. **Shigematsu, T., K. Yumihara, Y. Ueda, S. Morimura, and K. Kida.** 2003. Purification and gene cloning of the oxygenase component of the terephthalate 1,2-dioxygenase system from *Delftia tsuruhatensis* strain T7. FEMS Microbiol. Lett. **220:**255-260.

107. **Shigematsu, T., K. Yumihara, Y. Ueda, M. Numaguchi, S. Morimura, and K. Kida.** 2003. *Delftia tsuruhatensis* sp. nov., a terephthalate-assimilating bacterium isolated from activated sludge. Int. J. Syst. Evol. Microbiol. **53:**1479-1483.

108. **Shintani, M., M. Urata, K. Inoue, K. Eto, H. Habe, T. Omori, H. Yamane, and H. Nojiri.** 2007. The *Sphingomonas* plasmid pCAR3 is involved in complete mineralization of carbazole. J. Bacteriol. **189:**2007-2020.

109. **Sonoki, T., T. Obi, S. Kubota, M. Higashi, E. Masai, and Y. Katayama.** 2000. Coexistence of two different O demethylation systems in lignin metabolism by *Sphingomonas paucimobilis* SYK-6: cloning and sequencing of the lignin biphenyl-specific O-demethylase (LigX) gene. Appl. Environ. Microbiol. **66:**2125-2132.

110. **Stingley, R. L., B. Brezna, A. A. Khan, and C. E. Cerniglia.** 2004. Novel organization of genes in a phthalate degradation operon of *Mycobacterium vanbaalenii* PYR-1. Microbiology **150:**3749-3761.

111. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24:**1596-1599.

112. **Tarasev, M., and D. P. Ballou.** 2005. Chemistry of the catalytic conversion of phthalate into its *cis*-dihydrodiol during the reaction of oxygen with the reduced form of phthalate dioxygenase. Biochemistry **44:**6197-6207.

113. **Tarasev, M., A. Pinto, D. Kim, S. J. Elliott, and D. P. Ballou.** 2006. The "bridging" aspartate 178 in phthalate dioxygenase facilitates interactions between the Rieske center and the iron(II)-mononuclear center. Biochemistry **45:**10208-10216.

114. **Tarasev, M., F. Rhames, and D. P. Ballou.** 2004. Rates of the phthalate dioxygenase reaction with oxygen are dramatically increased by interactions with phthalate and phthalate oxygenase reductase. Biochemistry **43:**12799-12808.

115. **Tatusova, T. A., and T. L. Madden.** 1999. BLAST 2 S, a new tool for comparing protein and nucleotide sequences. FEMS Microbiol. Lett. **174:**247-250.

116. **Tierney, D. L., G. T. Gassner, C. Luchinat, I. Bertini, D. P. Ballou, and J. E. Penner-Hahn.** 1999. NMR characterization of substrate binding in the phthalate dioxygenase system. Biochemistry **38:**11051-11061.

117. **Tropel, D., and J. R. van der Meer.** 2004. Bacterial transcriptional regulators for degradation pathways of aromatic compounds. Microbiol. Mol. Biol. Rev. **68:**474-500.

118. **Tsang, H. T., C. J. Batie, D. P. Ballou, and J. E. Penner-Hahn.** 1989. X-ray absorption spectroscopy of the [2Fe-2S] Rieske cluster in *Pseudomonas cepacia* phthalate dioxygenase. Determination of core dimensions and iron ligation. Biochemistry **28:**7233-7240.

119. **Tyler, S. D., K. R. Rozee, and W. M. Johnson.** 1996. Identification of IS1356, a new insertion sequence, and its association with IS402 in epidemic strains of *Burkholderia cepacia* infecting cystic fibrosis patients. J. Clin. Microbiol. **34:**1610-1616.

120. **van Herwijnen, R., D. Springael, P. Slot, H. A. Govers, and J. R. Parsons.** 2003. Degradation of anthracene by *Mycobacterium* sp. strain LB501T proceeds via a novel pathway, through *o*-phthalic acid. Appl. Environ. Microbiol. **69:**186-190.

121. **Vila, J., Z. Lopez, J. Sabate, C. Minguillon, A. M. Solanas, and M. Grifoll.** 2001. Identification of a novel metabolite in the degradation of pyrene by *Mycobacterium* sp. strain AP1: actions of the isolate on two- and three-ring polycyclic aromatic hydrocarbons. Appl. Environ. Microbiol. **67:**5497-5505.

122. **Wang, Y., M. Rawlings, D. T. Gibson, D. Labbé, H. Bergeron, R. Brousseau, and P. C. K. Lau.** 1995. Identification of a membrane protein and a truncated LysR-type regulator associated with the toluene degradation pathway in *Pseudomonas putida* F1. Mol. Gen. Genetics **V246:**570-579.

123. **Wang, Y. Z., Y. Zhou, and G. J. Zylstra.** 1995. Molecular analysis of isophthalate and terephthalate degradation by *Comamonas testosteroni* YZW-D. Environ. Health Perspect. **103:**9-12.

124. **Wauters, G., T. De Baere, A. Willems, E. Falsen, and M. Vaneechoutte.** 2003. Description of *Comamonas aquatica* comb. nov. and *Comamonas kerstersii* sp. nov. for two subgroups of *Comamonas terrigena* and emended description of *Comamonas terrigena*. Int. J. Syst. Evol. Microbiol. **53:**859-862.

125. **Wen, A., M. Fegan, C. Hayward, S. Chakraborty, and L. I. Sly.** 1999. Phylogenetic relationships among members of the *Comamonadaceae*, and description of *Delftia acidovorans* (den Dooren de Jong 1926 and Tamaoka et al. 1987) gen. nov., comb. nov. Int. J. Syst. Bacteriol. **49:**567-576.

126. **Wu, J. H., W. T. Liu, I. C. Tseng, and S. S. Cheng.** 2001. Characterization of microbial consortia in a terephthalate-degrading anaerobic granular sludge system. Microbiology **147:**373-382.

127. **Zylstra, G., and D. Gibson.** 1989. Toluene degradation by *Pseudomonas putida* F1. Nucleotide sequence of the *todC1C2BADE* genes and their expression in *Escherichia coli*. J. Biol. Chem. **264:**14940-14946.

Curriculum Vita

Education

2002 - 2008   Rutgers, The State University of New Jersey

Joint Program of Molecular Biosciences at Rutgers and UMDNJ

Biotech center for Agriculture and the Environment

Department of Biochemistry and Microbiology

Ph.D. in Microbiology and Molecular Genetics

1996 - 1999   Zhejiang University, China

M.S. in Environmental Microbiology

1989 - 1993   Zhejiang Agricultural University, China

B.S. in Plant Protection

Awards

2005 - 2008   Graduate Assistantship

2006   Ivarson Student Assistant Fund

2006   ASM Student Travel Grant

2004 - 2005   Teaching Assistantship, Microbiology

2003 - 2004   Graduate Assistantship

2002 - 2003   H. Boyd and Jeanette I. Woodruff Microbiology Fellowship