GENETIC AND FUNCTIONAL CHARACTERIZATION OF BIPHENYL, DIPHENYLMETHANE, AND DIPHENYLETHER DEGRADATION

By

AAKANSHA JESSICA ROBERTS

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Microbial biology

Written under the direction of

Gerben Zylstra

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

January, 2020

ABSTRACT OF THE DISSERTATION

Genetic and Functional Characterization of Biphenyl,

Diphenylmethane, and Diphenylether Degradation

By AAKANSHA JESSICA ROBERTS

Dissertation Director:

Gerben J. Zylstra

Diphenylmethane, biphenyl, and diphenylether are common contaminants in a former industrial site along the Passaic River with similar structures that are of interest due to their toxicity. This work identifies strains involved in their degradation, and the genes involved in the pathways for their degradation. Based on 16s rRNA analysis three phylogenetically diverse *Pseudomonas* species were isolated from the contaminated site based on their ability to grow on diphenylmethane. They were named *Pseudomonas* sp. AJR09, *Pseudomonas stutzeri* AJR13, and *Pseudomonas* sp. AJR20. The strains were later tested on biphenyl and were found to be able to metabolize it using the same pathway as for diphenylmethane degradation. All three strains were found to contain identical dioxygenase gene sequences. The presence of the identical genes in three diverse species led us to the conclusion that the genes must be horizontally transferred in the environment. Based on the phenotype conferred by the element and the fact that one strain had previously lost the ability to grow on the substrates tested, we hypothesized that the element was an Integrative and Conjugative Element. Integrative and Conjugative Elements (ICE) are a family of mobile genetic elements that can be

transferred between different cells/organisms, and once in the recipient they integrate into the host's chromosome using specific recombination sites. One of the three strains, the *P. stutzeri* AJR13, was successfully mated with the well characterized *P. putida* KT2440 which subsequently gained the ability to grow on biphenyl, diphenylmethane, and salicylate. This demonstrated the ability of the genes to self-mobilize leading us to believe that the genes for degradation of the three compounds must be horizontally transferred in the environment. The whole genomes of the three Passaic River strains and the KT2440 recipient were sequenced and assembled to reveal that the degradative genes are indeed present on an ICE. The ICE is about 128 kb in length and inserts at a short sequence at the end of a tRNA Gly (CCC). It contains an integrase and other genes involved in the transfer of the ICE, and genes for diphenylmethane/biphenyl and salicylate degradation. It also contains a number of repeated sequences. We noticed that mutations in the ICE were necessary in order to demonstrate growth similar to the wildtype. Our work demonstrates that integrative and conjugative elements play a large role in the spread of biodegradative genes in the environment. Another aim was to characterize a dioxygenase gene involved in the degradation of diphenylether. Only a few bacteria have been isolated for growth on diphenylether (DPE) as the sole carbon and energy source. *Sphingobium* sp. strain SS3 is perhaps the best studied diphenylether degrading strain with characterization of the catabolic pathway in the 1990s by investigators at the University of Hamburg. The DPE catabolic pathway is initiated by a dioxygenase attack resulting in the formation of catechol and phenol. We have sequenced the SS3 genome and identified 13 genes encoding Rieske type oxygenases. Analysis of the genome environments and comparison to related enzymes in the database identified

promising candidate genes that could possibly be involved in the pathway for DPE degradation. We carried out an RT-PCR experiment to identify which of the dioxygenase genes are expressed when SS3 is growing in the presence of DPE. A gene encoding a putative benzoate dioxygenase was upregulated most likely because the gene was situated in an operon encoding the catechol branch of the beta-ketoadipate pathway. A gene encoding a putative DPE dioxygenase was also upregulated. Gene knockout experiments and heterologous expression of the angular dioxygenase confirmed its catalytic activity. This adds to our knowledge of angular dioxygenase and helps us genetically characterize only the second diphenylether degrading strain.

ACKNOWLEDGEMENTS

This work would not have been possible without the guidance of my advisor Dr. Gerben J. Zylstra. I am grateful for his advice and availability whenever I needed assistance, both research-related and otherwise, especially when I was new to the country. I appreciate the opportunity to work in and complete my Ph.D. in his lab. I am also grateful for his support in my endeavor to train myself in bioinformatics skills by allowing me to attend multiple workshops and conferences.

I would like to thank my committee members Dr. Tamar Barkay, Dr. Jerome Kukor and Dr. Donna Fennell for accepting the invitation to be on my Thesis committee. I would also like to thank Dr. Lee Kerkhof who as a collaborator provided me with guidance on the sequencing component of the project. I am grateful to the Department of Biochemistry and Microbiology and the Microbial Biology Graduate Program for the Teaching assistantship, awards and financial support that have allowed me to attend multiple conferences leading to my scientific and professional development.

I am extremely grateful to the past and present members of the Zylstra Lab especially Hung-Kuang, Rayan, Suha and Thamer for the multiple discussions about science and otherwise that made the lab environment enjoyable and helped me become a better scientist.

I am also fortunate to have met some amazing people along the way, some friends that stayed others who did not but nonetheless played their part in my journey. I am grateful to Nicole, Spencer, Fatima, Ashley, and Ann for the part you played in helping me continue on through the years. And most importantly Jonathan for being by my side and helping me over the finish line. Your support and encouragement made this experience easier, more enjoyable and absolutely worth it.

And last but not the least this work would not have been possible without the continued love, support and confidence of my parents Annie and Clarence, my brother Yohann, and the rest of my family especially my grandmother Jessie. You helped me believe in myself at all times.

DEDICATION

To my family for their constant love, support, encouragement and for simply being awesome.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Hazardous compounds in nature

Biphenyl, dibenzo-*p*-dioxin, and dibenzofuran, and their chlorinated forms have long been known to cause contamination in the environment. These compounds are lipophilic in nature, chemically similar in structure and have similar physicochemical behaviors and toxicology effects (1). Rivers and other water bodies were often used to release industrial waste until the 70s where the toxic industrial pollutants and contaminants would interact with aquatic ecosystems often leading to bioaccumulation and biomagnification (2). The Passaic River is one such contaminated river. The site of interest on the banks of the river is the site of the Diamond Alkali site, also known as the Lister Avenue site in Newark, New Jersey, USA, located 3.1 miles upstream of where the Passaic meets Newark Bay (3). This was also the former site of Diamond Shamrock Chemicals Company and its predecessors between 1948 and 1969 that manufactured pesticides and herbicides including one of the main components of Agent Orange, a military defoliant. Due to the presence of waste from these industries in the river, the site was added to the Superfund National Priorities List in September 1984 on account of the level of 2,3,7,8-tetrachlorodibenzodioxin contamination (3). A recent study examined the levels of polybrominated diphenylethers, polychlorinated biphenyls and polychlorinated dibenzo-*p*-dioxins/furans in the sediment, porewater, river water and organisms like blue crab and different finfish species from the lower Passaic River to reveal bioaccumulation of the toxic compounds in higher organisms (4). By studying the degradation of these related compounds, we aim to learn more about the genes involved in the pathways for

degradation. The three primary compounds studied in this thesis are biphenyl, diphenylmethane, and diphenylether (Figure 1). All three consist of two aromatic rings linked together either directly or by a methylene bridge or by an ether bridge. The general pathways for degradation are similar in these types of compounds and normally involves a dioxygenase enzyme.

Relevance of compounds and their toxicity

While dioxin/dibenzofuran related compounds are generally produced as useless byproducts in the process of waste combustion, petroleum refining, metallurgy production, and synthesis of some chlorinated chemicals (1), the compounds being studied had a number of purposes. Polychlorinated biphenyl compounds were produced as hydraulic fluids, plasticizers, dyes, and adhesives (5). They were also previously used as transformer oil and paint ingredients (1). These products were produced widely until the 1970s and 1980s when they were widely banned as they were found to accumulate in the environment and bioaccumulate in the food chain (6). Some studies have shown that polychlorinated biphenyls are involved in metabolic diseases like diabetes, and endocrine disruption particularity the thyroid and reproductive organs (6). Biphenyl has often been used as a mineralizable analog for polychlorinated biphenyl degradation studies (7) and we will be doing the same in this study.

Diphenylmethane on the other hand is naturally found as a component of crude oil (8). It is a major component of polyurethane in the form of diphenyl methane diisocyanate, and a central building block of the disinfectant hexachlorophene, and a

number of anti-allergenic agents (9). It also has a structure similar to the insecticide DDT that is known to persist in the environment (10).

Diphenylether in the polybrominated form has previously been used as a flame retardant in a number of consumer products like electronics, textiles, and foam however this has been greatly reduced since it was determined that this compound also bioaccumulates and is toxic (11). There are also a number of diphenyl ether based herbicides. The diarylether structure is also an important part of the structure of hard coal (12). Diphenylether derivatives are also used as heat transfer liquids and perfume additives (13). It has been linked to endocrine disruption, reproductive/developmental toxicity including neurotoxicity, and cancer in mammals (14).

The above-mentioned dioxin like compounds are toxic to vertebrates and this toxic effect is mediated primarily via the aryl hydrocarbon receptor (AHR, or "dioxin receptor"). This is a transcription factor that when activated by the appropriate ligand works with the aryl hydrocarbon receptor nuclear translocator (ARNT) to alter the expression of its target genes (15).

Degradation of Hazardous Aromatic Compounds

Polycyclic aromatic compounds are ubiquitous in nature due to the production of numerous natural flavonoids, dibenzofurans, cresols, xanthones and lignin by plants (16). Thus, bacteria have evolved to carry out bioremediation over a long time. Bioremediation involves either the biotransformation of toxic compounds into harmless metabolites or their mineralization into carbon dioxide and water (17). Aromatic compounds are usually reduced compounds that can be made more susceptible to degradation by increasing the

oxidation state of the aromatic nucleus by the incorporation of molecular oxygen; this enables microorganisms to utilize them as a sole source of carbon and energy (16). Molecular oxygen being chemically inert requires the help of an oxygenase, a type of oxidoreductase for insertion into the aromatic ring (16) they also usually require co-substrates like FAD, NADH or NADPH for electron transport (18).

Oxygenase Enzymes

Oxygenase enzymes consist of two types of enzymes, monooxygenases and dioxygenases, based on the number of oxygen atoms used during the oxygenation step (19). Monooxygenases, also known as hydroxylases or mixed-function oxidases, incorporate one atom of molecular oxygen into the aromatic ring to form a monohydroxy compound (16). Monooxygenases are further classified into two subclasses on the basis of the cofactors used: a flavin dependent monooxygenase or a P450 monooxygenase (20). The former contains a flavin prosthetic group requiring NADH or NADPH as a coenzyme (21). The latter are heme-containing and are found in prokaryotes and eukaryotes. Some monooxygenases have also been described that are pterin or metal ion dependent while others do not require a cofactor (20).

Dioxygenases on the other hand can be classified as either ring hydroxylating to yield a dihydroxylated product or ring cleaving for aromatic diols effectively breaking the aromaticity of the nucleus (16). In order for ring cleavage to occur via the latter, double hydroxylation of the aromatic ring by the former is a prerequisite (22). Some well characterized ring hydroxylating dioxygenase are the biphenyl (23) and naphthalene (24) dioxygenases, that initiate degradation in the respective compounds through the

formation of a dihydrodiol intermediate (16, 25). These enzymes belong to the Rieske non-heme iron oxygenase superfamily (22). In general, these dioxygenases consist of three components, two of which are involved in electron transfer, in the case of naphthalene electrons are derived from NAD(P)H by the iron-sulfur flavoprotein reductase and transferred to a Rieske-type iron sulphur centre in a ferredoxin protein. They are then transferred to the catalytic mononuclear iron of the oxygenase (26, 27).

On the other hand ring cleaving dioxygenase can be differentiated based on the location of the cleavage: they can be either intradiol, if cleavage occurs between the hydroxyl groups on the aromatic ring, also known as ortho cleavage or extradiol, if the cleavage occurs between a hydroxyl group and a non-hydroxylated carbon on the aromatic ring, this is known as meta cleavage (19). Extradiol dioxygenases are most commonly found to house a nonheme iron Fe(II) in their active site but others were found to also be active with Mn(II) (19). Fe(III) is most commonly found in intradiol dioxygenases (25). The Fe(II) of extradiol dioxygenases is usually inert to molecular oxygen until it binds to the aromatic substrate, this reduces it and opens the site for the binding of molecular oxygen consequently transferring electron density to the oxygen, activating it (28).

The catalytic component of the ring oxidizing dioxygenases has a $\alpha 3\beta 3$ configuration with a mushroom-like quaternary structure. The α subunit is divided into the Rieske [2Fe-2S] cluster domain and the mononuclear iron domain (27). The β subunit on the other hand does not have any active sites and is assumed to be purely structural (27). For individual α subunits the distance between the mononuclear iron and Reiske cluster is too far apart for electron transfer (~45 Å) however this distance is reduced to

~12Å between the preceding subunits mononuclear iron and the succeeding subunits Rieske centers to permit electron transfer (27).

While multiple classification systems have been proposed based on the electron transport components (29) and the evolution of the enzymes (30), the easiest way to classify them is based on their native substrate and the amino acid sequence of their α subunit. They can thus be classified as belonging to the naphthalene, toluene/benzene, biphenyl, and benzoate/toluate dioxygenase classes (22).

The ability of an organism to degrade the substrates depends on the substrate specificity of the upper pathway enzymes, the induction of the upper pathway by the substrates, the resistance to the toxic effects of the substrates, and the ability to degrade the end products of the upper pathway. Degradation metabolic pathways are usually divided into an upper, lower and central carbon pathway. The upper pathway is usually induced by a specific parent substrate, the lower pathway is induced by intermediates of the upper pathway, and the central carbon metabolism is usually constitutive since they are housekeeping pathways (16).

Common Upper Pathway for Degradation

The upper pathway is the same for both biphenyl and diphenylmethane however the end products of the upper pathway (Figure 2) are benzoate and phenylacetate respectively (9) as well as 2-hydroxypenta-2,4-dienoic acid (31); this difference in intermediates leads to the utilization of a different lower catabolic pathway.

The biphenyl catabolic pathway like most aromatic dioxygenase systems consists of four enzymatic steps that lead to the conversion of biphenyl (or diphenylmethane) to

benzoate (or phenylacetate). First a multi-subunit biphenyl dioxygenase prepares the aromatic ring for attack by adding oxygen to it. Next a dihydrodiol dehydrogenase rearomatizes the ring. This is followed by a 2,3-dihydroxybiphenyl dioxygenase, which cleaves the ring at the 1,2 position (32) and has a ferrous iron as a prosthetic group (33). Finally, a 2,4-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid hydrolase (HOPDA hydrolase) hydrolyses the ring cleavage product producing benzoate and a 5-carbon fragment (34-36).

The three components that make up the dioxygenase enzyme are a Rieske-type dioxygenase which forms the catalytic unit (Figure 3). It is a heterohexamer that consists of the large alpha and the small beta subunit, three molecules of each make up the respective subunits (37). The C-terminal portion of the alpha subunit contains the [Fe_2-S_2] Rieske cluster and the mononuclear iron containing catalytic center (38). This is the major determinant of structural specificity. It was found that random mutagenesis of the biphenyl oxygenase C-terminal end resulted in changes in the substrate specificity and activity of the dioxygenase (39, 40). The other two components are a ferredoxin and a ferredoxin reductase. These are responsible for the transfer of electrons from the electron donor NADH to the Rieske center and then to the mononuclear iron catalytic unit (9, 41).

How are they transferred? Plasmids vs. transposons vs. ICEs

In the environment pathways for degradation of aromatics are constantly evolving to give rise to new catabolic pathways due to the acquisition of genes via horizontal transfer. The common elements responsible for this are conjugative plasmids, integrative plasmids, conjugative transposons, and integrative and conjugative elements. The

elements are diverse existing in the linear or circular form or integrated into the chromosome and can vary in size between a few kilobases to above 500 kb (42). Some well characterized catabolic plasmids are the toluene degrading TOL plasmid from *Pseudomonas putida* strain mt-2 that contains catabolic, transfer and regulatory genes (42), the OCT plasmid carries genes for *n*-octane degradation in *Pseudomonas oleovorans* (43) and the pDTG1 plasmid from *Pseudomonas putida* strain NCIB 9816-4 carries genes for naphthalene degradation through salicylate (44). Some biphenyl gene carrying plasmids have also be studied like the pSK4 plasmid in *Cupriavidus* sp. strain SK-4 (45) and the linear plasmid in *Rhodococcus erythropolis* TA421 (46).

There are also a number of well characterized catabolic transposons. Tn*4653* contains the genes for toluene degradation and genes involved in the transposition of the element. Tn*4651* is another toluene gene containing element (47). Tn*5271* carries the genes for 3- and 4-chlorobenzoate degradation and lies either in the plasmid or chromosome of *Alcaligenes* sp. strain BR60 (48). Tn*4371* is a well characterized biphenyl and 4-chlorobiphenyl catabolic gene containing transposon found in *Cupriavidus oxalaticus* (previously *Ralstonia eutropha*) A5 (49).

What are ICEs?

Integrative and conjugative elements (ICEs), also known as conjugative transposons, are a large family of mobile genetic elements. They are usually found integrated into the host's chromosome and replicate with it, and are thus inherited and transferred vertically. Certain conditions may cause them to excise, circularize into a circular covalently closed molecule, and transfer to a recipient’s chromosome (50). ICEs

are of interest as they combine the features of plasmids, transposons, and bacteriophages. They can integrate into and excise from the chromosome like a transposon, while using conjugation to be transferred to the recipient's chromosome like a plasmid, and they contain a recombination system similar to those seen in bacteriophages (Figure 4) (51).

In a study scanning a thousand genomes it was found that ICEs were the most abundant conjugative elements in prokaryotes (52). The first ICE (previously called a conjugative transposon) Tn*916* was discovered in a Gram positive organism (53) however it is suggested that ICEs first arose in Gram negative bacteria (diderms), specifically in Proteobacteria and then spread to other monoderms and archaea. This includes all the ICEs that are transferred as ssDNA. The proposal is based on the simplicity of the transfer systems present in monoderms. The requirement for fewer genes indicates that it could have arisen by deletion from a more complex gene machinery required in diderms (54).

ICEs are usually mosaic or modular in structure. They usually contain genes of similar function grouped together in the genome. These functional modules are often acquired separately (55).

Common components/genes of an ICEs

ICEs are known to consist of four essential functional modules or groups of genes involved in important functions: the integration/excision module, the replication/DNA processing module, the DNA secretion module, and the regulation module (50). The integration excision module is involved with intracellular mobility housing the *att* site and the integrase. The replication module consists of the *oriT* and the genes for the

cognate relaxase and the Type IV Coupling Protein involved in the ssDNA transfer. The DNA secretion module consists of the genes for the Type IV secretion system involved in the conjugative transfer, and finally the regulatory module, if present, can modulate the activity of other modules (50). In addition to these different ICEs contain a variable region that can confer the host with some selective advantage in its environment.

Classes/types of ICEs

There is no unified system of classification for ICEs nonetheless a number of classification systems have been proposed; however, the sheer diversity in size, insertion site, mechanism for mobility and transfer, and the cargo genes make a single system difficult. One classification proposes separation based on the conjugative systems present i.e. the relaxase (e.g. TraI), type IV secretion system proteins (e.g. VirB4) and the type IV coupling protein (e.g. TraG) (54). Another system has categories based on the integrase gene and the site for insertion (53). Some older systems for not only ICEs but a larger group of genomic islands classify them based on the selective genes being carried on the element.

Cargo Genes

ICEs vary widely in their size from the 11 kb pSAM2 in *Streptomyces ambofaciens* to the 674 kb PAISt in *Streptomyces turgidiscabies* (56). This size is primarily determined by the number of cargo genes being carried. The cargo genes can be highly variable even in ICEs having similar conjugation and recombination modules. These genes are not involved in the mobility of the ICE but confer the host with a

selective advantage. They may even lead to a lifestyle change such as resistance to antibiotics, heavy metals or phages, sucrose catabolism, bacteriocin synthesis, pathogenicity, or symbiosis (56).

The ICE Tn*916* initially found in *Enterococcus* (*Streptococcus*) *faecalis* was found to carry the genes for tetracycline resistance (57) and the SXT ICE in *Vibrio chlorae* O139 carries genes for resistance to multiple antibiotics (58). Other ICEs have been found to carry genes for carbon source utilization like the ICE*clc*^{B13} in *Pseudomonas knackmussii* B13 involved in 3-chlorobenzoate degradation (59), the *bph-sal* element in *Pseudomonas putida* KF715 involved in biphenyl and salicylate degradation (31), and the CTn*Scr94* in *Salmonella senftenberg* 5494-57 carrying the genes for sucrose fermentation (60). The ICE*Ml*Sym^{R7A} ICE carries the genes for nodule formation and nitrogen fixation in *Mesorhizobium loti* R7A (61). The ICE PAPI-1 in *Pseudomonas aeruginosa* PA14 is involved in both virulence and biofilm formation both important medically (62).

Degradation related/dioxygenase gene carrying ICEs

There is a great amount of divergence among the biphenyl catabolic (*bph*) genes in nature, pointing to the fact that these are probably very ancient genes and have thus evolved over the years (31). A number of ICEs carrying dioxygenase degradation pathway genes have been characterized. ICE*clc* is probably the most well-known ICE from *Pseudomonas knackmussii* B13 that contains the genes for chlorocatechol degradation (63). In ICE*clc*^{B13} the site of integration is an 18 bp sequence present at the 3'-end of the glycine tRNA gene (*attB*) and is identical to a region on the excised ICE itself. A gene 202 bp downstream of the integration site, a P4-family tyrosine

recombinase-type IntB13 integrase is responsible for mediating the integration. This process occurs via a staggered-cut followed by strand-swapping similar to that seen in a phage lambda system (59). ICE*clc*^{B13} has two promoters for the integrase *intB13* in different conformations, a weak promoter P_{int} is present upstream of the integrase in the chromosomal integrated form and a strong promoter P_{circ} is present in the circular form. The P_{int} is active in only about 3% of the cells in the stationary phase, especially when grown on 3-chlorobenzoate or fructose. However, in the circular excised form P_{circ} is constitutively active. This promoter faces outwards in the integrated form however on integration by recombination between the *attL* and *attR* sites, it changes its orientation to lie upstream of the P_{int}, thus voiding its activity. This constitutive expression thus ensures the integration of the ICE into the new host chromosome before the P_{circ} goes back to its previous orientation (64, 65). This mechanism thus decreases the chances of excision and increases the chances of integration when in the circular form, after entry into the host, respectively.

Another well characterized ICE was found to be involved in both biphenyl and salicylate catabolism in *Pseudomonas putida* KF715. The genes for salicylate catabolism converted salicylate to 2-hydroxymuconic semialdehyde. *P. putida* KF715 was observed to readily lose the 90 kilobase element, thus losing the ability to grow on both substrates, when grown in nutrient rich medium. The strain was also found to be able to transfer the large element to *P. putida* AC30 via filter mating (31). Another strain *Pseudomonas furukawaii* KF707 isolated from the same site was also found to contain a biphenyl and salicylate pathway containing ICE (66).

The ICE in *Acidovorax* sp. strain KKS102 called ICE$_{KKS102}$*4677* carries the polychlorinated biphenyl (PCB)/biphenyl degradation gene cluster. The ICE is ~62 kb and inserts into tRNA genes. In experiments to characterize the transfer of the ICE to three different recipient strains the frequencies for obtaining transconjugants was low, 5.8 $x10^{-10}$ on average i.e. about 1 colony after 4 days of incubation from about 2 x 10^9 donor cells (67).

A possible ICE is present in *Achromobacter* sp. BP3 that carries the biphenyl degradation genes. It contains an integrase with 98% similarity to the integrase of the *Pseudomonas putida* KF715 bph-sal ICE, and transposon related genes (68).

The integrase/recombinase involved in its transfer

The integrase enzyme is an important enzyme in any ICE as it is involved with the excision and recombination of the entire element. The process of recombination does not require any external energy and involves strand breakage, exchange, and rejoining (69). Integrases are site-specific recombinases and are divided into two recombinase superfamilies: tyrosine type recombinases and serine type recombinases (69). They are classified based on the amino acids in their active site. The tyrosine recombinase contains a 3'-phosphotyrosine that interacts with the cleaved DNA. It also causes recombination via the formation and resolution of a Holliday junction intermediate (70). The tyrosine type recombinases are further divided into two types: the complex unidirectional (tyrosine-type phage integrases) recombinase and the simple bidirectional recombinase (70). The tyrosine-type phage integrase requires the recognition of an attachment site on the bacterial chromosome called the *attB* site and a homologous region on the mobile

genetic element called the *attP* and gives rise to an *attL* and *attR* site on the left and right junction of the integrated fragment, respectively (71).

tRNA: the site for ICE insertion

A common site of insertion for a number of mobile genetic elements are tRNA genes. The same site of insertion is observed in a number of retroviruses (72) since the translational machinery associated is vital for viral replication. As the 3' end of the tRNAs are accessible, unlike other RNA molecules, they can be recognized and act as primers for their replication (73). This region is also transcriptionally active conferring the integrated element with an advantage. As previously mentioned, the integrase gene responsible for the recombination is similar to phage integrases which could explain the integration into different tRNA genes. A number of mobile genetic elements are known to integrate into a number of different tRNA genes. A 106 kb plasmid was found to integrate into a lysine tRNA gene in a *Pseudomonas aeruginosa* sp. (73), a number of plasmids into the leucine tRNA in *Haemophilus influenzae* (74), and the *bph-sal* ICE into the glycine tRNA in *Pseudomonas putida* KF715 (31) to name a few.

Mechanism of transfer: the Type IV secretion system and rolling circle replication for transfer by conjugation

A component deemed essential to ICEs in a number of studies are the conjugation machinery. The types of ICEs being described here utilize a type IV secretion system similar to those present in plasmids (55). The conjugative machinery consists of a

relaxase or integrase, a type IV coupling protein and a type IV secretion system, the only exceptions are for the ICEs found in *Actinobacteria* (75).

The type IV secretion system encodes a protein complex that forms a secretion channel spanning the donor's membranes which usually also includes an extracellular pilus. In Gram negative organisms this machinery spans both membranes, the periplasm and the cell wall (76). The genes are usually arranged in a single or a few operons. The *virB* operon usually consists of 11 genes (*virB1* to *virB11*) that encode the components of the Type IV secretion systems. The *virD* operon encodes 4 genes (*virD1* to *virD4*) that encode the relaxase and the type 4 coupling protein. The genes essential for building a Type IV secretion system consist of a type 4 coupling protein, the Type IV secretion system secretion channel consisting of energizing components i.e. VirB4 and VirB11, the transmembrane components responsible for the formation of translocation channels (VirB3, VirB6, and VirB8 in the inner membrane, and VirB7, VirB9, and VirB10 in the outer membrane), the hydrolase VirB1; and a pilus consisting of components VirB2 and VirB5 (77). The number and complexity of the conjugation system may vary based on the type of cell wall; whether a thicker cell wall in Gram positives or multiple membranes in Gram negatives. However the minimal conjugative system of *Tn916* is capable of transferring DNA across both (55).

In order to initiate the transfer of an ICE a relaxase binds covalently to and nicks the 5'-end of the DNA via a transesterification reaction at the origin of transfer (*oriT*) to create T-DNA. Also, a recombinase or integrase binds to the ICE bordering sequences. Factors then cause unwinding of the two strands. This nucleoprotein complex then interacts with the ICE encoded type 4 coupling protein (an ATPase) to target the mating

pore of the type IV secretion system, this enables transfer to the recipient cells. Once transferred the single stranded DNA is recircularized by the relaxase using the energy released while breaking the initial phosphodiester bond during strand nicking and is used as a template for DNA polymerase to reform the double stranded circular molecule by rolling circle replication (55, 78, 79).

Some ICEs use the same origin of replication and proteins for both conjugative transfer and rolling circle replication, while some have distinct origins and associated proteins (55). Studies of some ICEs demonstrate that they undergo autonomous rolling circle replication, this is of importance for the maintenance of the ICE in the population of cells (80). ICEs like ICE*Bs1* and ICE*Ml*Sym^{R7A} have a bifunctional *OriT* and DNA relaxase that enable the replication (80). However it is difficult to study due to the low rate of induction and excision for most ICEs (55).

Regulation

ICEs usually remain integrated in the host chromosome. The genes for conjugation are repressed under normal conditions as their constant expression can be detrimental to the cell (55). Certain mechanisms are involved in the control of the Type IV secretion system gene expression like feedback based on intracellular relaxase levels, cell envelope or environmental stress response effect on the expression, molecule copy number control and incompatibility similar to that found in plasmids, partitioning, or entry exclusion, and immunity systems like CRISPR and CRISPR associated genes, and transcription blocking by HN-S silencing (79, 81).

Under specific conditions the genes for excision and transfer are de-repressed to cause the ICE to excise from the chromosome; however this excision happens in a very small population of the cell population (55). Stress conditions like the *recA* mediated SOS response due to DNA damage in cells can lead to ICE activation. They can also be induced due to quorum sensing signaling the availability of potential mating partner recipients, or by the onset of the stationary phase of growth. Some others are induced by the phenotype they carry i.e. tetracycline induces the tetracycline resistance conferring Tn*916* ICE (78).

Once the ICE excises it must either reintegrate into the host chromosome or transfer to the recipient and integrate into the chromosome to prevent loss of the ICE from daughter cells and to ensure vertical inheritance (55). Thus, ICEs capable of autonomous replication have an advantage as they can decrease the odds of the ICE being lost in case the integration site is being replicated. Further multiple transfers of each copy to a recipient cell increases the frequency of transfer. Multiple ICE copies with an active partitioning system can also help ensure the repartitioning of the ICEs to the daughter cells appropriately (82). For all these reasons the number of ICE-free cells in a population is low. The rate of ICE excision from the host chromosome is very low to begin with (80) and thus the rate of transfer is also low, with only 10^{-2} to 10^{-7} transconjugants acquired per donor cell under laboratory conditions (81).

ICE containing organisms

The host range for ICEs can be very diverse. Most are known to transfer between related species if not within the same genus, while others like those belonging to the

Tn916 family of ICEs can transfer across a number of Gram positives and some Gram negatives like Proteobacteria (56). The well characterized degradation ICE *ICEclc*^{B13} is capable of transferring between betaproteobacteria and gammaproteobacterial in the environment (78). The limitations in host range are believed to arise from incompatibility of the different steps in the conjugation machinery and possibly the need for transient replication to form a double stranded molecule in the recipient host (56). Transfer of an ICE could also have other effects on the recipient. It is known that acquiring foreign DNA can temporarily affect host fitness. It has also been demonstrated for the ICE CTnDOT that the addition of genes on the ICE to the host *Bacteroides thetaiotaomicron* VPI-5482 can alter the expression of other genes, even those present on other mobile genetic elements (78).

This research aims to further our understanding about lateral dioxygenase genes by describing another variant involved in the degradation of both biphenyl and diphenylmethane in Chapter 2. We contribute to the knowledge about ICEs by the characterization of an Integrative and Conjugative Element that carries the genes involved in the above pathway in Chapter 3. And finally, we characterize a novel angular dioxygenase involved in the degradation of diphenylether in Chapter 4.

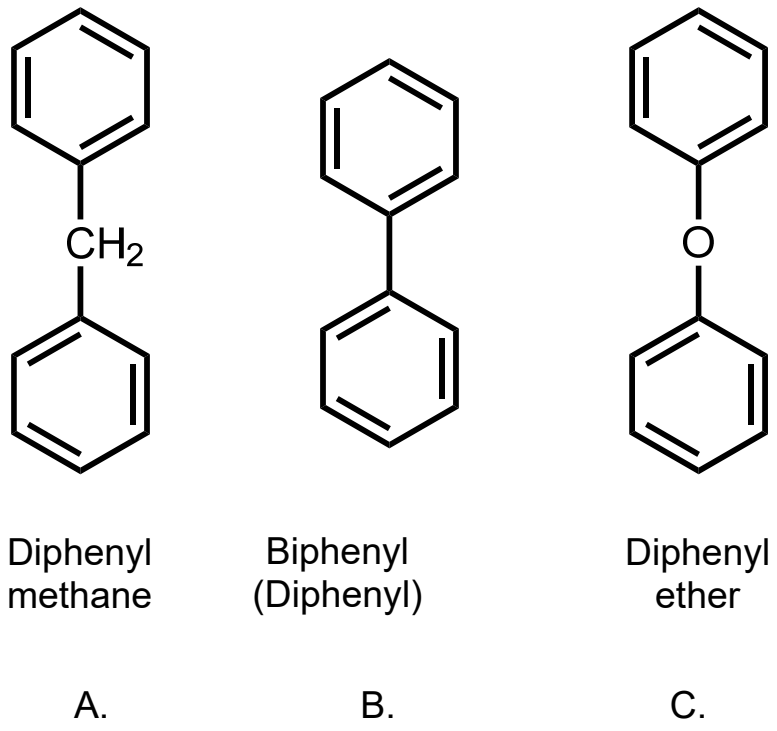

Figure 1. Bicyclic Compounds. (A) Diphenylmethane (B) Biphenyl (C) Diphenylether.

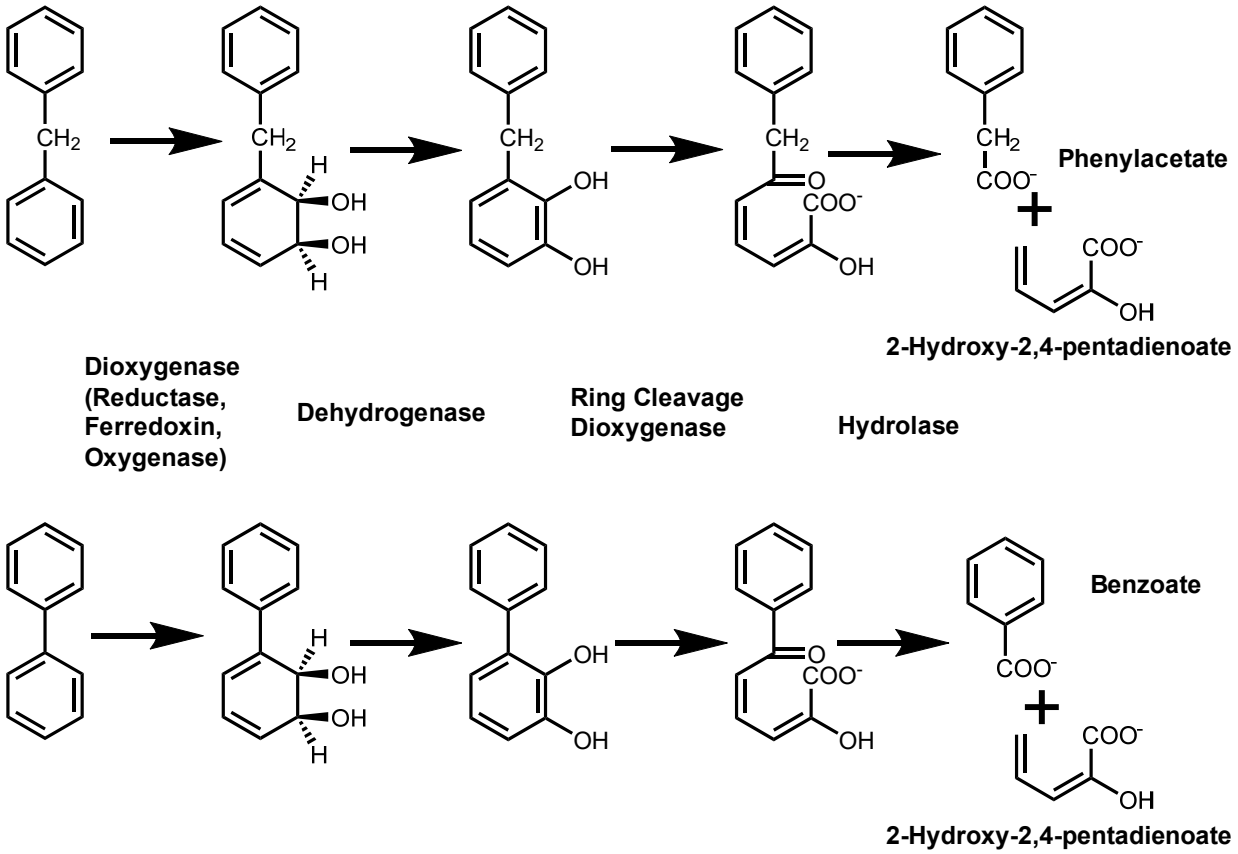

Figure 2. Pathway for degradation. The same enzymatic pathway is involved in the degradation of both diphenylmethane (top) and biphenyl (bottom).

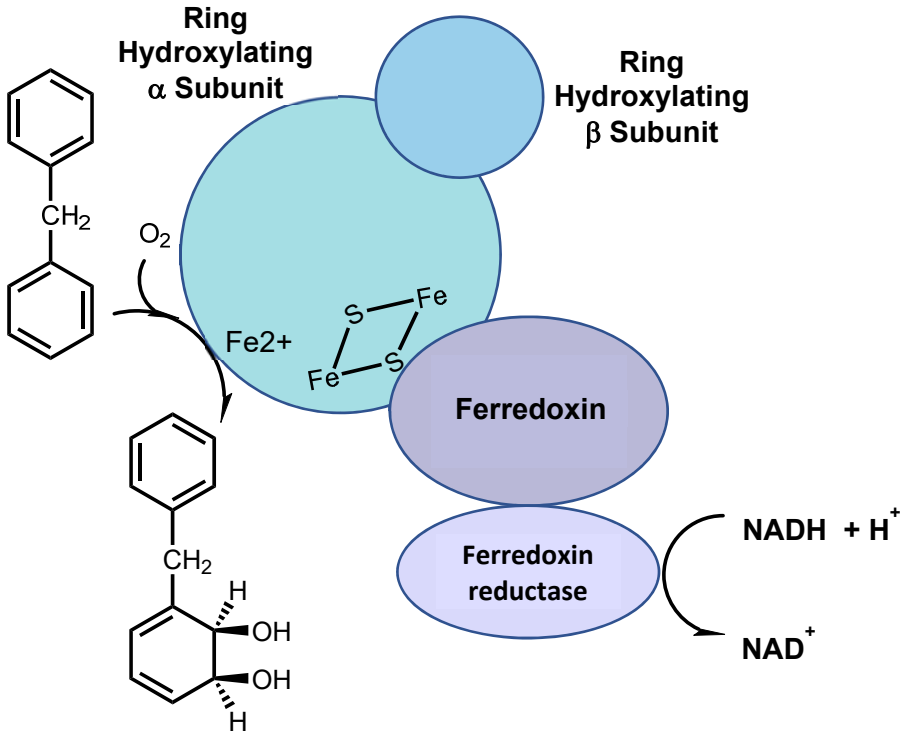

Figure 3. Biphenyl dioxygenase consists of a large alpha and a small beta subunit.

The electron transfer components of the enzyme transfer electrons from an NADH via a ferredoxin reductase and ferredoxin to the Rieske center which is then transferred to the catalytic center and involved in the addition of molecular oxygen to the aromatic ring.

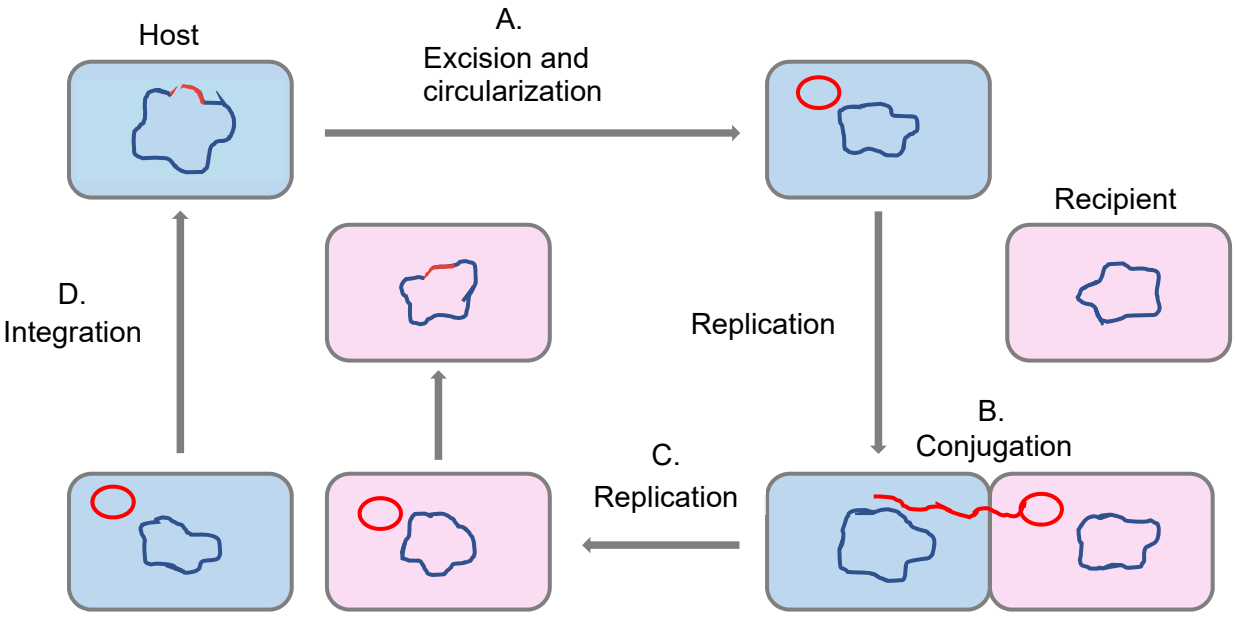

Figure 4. Life cycle of an integrative and conjugative element.

(A) The ICE excises from the chromosome and circularizes in response to certain conditions (B) The ICE is then transferred to a recipient via the conjugation machinery (C) Both in the host and recipient rolling-circle replication occurs to reform a double stranded DNA molecule (D) This then reintegrates into the host chromosome.

Chapter 2: Isolation and Characterization of Diphenylmethane and Biphenyl Degrading Bacteria from the New Jersey Passaic River

INTRODUCTION

Diphenylmethane and biphenyl are two hazardous and toxic compounds with a structure similar to each other. Both compounds consist of two aromatic rings either directly linked together (biphenyl) or linked together by a methylene bridge (Figure 1). Diphenylmethane and biphenyl are of environmental concern since they form a major component of a number of commonly found materials and compounds (9, 83), and are also intermediates in the breakdown of a number of more complex compounds (84). Bioconcentration of these compounds may occur in aquatic life-forms (85). These risks have led to our interest in identifying bacteria that can degrade these compounds.

A number of strains are known to degrade biphenyl via a well characterized catabolic pathway (9). Some, but not all, of these organisms are also capable of metabolizing diphenylmethane using the same enzymes as those that are involved in the biphenyl catabolic pathway. Since some biphenyl degrading bacteria can degrade diphenylmethane we hypothesized that some bacteria isolated for the ability to degrade diphenylmethane would also degrade biphenyl. The pathway for complete degradation of these two compounds involves an upper and a lower pathway (Figure 2). The upper pathway utilizes the same enzymes to metabolize both biphenyl and diphenylmethane. However the end products of the upper pathway are benzoate and phenylacetate respectively (9) as well as 2-hydroxypenta-2,4-dienoic acid (31). This thus requires that

the full metabolism of biphenyl and diphenylmethane involves different lower catabolic pathways.

The biphenyl catabolic pathway consists of four enzymatic steps that lead to the conversion of biphenyl (or diphenylmethane) to benzoate (or phenylacetate). First a multi-component enzyme, the dioxygenase adds both atoms of molecular oxygen to the aromatic ring to form a dihydrodiol compound. Then a dihydrodiol dehydrogenase rearomatizes the ring. The third step involves a 2,3-dihydroxybiphenyl ring cleaving dioxygenase, which cleaves the ring at the 1,2 position (32) and has a ferrous iron as a prosthetic group (33). Finally a 2,4-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid hydrolase (HOPDA hydrolase) hydrolyses the ring-cleavage product producing benzoate (or phenylacetate) and a 5 carbon fragment (34-36).

The initial enzyme in the pathway, biphenyl dioxygenase, is a very versatile enzyme that is known to act on a number of biphenyl analogs (9). Since the substrate specificity and the ability to degrade a compound is primarily dependent on the initial dioxygenase investigators often focus primarily on this first enzyme. This multicomponent enzyme involves the addition of both atoms of molecular oxygen to the aromatic ring (41) and normally consists of three components: a Rieske-type dioxygenase which forms the catalytic unit, (37), a ferredoxin and a ferredoxin reductase. The Rieske type oxygenase component is a heterohexamer that consists of three large alpha and three small beta subunits (37). The C-terminal portion of the alpha subunit contains the [Fe_2-S_2] Rieske cluster and the mononuclear iron containing catalytic center (38). This is the major determinant of structural specificity. It was found that random mutagenesis of the biphenyl oxygenase C-terminal end resulted in changes in the substrate specificity and

activity of the dioxygenase (39, 40). The other two components, the ferredoxin and ferredoxin reductase are responsible for the transfer of electrons from the electron donor NADH to the Rieske center and then to the mononuclear iron catalytic unit (9, 41). There is a great amount of divergence among the biphenyl catabolic (*bph*) genes in nature, pointing to the fact that these are probably very ancient genes and have thus evolved over the years (31).

The biphenyl genes are also found in a variety of genera and while some lie on the chromosome like in *Burkholderia xenovorans* LB400(86), others are often found to lie on a number of mobile genetic elements. These ensure that the genes can thus be easily transferred horizontally in the environment. *Tn*4371 is a well characterized catabolic transposon discovered in the betaproteobacteria *Alcaligenes eutrophus* A5, known for biphenyl degradation. (87) Plasmids containing the genes for biphenyl degradation have been found in plasmids like pRHL1 in *Rhodococcus* sp. (88, 89) and pSS5O in some *Alcaligenes* and *Acinetobacter* species (90). More recently a number of Integrative and Conjugative Elements (ICE) or conjugative transposons have been characterized that contain the genes for biphenyl degradation in *Acidovorax* sp. Strain KKS102 (67) and *Pseudomonas furukawaii* KF707 (66).

In this chapter three strains capable of growth on diphenylmethane and biphenyl were isolated and the biphenyl dioxygenase characterized. It was discovered that the same biphenyl dioxygenase genes were shared by the three diverse strains proving they were acquired by horizontal gene transfer. The biphenyl genes and their transfer are of interest to us because it contributes to our understanding of the degradation of toxic compounds by enabling the transfer of the genes for their degradation. Our work

identifies and characterizes a novel dioxygenase gene involved in the degradation of both diphenylmethane and biphenyl. The mobility of the dioxygenase gene was also observed between a number of strains. Thus, the presence of the genes for degradation of diphenylmethane and biphenyl on a mobile genetic element confers a selective advantage to recipient cells, allowing them to now grow in the contaminated environment.

METHODS

Strain isolation on diphenylmethane. The strains described were isolated from sediment samples obtained along the Passaic River in Northern New Jersey, USA. The sample site (N 40° 44.451 W 74° 07.657) was downstream from the Diamond Alkali Superfund Site, which is designated as a federal Superfund site under the Environmental Protection Agency (EPA) because of the high levels of contamination with 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) that it released (91).

Three 250 ml flasks with 50 ml of Mineral Salts Basal medium (92) were inoculated with 3 g each of the above sediment sample. Diphenylmethane (DPM) was supplied in the vapor form as the sole source of carbon and energy and incubated at 30ºC. On observing an increase in turbidity/growth, the flasks were sub-cultured by adding 1 ml of the original culture to fresh Mineral Salts Basal medium as mentioned above. This was repeated once more on an increase in turbidity to produce a second subculture. Next, three dilutions were prepared from each of the nine flasks and plated on Tryptic Soy Agar medium and incubated at 30ºC, this was done to ensure retrieval of a variety of organisms from fast to slow growers. Subsequently, a variety of morphologically distinct colonies

were selected from the nutrient rich medium and re-tested on Mineral Salts Basal medium plates with diphenylmethane as the sole carbon source again supplied in the vapor phase. The strains were then streaked for single colonies on nutrient rich medium and three isolated colonies from each plate were retested on Mineral Salts Basal medium with DPM. This was repeated twice more. The final strains were then frozen and retained for further analysis.

Testing on other substrates. The isolated strains were tested for their ability to grow on Mineral Salts Basal medium in addition to a number of aromatic compounds as a sole source of carbon and energy. The compounds tested were biphenyl which is similar in structure to diphenylmethane, phenylacetate and benzoate which are the products of the upstream pathway for degradation and some other compounds namely salicylate, 4-hydroxybenzoate, phthalate, phenantherene, naphthalene, and diphenylether. Diphenylmethane, biphenyl, naphthalene, and diphenylether were supplied in the vapor phase by placing it on the lid of the minimal medium plate, liquids were put in vials and placed on the lid. The others were added directly to the medium at a concentration of 10 mM after autoclaving the media.

Identification of strains by 16S rRNA gene sequencing. The strains were grown overnight on LB medium plates. The DNeasy UltraClean Microbial Kit (QIAGEN Cat No./ID: 12224) was used with an appropriate cell mass based on the manufacturers protocol for DNA isolation. In the final step the DNA was eluted with nuclease-free water (Sigma-Aldrich). It was then quantified using a Nanodrop 2000 Spectrophotometer (ThermoFischer Scientific) and stored at -20ºC until further use.

This DNA was then used as a template for PCR amplification using the ReadyMix™ Taq PCR Reaction Mix with $MgCl_2$ (Sigma-Aldrich, Cat no. P4600). The complete 16s DNA sequence was obtained using the 27f forward and the 1522r reverse primers (93). Methods were carried out in accordance with the manufacturers protocol with template DNA added to a concentration of 1 ng/µl. Half the reaction mixture was used to load a 0.8% agarose gel with a molecular marker to confirm the amplification of the expected product by size. The remaining PCR product was then purified using the MP Biomedicals™ GeneClean Kit (Fischer Scientific). The purified PCR product was then sequenced by the company GENEWIZ using Sanger sequencing. Based on the sequences obtained a phylogenetic tree was created using the ClustalW function of the DNAStar Lasergene MegAlign Pro program to illustrate the relationship between all the strains and some well characterized reference strains.

Characterization of dioxygenase genes in isolated strains. Next all of the strains were tested for the presence of the aromatic initial dioxygenase gene in the catabolic pathway, that is, the first enzyme in the pathway. The protocol was the same as above. We used the BPHD-f3, 5′-AACTGGAARTTYGCIGCVGA-3′ and the BPHD-r1, 5′-ACCCAGTTYTCICCRTCGTC-3′ primers (94) which are designed to match the conserved regions of a wide variety of dioxygenase. The amplified product was processed as above and then sequenced.

RT-PCR to determine if the same gene is used for the degradation of both diphenylmethane and biphenyl. A reverse transcriptase PCR experiment was carried out in order to confirm that the above dioxygenase gene was involved in the degradation of diphenylmethane, and also induced by both the substrates diphenylmethane and

biphenyl to initiate the respective upper pathways. This was carried out by growing cells on LB plates followed by 50 ml Mineral Salts Basal medium with 10 mM succinate. This was then used to inoculate 50 ml Mineral Salts Basal medium with either DPM, biphenyl, or succinate as a negative control. Cells were harvested during the exponential growth phase and used for total RNA isolation using the Qiagen RNeasy Mini Kit (Cat No./ID: 74104) followed by treatment with Invitrogen™ TURBO DNA-free™ Kit (Cat No. AM1907) to eliminate any residual contaminating DNA. The RNA quantity and quality was approximated using the Nanodrop 2000 Spectrophotometer (ThermoFischer Scientific).

The above was then used as template RNA for the RT-PCR reaction using QIAGEN OneStep RT-PCR Kit (Cat No: 210210), this allows the reverse transcription of the mRNA and the amplification of the cDNA in a single step reaction. The same well conserved dioxygenase primers BPHD-f3 and BPHD-r1 were used for cDNA amplification. Next 1% agarose gel electrophoresis was carried out to confirm the expression of the genes by the presence of the amplified product. DNA contamination controls were carried out by using the RNA template to carry out regular PCR.

RESULTS

Isolation of diphenylmethane and biphenyl degrading bacteria

The sample site on the banks of the Passaic River was selected due to a history of industrial pollution. Specifically, the sample site was downstream of a former pesticide company that released toxic contaminants, the most concerning of which is the highly

toxic 2,3,7,8-tetrachlorodibenzo-p-dioxin (2,3,7,8-TCDD), into the river via leaks and spills, wastewater discharges, and runoff (95). Decades of contamination have given rise to a wide variety of degradation capabilities of the bacteria in the Passaic River.

Diphenylmethane degrading bacteria were enriched in the sediment samples by incubating sediment samples with MSB containing diphenylmethane. Enrichment cultures were subcultured when growth was observed. After two subcultures the enrichment culture was plated at various dilutions on both TSA and MSB agar with diphenylmethane added in the vapor phase. In order to obtain the most number of different kinds of bacteria colonies were screened morphologically, physiologically, and genetically. Strains were tested for growth on biphenyl, benzoate, phenylacetate, salicylate, diphenylether, dibenzofuran, naphthalene, and phthalate. All the strains were able to utilize biphenyl as the sole source of carbon and energy. PCR amplification and sequencing of the 16S rRNA gene showed that the isolated bacteria grouped into three different species. A representative strain was selected from each group and used in all further experiments. The substrate range for growth of *Pseudomonas* sp. AJR09, *P. stutzeri* AJR13, and *Pseudomonas* sp. AJR20 are shown in Table 1 and an analysis of the 16S rRNA sequence is shown in Figure 5. Interestingly, *P. stutzeri* AJR13 did not grow on phenylacetate, which is an intermediate in the diphenylmethane degradation pathway, and salicylate. The 16S rRNA sequence of AJR09 is identical to that for the well-known strain *P. putida* KT2440. In the process of purifying the cultures AJR09 produced a variant strain (named AJR10) that lost the ability to grow on diphenylmethane, biphenyl, and salicylate.

Identification of the diphenylmethane and biphenyl dioxygenase genes

The identity of the diphenylmethane/biphenyl dioxygenase in the three strains was determined by PCR amplification using primers specific for the biphenyl dioxygenase family (94). All three isolated strains (AJR09, AJR13, AJR20) had the identical sequence for the 524 bp PCR fragment obtained (Figure 6). As expected, AJR10 did not yield a PCR product showing that the dioxygenase genes had indeed been lost. Comparison of the 524 bp sequence to other known dioxygenase sequences showed that it was most similar to that for the *Achromobacter* sp. BP3 biphenyl dioxygenase large subunit at a nucleotide sequence identity of 97.83% (Figure 6). The sequence does not cluster with other well characterized dioxygenase genes forming a new sub family.

Induction of the dioxygenase genes with diphenylmethane and biphenyl

In order to implicate that the same dioxygenase was involved in the pathway for both diphenylmethane and biphenyl degradation RT-PCR experiments were performed. The two compounds are similar in structure and thus could potentially utilize the same pathway for degradation. PCR amplification of mRNA using the same biphenyl dioxygenase primers as above showed a positive band following growth of AJR09, AJR13, and AJR20 on either diphenylmethane or biphenyl but did not show an amplified band from cells grown on succinate or in the no-RT controls (Figure 7). Nucleotide sequencing confirmed that the same fragment was obtained from all three strains and following growth on both diphenylmethane and biphenyl.

DISCUSSION

Three different species of *Pseudomonas* bacteria were isolated from the Passaic River for the ability to degrade diphenylmethane. All strains were also able to grow on biphenyl. The strains fell into three phylogenetic groups with AJR09 close to *P. plecoglossicida*, AJR20 close to *P. fulva*, and AJR13 falling into the genomovar 3 family of *P. stutzeri* (Figure 5). Based on PCR analysis all three strains were found to contain an identical dioxygenase gene. This gene formed a new sub family as it did not cluster with any other well characterized biphenyl dioxygenase genes. This leads us to hypothesize that the genes for diphenylmethane and biphenyl degradation must be horizontally transferred in the Passaic River environment.

All three strains express the same dioxygenase gene when grown on both diphenylmethane and biphenyl. This confirms that the strains utilize the same enzymatic upper pathway for the degradation of both substrates, leading to phenylacetate and benzoate, respectively. A large number of biphenyl dioxygenase have been studied however our current research is of importance since only a few biphenyl dioxygenases are capable of utilizing the same dioxygenase for the degradation of both diphenylmethane and biphenyl. *Pandoraea pnomenusa* B356 (9) is one other such strain that can oxidize both substrates using the same initial dioxygenase. It was interesting to note that AJR13 did not grow on phenylacetate however it did grow on diphenylmethane that forms phenylacetate. This leads us to believe that the strain must accumulate phenylacetate in the medium while possibly growing on the five carbon compound. AJR13 was also incapable of growth on salicylate in spite of carrying genes encoding for degradation on the ICE.

In the process of analyzing the diphenylmethane degrading strains a variant of AJR09 appeared (named AJR10) which had spontaneously lost the ability to grow on diphenylmethane, biphenyl and salicylate. This observation leads us to hypothesize that the genes for diphenylmethane and biphenyl degradation must lie on a mobile genetic element that could easily be lost. This is not uncommon as a number of biphenyl dioxygenase genes are known to be horizontally transferred in the environment. A number of strains have been isolated from biphenyl contaminated soil in Kitakyushu, Japan that contain the genes for both biphenyl and salicylate degradation (66). The best characterized of these, *Pseudomonas putida* KF715 (31) and *Pseudomonas furukawaii* KF707 (96) have been known to contain genes for both biphenyl and salicylate degradation lying in close proximity to each other in the genome allowing them to be easily mobilized together.

The knowledge about an upper catabolic pathway capable of degrading both diphenylmethane and biphenyl present on a mobile element gives us an insight into the selection occurring in-situ in Passaic River sediment and the mobility of genes which can be exploited in the future for the clean-up of extremely contaminated sites by the indigenous microbial population.

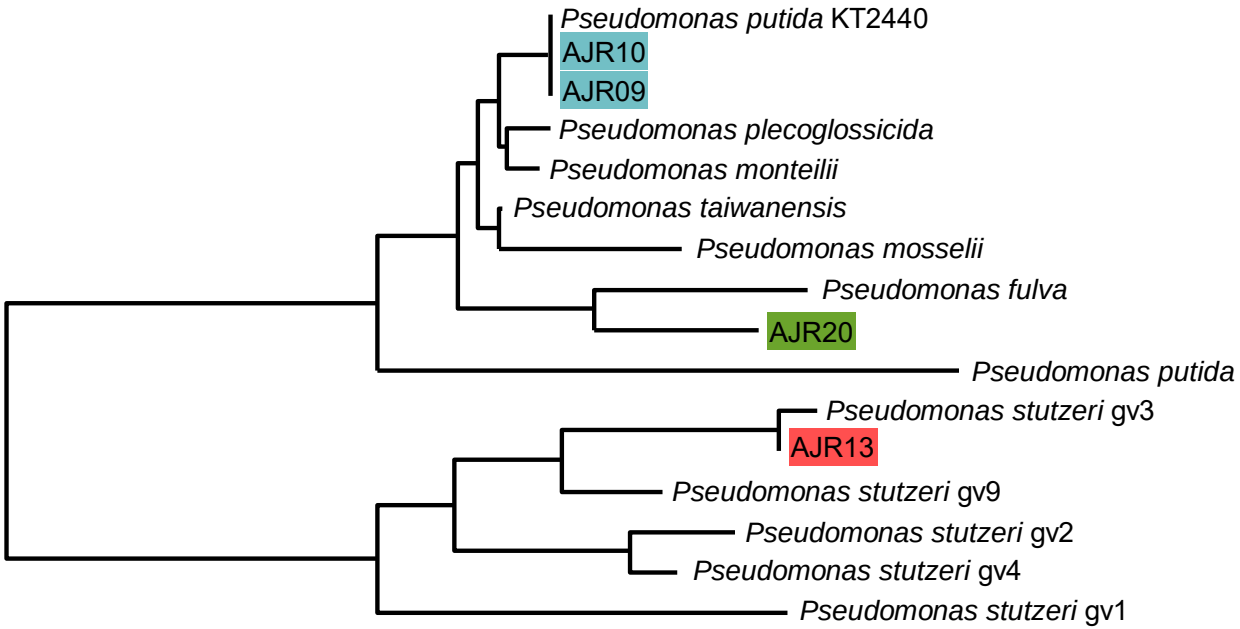

Figure 5. Phylogenetic characterization of the isolated strains AJR09, AJR10, AJR13, and AJR20 in comparison to Type Strain 16S rRNA sequences downloaded from RDP based on their 16s rRNA sequences.

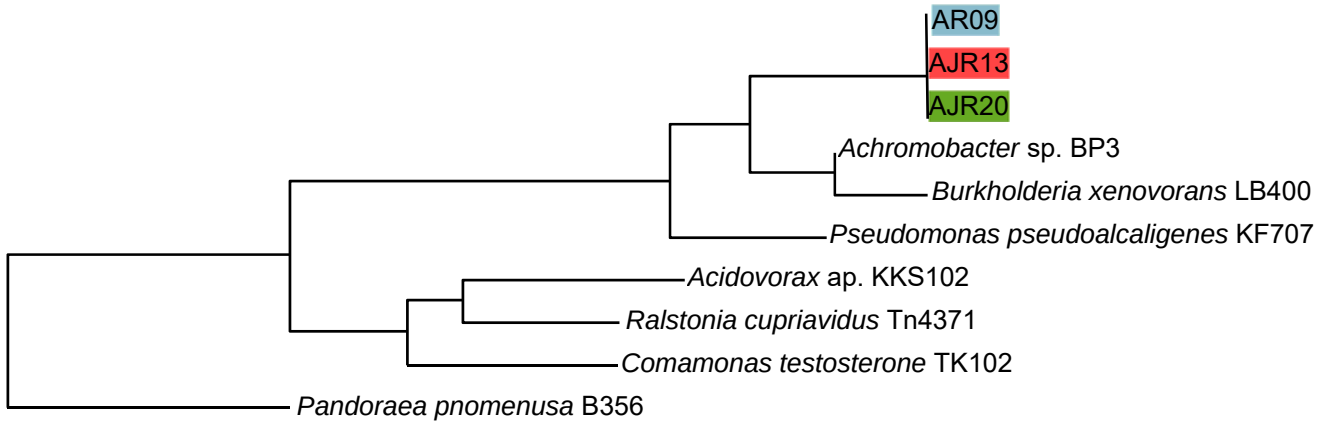

Figure 6. Relationship of the AJR09/AJR13/AJR20 diphenylmethane/biphenyl dioxygenase to other known oxygenases based on the partial sequence of the PCR product

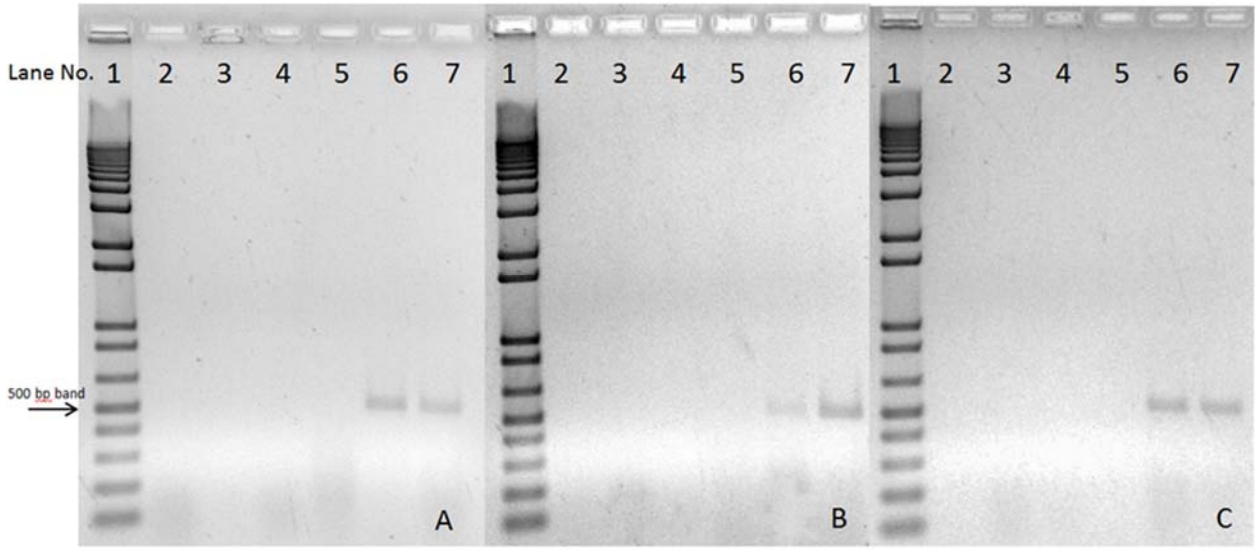

Figure 7. Reverse transcriptase PCR of mRNA following growth on various substrates.

A. *Pseudomonas* sp. AJR09, B. *P. stutzeri* AJR13, C. *Pseudomonas sp.* AJR20. Lane 1 is the Invitrogen 1 Kb plus DNA ladder. Lanes 2, 3, and 4 are the no reverse transcriptase negative controls testing for contaminating DNA in the RNA extracts from cultures grown on succinate, biphenyl, and diphenylmethane, respectively. Lanes 5, 6, and 7 are the reverse transcriptase PCR results, using RNA extracted from cells grown on succinate, biphenyl and diphenylmethane, respectively.

	AJR09	AJR13	AJR20	AJR10
Diphenylmethane	+	+	+	–
Biphenyl	+	+	+	–
Salicylate	+	–	+	–
Benzoate	+	+	+	+
Phenylacetate	+	–	+	+
Diphenylether	–	+	–	–
Napthalene	–	–	–	–
Dibenzofuran	–	–	–	–
Phthlate	–	–	–	–

Table 1. List of substrates tested for growth of the strains of interest.

Chapter 3: Characterization of an Integrative and Conjugative Element involved in Diphenylmethane and Biphenyl Degradation in *Pseudomonas* sp.

INTRODUCTION

Biphenyl and diphenylmethane are hazardous compounds that are similar in structure. The genes for their degradation, especially the pathway for biphenyl degradation, as well as the initial biphenyl dioxygenase has been well characterized (36). The biphenyl pathway genes are found both on the chromosome and on mobile genetic elements. *Burkholderia xenovorans* LB400 is a well characterized strain known to house the genes on its chromosome (86). Some other biphenyl degradation pathway genes are known to be readily transferable as is the case with the catabolic transposon *Tn*4371 (87); plasmids such as pRHL1 (89) in *Rhodococcus* sp. and pSS5O (90) in *Alcaligenes* and *Acinetobacter* species; and integrative and conjugative elements or ICEs like the ones in *Acidovorax* sp. strain KKS102 (67) and *Pseudomonas furukawaii* (previously *pseudoalcaligenes*) KF707 (66).

We had previously isolated strains *Pseudomonas* sp. AJR09, *Pseudomonas* sp. AJR20, and *Pseudomonas stutzeri* AJR13 from the historically contaminated Passaic River in New Jersey. All three strains were found to contain the identical versatile biphenyl dioxygenase gene that was involved in the degradation of not only biphenyl but also diphenylmethane. We also identified a variant of AJR09 (named AJR10) that had spontaneously lost the ability to degrade diphenylmethane, biphenyl, and salicylate. This led us to hypothesize that the genes for the degradation of these three subunits must be

horizontally transferred in the environment and lie on a readily transferable mobile genetic element.

Pseudomonas furukawaii KF707 and *Pseudomonas putida* KF715 were found to carry the genes for both biphenyl and salicylate catabolism with strain KF715 able to spontaneously lose or transfer the genes (31, 96). Thus, Nishi et al. (31) proposed that the *P. putida* KF715 genes for biphenyl and salicylate metabolism are on an Integrative and Conjugative Element (ICE). Characteristics similar to those in the biphenyl-salicylate element have been observed in other genomic islands, an example of this is the well characterized ICE*clc* containing genes for 3- and 4-chlorocatechol metabolism (59).

ICEs are of interest as they are mobile self-transmissible genetic elements. The sequence not only encodes the mechanism for conjugation but also the machinery that regulates and directs the excision of the element from the chromosome. An ICE has the combined features of a number of mobile genetic elements; they can integrate into and excise from the chromosome and can be transferred horizontally. Initially it was believed that ICEs were incapable of autonomous replication (78) however there is now evidence that ICEs may be capable of self-replication (55). These ICEs also do not confer resistance to the organism to other ICEs and thus often more than one ICE can exist in the same host (97).

In this chapter we validate the hypothesis that the genes for diphenylmethane, biphenyl, and salicylate degradation lie on a mobile genetic element and that this element is an ICE. Further we characterize the genes it contains and its ability to transfer the ICE to other strains. ICE applications interest us because it can revolutionize the degradation of toxic compounds by enabling the transfer of the genes for their degradation. The

movement of the ICE was also demonstrated conferring the new host strain with the ability to degrade the compounds. Thus, the presence of the genes for degradation of diphenylmethane and biphenyl on an ICE demonstrates its ability to move around in the environment conferring a selective advantage to the recipient cells that can now grow in the contaminated environment.

METHODS

Transfer of the mobile genetic element. *Pseudomonas putida* KT2440 was selected as the recipient strain for the experiment to determine if the genes could be transferred horizontally. *Pseudomonas stutzeri* AJR13 was selected as the donor. *P. putida* KT2440 is resistant to chloramphenicol while the donor is susceptible, this phenotype was used to select for the recipient while the salicylate genes were used to select for the transferred mobile genetic element. The two strains were grown on LB plates overnight. They were then mixed in a 10:1 donor: recipient ratio on an LB plate and incubated at 30ºC. After 16 hours the mixed culture was resuspended in 50 mM sodium/potassium phosphate buffer and plated on Mineral Salts Basal medium (92) supplemented with 10 mM salicylate and 33mg/ml chloramphenicol and incubated at 30ºC until colonies appear. Four colonies were then picked and streaked for single colonies on the same media as above. This was repeated twice more. The final culture was frozen at -70ºC for future work and named *Pseudomonas putida* KT2440-AJR13$_{ICE}$.

The *Pseudomonas putida* KT2440-AJR13$_{ICE}$ strain was then tested on Mineral salts basal medium supplemented with biphenyl and diphenylmethane, in the vapor phase

by placing them in the lids of the Petri plates. The transfer was also confirmed using PCR amplification using the BPHD-f3, 5′- AACTGGAARTTYGCIGCVGA-3′ and the BPHD-r1, 5′- ACCCAGTTYTCICCRTCGTC-3′ primers (94), which were used previously to characterize the dioxygenase gene.

Whole genome sequencing of four strains was carried out. In order to confirm the presence of and sequence of the possible ICE the four strains, *Pseudomonas* sp. AJR09, *Pseudomonas* sp. AJR20 and *Pseudomonas stutzeri* AJR13 (representative of the isolated strains) and *Pseudomonas putida* KT2440-AJR13$_{ICE}$, were selected for whole genome sequencing.

For DNA isolation, mid-log phase strains were isolated from LB broth incubated at 30^{o}C. The Genomic DNA Buffer Set (QIAGEN) and Genomic-tip 20/G columns (QIAGEN) were used. Isolation was carried out according to the kit instructions. The quality of the DNA was determined by running at 0.8% agarose gel at 25 volts for 16 hours. DNA was quantified using a Nanodrop 2000 Spectrophotometer (ThermoFischer Scientific). For sequencing, 100 µl of the samples were sent to MicrobesNG at the University of Birmingham (UK) (supported by the BBSRC; grant number BB/L024209/1) at a concentration of 30 ng/µl. The libraries were prepared and sequenced on the Illumina HiSeq using a 250bp paired end protocol. After sequencing the reads were adapter trimmed by MicrobesNG using Trimmomatic 0.30 (98) and de novo assembly was performed using SPAdes v3.7 (99). The contigs were annotated using Prokka v1.11 (100). The above analysis was repeated in the lab using Trimmomatic 0.36, FastQC v0.11.7 (https://github.com/s-andrews/FastQC), SPAdes v3.11.1 and Velvet version 1.2.10 (101). The resulting assembled contigs were then compared with the

original contigs and evaluated using QUAST v5.0.2 (102) and progressiveMauve on PATRIC web resources Version 3.5.43 (103).

Characterization of the Integrative and Conjugative Element. The SeqMan Pro™ Version: 10.1.2(20),419. DNASTAR. Madison, WI. was used to align all the previously assembled contigs from all of the above strains to each other to reveal the regions of similarity. PCR primers were designed and used to fill some gaps in the alignment due to repeat regions. The assembled contig element was compared to other known ICEs and the gene composition and relative alignment was used to determine the boundaries of the ICE.

The *Pseudomonas putida* KT2440-AJR13$_{ICE}$ genome was also compared to the *Pseudomonas putida* KT2440 genome in the database to find insertions and differences. Thus, confirming the sequence of the transferred ICE. The three ICE's AJR09$_{ICE}$, AJR13$_{ICE}$ and AJR20$_{ICE}$ were also compared to each other using Mauve for comparisons at the operon level and SeqMan Pro™ for comparison at the nucleotide level. The sequences of other well characterized ICEs like *Pseudomonas furukawaii* KF707 and *Acidovorax* sp. strain KKS102 were also compared using progressiveMauve through PATRIC and Easyfig v2.2.2. The raw reads were mapped back to the assembled ICEs to confirm the assembly using BWA-MEM (104) and SAMtools (105), and visualized using Tablet (106).

Demonstrating the loss of the ICE. The *Pseudomonas putida* KT2440-AJR13$_{ICE}$ strain was cultured on a nutrient rich LB plate at 30^{o}C overnight. This was then tested on Mineral salts basal medium with 10 mM salicylate to confirm the presence of the ICE before being used to inoculate 5 ml of LB broth in 20 ml tubes. The culture was allowed

to grow to log phase. A 10^6 dilution of the culture in 50 mM sodium/potassium phosphate buffer was plated on an LB plate and then 10 µl was transferred to another 5 ml LB broth tube. This was repeated daily. After growth 50 colonies were picked and tested on Mineral Salts Basal medium with 10 mM salicylate to test for the loss of the phenotype conferred by the ICE. Once the colonies were obtained, the loss of the ICE was confirmed by testing for growth on the Mineral Salts Basal medium with biphenyl supplied in the vapor phase. This was then confirmed by PCR amplification with the dioxygenase gene primers used previously. The strains were purified by transferring twice on LB plates and the culture was frozen for further use. The strain was designated *Pseudomonas putida* KT2440-AJR13.

Creating a ΔpobA strain of Pseudomonas putida KT2440 to create a selectable marker. Primers were designed to amplify the region flanking the Δ*pobA* gene, about a thousand bases upstream and downstream of the gene using the Phusion® High-Fidelity PCR Kit (New England Biolabs, Cat no. E0553S). The forward primer was named pobA_F2: 5'-CCCGGATCCcgctgtacatacagcagctg-3' (containing the BamHI restriction site) and the reverse primer was pobA_R2: 5'- CCCGAATTCgatacattcctcaactcgcc-3' (containing the EcoRI restriction site). The product was then purified using the MP Biomedicals™ GeneClean Kit (Fischer Scientific). This was inserted into the pGEM-T easy vector (Promega, Madison, WI) and transformed into *E. coli* DH5α (103). The construct was then sequenced by Genewiz (South Plainfield, NJ).

This plasmid construct was then isolated from the hosts using the NucleoSpin® Plasmid Kit (Macherey-Nagel). This was then digested using the restriction sites created on the ends of the PCR product. The product was then cleaned using the MP

Biomedicals™ GeneClean Kit (Fischer Scientific). The pRK415, low copy number plasmid with a tetracycline resistance antibiotic marker (107), was also digested with the same restriction enzymes. The two were then ligated using the T4 DNA Ligase and the corresponding buffer (New England Biolabs) according to the manufacturers protocol and then transformed into an *E. coli* DH5α host.

The construct was once again isolated and digested with a single restriction site as previously mentioned. The Km^r cassette from the plasmid p34S-Km3 (108) was also digested with the same enzyme before being ligated into the construct and selecting for both kanamycin and tetracycline. This was once again transformed into the *E. coli* DH5α.

The construct was then transferred into the *Pseudomonas putida* KT2440 strain by triparental mating (109) with the *E. coli* HB101 strain carrying the helper plasmid pRK2013 (110). The strain mix was plated on Minimal Salts Basal medium supplemented with tetracycline and succinate to select for the plasmid and the recipient respectively. The transconjugant strain of *Pseudomonas putida* KT2440 was subject to PCR amplification using the primers used initially pobA_F2 and pobA_R2, to confirm the transfer of the plasmid. This transconjugant was then transferred to 5 ml LB broth and grown to late log phase before being transferred again, this was done repeatedly while an appropriate dilution of cells were plated on LB plates. From there colonies were picked and tested on LB plates with kanamycin and tetracycline respectively. A colony containing a successful knockout would be kanamycin resistant but tetracycline sensitive. This colony was then subject to PCR amplification using the initially used primers to confirm that the knockout was successful and double crossover had occurred. The

knockout was then tested on the substrate, *p*-hydroxybenzoate and a lack of growth was observed.

Transferring the ICE from an ICE donor to a recipient. Other ICE transfers were carried out as mentioned above. In the first instance *Pseudomonas putida* KT2440-AJR13$_{ICE(M)}$, the mutated recipient was the ICE donor and *Pseudomonas putida* KT2440-AJR13 (strain cured of ICE) was the recipient. Selection was carried out on Minimal Salts Basal medium supplemented with 10 mM salicylate and chloramphenicol.
In the second transfer experiment, the donor was *Pseudomonas putida* KT2440-AJR13$_{ICE}$ (the original ICE recipient) and the recipient was the previously created *p*-hydroxybenzoate knockout Δ*pobA*-*Pseudomonas putida* KT2440. Here selection was carried out on Minimal Salts Basal medium with 10 mM salicylate and kanamycin.

Sequencing the mutant to determine the mutated gene. In order to carry out MinION sequencing DNA was isolated from log phase grown cells using a modified phenol-chloroform method (111). The DNA was then purified using the Agencourt AMPure XP beads (Beckman Coulter; Brea, CA, USA) and was later run on an 0.8% agarose gel in TAE buffer to confirm the quality and quantity. Library preparation was carried out using the Rapid Sequencing Kit SQK-RAD004 (Oxford Nanopore Technologies, Oxford, UK) followed by loading 100 µl of the sample on a flowcell and used for MinION sequencing. The raw fast5 files were basecalled using Guppy version 3.2.2. this was used for MinION sequencing.

The reads were assembled using Canu v1.8, with five rounds of correction followed by trimming and assembly, with the following parameters: minReadLength = 5000 minOverlapLength = 1000 correctedErrorRate = 0.12 corOutCoverage = 60

corMinCoverage = 4 corMhapSensitivity = normal. The assembly was then polished using Racon v1.4.3, five rounds of polishing were carried out on the on the Amarel cluster, Office of Advanced Research Computing (OARC) at Rutgers, The State University of New Jersey. The assembly was then annotated using prokka v1.14.0.

RESULTS

Transfer of the ICE *from P. stutzeri* AJR13 to *P. putida* KT2440

It was important to demonstrate the transferability of the ability to grow on diphenylmethane, biphenyl, and salicylate from one organism to another to confirm our hypothesis that the genes that encode these abilities are on an ICE. The well characterized *P. putida* KT2440 was selected as the recipient in the experiment since it lacks the ability to grow on diphenylmethane, biphenyl, and salicylate and does have the ability to grow on the products of the biphenyl and diphenylmethane upper pathway (benzoate and phenylacetate, respectively).

P. stutzeri AJR13 was selected as the donor strain to enable ease of selection. The organisms were mated biparentally and ICE recipients were selected based on their ability to grow on salicylate (a gene suspected to be present on the ICE) and chloramphenicol (the recipient was resistant). Salicylate was chosen to be the selective marker due to the higher growth rate compared to biphenyl and also since salicylate can be added directly to the media rather than as a volatile compound. Exconjugates were confirmed to be *P. putida* KT2440 (and not *P. stutzeri* AJR13) by 16S rRNA analysis. In

addition, transfer of the diphenylmethane/biphenyl dioxygenase gene was confirmed by PCR with primers specific for the diphenylmethane/biphenyl dioxygenase.

Exconjugates selected for growth on salicylate were cross examined for the ability to grow on diphenylmethane and biphenyl. Interestingly, all the exconjugates tested were able to immediately grow on diphenylmethane at rates comparable to *P. stutuzeri* AJR13. However, interestingly, the exconjugates did not immediately grow on biphenyl and it took three days for colonies to appear in the streaks on biphenyl plates. A few of the single colonies, which appeared to be possible mutants, were isolated and purified. These strains were once again tested on biphenyl and this time growth was immediate (overnight) and filled the entire streak. One mutant strain was saved and designated KT2440-AJR13$_{ICE(M)}$ while an original exconjugates was saved as KT2440-AJR13$_{ICE}$.

Genome sequencing of AJR09, AJR13, AJR20, and KT2440-AJR13$_{ICE}$

In order to determine the sequence of the mobile genetic element the four strains that contained the possible ICE were sequenced. Illumina short read sequencing led to draft genome sequences that were composed of over a hundred contigs. The assembly was repeated in the lab using SPAdes and compared to the assembly provided by MicrobesNG with the help of QUAST and progressiveMauve on PATRIC. The assemblies were found to be very similar and thus the original assembly was used for further analysis.

Contigs of the donor AJR13 and the recipient strain KT2440-AJR13$_{ICE}$ were examined for overlapping sequences. Contigs from both strains were also compared to the *P. furukuwai* KF707 ICE for biphenyl and salicylate degradation. KT2440-AJR13$_{ICE}$

was also compared to the KT2440 genome present in the database and noted which contig regions did not align with the KT2440 genome. PCR primers were then designed to bridge the gaps between the ICE contigs

ICE Characterization

The full length ICE was annotated using Prokka. As is typical for ICEs the insertion point was a gene for tRNA-glycine(CCC). By ascertaining the ends of the element, we were able to determine its size to be 127,965 bp. At the start of the ICE is an integrase gene of the tyrosine type of recombinase. This was followed by the operon encoding for the diphenylmethane and biphenyl degradation upper pathway. The ICE also contains genes for salicylate and benzoate degradation. The benzoate pathway was the meta pathway for degradation as opposed to the ortho pathway already present in the KT2440 chromosome.

ICE biphenyl degradation is not due to a host mutation

As mentioned above, transfer of the ICE from *P. stutzeri* AJR13 to KT2440 resulted in a strain that readily grew on salicylate and diphenylmethane but that did not readily grow on biphenyl. However, spontaneous mutants did arise that are able to efficiently grow on biphenyl. These mutants could be the result of host mutations or mutations in the ICE itself. If the mutation was in the host chromosome then curing a mutated strain of the ICE would leave a mutated chromosome. Transferring an ICE back in would lead to immediate growth on biphenyl.

A variant KT2440-AJR13$_{ICE(M)}$ strain that had lost the ICE was obtained by culturing the strain in the absence of any selective pressure. Colonies were tested for the loss of the Sal+ Bph+ phenotype and the loss was confirmed by PCR with the biphenyl dioxygenase primers. The cured strain was designated KT2440-AJR13.

The KT2440-AJR13 strain was mated biparentally with AJR13 and ICE recipients were selected for salicylate growth and chloramphenicol resistance. Transfer of the ICE was confirmed by PCR with the dioxygenase primers. Colonies were tested for growth on diphenylmethane and biphenyl with the observation that they immediately grew on diphenylmethane they did not grow on biphenyl. As demonstrated previously, colonies did appear in the streaks after three days which represented mutants that when re-streaked grew immediately on biphenyl (overnight). These observations suggest that the mutation to gain rapid growth on biphenyl is in the ICE and not in the KT2440 host genome.

ICE biphenyl degradation is due to an ICE mutation

In order to test the hypothesis that the ICE after transfer to KT2440 mutated to allow efficient growth on biphenyl the ICE was transferred from KT2440-AJR13$_{ICE(M)}$ to a KT2440 derivative. KT2440-AJR13$_{ICE(M)}$ was mated with KT2440ΔpobA with selection for growth on salicylate and resistance to kanamycin. Exconjugates were analyzed for the ability to grow on *p*-hydroxybenzoate, biphenyl, and diphenylmethane. As expected due to the recipient phenotype, exconjugate colonies did not grow on *p*-hydroxybenzoate. This strain was found to immediately grow on biphenyl and salicylate however it demonstrated slower growth on diphenylmethane, comparable to the growth initially observed on biphenyl. This demonstrated that the mutation could be transferred to the

new recipient, conferring it with the ability to immediately grow on the substrate biphenyl while displaying a different growth pattern on diphenylmethane.

DISCUSSION

Previously it was determined that identical genes for biphenyl/diphenylmethane degradation were present in three different species isolated from a contaminated Passaic River sediment. This led us to hypothesize that the degradation genes lie on a mobile genetic element possibly an Integrative and Conjugative Element (ICE). The ICE was found to be 127,965 bp in length. Two of the ICEs were found to be identical, AJR13$_{ICE}$ and AJR20$_{ICE}$ while AJR09$_{ICE}$ had a large inversion in the second half of the sequence. Barring the inversion the sequences were identical (Figure 8). Comparison was carried out using Easyfig v2.2.2 (112).

A common site of insertion for a number of strains is the tRNA gene. Insertion has been observed into the lysine tRNA (73), proline tRNA (113) and leucine tRNA (74) to name a few. However, the most common site for insertion in this type of ICE is the glycine tRNA as observed in *Pseudomonas aeruginosa* JB2 (114) and *Pseudomonas furukawai* KF707 (66). The AJR$_{ICE}$ also inserts into a tRNA^{gly} specifically the tRNA^{gly}(CCC) when transferred to the KT2440 strain. It has an 19 bp *att* insertion site similar to the 18 bp site found in *Pseudomonas furukawai* KF707 (66). However, the extra base is an extra cytosine at the end of the sequence. A comparison of the four insertion sequences revealed a sequence of 5'-TTCCCWYYRCCCGCTCCAC-3' this has been compared to other *att* sites in Figure 9. The ICE inserted at the end of the tRNA-gly

gene in the chromosome effectively duplicating the 9 bp sequence at either end ensuring that the function of the tRNA-gly gene was not disrupted. The ICE also encodes an integrase of the tyrosine type recombinase based on similarity to other sequences in the database. The 1854 bp integrase has 99.41% similarity with the ICE*bph-sal*KF707 and 99.08% similarity with the *Achromobacter* sp. BP3ICE.

The genes involved in the degradation of biphenyl are present in an operon at the start of the ICE following the integrase. It consists of genes encoding the complete biphenyl pathway beginning with *bphR* involved in regulation followed by the three component dioxygenase including electron its transport proteins. This is followed by the *bphB* dihydrodiol dehydrogenase and the *bphC* 2,3-dihydroxybiphenyl 1,2-dioxygenase (ring cleavage dioxygenase). There is then a cluster of genes involved in the metabolism of the 5 carbon compound produced by the upper pathway and finally the hydrolase *bphD*. The operon bears a high level of similarity to the sequences in *P. furukawai* KF707 and *Achromobacter* sp. BP3. On comparison it was determined that the ~11 kb operon has 98% similarity at the nucleotide level to the operon in *Achromobacter* sp. BP3 and *P. furukawaii* KF707 (Figure 10).

A total of 152 genes were annotated on the ICE a third of which seemed to be involved in degradation/metabolism pathways, a third were involved in the excision, replication and transport, and the final third were hypothetical proteins. The ICE carried the pathways for benzoate metabolism via the meta pathway and for salicylate metabolism, it also carried two copies of the pathway for catechol degradation. The catechol pathway included the genes for conversion of catechol like the catechol 2,3-dioxygenase and other downstream enzymes. On comparison of the duplicated ~7.5 kb

region (Figure 11) we observed that the initial ~5.1 kb of the duplicated region has an 86% similarity while the last ~2 kb is more diverse with a similarity of 80%. There is a 0.3 kb region between the two that has a high level of divergence.

The ICE also carries a number of transposase enzymes belonging to a number of families like *IS*4, IS401 and four *IS*1001 regions. Two *IS*1001 flank a 355 bp imperfect (98%) repeat. Between these two sequences lie a number of mobile element-associated genes that could possibly form a transposable element and mobilize genes. There is another repeat of 1337 bp that lies in the 60 and 94 kb regions respectively. The region between these repeats consists of primarily benzoate degradation related genes leading to the theory that a single sequence containing these genes may have been acquired by the ICE previously. The numerous repeat regions and mobilizing genes contribute to the dynamic nature of the ICE in the environment. Interestingly the ICE sequence is a 98% match for two other sequences in the database, the well characterized KF707$_{ICE}$ and an ICE like sequence in *Pseudomonas alcaliphila* JAB1. We compared the three sequences (Figure 10) and found that the major differences with JAB1 lie in the inversion of the central region flanked by the two catechol pathway repeat regions. The KF707$_{ICE}$ has a mobile element protein after the biphenyl pathway that is absent in the AJR13$_{ICE}$, in addition it lacks 6 genes possibly involved in transport of aromatics/benzoate but contains an extra transposase in the same region. The Proteome Comparison tool on PATRIC (102) was used to compare the KF707$_{ICE}$ at the protein level (Figure 12), irrespective of order we notice divergence in the degradation genes but similarity in the regions containing mobile element proteins i.e. the backbone. JAB1 is more similar than KF707 however they share similarity for a few genes.

The AJR13$_{ICE}$ was then compared to another well characterized ICE, the *Pseudomonas knackmussii* B13 ICE (71) (Figure 13) using blastn and Easyfig. While the degradative capacity of the two elements is different since the B13$_{ICE}$ carries the genes for chlorocatechol degradation there was still a level of similarity between the backbone of the ICE i.e. the mobile element genes. The catechol pathway also shared some similarity with the pathway in AJR13$_{ICE}$. The initial integrase as well as the integrase regulator is similar as is the second half of the ICE that contains genes for the type IV secretion system and ICE transfer.

The excision and loss of the ICE due to absence of a selective pressure was demonstrated. Its transfer was also demonstrated to *P. putida* KT2440. All the ICE genes were observed in the recipient strain and insertion was into the tRNA^{gly}(CCC) gene, as expected. The recipient strain acquired the *sal*+ and *dpm*+ phenotype however no growth was initially observed on biphenyl. It was determined that a mutation in the ICE sequence was necessary in order to be able to degrade biphenyl. This demonstrates that possessing the genotype may not be enough to confer a selective advantage. Mutations may be required for a functional gene.

One important difference was observed in the ICE in AJR13 and AJR09. While the first ~74 Kb are similar the end of the ICE seemed to be missing. On further inspection we determined that an inversion had occurred in the strain. The inverted region included a ~51 kb region followed by the ~62 kb right ICE flanking region. The inversion is flanked by a 127 bp identical *IS*4 type palindromic regions on either side of the inversion. The left flank region of the inverted sequence encodes a transposase InsG that is involved in the excision and insertion of these type of elements. This type of elements

is common to mobile DNA elements and is known to form composite transposons (115). It belongs to a class of DDE transposase that carries out transposition via a direct transesterification reaction. The class is named based on the amino acids that are catalytically important: aspartate (D), aspartate (D) and glutamate (E) (115). This IS seems to be absent in other ICE containing strains like *Pseudomonas furukawaii* KF707, *Achromobacter* sp. BP3 and *Acidovorax* sp. KKS102. The inverted region contains the benzoate pathway genes along with a number of mobile element genes and hypothetical proteins. This inversion could potentially inactivate the ICE. We attempted to transfer the ICE to the kanamycin resistant recipient with selection on both salicylate and later biphenyl however no exconjugates were obtained further supporting our hypothesis. This portrays the dynamic nature of ICEs in the environment. In the absence of the inversion the ICE in AJR09 is identical to the ICE in AJR13.

Attempts were made at identifying the mutated gene however all the candidate genes seemed to be unrelated to the pathway being studied and further analysis is required to peruse the genome and determine the gene of interest.

Future work should involve identification of the ICE mutation required to gain biphenyl degradation in KT2440. Since the mutated strain acquired wild type growth on biphenyl while demonstrating slower growth on diphenylmethane the mutation could possibly be in the genes responsible for substrate specificity, transport or metabolism of the downstream products. It will also be interesting to demonstrate the ICEs host range to other strains outside the *Pseudomonas* genus. This research demonstrates the horizontal transfer of the ICE and its ability to self-mobilize and confer a selective advantage in the toxic environment. It is of importance as it characterizes the movement of genes in the

Passaic River sediment. This could mean more degradation of the contaminants and faster clean-up of the contaminated sites. It was, however, observed that mutations may be needed in the ICE sequence before it can functionally express the phenotype. This work also demonstrates the variety of genes capable of being mobilized. In the future it would be interesting to see if the genes for degradation of more toxic compounds could be moved to the ICE to allow them to be more easily spread for clean-up of a contaminated site.

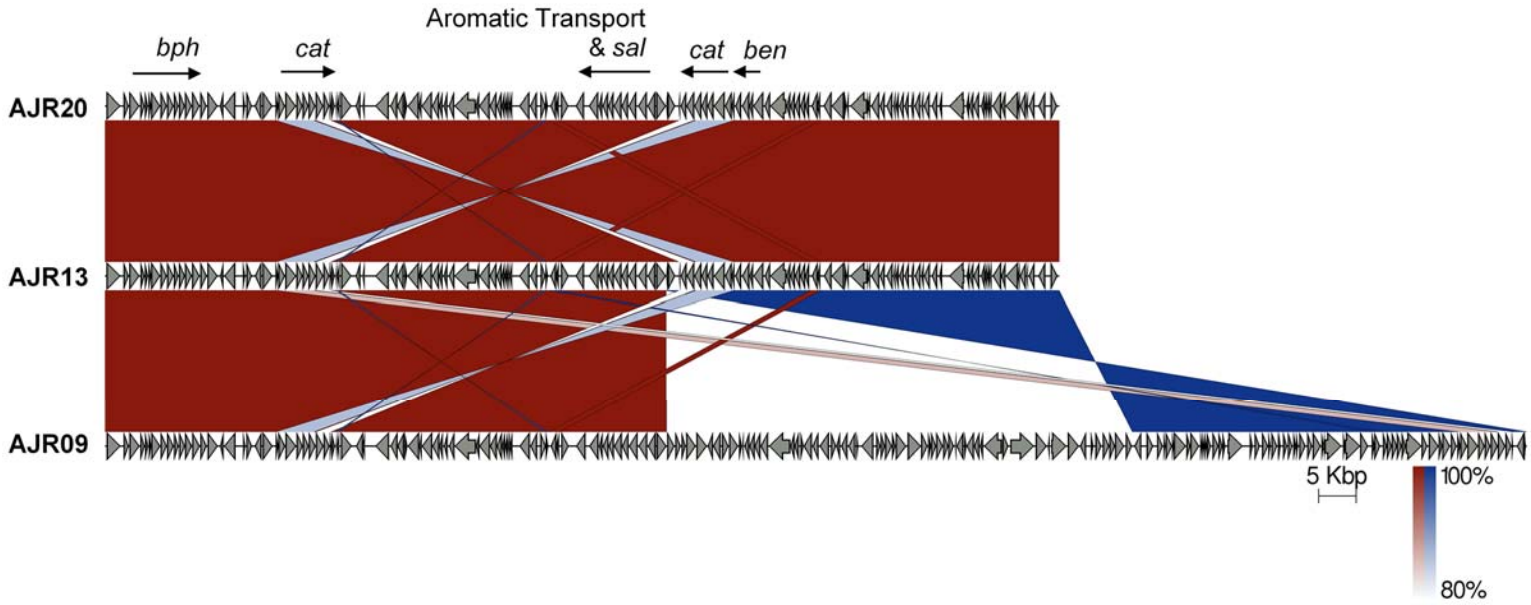

Figure 8. Comparison of ICEs.

AJR13$_{ICE}$ (center) is compared to AJR20 $_{ICE}$ (above) and AJR09 $_{ICE}$ (below). AJR20$_{ICE}$ and AJR13$_{ICE}$ are identical while an inversion is present in AJR09$_{ICE.}$ Designations: *bph*, genes coding for biphenyl degradation; *cat*, genes coding for catechol degradation; *sal*, genes coding for salicylate degradation; *ben*, genes coding for benzoate degradation. Red represents a direct match while blue represents an inverted match, the darker color represents a higher percent match.

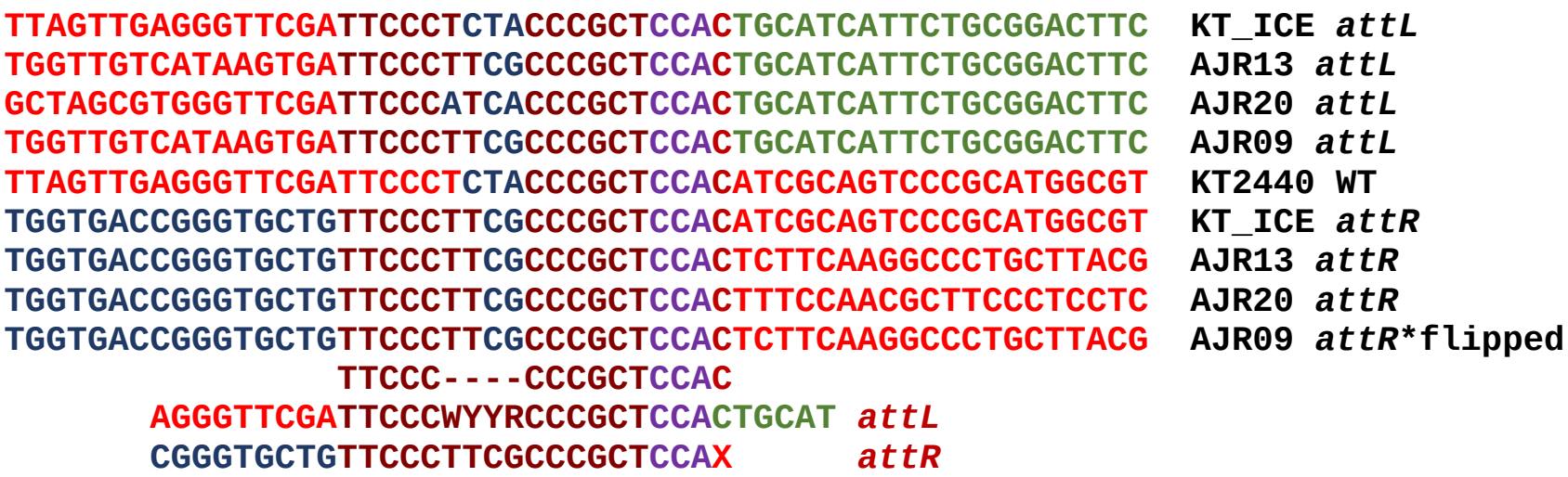

Figure 9. Comparison on the att ICE insertion sites.

Green, beginning of the ICE; blue, end of the ICE; red, genome outside the ICE; maroon, consensus insertion point (duplicated region); purple, CCA end of $tRNA^{Gly}$(CCC) (amino acid attachment point)

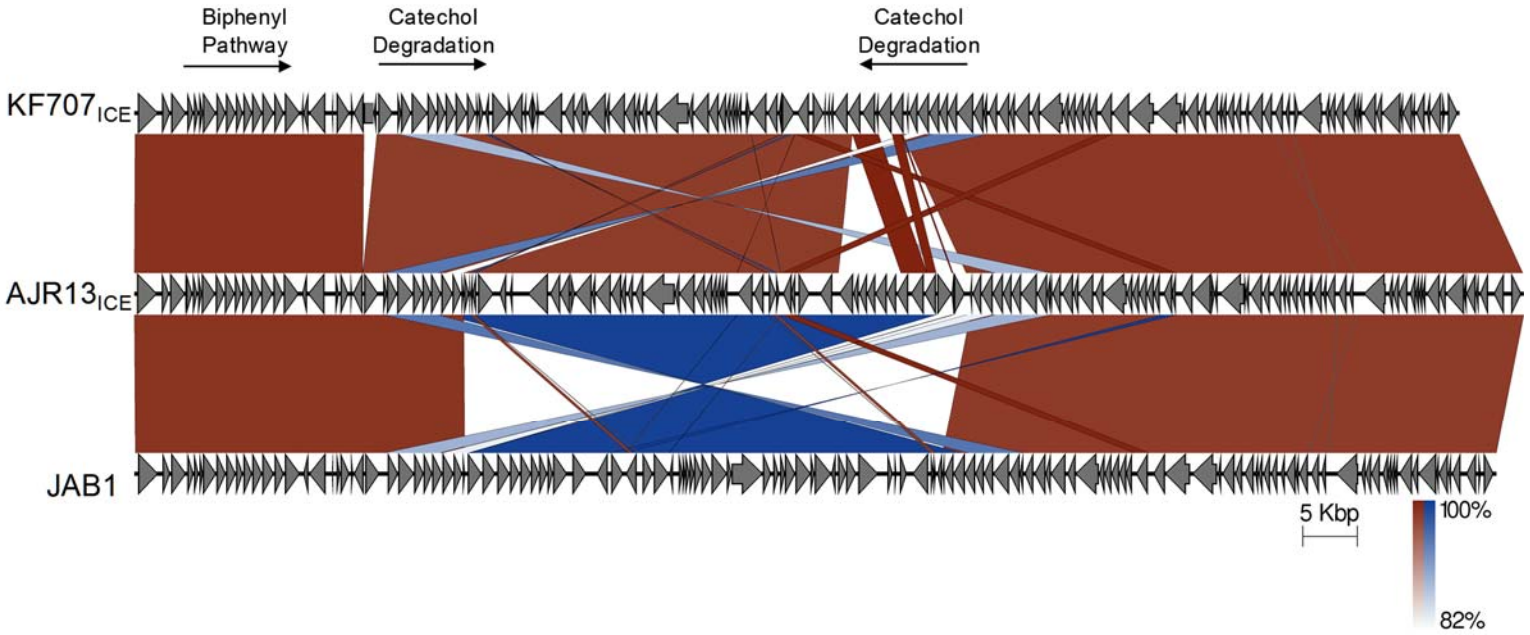

Figure 10. Comparison of AJR13ICE to closely related sequences.

KF707$_{ICE}$ while similar contains some deleted and inserted genes not present in AJR13$_{ICE}$ and JAB1 contains an inversion of the central segment of the possible ICE. Red represents a direct match while blue represents an inverted match, the darker color represents a higher percent match.

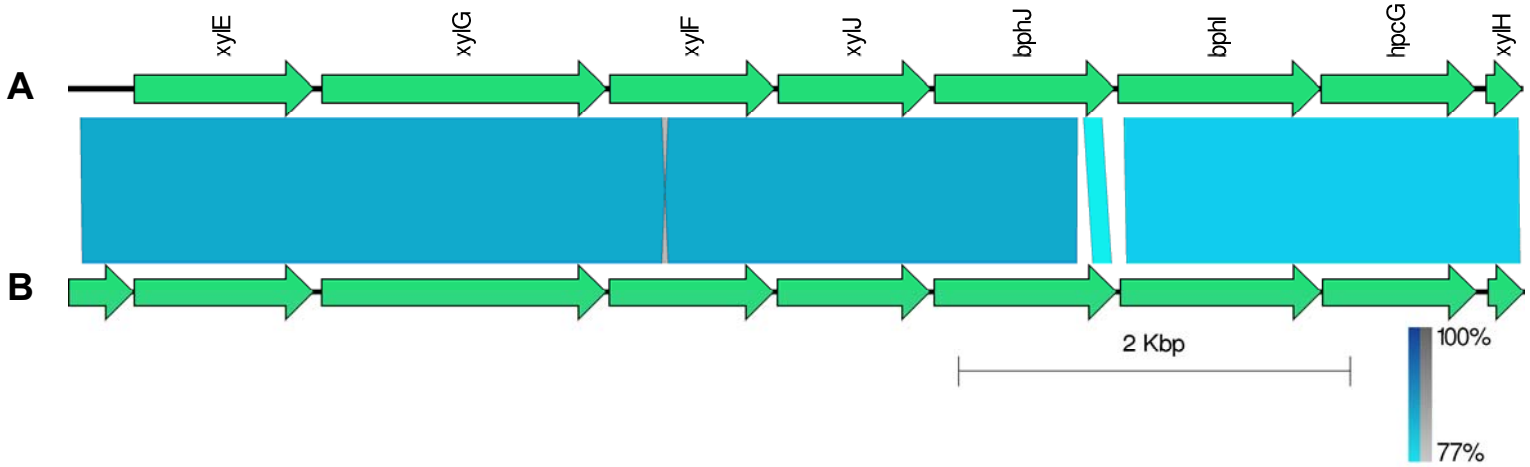

Figure 11. Comparison of the two catechol degradation pathways on the ICE.

(A) initial repeat at ~23 kb region (B) second copy of repeat at ~77 kb region.

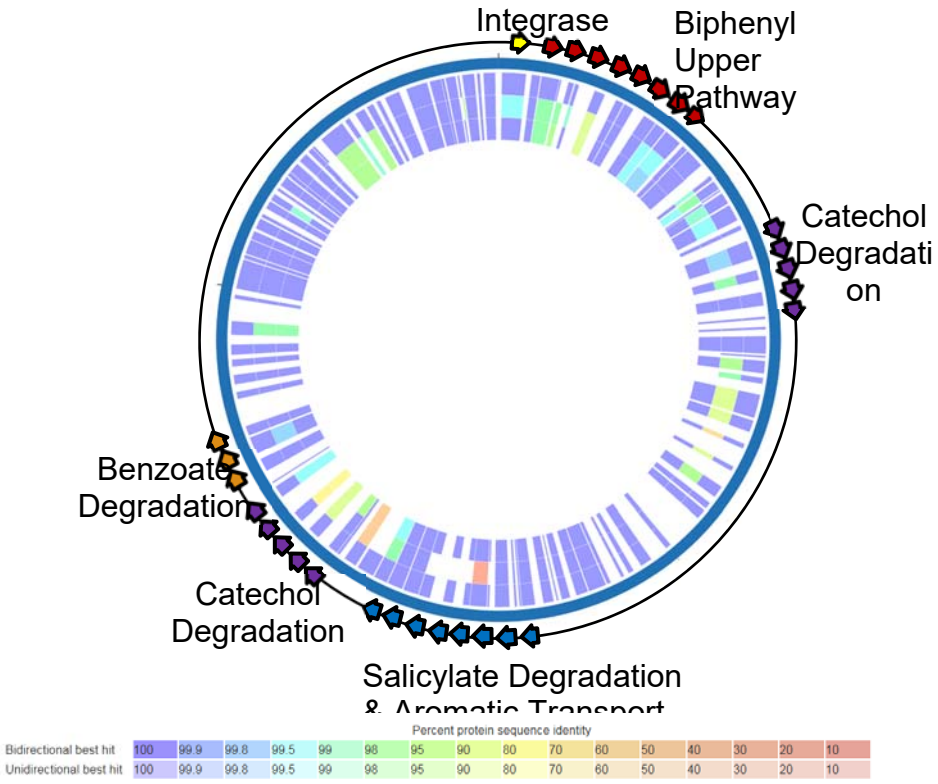

Figure 12. Predicted proteome comparison.

A higher degree of divergence is observed in the genes involved in degradation pathways as opposed to the backbone consisting of mobile element proteins. Form outside moving inwards AJR13$_{ICE}$, KF707$_{ICE}$ and proposed JAB1$_{ICE}$ region proteome.

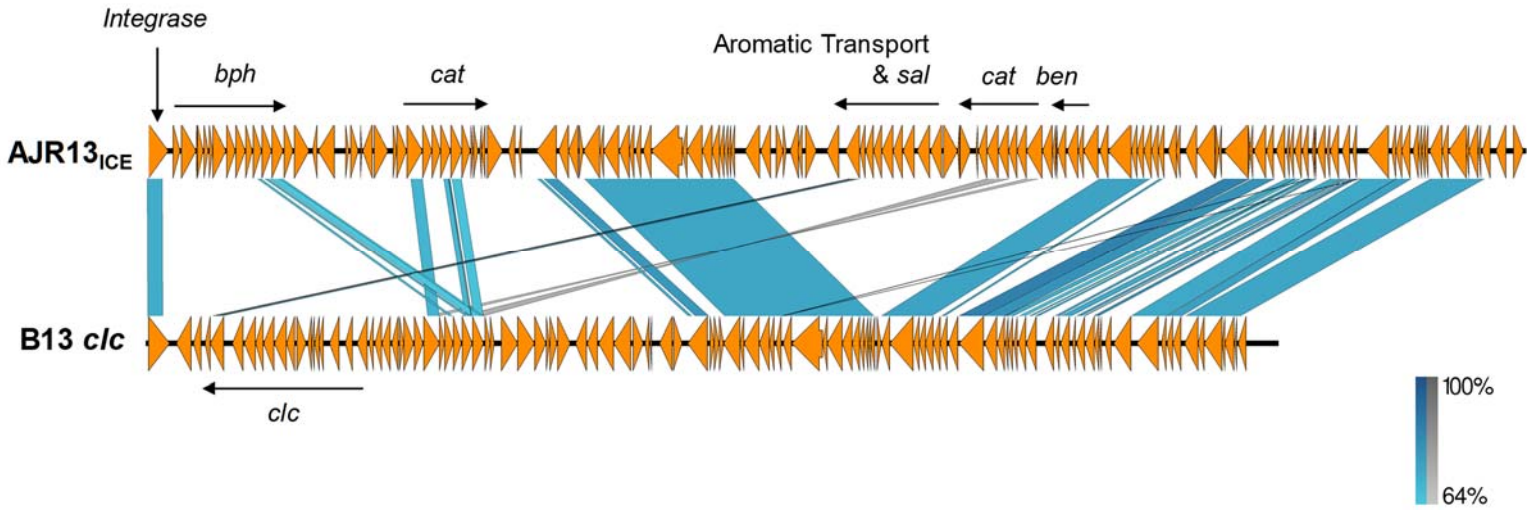

Figure 13. Comparison of AJR13ICE with the Pseudomonas knackmussii B13ICE.

Designations: *bph*, genes coding for biphenyl degradation; *cat*, genes coding for catechol degradation; *sal*, genes coding for salicylate degradation; *ben*, genes coding for benzoate degradation; *clc*, genes coding for chlorocatechol degradation.

Chapter 4: Identification and Characterization of a Novel Angular dioxygenase from *Sphingobium yanoikuyae* SS3 Involved in the Degradation of Diphenylether

INTRODUCTION

Oxygen bridged aromatic compounds are very stable in nature. Examples of these types of compounds are dibenzo-*p*-dioxins, dibenzofurans, and diphenylether type compounds. These types of compounds are widely distributed in nature based on the fact that they are resistant to microbial degradation (116). The diarylether structure is also an important part of the structure of hard coal (12). Diphenylether derivatives are used as heat transfer liquids and perfume additives (117). Brominated diphenylethers are commonly used as flame retardants to reduce the flammability of plastics, textiles, electronic appliances and polyurethanes, and hence are present all around us. Because of their low cost and effective functionality, more than 67,400 tons of brominated diphenyl ethers are produced annually (118) and several thousand tons of diphenylether have been produced annually since 1930 by industry (119). Diphenylether has been detected in drinking water and sea water (119). It is metabolized by vertebrates (119) and can lead to bioaccumulation in mammalian tissue. It is also capable of causing endocrine disruption (118) and hence its degradation is a major concern for us.

A number of bacteria are able to transform or degrade diphenylether (Figure 14). *Pseudomonas cepacia* Et4 (120) and more recently *Cupriavidus* sp. WS (121) yield phenol and 2-pyrone-6-carboxylate (12), *Sphingomonas* sp. PH-07 yields 2-hydroxy muconic acid and phenol (118), *Pseudomonas cruciviae* S93B1 yields 2-

phenoxymuconic acid (122), and *Erwinia* sp. strain CU3614 that was also found to be capable of degradation of a number of diphenylether related compounds (123, 124). Thus, most catabolic pathways studied to date are biphenyl-type pathways acting to produce phenol and a six carbon fragment. *Sphingomonas* sp. strain SS3 is an exception. Work in the early 1990s implicated an angular type of dioxygenase which directly results in the formation of phenol and catechol from diphenylether by this organism. *Sphingomonas* sp. strain SS3 has been shown to degrade diphenylether and its monohalogenated 4-fluoro, 4-chloro, and to a lower extent 4-bromo derivatives and use these compounds as the sole source of carbon and energy; phenol and catechol were the end products of the pathway and were then channeled into intermediary metabolism via the 3-oxoadipate pathway (119). Recently *Sphingobium phenoxybenzoativorans* SC_3 was found to utilize diphenylether as the sole source of carbon and energy utilizing a novel initial angular dioxygenase (125).

While work published in 1992 preliminarily defined the *Sphingomonas* sp. strain SS3 catabolic pathway for diphenylether degradation no work has been performed since to fully characterize the catabolic pathway and to identify the genes involved. This reaction is of importance as the initial dioxygenase in the degradation of dioxin and other aromatic compounds is often the rate limiting step (125). We also believe that elucidation of the angular dioxygenase involved in the degradation will give us more information about angular dioxygenases that attack and breakdown more toxic compounds like dioxin and dibenzofuran.

METHODS

Strain Acquisition and Confirmatory Testing. *Sphingomonas* sp. strain SS3 was obtained from the German Culture Collection DSMZ (119). The strain was grown on LB medium and then tested on Minimal Medium 465 medium (https://www.dsmz.de/microorganisms/medium/pdf/DSMZ_Medium465.pdf) with diphenylether as the sole source of carbon added directly to the flask to verify its ability to grow on diphenylether.

Genome sequencing. *Sphingomonas* sp. strain SS3was grown to log phase in LB broth at 30 degrees and DNA was extracted using the Genomic DNA Buffer Set (QIAGEN) and Genomic-tip 20/G columns (QIAGEN). DNA quality was determined by running a 0.8% agarose gel for 16 hours and it was quantified using a Nanodrop 2000 Spectrophotometer (ThermoFischer Scientific). A sample of concentration 30 ng/μL (100 μL) sent to MicrobesNG (Birmingham, UK) where libraries were prepared and sequenced by Illumina HiSeq using the 250bp paired-end protocol. These reads were then adapter trimmed using Trimmomatic 0.30 with a sliding window quality cutoff of Q15 (98). Preliminary de novo assembly was then performed using SPAdes version 3.7 (99), and contigs were annotated using Prokka 1.11. (100) QUAST version 5.0.2 was used to assess the quality of the assembly. (126) The assembly consisted of 228 contigs-140 contigs (>1,000bp), with the largest contig 360686 bp, an N50 of 106615 bp and an L50 of 18.

In order to close the genome nanopore sequencing was carried out. DNA was isolated from cells in the mid-log phase using a modified phenol-chloroform method (111). DNA purification was carried out using the Agencourt AMPure XP beads

(Beckman Coulter; Brea, CA, USA) and was later run on an 0.8% agarose gel in TAE buffer to confirm the quality and quantity. Library preparation was carried out using the Rapid Sequencing Kit SQK-RAD004 (Oxford Nanopore Technologies, Oxford, UK). The sample was then loaded on a flowcell and used for MinION sequencing. The raw data was basecalled using Guppy version 3.2.2.

The fastq reads were assembled using the Unicycler hybrid assembly pipeline version 0.4.8 (127) run on the Amarel cluster, Office of Advanced Research Computing (OARC) at Rutgers, The State University of New Jersey.

The raw reads were mapped back to the assembled genome using Minimap2 (128) for the long reads and using BWA for short reads (129) and manually checked for any mis-assemblies to assess the quality of the assembly.

QUAST version 5.0.2 (126) was used to evaluate the genome assembly. We obtained five contigs: the chromosome of size 5,056,138 bp, and four plamids/mobile genetic elements of sizes 524,654 bp, 224,120 bp, 127,790 bp and 37,643 bp as opposed to the number of plasmids previously predicted. (130) The average sequencing coverage depth was 321 and 98.6% of the reads were mapped to the assembly. The GC% was 63.87 with an N50 of 5056138 and an L50 of 1.

The genome was next annotated using Prokka version 1.14.0. (100) The genome was found to contain a number of dioxygenase genes that could possibly be involved in the environmental degradation of bicyclic compounds. Following annotation the SS3 genome was scanned with the standalone Blast+ tblastn program using a series of manually curated seed amino acids sequences of various known dioxygenase genes. The protein hits acquired were then extracted, placed in a single fasta file, and were used

again to scan the genome in order to ensure that we obtained all the possible candidate genes.

RT-PCR. Primers were designed to amplify the candidate dioxygenase genes. The strain was streaked on LB plates overnight and later transferred to LB broth. Log phase culture was then transferred to Minimal Medium 465 supplemented with succinate, log phase culture from this flask was then transferred to Minimal Medium 465 supplemented with either 10 mM succinate, 10 mM benzoate, or 10 mM diphenyl ether, as described previously. The cells were harvested during the log phase of growth and RNA extraction was carried out using the Monarch® Total RNA Miniprep Kit (New England Biolabs Inc.), this was followed by DNase treatment to prevent DNA contamination using the TURBO DNA-free™ Kit (Invitrogen). The resulting mRNA was used to set up the RT-PCR reactions with the OneTaq® One-Step RT-PCR Kit (New England Biolabs Inc.). This was then run on a 1% agarose gel for 1 hour at 90V. We also included a PCR negative control using the Taq 2X Master Mix (New England Biolabs Inc.) to confirm the absence of DNA contamination.

Construction of expression vectors. DNA was isolated from cells grown overnight on LB plates using the DNeasy UltraClean Microbial Kit (QIAGEN). Primers were designed to amplify the alpha and beta subunits of the candidate dioxygenases using the Phusion® High-Fidelity DNA Polymerase (New England Biolabs Inc.), followed by DNA purification using the MP Biomedicals™ GeneClean Kit (Fischer Scientific). These were then introduced into the expression vector pET-30a after digestion with restriction enzymes by New England Biolabs Inc. T4 DNA Ligase and the corresponding buffer (New England Biolabs) were used to ligate the insert to the vector and incubated for 2

hours at 16 degrees C. This was followed by transformation into competent *E. coli* BL21 cells by keeping the cells on ice followed by addition of the ligation mixture and a heat shock at 42 degrees C for 45 seconds. After recovery on ice, the cells were plated on LB medium substituted with kanamycin and incubated at 37 degrees C overnight to select for cells transformed with the construct. These strains were then stored at -70 degrees C. The plasmid was then re-isolated from the cells and run on a 0.8% agarose gel to confirm presence of the correct construct.

Degradation of Diphenylether by *E. coli* clones. The plasmid containing expression strains and a negative control (empty pET30a vector with no genes) were plated on LB substituted with kanamycin overnight and then used to inoculate LB broth with kanamycin (50 ml) overnight at 37 degrees C. The next day 100 ml LB broth with kanamycin was inoculated to make a 2% culture. This was incubated at 37 degrees C till an O.D. of 0.6 was reached at 600 nm. Isopropyl-β-d-thiogalactopyranoside (IPTG) was then added to the flasks to induce production of the proteins. After 2 hours the cells were washed twice with 50 mM sodium/potassium phosphate buffer (pH 6.8) and concentrated 3x in the same buffer and incubated for 24 hours with the addition of 20 mM glucose and 3 mM diphenylether. The cells were then harvested by centrifugation at 4 degrees C at 5000 rpm for 12 minutes. The supernatant was then collected and filtered using a 0.45-µm-pore-size filter before being run through a C_{18} column of the HPLC in the reverse phase with a 0 to 100% methanol gradient in water under 0.1% acetic acid conditions. The retention time of the different peaks were compared with those produced by the standards and the negative control.

Strain SS3 Gene Knockouts. To create knockouts of the gene of interest primers were designed for the regions flanking the alpha subunit of the dioxygenase. The sequence was amplified using PCR and inserted, as mentioned previously, into the broad host-range pRK415 plasmid vector (107). A Km^r cassette from p34S-Km3 (108) was then inserted into the previously assembled construct to form the vector used to create the knockout vector. The final construct was transformed into *E. coli* DH5α and was selected for based on its ability to grow on LB with kanamycin and tetracycline. This construct was then confirmed using PCR amplification of the insert and frozen at -70 degrees C.

This construct was then used to carry out triparental mating (109) with the donor strains constructed, the recipient strain *Sphingobium yanoikuyae* SS3, and *E. coli* HB101 carrying the helper plasmid pRK2013. The three strains were grown overnight and then mixed together on an LB plate and incubated overnight at 30 degree C. The mixed culture was then collected and resuspended in 50 mM sodium/potassium phosphate buffer (pH 7.25) before being plated on Mineral salts basal medium (92) supplemented with 10 mM phenylalanine and tetracycline to select for the recipient. After 4-5 days the transconjugant colonies were streaked for single colonies thrice to ensure it was a pure culture before being frozen at -70 degrees C. The transconjugants were transferred repeatedly in LB broth while simultaneously checking for the loss of the pRK415 plasmid by plating on LB containing tetracycline while containing the kanamycin resistance gene caused by a double crossover event and knockout. This strain was then purified and stored for further analysis.

Complementation of the loss of growth on diphenylether. In order to complement the knockout mutant the original alpha and beta dioxygenase genes were cloned into the

pRK415 plasmid as mentioned above. This was then transferred into the knockout mutant using triparental mating as mentioned above and selected for based on its ability to grow on tetracycline. This complemented strain was then purified and frozen at -70 degree C for further analysis.

RESULTS

SS3 Genome Sequence. *Sphingomonas* sp. strain SS3 was obtained directly from the DSMZ culture collection and confirmed to grow on diphenylether as the sole source of carbon and energy. Using a combination of Illumina short read and Oxford Nanopore long read technologies an almost complete genome was assembled. The strain contains a 5,056,138 bp (uncircularized) chromosome and four circular plasmids: 524,654 bp, 224,120 bp, 127,790 bp, and 37,643 bp. Comparison of the SS3 16s rRNA gene to sphingomonad type strains identified SS3 as *Sphingobium* (formerly *Sphingomonas*) *yanoikuyae* (Figure 15)

Rieske Dioxygenase Identification. Rieske type dioxygenases were identified in the SS3 genome sequence using the stand alone blastP and TblastN programs and a seed file of known aromatic dioxygenase enzyme amino acid sequences. Genes/enzymes were initially labelled as to the Illumina assembly contig (and position) where they were located. Thirteen Rieske alpha subunit encoding genes were identified in the genome (Figure 16). However, only four of these had beta subunit encoding genes nearby. While Rieske dioxygenases attacking aromatic compounds can be both alpha subunit only and

alpha/beta subunit types, most polycyclic aromatic hydrocarbon oxidizing Rieske oxygenases are of the alpha/beta type. Further analysis by comparison to selected known dioxygenases (especially angular dioxygenases attacking ether linked polycyclic compounds) led us to hypothesize that the dioxygenase encoded by Contig010 attacks benzoate and that the dioxygenase encoded by Contig095 attacks diphenylether (Figure 16). An alternative hypothesis is that the nearly identical dioxygenase encoded by Contigs 042 and 087 attacks diphenylether due to its low similarity to a carboxydiphenylether dioxygenase. Bolstering our hypothesis that the dioxygenase gene on Contig010 encodes a benzoate dioxygenase is the fact that the gene(s) are in an operon with the catechol pathway genes (see below).

RNA Analysis. RNA from strain SS3 was analyzed by RT-PCR under various growth conditions (Figure 17). PCR primers amplifying about 500 bp were designed for the Contig010, Contig095, Contig042, and Contig087 dioxygenase genes. Controls with no reverse transcriptase reaction did not show any amplification (data not shown). Primers for the dioxygenase gene on Contig042 showed no amplification with RNA from cells grown under any condition (data not shown) and the nearly identical gene on Contig087 showed no amplification from RNA from succinate or benzoate grown cells and showed a faint band (~800 bp) from RNA from diphenylether grown cells. However, using primers for Contig010 and Contig095 amplification was seen from RNA extracted from all three growth conditions with faint bands from succinate and benzoate RNA and a very heavy band from diphenylether RNA. The gene encoding the Contig010 dioxygenase is in an operon also containing all of the genes needed for catechol degradation (Figure 18).

Therefore, it is not surprising that the Contig10 benzoate dioxygenase is upregulated during SS3 growth on diphenylether since in this strain diphenylether is metabolized through catechol. Upregulation of the catechol pathway would also result in the upregulation of the co-operonic Contig010 benzoate oxygenase genes. Based on all the data (similarity to known dioxygenases, identity of adjacent operonic genes, and the RT-PCR results) the Contig010 dioxygenase is definitively proven to be a benzoate dioxygenase.

Heterologous Expression. The three dioxygenases (from Contig010, Contig087, and Contig095) were heterologously expressed in *E. coli* using the T7 based vector pET30. Incubation of induced cells with diphenylether followed by HPLC analysis of the supernatant showed that the heterologously expressed Contig095 dioxygenase produced catechol and phenol from diphenylether (Figure 19). Thus, the Contig095 dioxygenase has the expected enzymatic activity of a diphenylether dioxygenase.

Gene Knockout. Since the Contig095 dioxygenase has the expected activity the gene encoding this enzyme was knocked out from the SS3 genome. Strain SS3ΔN95 does not grow on diphenyl ether (Figure 20). Growth is restored when pRK415 containing the Contig095 dioxygenase genes is added to complement the chromosomal knockout however there is a lag in growth compared to the wildtype attributed to the tetracycline added to the medium in order to present the loss of the plasmid. This was confirmed by observing the growth pattern of the wildtype SS3 containing the empty pRK415 plasmid (Figure 20). Thus, based on (low) similarity to known angular dioxygenases, gene

induction during SS3 growth on diphenylether, heterologous expression in *E. coli*, and gene knockout, the Contig095 dioxygenase has been proven to be a diphenylether dioxygenase.

DISCUSSION

Sphingomonads are well known for their ability to degrade a wide range of unusual aromatic compounds. While several organisms are known to degrade diphenylether *Sphingobium yanoikuyae* SS3 is the best studied degrader from a biochemical and physiological standpoint (13, 119). However, despite a flurry of published studies in the 1990's not much in depth research on diphenylether degradation has been performed in the last 20 years. Our work thus provides the genetic basis for an historical strain's ability to degrade diphenylether. Comparison of the SS3 diphenylether dioxygenase alpha subunit to amino acid sequences in the GenBank database yields few matches showing that this enzyme is somewhat unique. There is a 100% match to a dioxygenase of unknown function encoded by the chloroacetanilide herbicide degrading *Sphingobium* sp. strain MEA3-1 (131, 132). Besides this blast match all other matches are below 55% identity. Recently one other diphenylether dioxygenase has been identified in *Sphingobium phenoxybenzoativorans* SC_3 (125) and this enzyme shows only 35% identity (49% similarity) to the diphenylether dioxygenase discovered in our work. For comparison, our diphenylether angular dioxygenase shows 42% identity to the well-characterized *Sphingomonas wittichii* RW1 dioxin dioxygenase (133).

Comparison of the SS3 region encoding the diphenylether angular dioxygenase and the similar region from MEA3-01 is shown in Figure 21. This is an active region for

insertions, deletions, and rearrangements. Three identical IS6 family insertion sequences are present in this SS3 diphenylether genome region with the genome region between two of them flipped relative to the MEA3-01 genome and a large deletion between the other two (present in MEA03-01 but not SS3). In addition, the MEA03-01 genome has an insertion sequence inserted directly upstream of the diphenylether genes that is not in the SS3 genome.

Interestingly, the SS3 diphenylether dioxygenase alpha and beta subunits ("oxygenase component") expressed in *E. coli* showed efficient transformation of diphenylether to phenol and catechol. Detection of the products was by HPLC directly from culture supernatants without extraction and/or concentration of the metabolites. This was even though no native ferredoxin or ferredoxin reductase was present. Indeed, inspection of the SS3 genome sequence shows that no genes encoding dioxygenase electron transport components were near the genes for the diphenylether dioxygenase oxygenase component alpha and beta subunits. It is well known that dioxygenases in sphingomonads typically share electron transport components and that these components are often scattered in the genome (133-135).

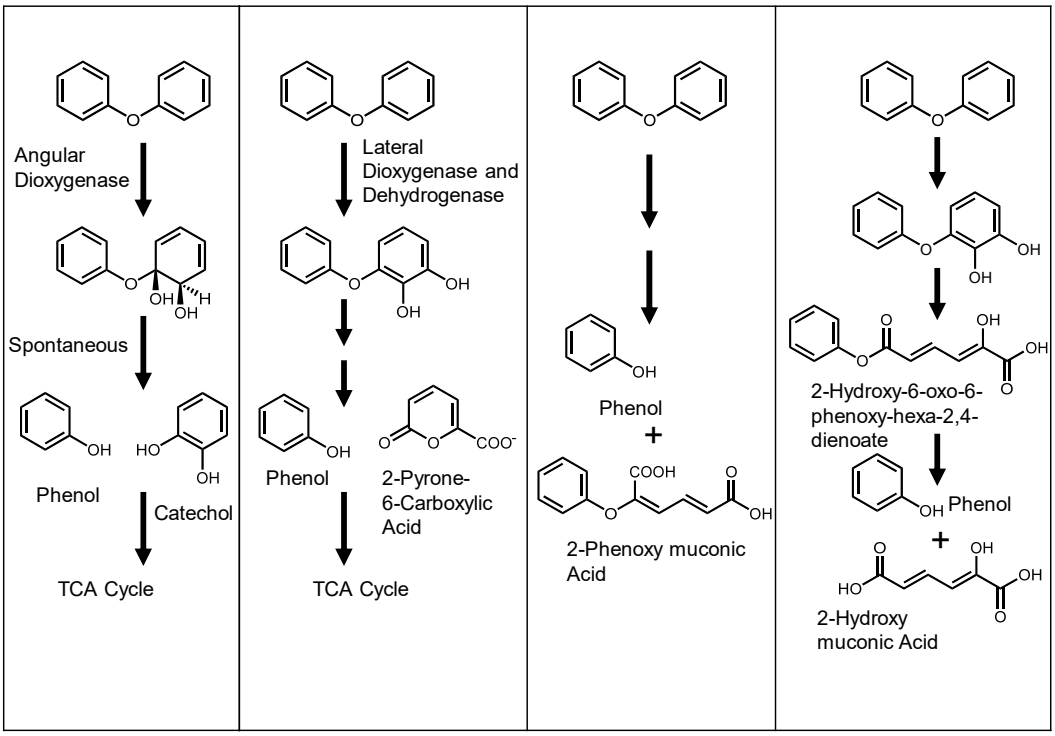

Figure 14. Proposed pathways for the degradation of diphenylether.

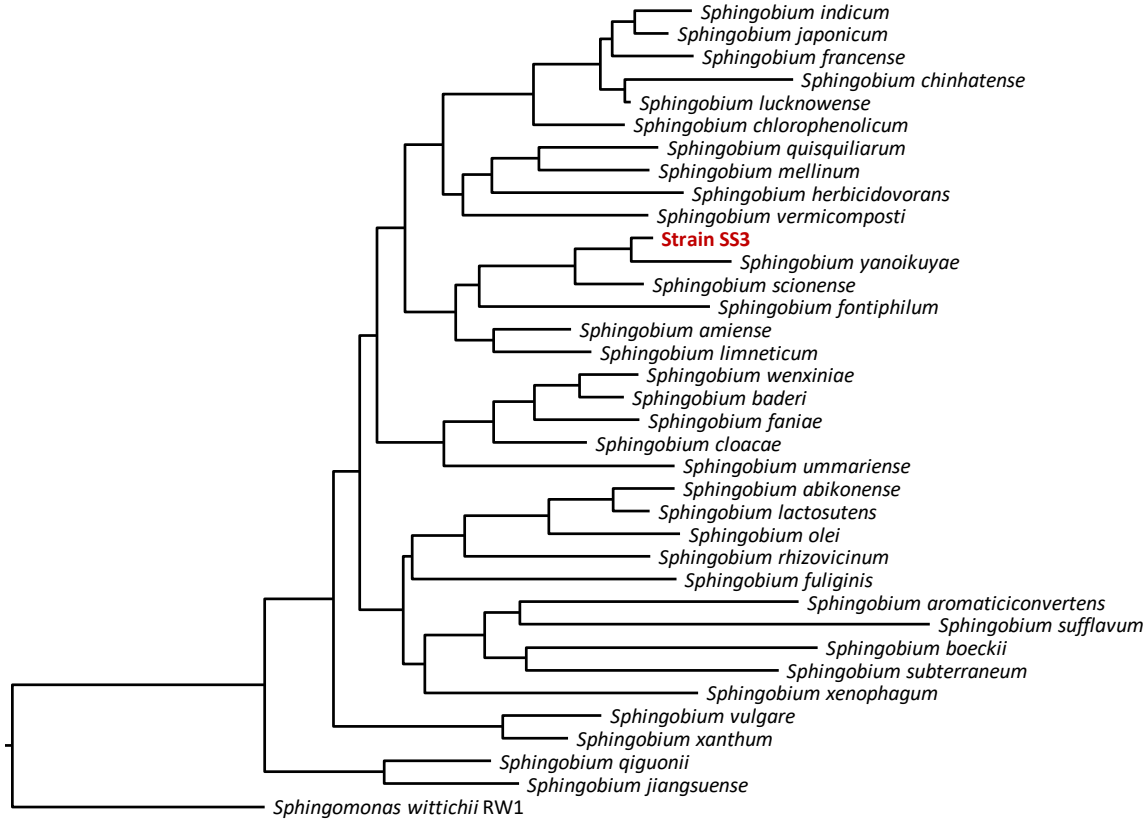

Figure 15. Comparison of the *Sphingomonas* sp. strain SS3 16S rRNA gene to sphingomonad type strains.

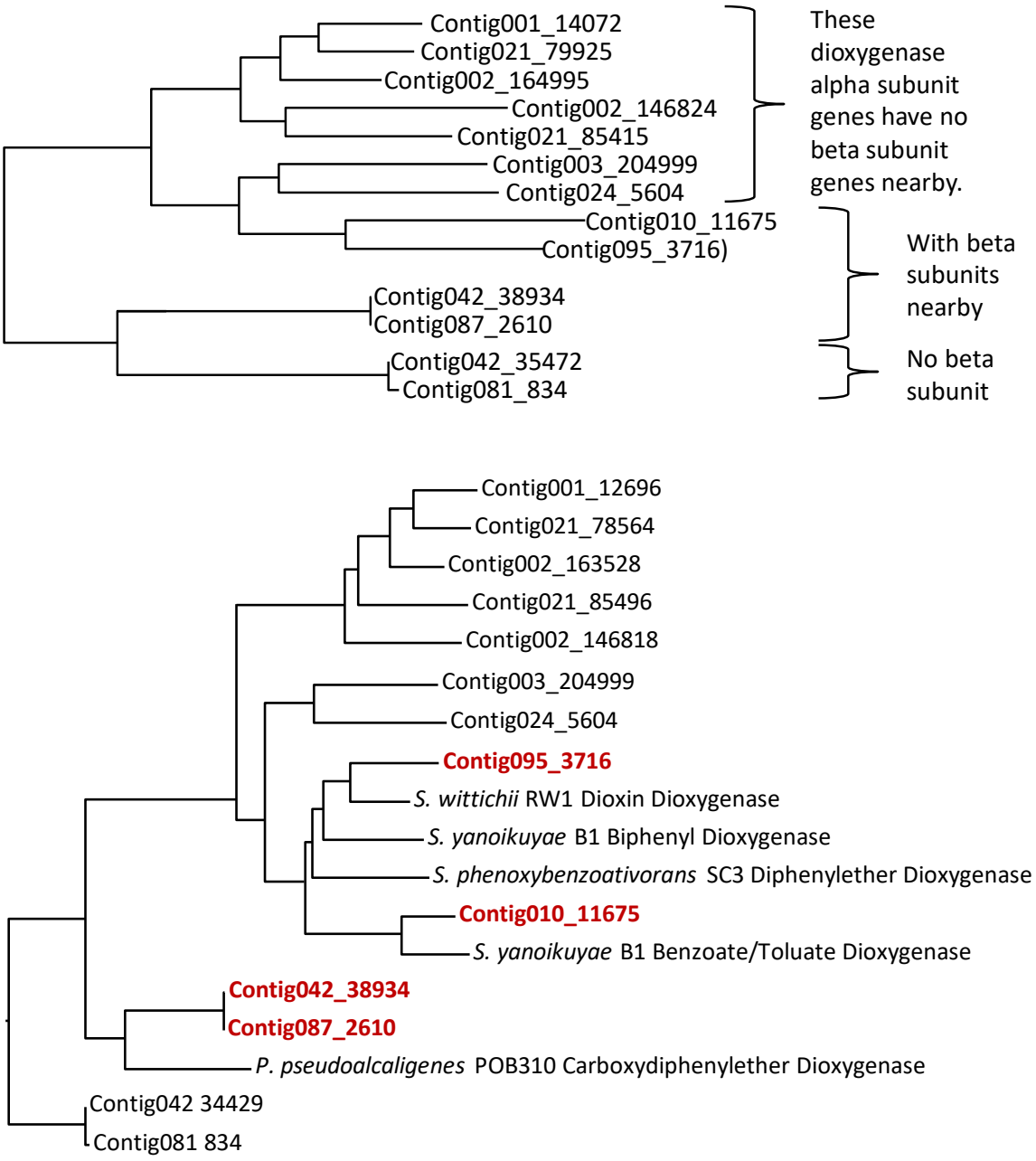

Figure 16.. Identification of thirteen Rieske dioxygenase encoding genes in the SS3 genome.

Top: Relationship to each other and characteristics. Bottom: Relationship to selected enzymes

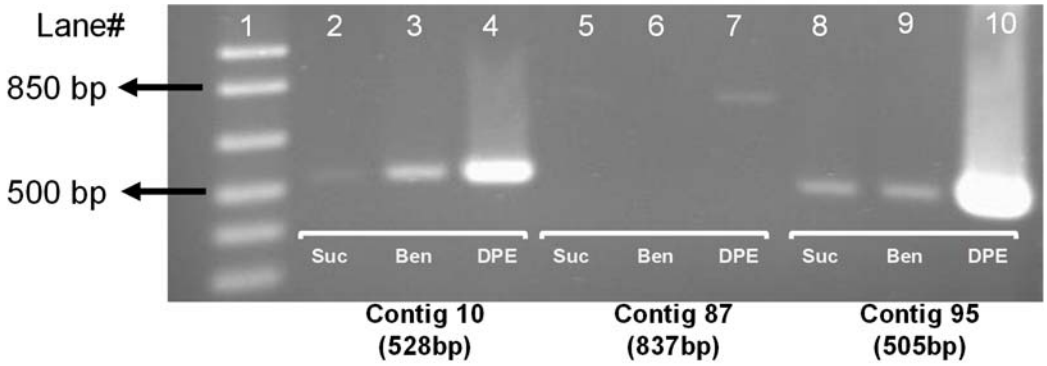

Figure 17. RT-PCR of RNA from strain SS3 grown on succinate, benzoate, or diphenylether.

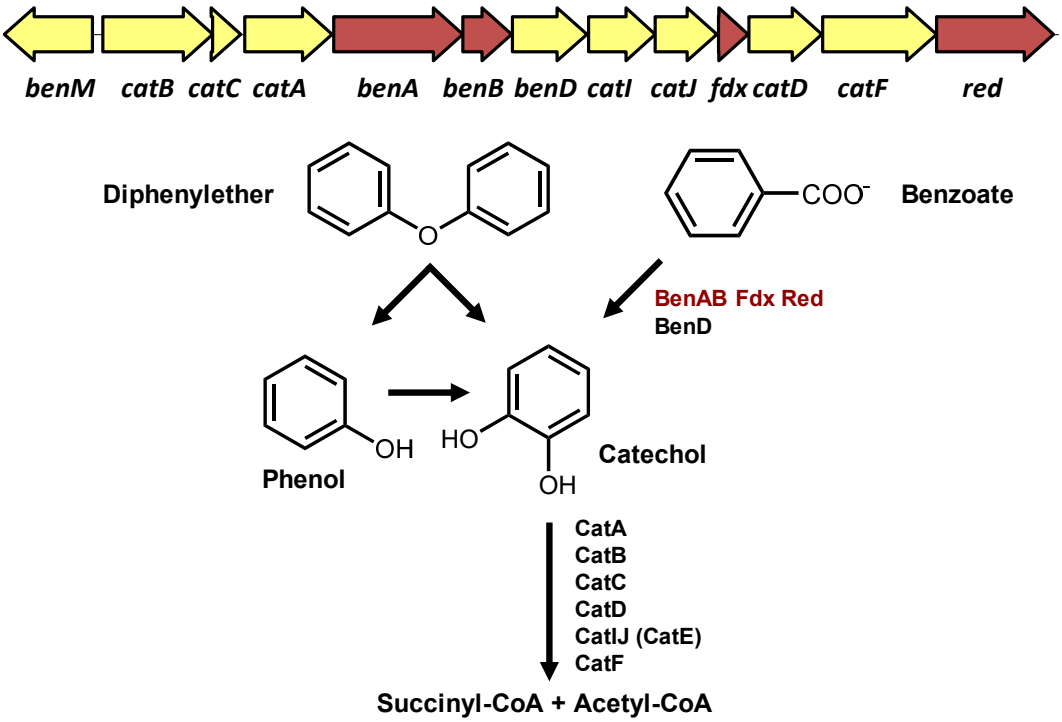

Figure 18. Contig010 benzoate/catechol degradation operon.

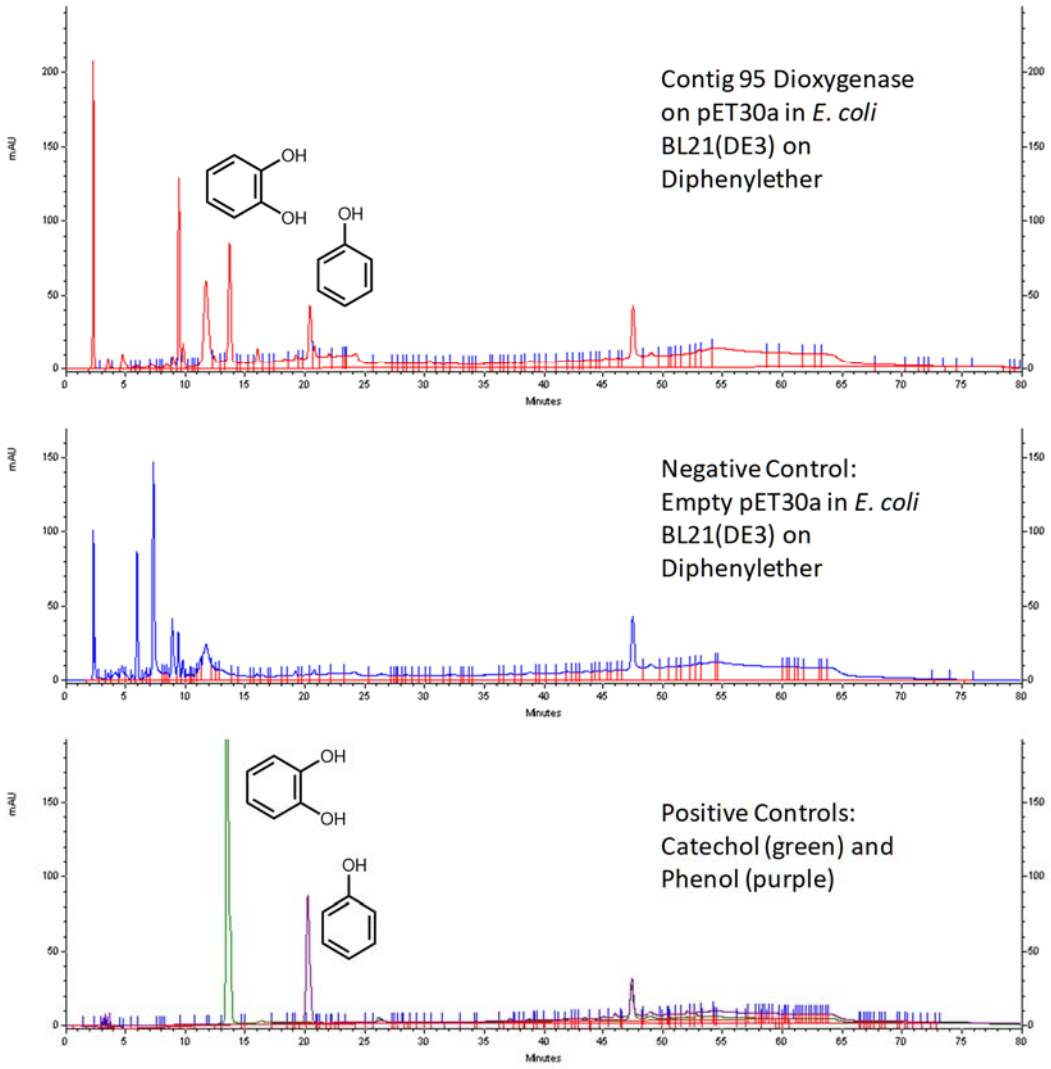

Figure 19. HPLC analysis of culture supernatants compared to catechol and phenol controls.

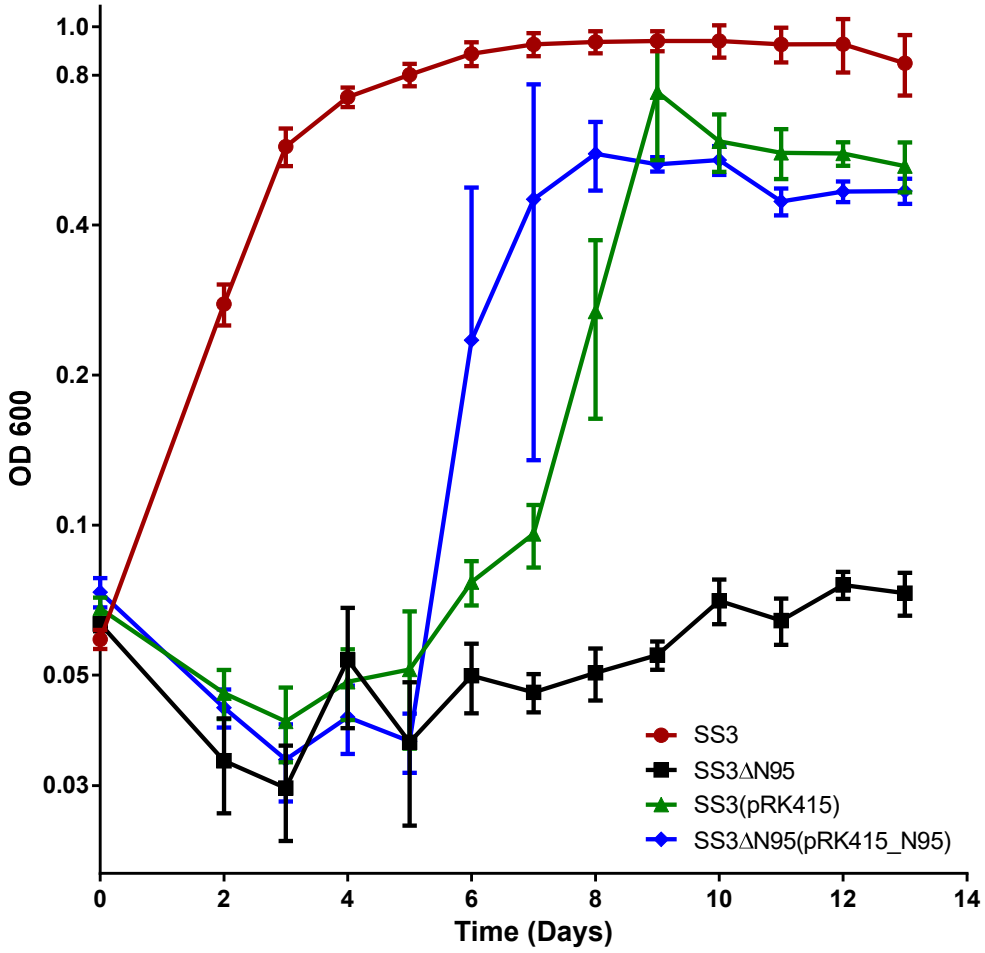

Figure 20. Growth of *Sphingobium yanoikuyae* SS3 and an N95 knockout mutant on diphenylether.

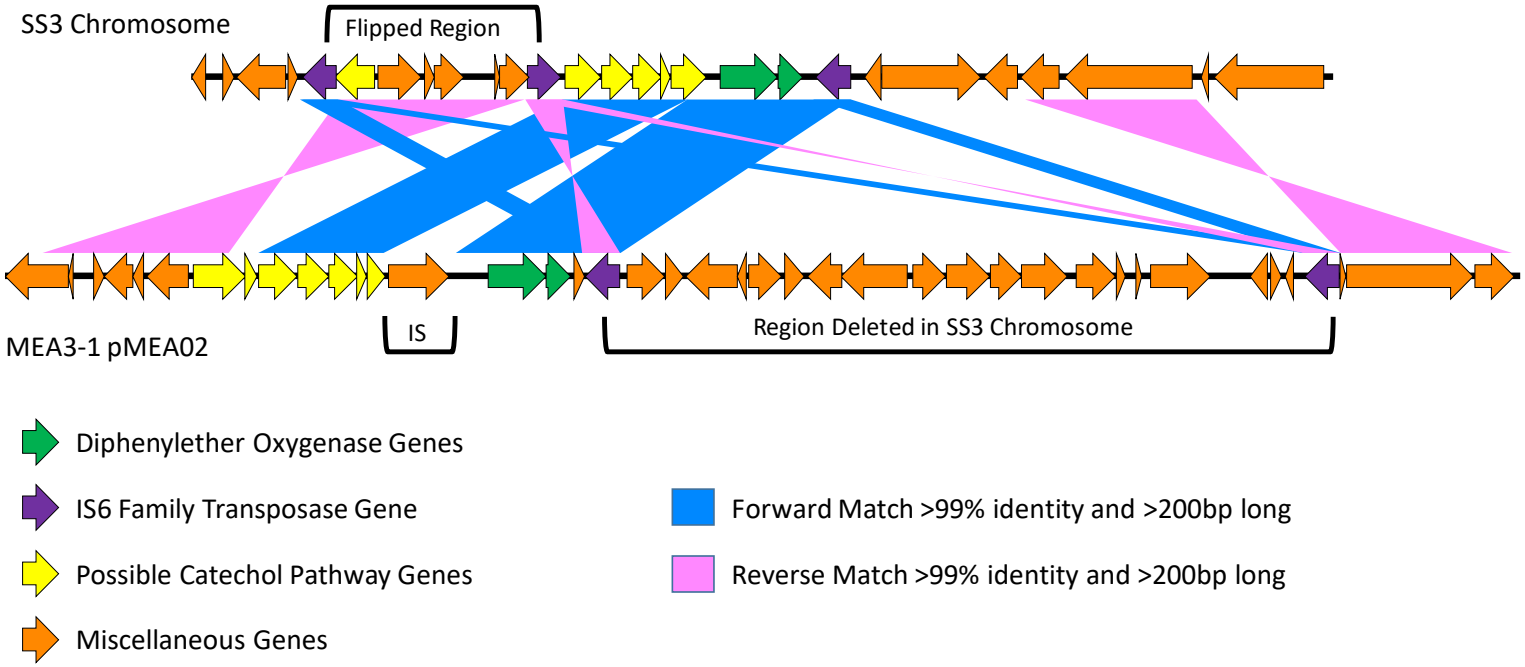

Figure 21. Comparison of the *S. yanoikuyae* SS3 genome region containing the genes for diphenylether dioxygenase to the analogous genome region in *Sphingobium* sp. strain MEA3-01

References

1. Shen H, Guan R, Ding G, Chen Q, Lou X, Chen Z, Zhang L, Xing M, Han J, Wu Y. 2017. Polychlorinated dibenzo-*p*-dioxins/furans (PCDD/Fs) and polychlorinated biphenyls (PCBs) in Zhejiang foods (2006-2015): Market basket and polluted areas. Sci Total Environ 574:120-127.
2. Brown SM, Heguy A, Zappile P, Chen H, Goradia A, Wang Y, Hao Y, Roy NK, Vitale K, Chambers RC, Wirgin I. 2017. A dramatic difference in global gene expression between TCDD-treated Atlantic tomcod larvae from the resistant Hudson River and a nearby sensitive population. Genome Biol Evol 9:2251-2264.
3. Quadrini JD, Ku W, Connolly JP, Chiavelli DA, Israelsson PH. 2015. Fingerprinting 2,3,7,8-tetrachlorodibenzodioxin contamination within the lower Passaic River. Environ Toxicol Chem 34:1485–1498.
4. Khairy MA, Noonan GO, Lohmann R. 2019. Uptake of hydrophobic organic compounds, including organochlorine pesticides, polybrominated diphenyl ethers, and perfluoroalkyl acids in fish and blue crabs of the lower Passaic River, New Jersey, USA. Environ Toxicol Chem 38:872-882.
5. Walker WJ, McNutt RP, Maslanka CK. 1999. The potential contribution of urban runoff to surface sediments of the Passaic River: sources and chemical characteristics. Chemosphere 38:363-377.
6. Zani C, Magoni M, Speziani F, Leonardi L, Orizio G, Scarcella C, Gaia A, Donato F. 2019. Polychlorinated biphenyl serum levels, thyroid hormones and endocrine and metabolic diseases in people living in a highly polluted area in

North Italy: a population-based study. Heliyon. 5:e01870. doi:10.1016/j.heliyon.2019.e01870.

7. Garrido-Sanz D, Manzano J, Martín M, Redondo-Nieto M, Rivilla R. 2018. Metagenomic analysis of a biphenyl-degrading soil bacterial consortium reveals the metabolic roles of specific populations. Front Microbiol. 9:232. doi:10.3389/fmicb.2018.00232.
8. Nowak P, Kucharska K, Kamiński M. 2019. Ecological and health effects of lubricant oils emitted into the environment. Int J Environ Res Public Health. 16:doi:10.3390/ijerph16163002.
9. Pham TT, Sylvestre M. 2013. Has the bacterial biphenyl catabolic pathway evolved primarily to degrade biphenyl? The diphenylmethane case. J Bacteriol 195:3563-3574.
10. Focht DD, Alexander M. 1970. Bacterial degradation of diphenylmethane, a DDT model substrate. Appl Microbiol 20:608-611.
11. Yang J, Zhao Y, Li M, Du M, Li X, Li Y. 2019. A review of a class of emerging contaminants: the classification, distribution, intensity of consumption, synthesis routes, environmental effects and expectation of pollution abatement to organophosphate flame retardants (OPFRs). Int J Mol Sci. 20:2874. doi:10.3390/ijms20122874.
12. Pfeifer F, Trüper HG, Klein J, Schacht S. 1993. Degradation of diphenylether by *Pseudomonas cepacia* Et4: enzymatic release of phenol from 2,3-dihydroxydiphenylether. Arch Microbiol 159:323-329.

13. Schmidt S, Wittich RM, Fortnagel P, Erdmann D, Francke W. 1992. Metabolism of 3-methyldiphenyl ether by *Sphingomonas* sp. SS31. FEMS Microbiol Lett 96:253-258.

14. Schecter A, Päpke O, Tung KC, Staskal D, Birnbaum L. 2004. Polybrominated diphenyl ethers contamination of United States food. Environ Sci Technol 38:5306-5311.

15. Hahn ME, Karchner SI, Shapiro MA, Perera SA. 1997. Molecular evolution of two vertebrate aryl hydrocarbon (dioxin) receptors (AHR1 and AHR2) and the PAS family. Proc Natl Acad Sci U S A 94:13743-13748.

16. Phale PS, Basu A, Majhi PD, Deveryshetty J, Vamsee-Krishna C, Shrivastava R. 2007. Metabolic diversity in bacterial degradation of aromatic compounds. OMICS 11:252-279.

17. Seo JS, Keum YS, Li QX. 2009. Bacterial degradation of aromatic compounds. Int J Environ Res Public Health 6:278-309.

18. Karigar CS, Rao SS. 2011. Role of microbial enzymes in the bioremediation of pollutants: a review. Enzyme Res. 2011:doi:10.4061/2011/805187.

19. Sidhu C, Solanki V, Pinnaka AK, Thakur KG. 2019. Structure elucidation and biochemical characterization of environmentally relevant novel extradiol dioxygenases discovered by a functional metagenomics approach. mSystems. 4:e00316-19. doi:10.1128/mSystems.00316-19.

20. Arora PK, Srivastava A, Singh VP. 2010. Application of monooxygenases in dehalogenation, desulphurization, denitrification and hydroxylation of aromatic compounds. J Bioremediat Biodegrad. 01:doi:10.4172/2155-6199.1000112.

21. van Berkel WJ, Kamerbeek NM, Fraaije MW. 2006. Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. J Biotechnol 124:670-689.

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

23. Imbeault NY, Powlowski JB, Colbert CL, Bolin JT, Eltis LD. 2000. Steady-state kinetic characterization and crystallization of a polychlorinated biphenyl-transforming dioxygenase. J Biol Chem 275:12430-12437.

24. Kauppi B, Lee K, Carredano E, Parales RE, Gibson DT, Eklund H, Ramaswamy S. 1998. Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene 1,2-dioxygenase. Structure 6:571-586.

25. Harayama S, Rekik M. 1989. Bacterial aromatic ring-cleavage enzymes are classified into two different gene families. J Biol Chem 264:15328-15333.

26. Boyd DR, Sharma ND, Allen CC. 2001. Aromatic dioxygenases: molecular biocatalysis and applications. Curr Opin Biotechnol 12:564-573.

27. Baratto MC, Lipscomb DA, Larkin MJ, Basosi R, Allen CCR, Pogni R. 2019. Spectroscopic characterisation of the naphthalene dioxygenase from *Rhodococcus* sp. strain NCIMB12038. Int J Mol Sci. 20:doi:10.3390/ijms20143402.

28. Lipscomb JD. 2008. Mechanism of extradiol aromatic ring-cleaving dioxygenases. Curr Opin Struct Biol 18:644-649.

29. Gassner GT, Ludwig ML, Gatti DL, Correll CC, Ballou DP. 1995. Structure and mechanism of the iron-sulfur flavoprotein phthalate dioxygenase reductase. FASEB J 9:1411-1418.

30. Batie C. 1992. Phthalate dioxygenase reductase and related flavin-iron-sulfur-containing electron transferase. Chemistry and Biochemistry of Flavoenzynes:543-566.

31. Nishi A, Tominaga K, Furukawa K. 2000. A 90-kilobase conjugative chromosomal element coding for biphenyl and salicylate catabolism in *Pseudomonas putida* KF715. J Bacteriol 182:1949-1955.

32. Ishigooka H, Yoshida Y, Omori T, Minoda Y. 1986. Enzymatic dioxygenation of biphenyl-2,3-diol and 3-isopropylcatechol. Agric Biol Chem 50:1045-1046.

33. Furukawa K, Arimura N. 1987. Purification and properties of 2,3-dihydroxybiphenyl dioxygenase from polychlorinated biphenyl-degrading *Pseudomonas pseudoalcaligenes* and *Pseudomonas aeruginosa* carrying the cloned *bphC* gene. J Bacteriol 169:924-927.

34. Taira K, Hirose J, Hayashida S, Furukawa K. 1992. Analysis of *bph* operon from the polychlorinated biphenyl-degrading strain of *Pseudomonas pseudoalcaligenes* KF707. J Biol Chem 267:4844-4853.

35. Omori T, Sugimura K, Ishigooka H, Minoda Y. 1986. Purification and some properties of 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid hydrolyzing enzyme from *Pseudomonas cruciviae* S93 B1 involved in the degradation of biphenyl. Agric Biol Chem 50:931-937.

36. Hayase N, Taira K, Furukawa K. 1990. *Pseudomonas putida* KF715 *bphABCD* operon encoding biphenyl and polychlorinated biphenyl degradation: cloning, analysis, and expression in soil bacteria. J Bacteriol 172:1160-1164.

37. Misawa N, Shindo K, Takahashi H, Suenaga H, Iguchi K, Okazaki H, Harayamab S, Furukawa K. 2002. Hydroxylation of various molecules including heterocyclic aromatics using recombinant *Escherichia coli* cells expressing modified biphenyl dioxygenase genes. Tetrahedron 58:9605-9612.

38. Gómez-Gil L, Kumar P, Barriault D, Bolin JT, Sylvestre M, Eltis LD. 2007. Characterization of biphenyl dioxygenase of *Pandoraea pnomenusa* B-356 as a potent polychlorinated biphenyl-degrading enzyme. J Bacteriol 189:5705-5715.

39. Barriault D, Sylvestre M. 2004. Evolution of the biphenyl dioxygenase BphA from *Burkholderia xenovorans* LB400 by random mutagenesis of multiple sites in region III. J Biol Chem 279:47480-47488.

40. Zielinski M, Kahl S, Standfuss-Gabisch C, Cámara B, Seeger M, Hofer B. 2006. Generation of novel-substrate-accepting biphenyl dioxygenases through segmental random mutagenesis and identification of residues involved in enzyme specificity. Appl Environ Microbiol 72:2191-2199.

41. Ní Chadhain SM, Norman RS, Pesce KV, Kukor JJ, Zylstra GJ. 2006. Microbial dioxygenase gene population shifts during polycyclic aromatic hydrocarbon biodegradation. Appl Environ Microbiol 72:4078-4087.

42. Nojiri H, Shintani M, Omori T. 2004. Divergence of mobile genetic elements involved in the distribution of xenobiotic-catabolic capacity. Appl Microbiol Biotechnol 64:154-174.

43. Chakrabarty AM. 1974. Dissociation of a degradative plasmid aggregate in *Pseudomonas*. J Bacteriol 118:815-820.

44. Dennis JJ, Zylstra GJ. 2004. Complete sequence and genetic organization of pDTG1, the 83 kilobase naphthalene degradation plasmid from *Pseudomonas putida* strain NCIB 9816-4. J Mol Biol 341:753-768.

45. Ilori MO, Picardal FW, Aramayo R, Adebusoye SA, Obayori OS, Benedik MJ. 2015. Catabolic plasmid specifying polychlorinated biphenyl degradation in *Cupriavidus* sp. strain SK-4: mobilization and expression in a pseudomonad. J Basic Microbiol 55:338-345.

46. Kosono S, Maeda M, Fuji F, Arai H, Kudo T. 1997. Three of the seven *bphC* genes of *Rhodococcus erythropolis* TA421, isolated from a termite ecosystem, are located on an indigenous plasmid associated with biphenyl degradation. Appl Environ Microbiol 63:3282-3285.

47. Sota M, Endo M, Nitta K, Kawasaki H, Tsuda M. 2002. Characterization of a class II defective transposon carrying two haloacetate dehalogenase genes from *Delftia acidovorans* plasmid pUO1. Appl Environ Microbiol 68:2307-2315.

48. Nakatsu C, Ng J, Singh R, Straus N, Wyndham C. 1991. Chlorobenzoate catabolic transposon Tn*5271* is a composite class I element with flanking class II insertion sequences. Proc Natl Acad Sci U S A 88:8312-8316.

49. Springael D, Ryngaert A, Merlin C, Toussaint A, Mergeay M. 2001. Occurrence of Tn*4371*-related mobile elements and sequences in (chloro)biphenyl-degrading bacteria. Appl Environ Microbiol 67:42-50.

50. Carraro N, Burrus V. 2014. Biology of three ICE families: SXT/R391, ICEBs1, and ICESt1/ICESt3. Microbiol Spectr. 2:doi:10.1128/microbiolspec.MDNA3-0008-2014.

51. Wawrzyniak P, Płucienniczak G, Bartosik D. 2017. The different faces of rolling-circle replication and its multifunctional initiator proteins. Front Microbiol. 8:2353. doi:10.3389/fmicb.2017.02353.

52. Guglielmini J, Quintais L, Garcillán-Barcia MP, de la Cruz F, Rocha EP. 2011. The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. PLoS Genet. 7:e1002222. doi:10.1371/journal.pgen.1002222.

53. Burrus V, Pavlovic G, Decaris B, Guedon G. 2002. Conjugative transposons: the tip of the iceberg. Mol Microbiol 46:601-10.

54. Guglielmini J, de la Cruz F, Rocha EP. 2013. Evolution of conjugation and type IV secretion systems. Mol Biol Evol 30:315-331.

55. Johnson CM, Grossman AD. 2015. Integrative and conjugative elements (ICEs): what they do and how they work. Annu Rev Genet 49:577-601.

56. Bellanger X, Payot S, Leblond-Bourget N, Guédon G. 2014. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. FEMS Microbiol Rev 38:720-60.

57. Franke AE, Clewell DB. 1981. Evidence for a chromosome-borne resistance transposon (Tn*916*) in *Streptococcus faecalis* that is capable of "conjugal" transfer in the absence of a conjugative plasmid. J Bacteriol 145:494-502.

58. Wozniak RA, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, Garriss G, Déry C, Burrus V, Waldor MK. 2009. Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. PLoS Genet. 5:e1000786. doi:10.1371/journal.pgen.1000786.

59. Sentchilo V, Czechowska K, Pradervand N, Minoia M, Miyazaki R, van der Meer JR. 2009. Intracellular excision and reintegration dynamics of the ICE*clc* genomic island of *Pseudomonas knackmussii* sp. strain B13. Mol Microbiol 72:1293-1306.

60. Hochhut B, Jahreis K, Lengeler JW, Schmid K. 1997. CTn*scr94*, a conjugative transposon found in enterobacteria. J Bacteriol 179:2097-2102.

61. Ramsay JP, Sullivan JT, Stuart GS, Lamont IL, Ronson CW. 2006. Excision and transfer of the *Mesorhizobium loti* R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS. Mol Microbiol 62:723-734.

62. Carter MQ, Chen J, Lory S. 2010. The *Pseudomonas aeruginosa* pathogenicity island PAPI-1 is transferred via a novel type IV pilus. J Bacteriol 192:3249-3258.

63. van der Meer JR, Ravatn R, Sentchilo V. 2001. The *clc* element of *Pseudomonas* sp. strain B13 and other mobile degradative elements employing phage-like integrases. Arch Microbiol 175:79-85.

64. Miyazaki R, van der Meer JR. 2013. A new large-DNA-fragment delivery system based on integrase activity from an integrative and conjugative element. Appl Environ Microbiol 79:4440-4447.

65. Sentchilo V, Zehnder AJ, van der Meer JR. 2003. Characterization of two alternative promoters for integrase expression in the *clc* genomic island of *Pseudomonas* sp. strain B13. Mol Microbiol 49:93-104.

66. Hirose J, Fujihara H, Watanabe T, Kimura N, Suenaga H, Futagami T, Goto M, Suyama A, Furukawa K. 2019. Biphenyl/PCB degrading *bph* genes of ten bacterial strains isolated from biphenyl-contaminated soil in Kitakyushu, Japan:

comparative and dynamic features as integrative conjugative elements (ICEs). Genes (Basel). 10:404. doi:10.3390/genes10050404.

67. Ohtsubo Y, Ishibashi Y, Naganawa H, Hirokawa S, Atobe S, Nagata Y, Tsuda M. 2012. Conjugal transfer of polychlorinated biphenyl/biphenyl degradation genes in *Acidovorax* sp. strain KKS102, which are located on an integrative and conjugative element. J Bacteriol 194:4237-4248.

68. Hong Q, Dong XJ, He LJ, Jiang X, Li SP. 2009. Isolation of a biphenyl-degrading bacterium, *Achromobacter* sp BP3, and cloning of the *bph* gene cluster. Int Biodeter Biodegr 63:365-370.

69. Wang J, Liu Y, Du K, Xu S, Wang Y, Krupovic M, Chen X. 2018. A novel family of tyrosine integrases encoded by the temperate pleolipovirus SNJ2. Nucleic Acids Res 46:2521-2536.

70. Hirano N, Muroi T, Takahashi H, Haruki M. 2011. Site-specific recombinases as tools for heterologous gene integration. Appl Microbiol Biotechnol 92:227-239.

71. Ravatn R, Studer S, Zehnder AJ, van der Meer JR. 1998. Int-B13, an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase *clc* element of *Pseudomonas* sp. strain B13. J Bacteriol 180:5505-5514.

72. Hou YM. 1999. Transfer RNAs and pathogenicity islands. Trends Biochem Sci 24:295-298.

73. Kiewitz C, Larbig K, Klockgether J, Weinel C, Tümmler B. 2000. Monitoring genome evolution ex vivo: reversible chromosomal integration of a 106 kb

plasmid at two tRNA(Lys) gene loci in sequential *Pseudomonas aeruginosa* airway isolates. Microbiology 146:2365-2373.

74. Dimopoulou ID, Russell JE, Mohd-Zain Z, Herbert R, Crook DW. 2002. Site-specific recombination with the chromosomal tRNA(Leu) gene by the large conjugative *Haemophilus* resistance plasmid. Antimicrob Agents Chemother 46:1602-1603.

75. Carraro N, Poulin D, Burrus V. 2015. Replication and active partition of integrative and conjugative elements (ICEs) of the SXT/R391 family: The line between ICEs and conjugative plasmids is getting thinner. PLoS Genet. 11:e1005298. doi:10.1371/journal.pgen.1005298.

76. Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, Cascales E. 2005. Biogenesis, architecture, and function of bacterial type IV secretion systems. Annu Rev Microbiol 59:451-485.

77. Chandran Darbari V, Waksman G. 2015. Structural biology of bacterial type IV secretion systems. Annu Rev Biochem 84:603-629.

78. Wozniak RA, Waldor MK. 2010. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. Nat Rev Microbiol 8:552-563.

79. Alvarez-Martinez CE, Christie PJ. 2009. Biological diversity of prokaryotic type IV secretion systems. Microbiol Mol Biol Rev 73:775-808.

80. Lee CA, Babic A, Grossman AD. 2010. Autonomous plasmid-like replication of a conjugative transposon. Mol Microbiol 75:268-279.

81. Delavat F, Miyazaki R, Carraro N, Pradervand N, van der Meer JR. 2017. The hidden life of integrative and conjugative elements. FEMS Microbiol Rev 41:512-537.

82. Carraro N, Burrus V. 2015. The dualistic nature of integrative and conjugative elements. Mob Genet Elements 5:98-102.

83. Gesell M, Hammer E, Specht M, Francke W, Schauer F. 2001. Biotransformation of biphenyl by *Paecilomyces lilacinus* and characterization of ring cleavage products. Appl Environ Microbiol 67:1551-1557.

84. Focht DD, Alexander M. 1970. DDT metabolites and analogs: ring fission by *Hydrogenomonas*. Science 170:91-92.

85. Environmental Protection Agency W, DC. Office of Toxic Substances. 1995. Biodegradation and bioaccumulation data of existing chemicals based on the Cscl Japan, with cover letter dated 01/19/95.

86. Erickson BD, Mondello FJ. 1992. Nucleotide sequencing and transcriptional mapping of the genes encoding biphenyl dioxygenase, a multicomponent polychlorinated-biphenyl-degrading enzyme in *Pseudomonas* strain LB400. J Bacteriol 174:2903-2912.

87. Springael D, Kreps S, Mergeay M. 1993. Identification of a catabolic transposon, Tn*4371*, carrying biphenyl and 4-chlorobiphenyl degradation genes in *Alcaligenes eutrophus* A5. J Bacteriol 175:1674-1681.

88. Warren R, Hsiao WW, Kudo H, Myhre M, Dosanjh M, Petrescu A, Kobayashi H, Shimizu S, Miyauchi K, Masai E, Yang G, Stott JM, Schein JE, Shin H, Khattra J, Smailus D, Butterfield YS, Siddiqui A, Holt R, Marra MA, Jones SJ, Mohn WW,

Brinkman FS, Fukuda M, Davies J, Eltis LD. 2004. Functional characterization of a catabolic plasmid from polychlorinated- biphenyl-degrading *Rhodococcus* sp. strain RHA1. J Bacteriol 186:7783-7795.

89. Iwasaki T, Miyauchi K, Masai E, Fukuda M. 2006. Multiple-subunit genes of the aromatic-ring-hydroxylating dioxygenase play an active role in biphenyl and polychlorinated biphenyl degradation in *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol 72:5396-5402.

90. Hooper SW, Dockendorff TC, Sayler GS. 1989. Characteristics and restriction analysis of the 4-chlorobiphenyl catabolic plasmid, pSS50. Appl Environ Microbiol 55:1286-1288.

91. Yuan Z, Courtenay S, Chambers RC, Wirgin I. 2006. Evidence of spatially extensive resistance to PCBs in an anadromous fish of the Hudson River. Environ Health Perspect 114:77-84.

92. Stanier RY, Palleroni NJ, Doudoroff M. 1966. The aerobic pseudomonads: a taxonomic study. J Gen Microbiol 43:159-271.

93. Johnson JL. 1994. Similarity analysis of rRNAs. . *In* Gerhardt P, Murray RGE, Krieg NR, Wood WA (ed), Methods for General and Molecular Bacteriology. American Society for Microbiology, Washington, DC.

94. Iwai S, Chai B, Sul WJ, Cole JR, Hashsham SA, Tiedje JM. 2010. Gene-targeted-metagenomics reveals extensive diversity of aromatic dioxygenase genes in the environment. ISME J 4:279-285.

95. Parette R, Velinsky DJ, Pearson WN. 2018. Reconstruction of historical 2,3,7,8-tetrachlorodibenzo-*p*-dioxin discharges from a former pesticide manufacturing plant to the Lower Passaic River. Chemosphere 212:1125-1132.

96. Fujihara H, Yoshida H, Matsunaga T, Goto M, Furukawa K. 2006. Cross-regulation of biphenyl- and salicylate-catabolic genes by two regulatory systems in *Pseudomonas pseudoalcaligenes* KF707. J Bacteriol 188:4690-4697.

97. Tsuda M, Tan HM, Nishi A, Furukawa K. 1999. Mobile catabolic genes in bacteria. J Biosci Bioeng 87:401-410.

98. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.

99. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455-77.

100. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-2069.

101. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821-829.

102. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia F, Yoo H, Stevens RL. 2017.

Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Res 45:D535-D542.

103. Sambrook J, Fritsch EF, Maniatis T. 1987. Molecular cloning, a laboratory manual, 2ed. ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

104. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589-595.

105. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup GPDP. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079.

106. Milne I, Stephen G, Bayer M, Cock PJ, Pritchard L, Cardle L, Shaw PD, Marshall D. 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform 14:193-202.

107. Keen NT, Tamaki S, Kobayashi D, Trollinger D. 1988. Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. Gene 70:191-197.

108. Dennis JJ, Zylstra GJ. 1998. Improved antibiotic-resistance cassettes through restriction site elimination using *Pfu* DNA polymerase PCR. Biotechniques 25:772-776.

109. Ditta G, Stanfield S, Corbin D, Helinski DR. 1980. Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc Natl Acad Sci U S A 77:7347-7351.

110. Figurski DH, Helinski DR. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc Natl Acad Sci U S A 76:1648-1652.

111. Vetriani C, Tran HV, Kerkhof LJ. 2003. Fingerprinting microbial assemblages from the oxic/anoxic chemocline of the Black Sea. Appl Environ Microbiol 69:6481-6488.

112. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27:1009-1010.

113. Semsey S, Blaha B, Köles K, Orosz L, Papp PP. 2002. Site-specific integrative elements of rhizobiophage *16-3* can integrate into proline tRNA (CGG) genes in different bacterial genera. J Bacteriol 184:177-182.

114. Obi CC, Vayla S, de Gannes V, Berres ME, Walker J, Pavelec D, Hyman J, Hickey WJ. 2018. The integrative conjugative element *clc* (ICEclc) of *Pseudomonas aeruginosa* JB2. Front Microbiol. 9:1532. doi:10.3389/fmicb.2018.01532.

115. De Palmenaer D, Siguier P, Mahillon J. 2008. IS*4* family goes genomic. BMC Evol Biol. 8:18. doi:10.1186/1471-2148-8-18.

116. Rahman F, Langford KH, Scrimshaw MD, Lester JN. 2001. Polybrominated diphenyl ether (PBDE) flame retardants. Sci Total Environ 275:1-17.

117. Schmidt S, Wittich RM, Fortnagel P, Erdmann D, Francke W. 1992. Metabolism of 3-methyldiphenyl ether by Sphingomonas sp. SS31. FEMS Microbiol Lett 75:253-8.

118. Kim YM, Nam IH, Murugesan K, Schmidt S, Crowley DE, Chang YS. 2007. Biodegradation of diphenyl ether and transformation of selected brominated congeners by *Sphingomonas* sp. PH-07. Appl Microbiol Biotechnol 77:187-94.

119. Schmidt S, Wittich RM, Erdmann D, Wilkes H, Francke W, Fortnagel P. 1992. Biodegradation of diphenyl ether and its monohalogenated derivatives by *Sphingomonas* sp. strain SS3. Appl Environ Microbiol 58:2744-50.

120. Pfeifer F, Schacht S, Klein J, Trüper H. 1989. Degradation of diphenylether by *Pseudomonas cepacia*. Arch Microbiol 152:515-519.

121. Wang S, Bai N, Wang B, Feng Z, Hutchins WC, Yang CH, Zhao Y. 2015. Characterization of the molecular degradation mechanism of diphenyl ethers by *Cupriavidus* sp. WS. Environ Sci Pollut Res Int doi:10.1007/s11356-015-4854-3.

122. Takase I, Omori T, Minoda Y. 1986. Microbial degradation products from biphenyl-related compounds. Agric Biol Chem 50:681-686

123. Hiraishi A. 2003. Biodiversity of dioxin-degrading microorganisms and potential utilization in bioremediation. Microbes and Environments 18:105-125.

124. Liaw HJ, Srinivasan VR. 1989. Molecular cloning and expression of an *Erwinia* sp. gene encoding diphenyl ether cleavage in *Escherichia coli*. Appl Environ Microbiol 55:2220-5.

125. Cai S, Chen LW, Ai YC, Qiu JG, Wang CH, Shi C, He J, Cai TM. 2017. Degradation of diphenyl ether in *Sphingobium phenoxybenzoativorans* SC_3 Is initiated by a novel ring cleavage dioxygenase. Appl Environ Microbiol 83.

126. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 34:i142-i150.

127. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595.

128. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094-3100.

129. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-60.

130. Basta T, Keck A, Klein J, Stolz A. 2004. Detection and characterization of conjugative degradative plasmids in xenobiotic-degrading *Sphingomonas* strains. J Bacteriol 186:3862-72.

131. Dong W, Chen Q, Hou Y, Li S, Zhuang K, Huang F, Zhou J, Li Z, Wang J, Fu L, Zhang Z, Huang Y, Wang F, Cui Z. 2015. Metabolic pathway involved in 2-methyl-6-ethylaniline degradation by *Sphingobium* sp. strain MEA3-1 and cloning of the novel flavin-dependent monooxygenase system *meaBA*. Appl Environ Microbiol 81:8254-64.

132. Hou Y, Dong W, Wang F, Li J, Shen W, Li Y, Cui Z. 2014. Degradation of acetochlor by a bacterial consortium of *Rhodococcus* sp.T3-1, *Delftia* sp.T3-6 and *Sphingobium* sp.MEA3-1. Lett Appl Microbiol 59:35-42.

133. Armengaud J, Happe B, Timmis KN. 1998. Genetic analysis of dioxin dioxygenase of *Sphingomonas* sp. strain RW1: catabolic genes dispersed on the genome. J Bacteriol 180:3954-66.

134. Ni Chadhain SM, Moritz EM, Kim E, Zylstra GJ. 2007. Identification, cloning, and characterization of a multicomponent biphenyl dioxygenase from *Sphingobium yanoikuyae* B1. J Ind Microbiol Biotechnol 34:605-13.

135. Zylstra GJ, Kim E. 1997. Aromatic hydrocarbon degradation by *Sphingomonas yanoikuyae* B1. J Ind Microbiol Biotechnol 19:408-14.